https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

Soil water forecasting for Australian dryland agriculture

Qazi Muqeet Amir[1], Floris F. Van Ogtrop[1], and Thomas F.A. Bishop[1,2]

[1]Precision Agriculture, Hydrology & Geoinformation Science Laboratory; School of Life Environmental Sciences; Sydney Institute of Agriculture; The University of Sydney; Sydney, NSW 2015
[2]Sydney Informatics Hub; The University of Sydney; Sydney, NSW 2042

Correspondence: Qazi Muqeet Amir (muqeet.amir@sydney.edu.au)

Abstract. Soil water availability is a critical constraint to agricultural productivity. While soil water forecasting has previously been conducted in the literature for irrigated fields and surface soil (0-10 cm), there is a limited understanding of subsurface soil water in dryland paddocks, particularly in Australia. Due to rooting depth typically extending below 10 cm, subsurface soil water forecasts in cropping systems, for example, enable decision-making relating to sowing, fertiliser use, and seasonal yield potential. This study aimed to identify the best performing methods, and the underlying variables that affect accuracy when forecasting subsurface soil water (30-100 cm). The methods appraised in this study are Random Forest, XGBoost, Multi-Layer Perceptron (MLP), Long Short-Term Memory (LSTM), and an Encoder-decoder LSTM (ENCDEC). Various in situ and remotely sensed meteorological data were used to forecast soil water up to multiple months ahead, at 54 probe sites across Australia.

We find all models can perform with relatively similar accuracy for sub-monthly forecasts. For longer lead times, MLP and XGBoost performed better for sites with uniform rainfall throughout the year, and the memory-based models, LSTM and ENCDEC, better at sites with seasonally-dominant rainfall. Furthermore, we determine an upper limit to model utility, as accuracy can become poor many months ahead, and can be replaced by a historic average as an estimate. We find our models to be 'useful' up to 4 months ahead on average, after which accuracy is too low, or a historic average outperforms a model. We find rainfall intensity and seasonality, and model choice, to be key drivers of forecast accuracy at a new site. This study highlights the capability of elementary forms of each ensemble and deep learning model to provide sufficiently accurate soil water forecasts, particularly in the absence of a rainfall forecast feature.

1 Introduction

Consistent fluctuation of droughts and floods have been observed across Australia over the past century (Nicholls et al., 1997). These fluctuations have been attributed to the influence of large-scale ocean-atmosphere modes of variability such as the El-Nina Southern Oscillation, Indian Ocean Dipole, and Southern Annular Mode in Australia (Freund et al., 2017). As a result, the continent experiences a frequent occurrence of extreme temperatures and variable inter-annual rainfall (Jakob and Walland, 2016; Scanlon and Doncon, 2020). In sowing months, the Australian Bureau of Meteorology classify the inter-annual rainfall as moderately high – highly variable in most grain growing areas of Australia (Meteorology, b). Most of Australia's agriculture is produced under dryland conditions and so is contingent on rainfall for water input. The variability in annual rainfall creates a

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

complex landscape for growers to make decisions relating to sowing time, field trafficability, fertiliser inputs, frost, and yield. For example, Flohr et al. (2021) report inter-annual Autumn rainfall variability contributing to grain yield losses of 1–7% per week of delay past the optimal establishment time. Decisions such as this can be informed by reliable daily to seasonal paddock-specific soil water forecasts.

While rainfall forecasts can be informative, they do not account for infiltration, effect of land cover, and soil-water dynamics. Soil water forecasts can inform sowing time, which is crucial for dryland crops as germination and growth can only occur once there is sufficient water in the topsoil (Unkovich, 2010). Additionally, forecasts can give growers insight into soil water availability to determine optimum sowing dates for maximising yield. Studies show that wheat and maize yield are significantly influenced by deep soil water before sowing as the water stored during fallow periods provide early crop growth support (Fang et al., 2021; Zhang et al., 2021). In the instance of wheat, sowing too early risks frost damage, while sowing too late can reduce yield, as flowering could occur in suboptimal hot and dry conditions (Dadrasi et al., 2024). Since soil water influences the heat storage capacity of a soil, forecasts can be utilised to anticipate the degree of frost damage to grains expected from a potential frost event; and can inform crop variety selection to minimise yield loss (GRDC, 2016). Furthermore, Hochman and Horan (2018) estimate half of the average yield gap of Australian wheat can be attributed to low nitrogen fertiliser input. Soil water forecasts can be informative in determining fertiliser rates and timing to optimise absorption and utilisation of nutrients in the soil to meet the seasonal yield potential (Dong et al., 2023; Liu et al., 2023).

The current studies on soil water forecasting have been conducted across various countries, methods, depths, and lead times. It is common for research with this focus to appraise multiple machine learning methods as each method utilises a different mechanism and may suit the target domain better than another. For example, Brinkhoff et al. (2019), compared a variety of conventional machine learning methods (Lasso Regression, Linear Regression, Support Vector Machine, Decision Tree, and Random Forest) in forecasting soil water at five irrigated fields at a site near Whitton, Australia. Each ML method employed a unique mechanism, and they found different models perform best in each field. However, in overall performance, Random Forest performed the best. The use of Artificial Neural Networks (ANNs) has becoming increasingly popular for time-series forecasting due to the ability of ANNs to model complex non-linear relationships and be designed with specific model architectures that caters to domain-specific challenges (Casolaro et al., 2023; Egrioglu and Bas, 2024). ANNs allows for more complex feature combinations, identification of hidden feature attributes and correlations, memory persistence, and high generalisability. The most popular ANN in soil water forecasting literature is the Multi-Layer Perceptron (MLP). Despite being the most elementary ANN model, MLPs consistently outperforms conventional machine learning methods in forecasting soil water (Adeyemi et al., 2018; Cai et al., 2019; Dubois et al., 2021; Granata et al., 2023).

Recurrent Neural Networks (RNNs) are a class of ANNs which make use of sequential data and can account for long term dependencies in the input sequence. For example, Adeyemi et al. (2018) compare a MLP to an RNN called Long Short-Term Memory (LSTM) to forecast soil water 1 day ahead for irrigation scheduling at 3 sites across the United Kingdom. They found that LSTM notably outperforms the MLP model due to the "learned past time period" that the LSTM accounts for. LSTMs in particular have been frequently noted to perform well in a soil moisture forecasting application (Datta and Faroughi, 2023). LSTMs are frequently modified to improve certain aspects of the RNN architecture. LSTMs with slight adjustments such as

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

using PCA features and multihead units have been observed to improved variations of the LSTM in many instances (Jiang et al., 2023; Prakash et al., 2018). Various novel LSTM based models have been introduced in the literature too. For example, Ahmed et al. (2021) forecasted surface soil water (2 cm) with high accuracy at lead times up to 30 days using a novel variant of LSTM first presented by Cho et al. (2014), a Gated Recurrent Unit model (GRU), at four dryland sites across the Murray-Darling basin in Australia, and found it to outperform conventional LSTM. The use of encoder-decoder and attention-based LSTM models tend to perform highly. For example, Li et al. (2022a) use an axial attention ConvLSTM to conduct a one day ahead spatial forecast of 0-5 cm soil moisture across the contiguous United States of America. They report improved accuracy over Random Forest and traditional LSTM. Similarly, Li et al. (2022b) develop an encoder-decoder LSTM based on residual learning principles to predict 5 cm depth soil moisture at 13 flux sites across different countries at lead times up to 10 days. They develop an architecture that pass their predictors and a LSTM encoded-decoded forecast of soil moisture at the preceding time steps to the target lead time into an LSTM. They report notable improvement over LSTM models and a slight improvement over a conventional encoder-decoder LSTM; however, their proposed model does underperform in some instances. Li et al. (2022c) propose a novel multi-headed attention based LSTM model that creates information retention across LSTM cells along features and time. This model is tested at 10 sites across the northern hemisphere at 5 cm depth for up to 7 days ahead. They report an improvement over LSTM, conventional attention-based LSTM, and Random Forest in most cases.

To leverage each model's unique mechanics and strengths, feeding the output from one model into another (i.e. stacking) has proven to be successful. For example, Granata et al. (2023) compare Random Forest, Support Vector Machine and MLP individually, and as one stacked model. They found the stacked model performs the best, while individually, the MLP outperforms the other models. Yu et al. (2020) forecasted soil water in irrigated maize fields in China using Convolutional Neural Networks (CNN) and a GRU, with various meteorological variables as features. They found the stacking of both models was most accurate, while individually, the GRU outperformed the CNN. Togneri et al. (2022) forecast daily minimum and maximum soil water up to 2 days ahead across 12 fields in Brazil. They compare Linear Regression, Decision Tree, Random Forest, LSTM, MLP, StemGNN, and a stacked model. They found a stacked model architecture with the outputs from all models as features passed to LightGBM, performed the best. Amongst the individual models, LightGBM performed best and was most stable across every field.

Amongst the current research, soil water forecasting has most often been conducted for irrigated fields (Adeyemi et al., 2018; Brinkhoff et al., 2019; Cai et al., 2019; Chatterjee et al., 2018; Dubois et al., 2021; Granata et al., 2023; Togneri et al., 2022; Yu et al., 2020). In irrigated fields, short-term forecasts are of use to a grower as they serve as a guide for irrigation timing. For this reason, shorter lead times are aimed for without motivation to forecast past one week. Irrigated fields differ from dryland fields in their forecast capabilities as there are regular and controlled inputs of water to the soil in irrigated fields that add short-term seasonality, making them more predictable. Likewise, research conducted in natural systems do not experience the same root-water interactions and land cover changes observed in paddocks, making their findings less applicable to agricultural sites (Datta and Faroughi, 2023; Li et al., 2022a, c). The research conducted using modelled and remotely sensed soil moisture is limited by depth (0-10 cm) and hence are typically smoother as they do not include the effect of root interactions, changing land use, and storage (Ahmed et al., 2021; Li et al., 2022a; Prasad et al., 2018; Wang et al., 2024). Remote sensed soil moisture is of

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

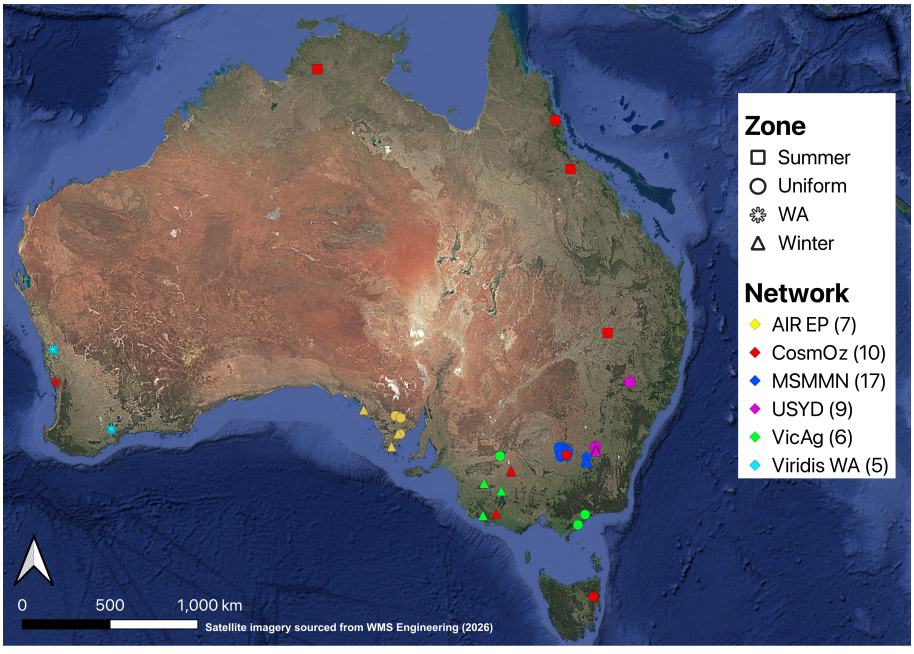

Figure 1. Sites used in this study, by source network and zone (WMS Engineering, 2026)

limited utility in agricultural paddocks as crop rooting depth can be as deep as 100 cm (Fan et al., 2016). Hence, subsurface soil moisture is crucial to understanding crop-water interactions, which probe data makes available. Furthermore, there is a lack of analysis on the underlying meteorologic and soil conditions that affect the accuracy of soil water forecasts at a site, especially as the current extent of studies are typically limited to a paddock or regional scale.

