



Article

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

A. A. Masrur Ahmed ¹, Ravinesh C Deo ^{1,*}, Nawin Raj ¹, Afshin Ghahramani ¹, Qi Feng ², Zhenliang Yin ² and Linshan Yang ²

¹ School of Science, University of Southern Queensland, Springfield, QLD 4300, Australia; abulabramasrur.ahmed@usq.edu.au (A.A.M.A.); Nawin.raj@usq.edu.au (N.R.); afshin.ghahramani@usq.edu.au (A.G.)

² Key Laboratory of Ecohydrology of Inland River Basin and Northwest, Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China; qifeng@lzb.ac.cn (Q.F.); yinzhenliang@lzb.ac.cn (Z.Y.); yanglsh08@lzb.ac.cn (L.Y.)

* Correspondence: ravinesh.deo@usq.edu.au; Tel.: +61-734704430



Citation: Ahmed, A.A.M.; Deo, R.C.; Raj, N.; Ghahramani, A.; Feng, Q.; Yin, Z.; 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 Sens.* **2021**, *13*, 554. <https://doi.org/10.3390/rs13040554>

Academic Editor:

Nemesio Rodriguez-Fernandez

Received: 16 December 2020

Accepted: 27 January 2021

Published: 4 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract: Remotely sensed soil moisture forecasting through satellite-based sensors to estimate the future state of the underlying soils plays a critical role in planning and managing water resources and sustainable agricultural practices. In this paper, Deep Learning (DL) hybrid models (i.e., CEEMDAN-CNN-GRU) are designed for daily time-step surface soil moisture (SSM) forecasts, employing the gated recurrent unit (GRU), complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), and convolutional neural network (CNN). To establish the objective model's viability for SSM forecasting at multi-step daily horizons, the hybrid CEEMDAN-CNN-GRU model is tested at 1st, 5th, 7th, 14th, 21st, and 30th day ahead period by assimilating a comprehensive pool of 52 predictor dataset obtained from three distinct data sources. Data comprise satellite-derived Global Land Data Assimilation System (GLDAS) repository a global, high-temporal resolution, unique terrestrial modelling system, and ground-based variables from Scientific Information Landowners (SILO) and synoptic-scale climate indices. The results demonstrate the forecasting capability of the hybrid CEEMDAN-CNN-GRU model with respect to the counterpart comparative models. This is supported by a relatively lower value of the mean absolute percentage and root mean square error. In terms of the statistical score metrics and infographics employed to test the final model's utility, the proposed CEEMDAN-CNN-GRU models are considerably superior compared to a standalone and other hybrid method tested on independent SSM data developed through feature selection approaches. Thus, the proposed approach can be successfully implemented in hydrology and agriculture management.

Keywords: deep learning algorithm; MODIS; gated recurrent unit; satellite models of soil moisture



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The precise requirements for water resource supply, constant monitoring, and forecasting are changing continuously with population growth, agricultural and human activities. Any variations in weather and perturbations in climate patterns due to anthropogenically-induced factors affect usable water distribution and accessibility. Instead of precipitation playing a paramount role, the terrestrial water basin tends to dominate the actual functioning of the hydrological, ecological, and inter-coupled socio-economic systems [1]. Notably, the knowledge of fundamental components of water reservoirs, e.g., soil moisture (SM) and streamflow, is essential for an effective water resources management strategy. SM also governs the physical interactions between land and the atmosphere [2,3] and acts as

a driver to feed irrigation systems [4], grazing and crop yield predictions [5]. A decline in groundwater reduces soil water content and the storage volume in underlying soils. A lack of soil moisture can affect agricultural and hydro-meteorological processes. Therefore, predictive models providing prior information on monitoring and forecasting water, such as in this study, are critical to soil moisture forecasts as a principal regulating factor in groundwater hydrology to understand the soil's future state.

With increasing computer power, researchers are developing intelligent models to extract features in historical data (e.g., SM). Such models demonstrate acceptable skills in forecasting hydro-metrological variables, e.g., precipitation [6–9], drought [10], streamflow [11,12], runoff [13,14], floods [15,16], soil moisture [17], water demand and water quality [18–21]. However, very few studies have focused on the prediction of soil moisture, with most examples being the artificial neural networks (ANN) [22] and the extreme learning machines (ELM) [23]. Irrespective of the model type and domain of applications, accurately forecasted soil moisture presents a greater understanding of water resources and agricultural management, leading to more sustainable decisions. Intelligent systems based on deep learning