Hence, this study focuses on understand forecasting capabilities for dryland agricultural sites in Australia, with respect to model choice, lead time, and site-specific climatic and soil properties. We aim to provide an understanding of model preference and expected accuracy for future forecasting efforts at Australian dryland agricultural sites. We achieve this by forecasting soil moisture at 54 probes at dryland sites across Australia for various lead times. We compare 5 distinct machine learning models, Random Forest, XGBoost, Multilayer Perceptron, Long Short Term Memory, and an encoder-decoder variant of Long Short Term Memory. We highlight the best performing models for various geographic and rainfall regions and analyse the effect of various site conditions that influence forecast accuracy at a site. We also make recommendations regarding the necessity of a modelled forecast where the accuracy is too poor, or a historic average estimate may suffice.

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

2 Data

2.1 Domain

We use soil moisture probe data at 54 sites from 6 publicly available soil moisture monitoring networks (with the exception of the USYD network) in Australia (Agriculture Victoria; AIR EP, 2019; Hawdon et al., 2014; Smith et al., 2012; Viridis Ag). The USYD soil moisture network is operated by the University of Sydney with sites across the Muttama catchment in New South Wales. Across all networks the probes measure volumetric water content (%VWC) at depths ranging from 10 cm to 110 cm. The sites chosen from each network were selected on the basis of quality and completeness of data. Furthermore, these sites are mostly used for various cropping and grazing practices, however, are all dryland sites with no regular irrigation input. All probe data was generally acquired from the oldest to most recent dates available, unless recording errors were noted in the source data.

2.2 Preprocessing

All probe data was pre-processed by identifying and removing erroneous readings and standardising for consistent temporal resolution and depth. Erroneous readings were identified as outlined by Dorigo et al. (2013). In particular the conditions for spike and break detection algorithms were followed (where the temporal resolution was greater than hourly, the shortest available Savitzky-Golay filter length was used). This was followed by a remote calibration of the soil moisture data at every site to ensure the readings were within a plausible range of values through a method referred to here as PTF scaling. PTF scaling involved scaling the uncalibrated soil moisture measurements within the theoretical lower and upper limits of soil moisture, the crop lower limit and drained upper limit, respectively (Gasch et al., 2017). With inputs of soil fractions, organic carbon, and depth into a Random Forest model trained by Hoskin et al. (2025), the crop lower limit and drained upper limit are calculated. For this study we source the required inputs for the Random Forest model as estimated by the Soil Landscape Grid Australia Malone et al. (2025). The soil moisture readings are then scaled within the PTF estimated limits. Further details on the PTF scaling based calibration for this data can be found in Amir et al. (2026). All readings are then resampled to a daily resolution. Finally, the readings across all available depths are averaged into one of two ranges: 30-60 cm and 60-100 cm.

2.3 Zones

We categorise each site, irrespective of network, on the basis of geography and rainfall seasonality, referred to in this study as ‘zones’. The rainfall seasonality dimension is defined by the annual dominant period of rainfall at a site (from 1991 to 2020), as classified by the Australian Bureau of Meteorology (Meteorology, a). We follow a simplified classification of dominant rainfall periods as follows: ‘Uniform’ rainfall sites experience a uniform distribution of rainfall throughout the year, whereas ‘Summer’ and ‘Winter’ dominant rainfall sites experience rain primarily in summer (December, January, and February) and winter (June, July, August), respectively. The geographic dimension is introduced to represent the extremely variable soil properties across Australia. Particularly the soils of Western Australia (WA), which are distinctly sandy with a low water holding capacity,

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

especially relative to the sites across the south-east, south, and east of Australia (McKenzie et al., 2004; Tilse et al., 2025; Wilford, 2012). In this study we create a distinct zone for the sites in Western Australia, known as 'WA', irrespective of its rainfall pattern (even though all WA sites in this study experience winter dominated rainfall). Hence our zones for this study are: 'Uniform', 'Winter', 'Summer', and 'WA'. These zones group sites based on a combination of soil type and rainfall seasonality. These zones will underlie our analysis to understand the effect of rainfall seasonality and soil type on forecasting accuracy.

2.4 Features

We use rainfall, maximum air temperature, solar radiation, evapotranspiration, EVI, and a long-term soil moisture average to forecast volumetric soil water at each site. Each feature is resampled to a daily average. The maximum air temperature and solar radiation are sourced from the SILO daily gridded product, which features Australia wide climatic data at a 5 km resolution (Jeffrey et al., 2001). Solar radiation and temperature were both included as they affect evapotranspiration and land cover distinctly. The rainfall is calculated as the maximum between the in situ measured rainfall and the SILO modelled estimated rainfall at that site. The maximum is considered to ensure no extreme rainfall event is underestimated and the model is able to account for its maximum possible rainfall intensity. The in situ measured rainfall is susceptible to mechanical errors, as noted in our data cleaning, where many soil moisture spikes did not correspond to in-situ measured rainfall event or there were large periods of missing rainfall data. The evapotranspiration (ET) was sourced from the CMRSET actual evapotranspiration product (Guerschman et al., 2022). The land cover was represented through a green-ness measure, EVI, calculated through Sentinel 2 Red, Blue and Near-Infrared bands. The EVI is natively calculated at a 5-day temporal resolution, but is resampled through a linear interpolation to a daily resolution. The long-term soil moisture average feature was calculated by an average of the soil moisture at the target date (unlike the rest of the features) in preceding years. Any periods of missing data were omitted.

3 Methods

3.1 Models

3.1.1 Random Forest

Random Forest is one of the most popular machine learning algorithms in soil water forecasting literature (Brinkhoff et al., 2019; Dubois et al., 2021; Granata et al., 2023; Togneri et al., 2022). It is an ensemble method that creates multiple decision trees, each constructed from a random subset of the training data (with replacement), known as bootstrapping. A prediction is made by aggregating and averaging the results of the trees. When ensemble models bootstrap their input data and aggregate the results of their models to produce a prediction, it is known as 'bagging'. Bagging creates a separate subset of the training data for every model, each with an identical distribution, which keeps variance across each tree low, creating more generalised models that are less likely to overfit. Methods such as Random Forest that employ bagging generally produce stable results, can handle high dimensionality, and limit overfitting (Tyralis et al., 2019). Random Forest only uses a subset of features in each tree and typically produce a low pairwise correlation between each tree, which further minimises model variance when

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

aggregated (Hastie et al., 2009). In this study we produced separate Random Forest models for each site. The hyperparameters optimised for the Random Forest model were:

- min_sample_leaf: minimum samples required to split an internal node (default = 2);
- n_estimators: Number of trees in the forest (default = 100);
- max_samples: Proportion of all samples to be used to train one tree (bootstrap = True).

The hyperparameter optimisation procedure for each model is detailed under Section 4.3.

3.1.2 XGBoost

Extreme Gradient Boosting (XGBoost) is a gradient-boosted decision tree algorithm. The XGBoost algorithm sequentially creates decision trees (known as 'boosting'), where each tree aims to reduce the residuals from the previous tree. The final prediction is a weighted sum of every tree. XGBoost is expected to perform well on large datasets as it incrementally reduces residuals across trees and weighs samples with greater residuals higher. Boosting algorithms can tend to overfit on highly noisy data as they may sequentially over-emphasize noisy samples (Maclin and Opitz, 1999). However, XGBoost includes hyperparameters that minimising overfitting such as regularisation parameters ('lambda' and 'alpha'), learning rate ('eta') and a tree loss-reduction parameter ('gamma'). In this study we produced separate XGBoost models for each site. The hyperparameters optimised for XGBoost in this study were:

- max_depth: Maximum depth of a tree;
- subsample: Proportion of all samples to consider when growing a tree. (default = 1);
- n_estimator: Number of boosting trees;
- eta: Step size shrinkage factor in tree boosting process (default = 0.3).

XGBoost is similar to Random Forest as both models use additive decision trees to make predictions, and the feature input is tabulated rather than sequentially structured. The key difference is in the learning methods, bagging and boosting; and how they deal with minimising model bias and variance, limiting overfitting, and deal with noisy data. While both learning methods are popular in literature, the comparison of the two can inform which learning method is more suited for modelling soil water behaviour.

3.1.3 Multi-Layer Perceptron

The Multi-Layer Perceptron (MLP) model consists of a collection of nodes, called a 'layer', that is connected to multiple other layers sequentially. The first layer consists of the feature inputs (one node per feature). Each node value is multiplied to a weight then summed together and added to a bias value that is passed through a non-linear activation function, to output a

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

resulting node. This process is repeated to produce a new layer of nodes, each with a new set of randomly generated weights and biases. This process is then repeated over the layer of resulting nodes and is continued on until the layers converge into a single node, which is the model output. The loss is computed by calculating the mean square error (MSE) between the output and observation. The gradient of the loss relative to each weight and bias in the network is calculated so each weight and bias can be adjusted in the direction of minimising loss. The weights and biases are recomputed and minimised over every example until the loss is at an acceptable level. The activation function applied between each layer was ReLU, besides between the final hidden layer and the output layer, where a sigmoid function is used. While the number of fully connected layers is constant across every site, the following network characteristics were optimised for each MLP model:

- Epochs: Number of times the training data was iterated over in training the model;
- Learning rate: Shrinkage factor of change in parameters in the direction of decreasing loss;
- Width: Number of neurons per fully connected layer.

3.1.4 Long Short-Term Memory

The Long Short-Term Memory (LSTM) unit is a type of Recurrent Neural Network (RNN) architecture that takes a series of input values to create a prediction. RNNs are popular in time-series forecasting as they take sequential data and can capture long term dependencies and temporal patterns. While RNNs are disadvantaged by vanishing/exploding gradients, whereby dependencies over a long sequence of data are lost, LSTMs are not inhibited by this problem. The LSTM unit keeps two rolling variables called the hidden state and cell state to accumulate the short-term and long-term memory of the sequence, respectively, without losing crucial information as the sequence unrolls. This is possible due to the LSTM architecture updating the long and short-term memory at every input value and limiting the change in gradient at each iteration (Lindemann et al., 2021).

The LSTM unit is connected to a fully connected layer and then to the output node. A separate LSTM model was optimised for each site with the following network characteristics optimised for each model:

- Epochs: Number of times the training data was iterated over in training the model;
- Learning rate: Shrinkage factor of change in parameters in the direction of decreasing loss;
- Sequence length: Number of days of each feature to include in one example;
- Hidden size: Number of neurons in fully connected layer between LSTM unit output and output node.

3.1.5 Encoder-Decoder LSTM

The encoder decoder LSTM (referred to as ENCDEC) is a variant of the LSTM model which features two separate LSTM units, an encoder and a decoder. The feature sequences are passed into two LSTM layers and are encoded through them as

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

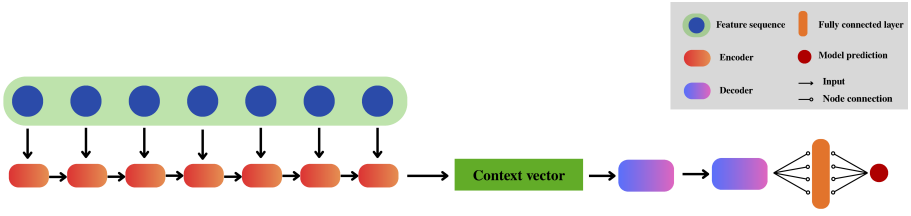

Figure 2. Encoder decoder LSTM (ENCDEC) architecture

a context vector. The context vector temporally connects and represents the features. The context vector is then passed to the decoder, which converts the context vector embeddings into a sequence of embedded predictions. The final sequence of embedded predictions are connected to a fully connected layer which outputs the model prediction. Figure 2 show the encoder decoder architecture. We use two LSTM layers each in the encoder and decoder. A separate model was optimised for each site. The network characteristics optimised for each ENCDEC model were the same as the LSTM, with the exception of sequence length, which is kept constant at 7 days for all sites.

3.2 Regression

To understand the factors affecting model quality we employ a multiple linear regression. This was done to understand the influence of geography, rainfall intensity and seasonality, soil properties, depth, training data, and modelling method on the overall performance of a forecast. A linear regression is fit with the RMSE of the forecasts as the target variable and the following co-variates:

- Zone: WA, Summer, Winter, Uniform (one-hot encoded);
- Method: Random Forest, XGBoost, MLP, LSTM, ENCDEC (one-hot encoded);
- Depth: 30-60 cm, 60-100 cm (label encoded);
- Clay-Sand fraction: Ratio of clay fraction to sand fraction (-) (numeric);
- Soil moisture range: Difference between minimum and maximum soil moisture (%VWC) (numeric);
- Average Rain: Mean annual rainfall (mm) (numeric);
- Training data: Number of training observations (numeric).

We employ backwards covariate elimination to only include significant covariates (significance level of 0.05) and ensured a low correlation ($|r|<0.5$) between covariates to minimise heightened effect size due to collinearity. The soil properties are estimated for each depth range through SLGA, and the ratio of the clay and sand was calculated to highlight sand or clay dominant soils. The covariates are z-scaled normalised to allow for standardised regression coefficients that the relative scale of effects is comparable with one another. We assess the range of variances across the z-scaled covariates graphically and find

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

a relatively even distribution that should not heighten effect size (Goldstein-Greenwood, 2023). The coefficients represent the average change in standard deviation of the RMSE given a one standard deviation change in the covariate of interest, given all other covariates remain constant.

3.3 Model utility

To understand the utility of a forecasting model over time, we run each model with weekly-increment increasing lead times until one of the following criteria is met:

1. The RMSE of the prediction exceeded 5 %VWC ;
2. The historic average was a more accurate predictor than the modelled estimate for at least three consecutive weeks;
3. The model forecasted 52 weeks ahead without either of the preceding criteria being met.

Criteria 1 defines our upper limit for a 'useful' model. We set this upper limit at 5 %VWC (15mm at 30-60 cm, and 20 mm at 60-100 cm). This criteria is framed by the Soil Moisture Active Passive Level 1 mission product accuracy requirement of 4 %VWC but with a greater leniency (Chan et al., 2016). At an RMSE greater than 5 %VWC we determine the model error too high to provide practical value. Criteria 2 is set to benchmark model accuracy against the accuracy of using the average historic soil moisture on the day of the forecast. Once the historic average becomes a more accurate estimate, we determine it to be a more suitable option than training a machine learning model. This limit is set to define an unnecessary use of a model when a historic average would suffice. Criteria 3 is checked as an upper limit to the model lead time for practical purposes. We omit Random Forest from this analysis due to its general low performance at high lead times (Figures 4 & 5).

4 Model training and validation

4.1 Training and testing strategy

At each site and for every lead time, an individual model was trained and tested. Our training and testing strategy involved hyperparameter tuning and training on a mutual training set and validating on a test set. The models were tested on the final 2 years of data available at each site. The remaining data was used in training and hyperparameter tuning. For sites with less than 3 years of total data, the testing was done on the final year of available data, to ensure at least 2 years of training data were available. The training and testing data were not shuffled to allow for persistence of the feature space for the memory-based models. For each model at a site, the seed was set to be the same to ensure the identical model initialisation, irrespective of lead time.

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

4.2 Model performance metrics

4.2.1 Validation

In this study model performance is defined on the basis of minimising the difference between the observed and predicted values. This model quality is best quantified by the Root Mean Square Error (RMSE)

The RSME is a measure of error that aggregates the squared difference of every prediction relative to its corresponding observation. The squared error term penalises large outliers and quantifies error irrespective of it being an under or over estimation. The RMSE matches domain units and is in a range of [0, ∞], where a perfect prediction is 0. An advantage of using RMSE is it being a standard validation metric across multiple bodies of work on soil water forecasting, which creates comparability between our work and past work. The RMSE is defined as,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{i=N}(x_i - \hat{o_i})}{N}} \tag{1}$$

where, x_i = Prediction, o_i = Observation, N = Number of testing days.

4.2.2 Optimisation

Hyperparameter optimisation was conducted for each model, where the ideal combination of hyperparameters is the combination which maximises an objective function. The hyperparameters tuned for each model are defined in Table 1. The objective function utilised for model optimisation is a scaled variation of the modified Kling Gupta Efficiency (KGE'), referred to here as KGE'$_s$. The hyperparameter combinations for each model with the highest KGE'$_s$ were selected to produce forecasts at all lead times.

KGE' is a multi-component criterion that represents the Euclidian distance of a prediction to the observation in three-dimensional space defined by 3 metrics of modelling error: correlation, bias, and variance (Liu, 2020). The bias and variance components (α and β) represent the agreement of the modelled and observed mean, and variation magnitude (referred to as 'distance'). The correlation (r) represents the shape and timing of variation (referred to as 'timing'). The optimal value of KGE' is at 1 and occurs when the mean and variance of the observed and predicted points are equal, and the correlation is equal to 1. These 3 components indicate the model's accuracy in regard to modelling the mean value of an observation sequence and capture the magnitude and instantaneity of variations.

As detailed by Cinkus et al. (2023), the KGE' has inherent counterbalancing errors, in which over and under estimation of the observations by the model cancel each other out, leading to the KGE' overestimating model accuracy. The susceptibility of KGE' to counterbalancing errors is due to the distance terms, α and β , being functions of the mean; an aggregation of predictions, without squaring or absolution. The scaled variation of KGE', KGE'$_s$, conceptually first introduced by Kling et al. (2012), scales the r:α:β to 2:1:1. Where the unscaled KGE' gives two-thirds weight to the mean-centered terms, α and β , and only a third to the correlation; KGE'$_s$ gives equal weighting to the distance and timing terms. The increased weighting

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

of r reduces the weighting of the distance terms, which in turn lowers the effect of counter balancing on the KGE'$_s$ value, in comparison to the KGE'. Additionally, the KGE'$_s$ aligns with our model validation methodology of giving equal priority to prediction instantaneity. The timing of soil water forecasts is crucial as critical decisions are made on a daily basis and require predictions of soil water be accurate in predicting event timing. The KGE'$_s$ is defined as,

$$KGE_s = 1 - \sqrt{\mathrm{S_a}(\alpha-1)^2 + \mathrm{S_b}(\beta-1)^2 + \mathrm{S_c}(r-1)^2} \tag{2}$$

where,

$$\alpha = \frac{\sigma_x}{\sigma_o} \tag{3}$$

$$\beta = \frac{\bar{x} - \bar{o}}{\sigma_o} \tag{4}$$

σ_x = Standard deviation of predictions,

σ_o = Standard deviation of observations,

$\mathrm{S_a} = 0.75,\ \mathrm{S_b} = 0.75,\ \mathrm{S_c} = 1.5.$

4.3 Hyperparameter tuning

Hyperparameter tuning is essential in machine learning applications as hyperparameters directly control training behaviour and can have a significant effect on model performance (Wu et al., 2019). Hyperparameter tuning was conducted by first conducting a grid search of all hyperparameter combinations reported in Table 1 for each model to identify the combinations that maximise KGE'$_s$ performance for a 30-day lead time. However, while a particular hyperparameter combination may perform the best at a 30-day lead time, due to the potential for overfitting, that combination may not perform as well for other lead times. Therefore, selecting the hyperparameter combination that maximises performance at a particular lead time is not appropriate, as it does not guarantee generalisability (Guan and Burton, 2022). To address generalisability concerns, each of the high performing hyperparameter combination were tested across all lead times and the most generally applicable (i.e. lowest KGE'$_s$ variance across lead time) was selected. Table 1 shows the hyperparameter values tested and selected for each model.

5 Results

5.1 Training and testing period

Figure 3 shows the relationship between the inter-quartile range (IQR) of the soil moisture in the training and testing data at each site. The rainfall intensity is highlighted too, where Dry indicates rainfall experienced at that site to be below the 25th

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

Table 1. Hyperparameters tested for each model

Model	Hyperparameter	Tested values
Encoder-Decoder LSTM	Epochs	100, 150, 200, 250, 300
	Hidden size	3, 5, 10, 15, 20
	Learning rate	0.001, 0.005
LSTM	Epochs	125, 150, 160, 180
	Hidden size	3, 5, 7, 9
	Learning rate	0.0005, 0.001, 0.005
	Sequence length	1, 2, 3, 4, 5, 7
MLP	Epochs	125, 150, 160, 180
	Width	5, 7, 9, 10, 12
	Learning rate	0.0005, 0.001, 0.005
XGBoost	eta	0.1, 0.2, 0.3, 0.4
	max_depth	1, 2, 3, 4, 5, 6
	n_estimators	20, 35, 30, 50
	subsample	0.75, 0.9, 1
Random Forest	min_sample_leaf	2, 3, 5, 7
	max_samples	0.75, 0.9, 0.1
	n_estimators	50, 75, 100, 120

335 percentile, and Wet indicating rainfall above the 75th percentile, of the annual average rainfall experienced across all the sites in this study. Only the 30-60 cm range IQR ratios are visualised in Figure 3 for conciseness as a similar relationship is exhibited in the 60-100 cm range. The IQR here represents the typical range of soil moisture experienced by the soil in the training or testing period. The purple line indicates an equal ratio between the IQR of the training and testing periods. The grey 'Similar' region indicates a similar IQR between the two periods, where the ratio of the IQR of both periods is less than two. Points 340 outside the grey region represent sites where one period accounted for a greater range of soil water than the other. For example points outside the grey region and below the line represent sites where the site experienced more drying and wetting events for the model to be trained on than tested on.

Most of our sites experience a similar range of soil water between the training and testing period. Sites in the Summer and Uniform zones in particular have a similar IQR between periods. However, some Uniform sites have a greater IQR in the testing 345 period than training, which indicate our models are being tested outside the bounds of their training data. The Winter sites on the other hand are mostly trained across a greater range of soil water than they are tested on, indicating a lower likelihood of the model being validated on an anomalously wet or dry period. WA sites exhibit a distinctly higher IQR in the training period than the testing period, despite experiencing high rainfall. Our Summer and WA sites exhibit a lower overall IQR, which is

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

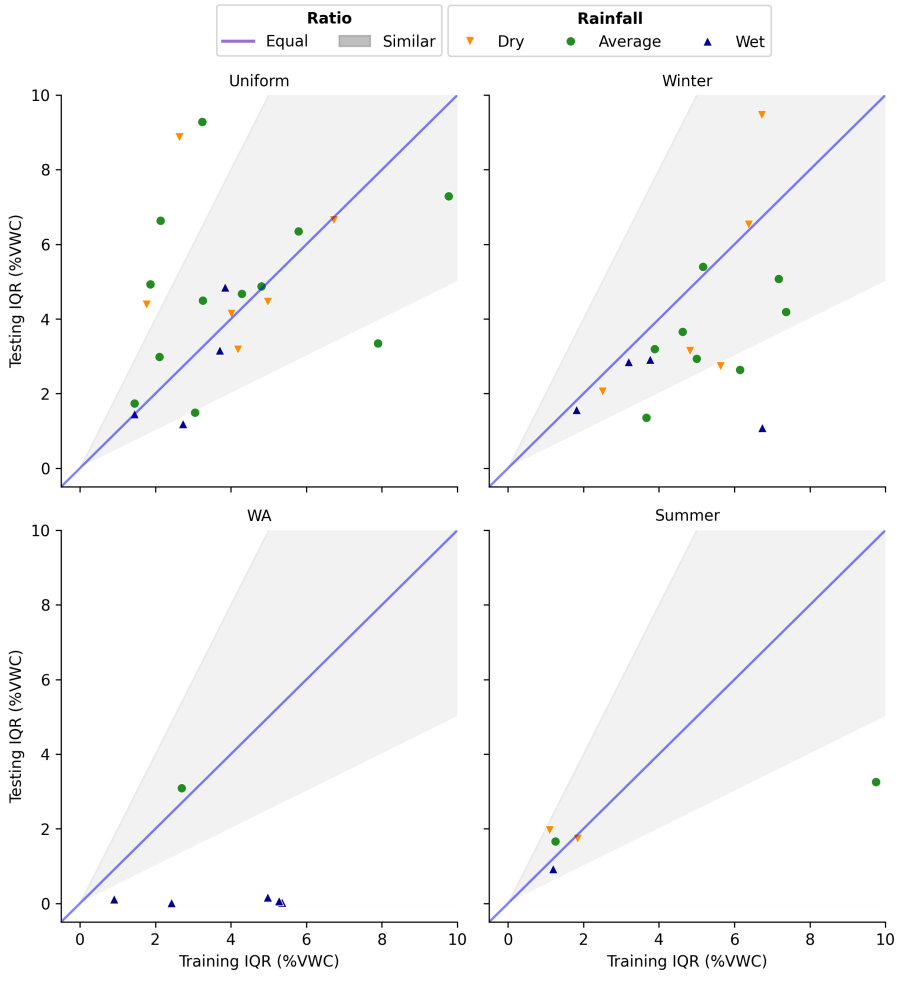

Figure 3. Relationship of training and testing period soil moisture interquartile range (IQR). The purple line indicates and equal IQR in the two periods; while the grey 'Similar' region indicates an IQR ratio between the two periods being within 0.5-2.

indicative of the soils generally not storing or losing water as frequently, even while experiencing relatively high rainfall at most sites.

5.2 Multi-week forecast

The accuracy of every method was analysed for lead times of 1 week, 1 month, 2 months, and 3 months. These lead times were selected as decisions for dryland inputs are typically made at weekly scale such as in relation to sowing timing or at a longer seasonal scale for yield potential and fertiliser input. Here, we analyse both depth ranges individually to understand the effect depth has on model accuracy, particularly as agricultural soils are subject to varying root interactions with depth.

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

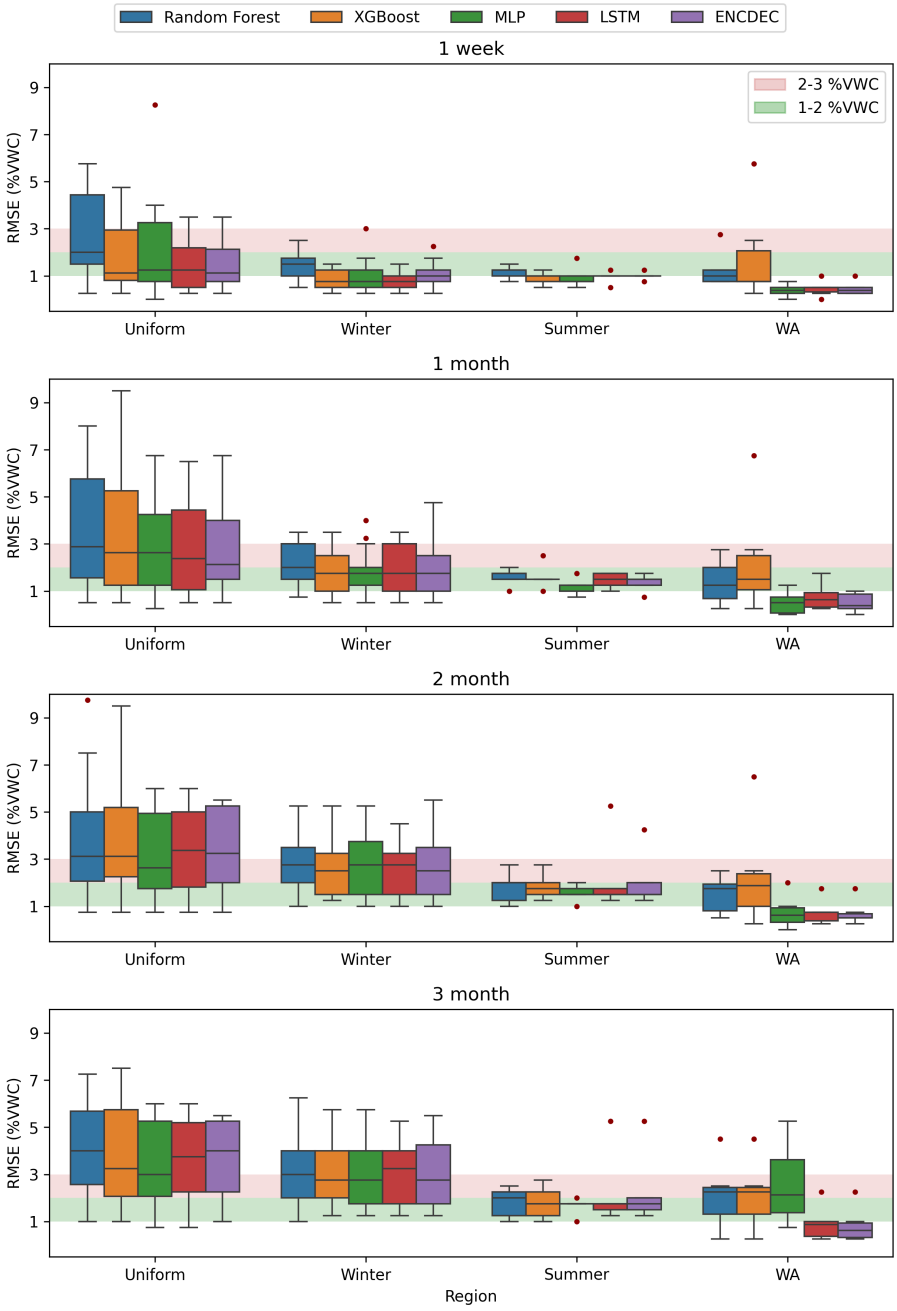

Figure 4. Performance of each method across lead time and zones (30-60 cm)

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

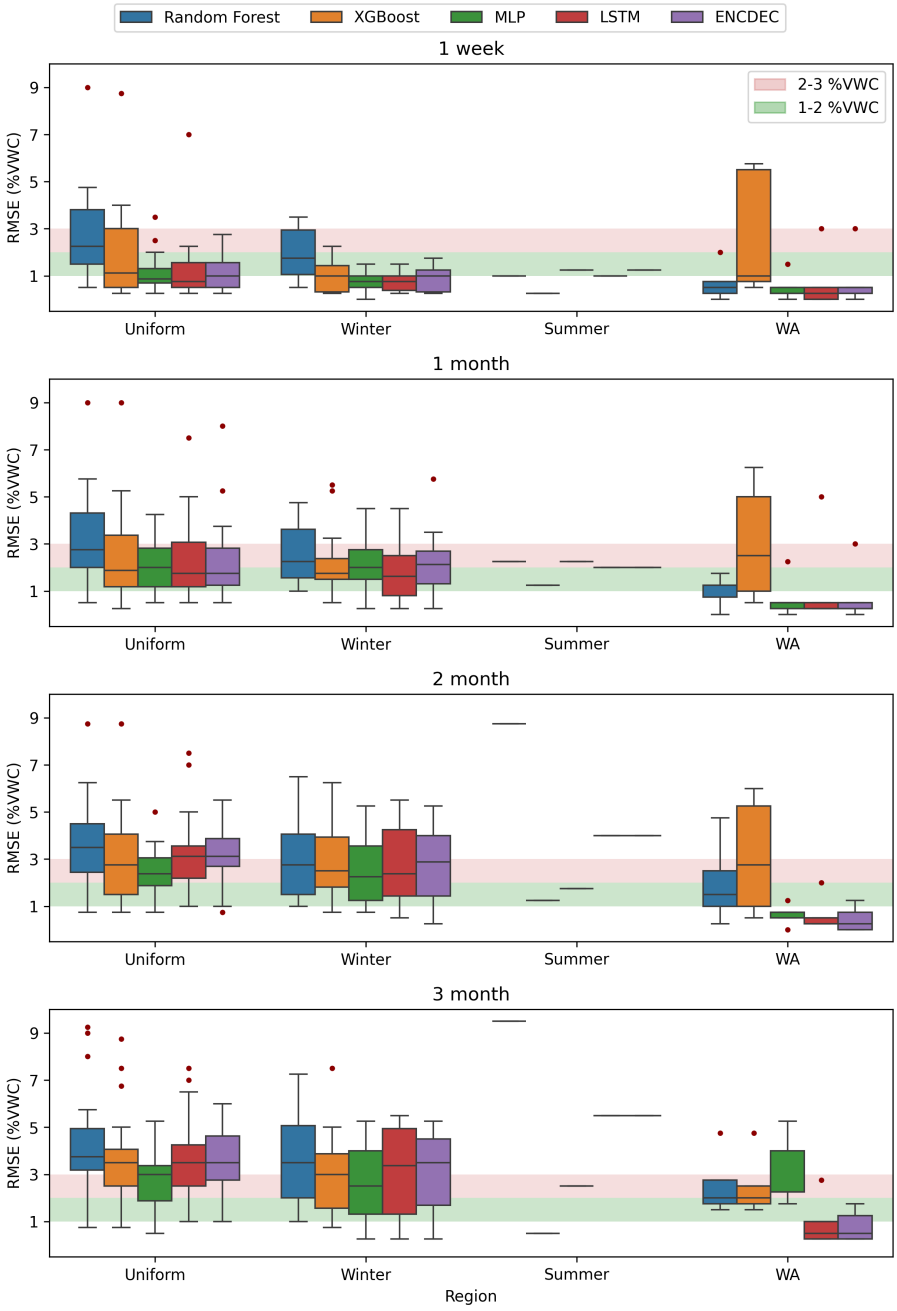

Figure 5. Performance of each method across lead time and zones (60-100 cm)

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

Figures 4 and 5 shows for a 1 week lead time, irrespective of depth, most sites are forecasted at an accuracy below 1 %VWC (3 mm in 30-60 cm depth range and 4 mm in the 60 – 100 cm depth range). Each zone is forecasted at a similar accuracy through depth, except for the Uniform zone sites, which are forecasted slightly less accurately by ever method in the shallower depth range. The WA sites are the most accurately forecasted zone, particularly by the deep learning methods. At a 1 week lead time, on the basis of median accuracy, model choice is less critical as using XGBoost or any of the deep learning methods produces forecasts at high and comparable accuracy to other methods, across all zones.

Most zone-method combinations exhibit an accuracy near 2 %VWC for a 1 month lead time. A general decrease in accuracy of 1-1.5 %VWC is noted as lead time increases from 7 days to 30 days, with the exception of WA sites. The WA sites are forecasted to less than 1 %VWC accuracy at both depths using any of the deep learning methods. In the 30-60 cm range, Uniform sites decrease in median accuracy by 1.5 %VWC (4.5 mm) and Winter sites decrease in median accuracy by 1 %VWC. A similar decrease in accuracy is observed in the 60-100 cm range. The notable increase in height of the upper whiskers at this higher lead time in Uniform and Winter zones indicates that some sites disproportionately decrease in accuracy over increasing lead times. The Summer sites do not experience a strong decrease in accuracy in the 30-60 cm range, however the Summer site at the 60-100 cm ranges does reduce in accuracy by approximately 1 %VWC for all methods.

There is a lesser median accuracy decrease of 0.5-1 %VWC as lead time increases from 1 month to 2 months, most notably at the deeper depth range. The median accuracy of Uniform and Winter sites are in the 2.5-3 %VWC range. The Summer and WA sites are forecasted at similar accuracy as at a 1 month lead time, of under 2 %VWC and 1 %VWC, respectively. The decrease in accuracy is less between a 2 to 3 month lead time increase than at a 1 to 2 month increase. While the accuracy is lower for the 3 month lead time, the accuracy decreases by at most 0.5 %VWC in most instances, particularly at Uniform and Winter sites. The most notable decrease in accuracy is in MLP for the WA sites, which decreases in accuracy by approximately 1.25 %VWC in the shallower and 2 %VWC in the deeper range. Besides this, the Summer and WA sites are forecasted at a very similar accuracy in both depth ranges.

There is no optimal model for forecast accuracy across all zones, but some patterns emerge across lead time for each zone. At Uniform sites, all methods except Random Forest perform similarly for the 1 week and 1 month lead times, with both LSTM-based methods performing the best (w.r.t median RMSE and the interquartile range). However, for the longer lead times, XGBoost and MLP are more accurate. The opposite is noted in WA, where initially all deep learning methods perform similarly, but over time the LSTM-based methods tend to be most accurate. At Winter sites XGBoost and MLP consistently perform the best across all lead times, but MLP outperforms XGBoost for the longer lead times in the 60-100 cm range. The Summer sites are best forecasted by MLP and XGBoost in the 30-60 cm range, while XGBoost is prominently most accurate for the Summer site at 60-100 cm. Random Forest is overall the least reliable method, exhibiting the poorest performance for most zones and lead times. However, it is a preferred choice over XGBoost in WA. The Uniform sites were overall forecasted at the lowest median accuracy, and the largest range of accuracies, irrespective of lead time. The Winter sites are forecasted at to a higher median accuracy but in a similar range to the Uniform sites, indicating less reliability of one specific model in capturing the dynamics of either zone. The Summer and WA site are consistently forecasted to a higher accuracy and a lower

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

range than the other zones, implying the selected model will perform reliably at a new site. There are no clear performance patterns with depth as model accuracy does not vary with any discernible pattern.

5.3 Sub-monthly method comparison

The heat maps in Figure 6 shows the proportion of the most accurate method at all sites in a zone up to 30 days of lead time. These heat maps aim to highlight how critical model choice is for lead times up to one month. This contextualises our findings within the broader literature on model selection which has typically conducted for lead times under one month. We analyse each zone separately to inform best model selection for a new site based on its rainfall-geographic nature. In this section, we refer to lead times as in the following terminology: 1-10 days as 'short-range', 11-20 days as 'medium range', and 21-30 days as 'long range'. This terminology is specific to the analysis of Figure 6.

For short-range lead times at Uniform sites, the shallow interval shows LSTM and XGBoost to perform the best most often, while the deeper interval exhibits a more uniform distribution of model choice. For medium-range forecasts, LSTM and XGBoost forecast most accurately most often for both depth ranges. However, for long-range forecasts, both depth intervals show a more uniform distribution across each model, however, LSTM and XGBoost remains the largest proportions. In selecting a model for Uniform zones for sub-monthly forecasts, LSTM or XGBoost are a reliable choice for all lead time. However, model choice is less critical as there is a relatively uniformly distributed model preference, particularly as lead time increases. This is affirmed by Figures 4 and 5 which show a similar median across methods for 1-week and 1-month forecasts, but the wide range of each boxplot may denote forecast accuracy is not distinctly captured better by one method over than another.

The Winter sites exhibit a majority proportion to XGBoost and MLP, across all lead times at shallower depth. However, there is a strong preference for MLP particularly for long-range forests. A similar proportion of XGBoost and MLP is seen at the deeper interval, but a fairly even distribution is observed between XGBoost, MLP, and LSTM for long-range forecasts. The Winter sites seem relatively less optimised by the memory-based models, as XGBoost and MLP seem to provide the most accurate estimates most frequently. However as seen in Figures 4 and 5, while memory-based models are less frequently the best, their accuracy is very similar to MLP and XGBoost for sub-monthly forecast. Overall, however MLP and XGBoost are the most reliable choice in selecting a model at a new Winter site.

The WA sites vary the most between depths in model preference. The shallower interval is modelled best by MLP and ENCDEC for the short-range forecasts, with ENCDEC and LSTM being viable options too for medium range forecasts. MLP is the most reliable choice for long term forecasts. The deeper interval is relatively evenly distributed between Random Forest, MLP, and LSTM for most short and medium range forecasts. However model preference is less consistent across long-range lead times as each Random Forest, MLP and ENCDEC each take a different portion for different lead times. As seen in Figures 4 and 5, the deep learning methods are still a more reliable choice as they perform well on average. Overall, between both depths, deep learning methods are preferred. Random Forest in many instances in WA can perform well in the 60-100 cm range. The only available Summer site at 60-100 is best captured by XGBoost in all instances. However, in the shallower interval, XGBoost is a reliable option at some short-range lead times, but overall MLP is most reliable in medium-range and

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

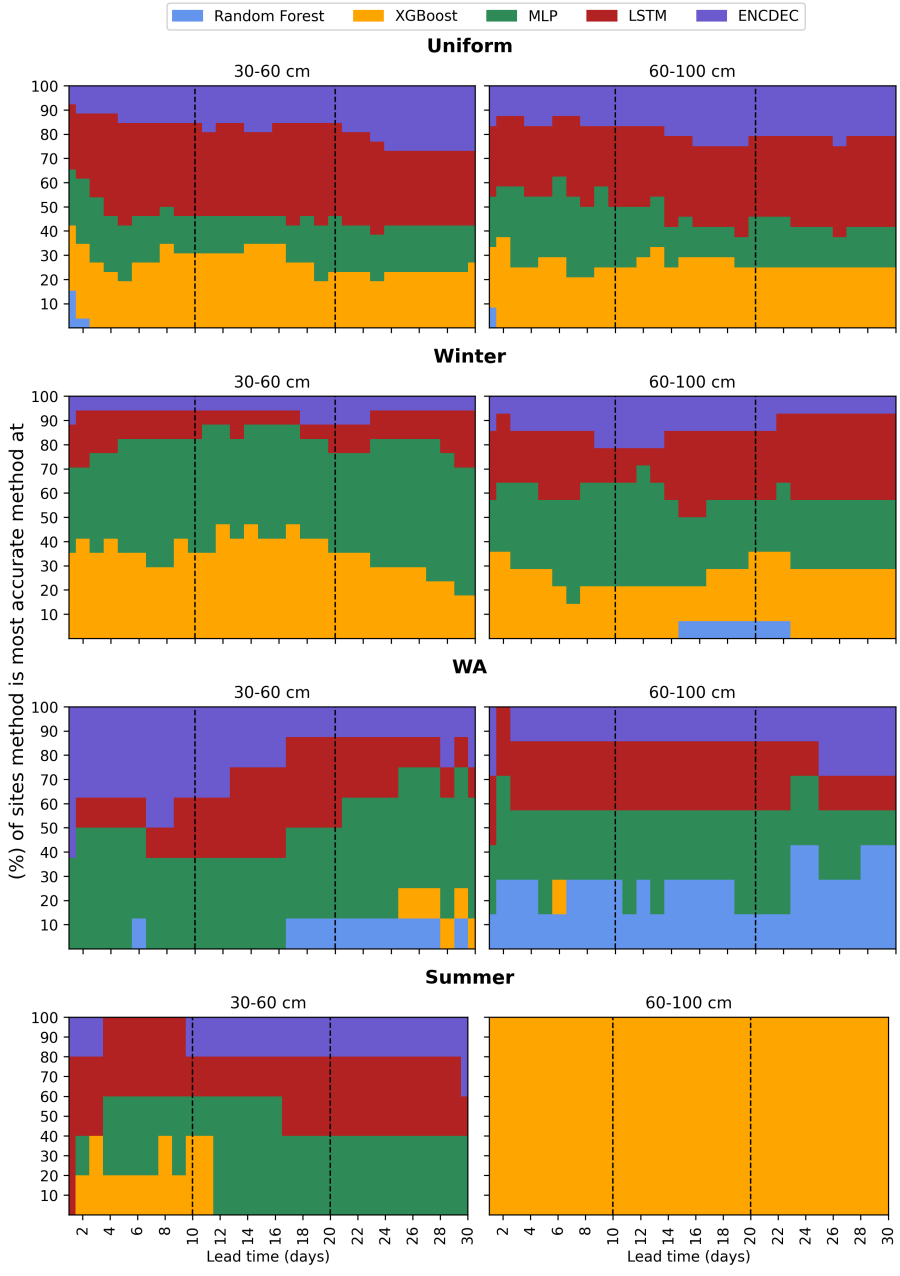

Figure 6. Proportion of sites each method was most accurate for at 1-30 day lead time. A greater area represents a more frequently high performing model.

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

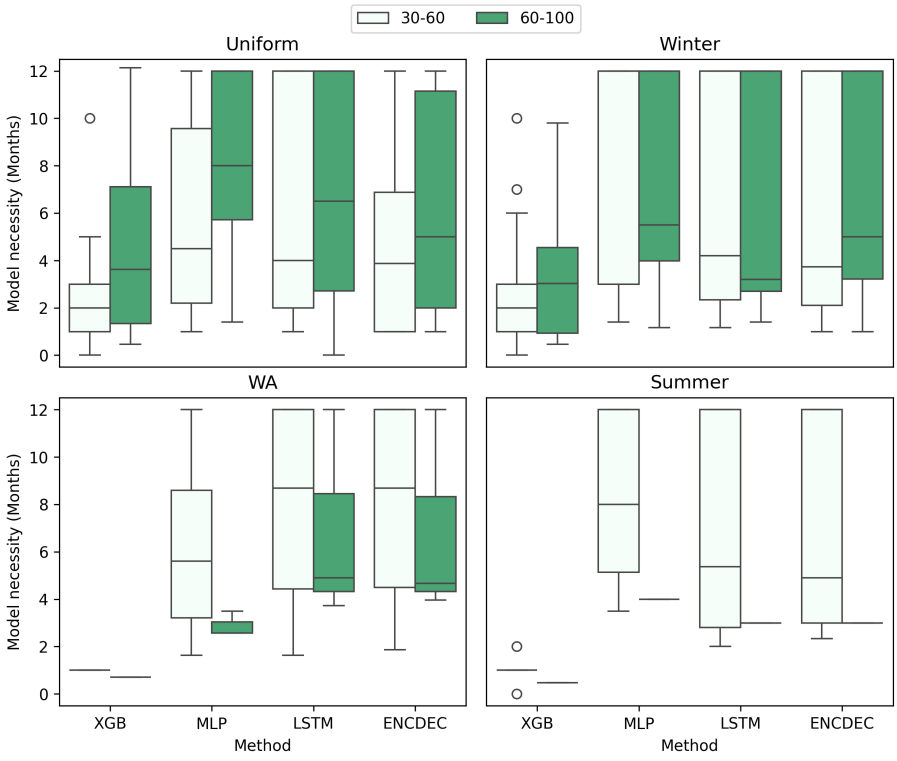

Figure 7. Model necessity, i.e, how far ahead the model forecast is more accurate than a long-term average prediction or 5 %VWC

LSTM and MLP are both equally reliable for long range forecasts. Overall, deep learning method should be preferred for all sub-monthly lead times at Summer sites.

5.3.1 Model utility

For Uniform sites we find models to forecast further ahead in the 60-100 cm range more often than in the 30-60 cm range. In the shallower interval, all deep learning models can be used at Uniform sites for a median period of approximately 4 months. For the deeper interval, MLP can be used for the longest, for 8 months on average, while LSTM and ENCDEC are viable for an average 5-7 months. However, many sites in the Uniform zone can be forecasted a full year ahead by deep learning models. XGBoost forecasts 2 months ahead on average in the 30-60 cm range and under 4 months at 60-100 cm. However, some sites are able to be forecasted as far as 10-12 months ahead. The sites which are forecasted through the full year are mostly not the same sites for every method, showing that some sites are not inherently forecasted better than others. Table 2 highlights the proportion of sites meeting each criterion were generally equal across all criteria, however exceeding the benchmark was most common limiting factor.

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

Table 2. Proportion of criteria met

		%		
Zone	Depth (cm)	Benchmark	Climatology	Year
Uniform	30-60	44	38	17
	60-100	39	35	26
Winter	30-60	7	63	29
	60-100	34	38	29
WA	30-60	0	69	31
	60-100	0	83	17
Summer	30-60	20	50	30
	60-100	100	0	0

The Winter sites were forecasted viably to a similar lead time across depth when using LSTM and ENC DEC. In comparison to the Uniform sites, a greater proportion of sites are forecasted a full year ahead. Using MLP to forecast in the 30-60 cm depth has a notably higher median of 12 months. MLP here proves to be a very viable option for long-term forecasts, as other models in this depth range for Winter sites are viable till approximately 4-5 months on average. The shallower region was mostly either limited by the historic average or able to be forecasted a whole year ahead. The deeper interval shows a greater shift towards the benchmark being hit, as is noted in Figures 4 and 5 too, where the upper whiskers for the winter sites have begun to exceed 5%VWC for winter sites and the median RMSE is nearer to 5%VWC here than at the 30-60 cm interval.

The WA sites exhibit a distinctly higher maximum lead time at the shallower depth range than the deeper regions. The memory based models are able to forecast on average 9 months ahead, while MLP can forecast between 5 and 6 months ahead. The 60-100 cm range is still reasonably forecasted ahead by 5 months by the LSTM-based models, and approximately 3 months for the MLP. Using XGBoost, all sites are forecasted to a maximum of one month ahead as the historic average becomes a better predictor earlier on than using any other method. The WA sites are generally forecasted below 5%VWC (Figures 4 & 5) and as seen in Table 2, the benchmark is not hit at any site. The historic average performs the best in WA as it eventually outperforms the model in most instances. The Summer sites are forecasted for the longest by MLP, up to 8 months on average, and the memory-based models up to 6 months. The summer sites are mostly limited by the historic average.

5.4 Regression Analysis

Our regression models are able to explain between 36-45% of variation in RMSE. These models are better predictors of accuracy than an intercept-only model, as indicated by the P(F-statistic) being below 0.05. The remaining variation is unknown and could be due to more complex soil water dynamics currently unaccounted for. However, some notable trends are visible across the variation the models can explain.

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

Lead Time	1 week			Lead Time	1 Month			Lead Time	3 Months		
Adj. R-squared:	0.45			Adj. R-squared:	0.47			Adj. R-squared:	0.36		
F-statistic:	45.1			F-statistic:	55.3			F-statistic:	55.53		
P(F-statistic):	1.74E-58			P(F-statistic):	2.01E-63			P(F-statistic):	1.41E-45		
Covariate	**Coefficient**	**2.5th CI**	**97.5th CI**	**Covariate**	**Coefficient**	**2.5th CI**	**97.5th CI**	**Covariate**	**Coefficient**	**2.5th CI**	**97.5th CI**
SM Range	0.73	0.63	0.83	SM Range	0.99	0.87	1.12	SM Range	0.84	0.69	0.99
Average Rain	0.11	0.02	0.21	Average Rain	0.22	0.10	0.34	Average Rain	0.35	0.21	0.49
Training Length	-0.11	-0.20	-0.01	Training Length	-0.23	-0.34	-0.11	Training Length	-0.36	-0.50	-0.23
Region: Uniform	0.17	0.08	0.27	Region: Summer	-0.13	-0.25	-0.01	Region: Uniform	0.24	0.09	0.38
Region: WA	0.19	0.08	0.30	Region: Uniform	0.13	0.01	0.26	Method: Random Fores	0.25	0.12	0.38
Method: ENCDEC	-0.19	-0.29	-0.08	Method: Random Fores	0.22	0.11	0.33				
Method: LSTM	-0.20	-0.31	-0.09	Method: XGBoost	0.16	0.05	0.27				
Method: MLP	-0.16	-0.27	-0.06	Depth	-0.15	-0.26	-0.04				
Method: Random Fores	0.23	0.12	0.34								

Figure 8. Regression results showing variables with a significant effect on forecast accuracy for varying lead times

The range of soil moisture is the most significant covariate at all lead times. Sites with a greater difference between the minimum and maximum soil water were forecasted at a lower accuracy. This can be anticipated as generally sites with more intense extremes (peaks or depressions) are typically most difficult to forecast due to their highly variable nature not being accounted for in the training data. The annual average rainfall is also significant across lead times, exhibiting a mild positive relationship with RMSE. For a 1 week lead time, annual average rainfall is less significant than at longer lead times as the 2.5th confidence interval is near 0. For the longer lead times the amount of rainfall has a stronger effect on the prediction quality, particularly at sites with greater annual average rainfall. While soil water range and the average rainfall exhibit a weak correlation (r ~ 0.25), the greater effect of rainfall on accuracy at longer lead times is due to more variation in soil moisture at sites with greater rainfall, even if not extreme, as rainfall is sole driver of increases in soil moisture in dryland fields. If there is more rainfall throughout the year, the soil moisture will likely vary more on average and be more unpredictable at longer lead times, where the feature space and target date are less related.

Training length is significant across all lead times, however to a stronger effect at lead times longer than 1 week (97.5th confidence interval is near 0 at a 1 week lead time). Greater training lengths typically improve forecast accuracy, especially at longer lead times as more training data allows the model to better learn soil water dynamics, particularly as they can change greatly over the span of months. The final covariate that is significant at all lead times is the binary covariate of Uniform zone sites. There is a mild effect of Uniform sites being less accurately forecasted across all lead times, as is explored in more detail in Figures 4 and 5. Notably, for a 1 week lead time WA zone sites exhibits a positive relationship with RMSE, to say WA sites perform poorer, which is contradictory with our findings in Figures 4 and 5. However, this is likely seen as the performance of all methods is considered in the regression analysis and the performance of XGBoost is likely swaying the overall high accuracy of the WA sites. Across lead times, where the method covariates are significant, deep learning methods all exhibit a negative relationship with RMSE (i.e. RMSE reduces when a deep learning method is picked). However, as we note in Figures 4, 5, and 6, the model choice is more variable for shorter lead times in many instances. This is supported in our regression analysis where the method covariates are all significant for the 1 week lead time than at the longer lead times, indicating less model generalisability for longer lead times.

To support our regression analysis, we plot the sites with lowest, median, and maximum rainfall sites from each zone as an example of typical soil moisture behaviour in each zone for a varying intensity of rainfall (Figure 9). The winter and summer

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

periods are highlighted to show when the dominant season of rainfall is expected for each zone. The Uniform and Winter sites experience a wider range of soil moisture values than the WA and Summer sites, irrespective of rainfall intensity. An increase in range is noted across all regions with increasing rainfall intensity, most notably at Uniform and Winter sites. This observation aligns with our findings from the regression analysis, where we determined the accuracy of forecasts to decrease as the range of soil moisture values increases or as rainfall intensity increases. In addition, this may explain the greater RMSE across Winter and Uniform sites noted in Figures 4 and 5. The WA sites have the lowest range consistently, and increases in soil moisture appear to correlate strongly with the dominant rainfall season occurring. The non-dominant rainfall periods are notably flatter. While the WA sites seem to be sensitive to the winter seasonal rainfall, the Winter sites appear to be affected by rainfall year long. This could be due to misidentification of rainfall dominant period at the site or the high water retaining soils in the Winter zone being more sensitive to rainfall.

6 Discussion

In this study we explored the impact of model selection, geography, rainfall, depth, and lead time on forecasting accuracy of soil moisture in Australian dryland fields. Here, we forecast soil moisture at 54 sites across Australia at two depth intervals (30-60 cm and 60-100 cm), for lead times ranging up to 52 weeks ahead. We conduct the first Australia-wide assessment of soil water forecasting accuracy to improve dryland agricultural decision making by providing a reference for future region-specific soil water forecasting efforts. In particular, informing model choice, expected accuracy, and the necessity for a model for forecasting at a new Australian dryland site. While previous studies have forecasted soil moisture for lead times typically under 10 days, this study aimed to understand the underlying factors affecting short and long term forecasts. We did so by using five machine learning models that employ distinct mechanisms to forecast soil moisture and inform our understanding of soil water dynamics. We appraise the most elementary forms of the best performing models amongst the literature.

We demonstrate the ability to forecast soil moisture to a reasonable accuracy at least 2 months and up to 12 months ahead, without a rainfall forecast, for soil moisture in a depth range of 30-100 cm. We are able to achieve this with elementary forms of each model, without any novel modification to the architectures. With respect to the necessity of using a machine learning model for long-term forecasts with the alternative of historic average estimates available and a reasonable degree of accuracy considered, we find the deep learning models to be useful for forecasts at our sites across the east and south coast of Australia up to 5 months ahead on average in the 30-60 cm interval, and at least 6 months ahead on average in the 60-100 cm depth. Whereas sites in WA can be forecasted on average 8 months ahead in the 30-60 cm interval but 5 months at the 60-100 cm interval. However, this is very variable as some sites were forecasted to a suitable accuracy by the models up to 1 year ahead, but others as poor as one month ahead before being outperformed by the historic average. We recommend determining an upper limit on lead time when performing long term soil moisture forecasts, where after a particular lead time, a historic average estimate may suffice, or the modelled estimate may have too much uncertainty to be of use.

Furthermore, we explored the factors affecting accuracy, of which rainfall intensity and seasonality were the greatest drivers of accuracy on the basis of rainfall being the only input of water into dryland soil. Rainfall quantity is a key predictor of forecast

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

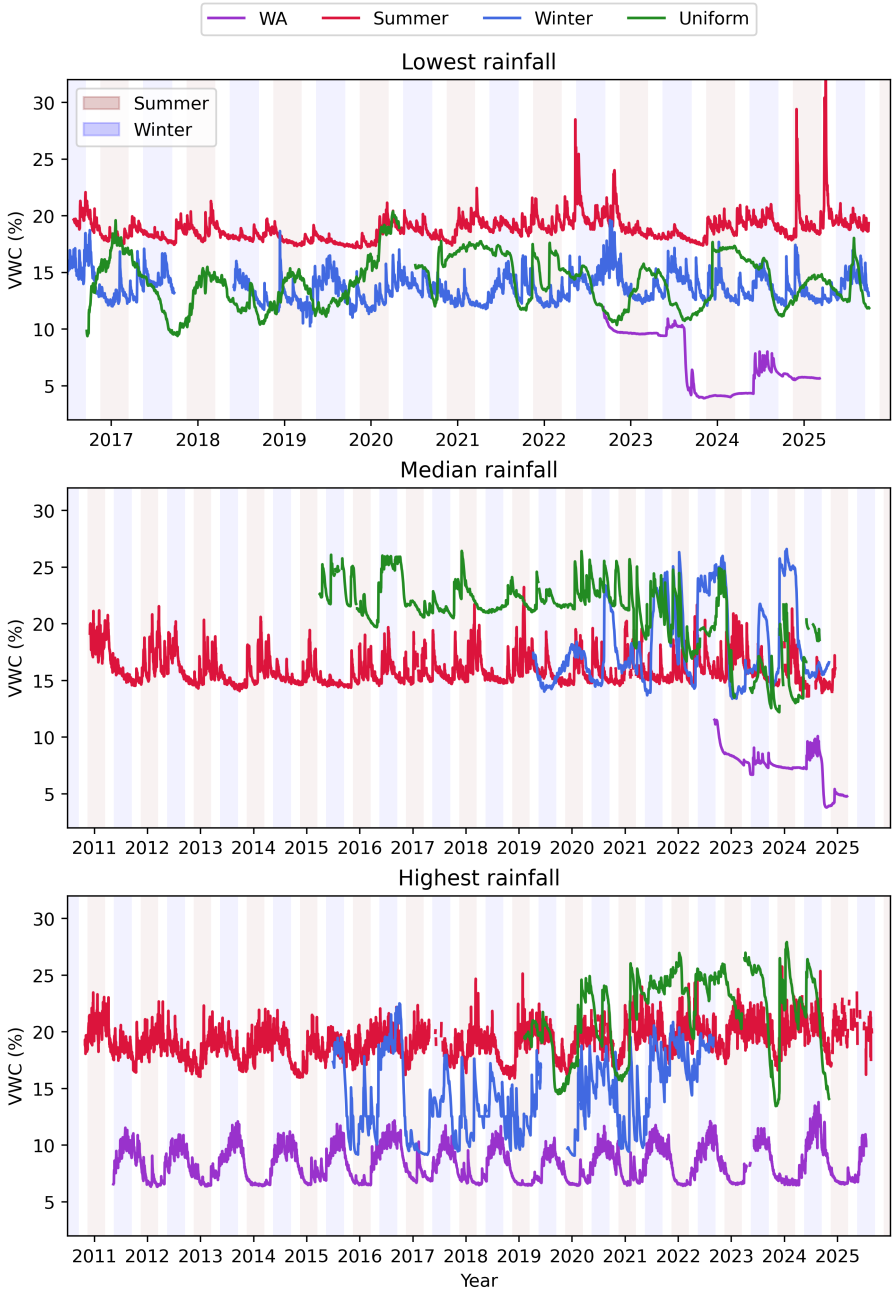

Figure 9. Soil moisture over time at selected sites across different regions. Dominant rainfall periods highlighted.

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

accuracy as it determines the daily differential of soil water. We find rainfall patterns to be an equally, if not more, significant predictor of accuracy as seasonal patterns determine the timing and intensity of large influx events that are often most difficult to forecast (Figure 8). We find that site experiencing uniformly distributed rainfall (spread out through south-east and south Australia) are less accurately forecasted in comparison to sites in other rainfall regions. This could be due to uniform rainfall areas experiencing scattered rainfall events that produce individual soil moisture spikes throughout the year, as opposed to the more condensed rainfall events in seasonal rainfall areas that lead to sustained elevated soil moisture levels over periods time. Spikes in soil moisture are difficult to forecast as their timing and magnitude are dependent on rainfall on that day, which the model does not know. Figure 9 show the greatest difference between observed and predicted soil moisture is in the spikes, as compared to any other event. Seasonal rainfall can mean less variability in soil moisture during the non-dominant periods, which is easier to forecast due to smaller daily differentials. This may explain why WA is the most accurately forecasted region at all lead times up to 3 months. The low water holding capacity of the sandy Western Australian soils have a low likelihood of a large daily differential in soil moisture, creating a more predictable temporal relationship for the model to learn (McKenzie et al., 2004; Wilford, 2012). Efforts for forecasting at a new site should consider the range of soil moisture values and the patterns of rainfall. Areas with a larger range of soil moisture, or intermittent and intense periods of rainfall can be most difficult to forecast (in the absence of a rainfall forecast predictor).

With respect to the models tested across all rainfall zones, we do not find one model to perform best across all sites, but rather certain models perform better than others across zones. This is particularly noted for leads times longer than one month, as we find model preference within each region become distinct at longer lead times. Model performance is less distinct at lead times less than one month, as supported by findings in the literature. For example, Adeyemi et al. (2018) report LSTM and MLP to perform nearly identically (Δ RMSE < 0.5 %VWC) for 1-day ahead forecasts across 3 sites in England. They do however note that LSTM was able to achieve a comparable performance with MLP with less input preprocessing. Whereas, Li et al. (2022a) and Han et al. (2021) find LSTM to outperform MLP at lead times under one week across multiple sites globally. Togneri et al. (2022) find a gradient boosting method (similar to XGBoost), LightGBM, to outperform both LSTM and MLP, and MLP to outperform LSTM for one-day ahead forecasts at 12 irrigated sites in Brazil. However, there is not much work on analysing model accuracy over longer lead time in the current literature and so the variability of model accuracy for longer lead times is not comparable. Our results align with the variability in model choice reported across the existing research, showing low sensitivity of model selection to lead times under one month. This affirms the need for model selection not only on the basis of general accuracy in other geographic regions or target domains but based on greater geographic and rainfall based classifications.

While RNNs are expected to perform well due to their high performance in some soil water forecasting studies, a key consideration is these studies forecast soil moisture at forested sites (with lesser effect of land cover and root-water interactions of cropping), at shallow depths (<10cm), or with the aided regularity of irrigation as an input (Datta and Faroughi, 2023; Han et al., 2021; Li et al., 2022a, b, c). All these factors add an element of seasonality to the inputs of water that are able to be learned by a memory-based model such as an LSTM. We do not find RNNs to be better at forecasting even at sites with greater autocorrelation in soil moisture or greater effect of monthly seasonality. This can be due to the highly variable nature of soil

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

moisture at the depth intervals explored in this study, the effect of crop roots on soil water dynamics, and the irregular input of water into the soil. The only sites at which LSTM notably outperformed others is at WA sites, where there are frequent periods of little change in soil moisture over time due to lower water holding capacity in their distinctly sandy soils.

7 Conclusion

Future soil water forecasting efforts can utilise these findings in determining an optimal modelling technique and in estimating accuracy based on the soil properties and meteorological landscape of their location of interest. Method selection can be further developed on with differing model architectures or hyperparameters, underpinned by our recommendations based on rainfall patterns and geography. Our recommendations can serve to inform model development with respect to overall model structure (ensemble versus RNN, for example) depending on aforementioned site characteristics. From there, a more novel architecture can be employed based on knowledge of the forecasting site of interest. For example, a larger MLP network can be considered at sites with more data points, or a longer feature period for the LSTM-based models at sites with hourly measured soil moisture.

A potentially productive addition to our method could be the inclusion of a rainfall forecast as a feature. Rainfall forecasts can provide foresight into the occurrence soil moisture spikes, which are most difficult for models to forecast. To a similar effect, features to highlight the likelihood of exceptionally high or low soil moisture may provide better estimate for extremes. For example, Togneri et al. (2022) present an 'anomaly index' to represent periods of extremes using an autoencoder to exemplify anomalous conditions across the feature space.

As the current body of work has been able to establish methods for high quality forecasts, more research into the effect of climate, geography, land cover, and rainfall patterns can provide benefit in understanding the limiting factors of forecasting accuracy. We are not able to consider a more thorough analysis of the effect of soil properties on accuracy, as the publicly available soil moisture at the time of writing do not include site measured soil properties, which are more representative than the coarse resolution modelled estimates used in this study. Likewise, more precise knowledge of land use and its patterns, such as information on sowing and harvesting patterns, could greatly improve forecasts as land use effects the water infiltration and evapotranspiration from the soil.

Data availability. **Murrumbidgee Soil Moisture Monitoring Network:** https://www.oznet.org.au,
Agriculture Victoria Soil Monitoring Network: www.extensionaus.com.au/soilmoisturemonitoring/,
Australian Cosmic-ray Neutron Soil Moisture Monitoring Network (CosmOz): www.cosmoz.ternlandscapes.net.au/sites,
AIR EP: www.airep.com.au/research/moisture-probes-weather-stations/

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

Author contributions. **Qazi Muqeet Amir:** Conceptualisation, Methodology, Software, Formal analysis, Writing - Original Draft and Editing. **Floris F. Van Ogtrop:** Methodology, Writing - Editing and Review, Supervision. **Thomas F.A. Bishop:** Conceptualisation, Methodology, Writing - Editing and Review, Supervision, Funding acquisition

Competing interests. The authors declare that they have no conflict of interest.

Disclaimer. This work was fully supported by the Australian Grains Research & Development Corporation (GRDC) project "Soil Water Nowcasting for the Grains Industry" (GRDC code: UOS2002–001RTX).

Acknowledgements. The authors would like to thank GRDC for their funding of the "Soil Water Nowcasting for the Grains Industry" project this research was made possible through. We thank the researchers at the Murrumbidgee Soil Moisture Monitoring Network, Agriculture Victoria, Ag Innovation and Development Eyre Peninsula, CSIRO, and Viridis Ag for establishing and continuing to make soil water probe data freely available. Finally, we thank Yi Yu, Ratneel Dao, Dhahi Al-Shammari, Penghui Wen, and the rest the members of Precision Agriculture, Hydrology, and Geoinformatics Lab (PAHGISL) at the University of Sydney for their constant support towards incremental improvements to this research.

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

References

Adeyemi, O., Grove, I., Peets, S., Domun, Y., and Norton, T.: Dynamic Neural Network Modelling of Soil Moisture Content for Predictive Irrigation Scheduling, Sensors, 18, 3408, https://doi.org/10.3390/s18103408, number: 10, 2018.

Agriculture Victoria: Soil Moisture Monitoring, https://extensionaus.com.au/soilmoisturemonitoring/.

Ahmed, A. A. M., Deo, R. C., Raj, N., Ghahramani, A., Feng, Q., Yin, Z., and Yang, L.: Deep Learning Forecasts of Soil Moisture: Convolutional Neural Network and Gated Recurrent Unit Models Coupled with Satellite-Derived MODIS, Observations and Synoptic-Scale Climate Index Data, Remote Sensing, 13, 554, https://doi.org/10.3390/rs13040554, number: 4, 2021.

AIR EP: Resilient EP - for a Profitable Farming Future, https://airep.com.au/research/resilient-ep/, 2019.

Amir, Q. M., Moloney, J. P., Van Ogtrop, F., and Bishop, T. F.: Remote and retroactive soil moisture probe calibration, [Manuscript submitted for publication], 2026.

Brinkhoff, J., Hornbuckle, J., and Ballester, C.: Soil moisture forecasting for irrigation recommendation, IFAC Proceedings Volumes (IFAC Papers-OnLine), 52, 385–390, https://doi.org/10.1016/j.ifacol.2019.12.586, 2019.

Cai, Y., Zheng, W., Zhang, X., Zhangzhong, L., and Xue, X.: Research on soil moisture prediction model based on deep learning, PLoS ONE, 14, e0214 508, https://doi.org/10.1371/journal.pone.0214508, 2019.

Casolaro, A., Capone, V., Iannuzzo, G., and Camastra, F.: Deep Learning for Time Series Forecasting: Advances and Open Problems, Information, 14, 598, https://doi.org/10.3390/info14110598, number: 11, 2023.

Chan, S. K., Bindlish, R., O'Neill, P. E., Njoku, E., Jackson, T., Colliander, A., Chen, F., Burgin, M., Dunbar, S., Piepmeier, J., Yueh, S., Entekhabi, D., Cosh, M. H., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., McNairn, H., Thibeault, M., Martínez-Fernández, J., González-Zamora, , Seyfried, M., Bosch, D., Starks, P., Goodrich, D., Prueger, J., Palecki, M., Small, E. E., Zreda, M., Calvet, J.-C., Crow, W. T., and Kerr, Y.: Assessment of the SMAP Passive Soil Moisture Product, IEEE Transactions on Geoscience and Remote Sensing, 54, 4994–5007, https://doi.org/10.1109/TGRS.2016.2561938, 2016.

Chatterjee, S., Dey, N., and Sen, S.: Soil Moisture Quantity Prediction using Optimized Neural Supported model for Sustainable Agricultural Applications, Sustainable Computing: Informatics and Systems, 28, https://doi.org/10.1016/j.suscom.2018.09.002, 2018.

Cho, K., Merrienboer, B. v., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y.: Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation, https://doi.org/10.48550/arXiv.1406.1078, arXiv:1406.1078, 2014.

Cinkus, G., Mazzilli, N., Jourde, H., Wunsch, A., Liesch, T., Ravbar, N., Chen, Z., and Goldscheider, N.: When best is the enemy of good – critical evaluation of performance criteria in hydrological models, Hydrology and Earth System Sciences, 27, 2397–2411, https://doi.org/10.5194/hess-27-2397-2023, 2023.

Dadrasi, A., Soltani, E., Makowski, D., and Lamichhane, J. R.: Does shifting from normal to early or late sowing dates provide yield benefits? A global meta-analysis, Field Crops Research, 318, 109 600, https://doi.org/10.1016/j.fcr.2024.109600, 2024.

Datta, P. and Faroughi, S. A.: A multihead LSTM technique for prognostic prediction of soil moisture, Geoderma, 433, 116 452, https://doi.org/10.1016/j.geoderma.2023.116452, 2023.

Dong, Z., Liu, Y., Li, M., Ci, B., Lu, X., Feng, X., Wen, S., and Ma, F.: Effect of different NPK fertilization timing sequences management on soil-petiole system nutrient uptake and fertilizer utilization efficiency of drip irrigation cotton, Scientific Reports, 13, 14 287, https://doi.org/10.1038/s41598-023-40620-9, 2023.

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

Dorigo, W., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A., Zamojski, D., Cordes, C., Wagner, W., and Drusch, M.: Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network, Vadose Zone Journal, 12, vzj2012.0097, https://doi.org/10.2136/vzj2012.0097, _eprint: https://acsess.onlinelibrary.wiley.com/doi/pdf/10.2136/vzj2012.0097, 2013.

Dubois, A., Teytaud, F., and Verel, S.: Short term soil moisture forecasts for potato crop farming: A machine learning approach, Computers and Electronics in Agriculture, 180, 105 902, https://doi.org/10.1016/j.compag.2020.105902, 2021.

Egrioglu, E. and Bas, E.: A new deep neural network for forecasting: Deep dendritic artificial neural network, Artificial Intelligence Review, 57, 171, https://doi.org/10.1007/s10462-024-10790-7, 2024.

Fan, J., McConkey, B., Wang, H., and Janzen, H.: Root distribution by depth for temperate agricultural crops, Field Crops Research, 189, 68–74, https://doi.org/10.1016/j.fcr.2016.02.013, 2016.

Fang, Q., Wang, Y., Uwimpaye, F., Yan, Z., Li, L., Liu, X., and Shao, L.: Pre-sowing soil water conditions and water conservation measures affecting the yield and water productivity of summer maize, Agricultural Water Management, 245, 106 628, https://doi.org/10.1016/j.agwat.2020.106628, 2021.

Flohr, B. M., Ouzman, J., McBeath, T. M., Rebetzke, G. J., Kirkegaard, J. A., and Llewellyn, R. S.: Redefining the link between rainfall and crop establishment in dryland cropping systems, Agricultural Systems, 190, 103 105, https://doi.org/10.1016/j.agsy.2021.103105, 2021.

Freund, M., Henley, B. J., Karoly, D. J., Allen, K. J., and Baker, P. J.: Multi-century cool- and warm-season rainfall reconstructions for Australia's major climatic regions, Climate of the Past, 13, 1751–1770, https://doi.org/10.5194/cp-13-1751-2017, 2017.

Gasch, C. K., Brown, D. J., Brooks, E. S., Yourek, M., Poggio, M., Cobos, D. R., and Campbell, C. S.: A pragmatic, automated approach for retroactive calibration of soil moisture sensors using a two-step, soil-specific correction, Computers and Electronics in Agriculture, 137, 29–40, https://doi.org/10.1016/j.compag.2017.03.018, 2017.

Goldstein-Greenwood, J.: The Shortcomings of Standardized Regression Coefficients | UVA Library, University of Virginia Library, https://library.virginia.edu/data/articles/the-shortcomings-of-standardized-regression-coefficients, 2023.

Granata, F., Di Nunno, F., Najafzadeh, M., and Demir, I.: A Stacked Machine Learning Algorithm for Multi-Step Ahead Prediction of Soil Moisture, Hydrology, 10, 1, https://doi.org/10.3390/hydrology10010001, number: 1, 2023.

GRDC: GRDC GROWNOTES: Managing frost risk, https://grdc.com.au/__data/assets/pdf_file/0027/208674/grdc-managing-frost-risk-tips-and-tactics-frost-050216-northen-southern-and-western-region.pdf.pdf, 2016.

Guan, X. and Burton, H.: Bias-variance tradeoff in machine learning: Theoretical formulation and implications to structural engineering applications, Structures, 46, 17–30, https://doi.org/10.1016/j.istruc.2022.10.004, 2022.

Guerschman, J. P., McVicar, T. R., Vleeshower, J., Van Niel, T. G., Peña-Arancibia, J. L., and Chen, Y.: Estimating actual evapotranspiration at field-to-continent scales by calibrating the CMRSET algorithm with MODIS, VIIRS, Landsat and Sentinel-2 data, Journal of Hydrology, 605, 127 318, https://doi.org/10.1016/j.jhydrol.2021.127318, 2022.

Han, H., Choi, C., Kim, J., Morrison, R. R., Jung, J., and Kim, H. S.: Multiple-Depth Soil Moisture Estimates Using Artificial Neural Network and Long Short-Term Memory Models, Water, 13, 2584, https://doi.org/10.3390/w13182584, 2021.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning, 0172-7397, Springer New York, NY, 2 edn., ISBN 978-0-387-84857-0, https://link.springer.com/book/10.1007/978-0-387-84858-7#bibliographic-information, 2009.

Hawdon, A., McJannet, D., and Wallace, J.: Calibration and correction procedures for cosmic-ray neutron soil moisture probes located across Australia, Water Resources Research, 50, 5029–5043, https://doi.org/10.1002/2013WR015138, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2013WR015138, 2014.

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

Hochman, Z. and Horan, H.: Causes of wheat yield gaps and opportunities to advance the water-limited yield frontier in Australia, Field Crops Research, 228, 20–30, https://doi.org/10.1016/j.fcr.2018.08.023, 2018.

Hoskin, N., Filippi, P., and Bishop, T. F. A.: How big is your bucket? Building soil-type specific pedotransfer functions to estimate soil water holding properties across Australia [Manuscript submitted for publication], Soil Research, 2025.

Jakob, D. and Walland, D.: Variability and long-term change in Australian temperature and precipitation extremes, Weather and Climate Extremes, 14, 36–55, https://doi.org/10.1016/j.wace.2016.11.001, 2016.

Jeffrey, S. J., Carter, J. O., Moodie, K. B., and Beswick, A. R.: Using spatial interpolation to construct a comprehensive archive of Australian climate data, Environmental Modelling & Software, 16, 309–330, https://doi.org/10.1016/S1364-8152(01)00008-1, 2001.

Jiang, S., Chen, G., Chen, D., and Chen, T.: Application and Evaluation of an Improved LSTM Model in the Soil Moisture Prediction of Southeast Chinese Tobacco-Producing Areas, Journal of the Indian Society of Remote Sensing, 51, 1843–1853, https://doi.org/10.1007/s12524-021-01438-y, 2023.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, Journal of Hydrology, 424-425, 264–277, https://doi.org/10.1016/j.jhydrol.2012.01.011, 2012.

Li, L., Dai, Y., Shangguan, W., Wei, N., Wei, Z., and Gupta, S.: Multistep Forecasting of Soil Moisture Using Spatiotemporal Deep Encoder–Decoder Networks, Journal of Hydrometeorology, 23, 337–350, https://doi.org/10.1175/JHM-D-21-0131.1, 2022a.

Li, Q., Li, Z., Shangguan, W., Wang, X., Li, L., and Yu, F.: Improving soil moisture prediction using a novel encoder-decoder model with residual learning, Computers and Electronics in Agriculture, 195, 106 816, https://doi.org/10.1016/j.compag.2022.106816, 2022b.

Li, Q., Zhu, Y., Shangguan, W., Wang, X., Li, L., and Yu, F.: An attention-aware LSTM model for soil moisture and soil temperature prediction, Geoderma, 409, 115 651, https://doi.org/10.1016/j.geoderma.2021.115651, 2022c.

Lindemann, B., Müller, T., Vietz, H., Jazdi, N., and Weyrich, M.: A survey on long short-term memory networks for time series prediction, Procedia CIRP, 99, 650–655, https://doi.org/10.1016/j.procir.2021.03.088, 2021.

Liu, D.: A rational performance criterion for hydrological model, Journal of Hydrology, 590, 125 488, https://doi.org/10.1016/j.jhydrol.2020.125488, 2020.

Liu, D., Shi, Z., Ma, Q., Zhang, Y., Cai, T., Zhang, P., and Jia, Z.: Strategy for matching fertilizer application with soil water before sowing can stabilize maize productivity under rainwater harvesting and mulching planting in dry areas: A six-year field experiment, Agricultural Water Management, 287, 108 452, https://doi.org/10.1016/j.agwat.2023.108452, 2023.

Maclin, R. and Opitz, D.: Popular Ensemble Methods: An Empirical Study, Journal of Artificial Intelligence Research, 11, 169–198, https://doi.org/10.1613/jair.614, arXiv:1106.0257 [cs], 1999.

Malone, B. P., Searle, R., Stenson, M., McJannet, D., Zund, P., Román Dobarco, M., Wadoux, A. M. J. C., Minasny, B., McBratney, A., and Grundy, M.: Update and expansion of the soil and landscape grid of Australia, Geoderma, 455, 117 226, https://doi.org/10.1016/j.geoderma.2025.117226, 2025.

McKenzie, N., Jacquier, D., Isbell, R., and Brown, K.: Australian Soils and Landscapes: An Illustrated Compendium, CSIRO Publishing, ISBN 978-0-643-10073-2, https://doi.org/10.1071/9780643100732, 2004.

Meteorology, s. A. c. G.-B. o.: Climate classification maps, Bureau of Meteorology, http://www.bom.gov.au/climate/maps/averages/climate-classification/?maptype=kpngrp, archive Location: Australia, a.

Meteorology, s. A. c. G.-B. o.: Rainfall variability maps, Bureau of Meteorology, http://www.bom.gov.au/climate/maps/averages/rainfall-variability/?period=apr, archive Location: Australia, b.

https://doi.org/10.5194/egusphere-2026-2438
Preprint. Discussion started: 20 May 2026

Nicholls, N., Drosdowsky, W., and Lavery, B.: Australian rainfall variability and change, Weather, 52, 66–72, https://doi.org/10.1002/j.1477-8696.1997.tb06274.x, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1477-8696.1997.tb06274.x, 1997.

Prakash, S., Sharma, A., and Sahu, S. S.: Soil Moisture Prediction Using Machine Learning, in: 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), pp. 1–6, https://doi.org/10.1109/ICICCT.2018.8473260, 2018.

Prasad, R., Deo, R. C., Li, Y., and Maraseni, T.: Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition, Geoderma, 330, 136–161, https://doi.org/10.1016/j.geoderma.2018.05.035, 2018.

Scanlon, T. T. and Doncon, G.: Rain, rain, gone away: decreased growing-season rainfall for the dryland cropping region of the south-west of Western Australia, Crop & Pasture Science, 71, 128–133, https://doi.org/10.1071/CP19294, 2020.

Smith, A. B., Walker, J. P., Western, A. W., Young, R. I., Ellett, K. M., Pipunic, R. C., Grayson, R. B., Siriwardena, L., Chiew, F. H. S., and Richter, H.: The Murrumbidgee soil moisture monitoring network data set, Water Resources Research, 48, https://doi.org/10.1029/2012WR011976, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2012WR011976, 2012.

Tilse, M. J., Bishop, T. F. A., and Filippi, P.: Predicting within-field grain protein content at scale using agronomic and remote sensing variables, and machine learning, Precision Agriculture, 26, 78, https://doi.org/10.1007/s11119-025-10267-9, 2025.

Togneri, R., Felipe dos Santos, D., Camponogara, G., Nagano, H., Custódio, G., Prati, R., Fernandes, S., and Kamienski, C.: Soil moisture forecast for smart irrigation: The primetime for machine learning, Expert Systems with Applications, 207, 117 653, https://doi.org/10.1016/j.eswa.2022.117653, 2022.

Tyralis, H., Papacharalampous, G., and Langousis, A.: A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources, Water, 11, 910, https://doi.org/10.3390/w11050910, number: 5, 2019.

Viridis Ag: Viridis Ag – Better Crops – Better Future, https://viridisag.com/.

Wang, J., Sun, Z., Yang, T., Wang, B., Dou, W., and Zhu, W.: Quantifying the effect of salinity on dielectric-based soil moisture measurements using COSMOS records, Journal of Hydrology, 643, 131 925, https://doi.org/10.1016/j.jhydrol.2024.131925, 2024.

Wilford, J.: A weathering intensity index for the Australian continent using airborne gamma-ray spectrometry and digital terrain analysis, Geoderma, 183-184, 124–142, https://doi.org/10.1016/j.geoderma.2010.12.022, 2012.

WMS Engineering: WMS-Engineering/AusMap, https://github.com/WMS-Engineering/AusMap, original-date: 2017-08-24T22:32:01Z, 2026.

Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H., and Deng, S.-H.: Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimizationb, Journal of Electronic Science and Technology, 17, 26–40, https://doi.org/10.11989/JEST.1674-862X.80904120, 2019.

Yu, J., Xin, Z., Xu, L., Dong, J., and Zhangzhong, L.: A hybrid CNN-GRU model for predicting soil moisture in maize root zone, Agricultural Water Management, 245, https://doi.org/10.1016/j.agwat.2020.106649, 2020.

Zhang, X., Dong, Z., Wu, X., Gan, Y., Chen, X., Xia, H., Kamran, M., Jia, Z., Han, Q., Shayakhmetova, A., and Siddique, K. H. M.: Matching fertilization with water availability enhances maize productivity and water use efficiency in a semi-arid area: Mechanisms and solutions, Soil and Tillage Research, 214, 105 164, https://doi.org/10.1016/j.still.2021.105164, 2021.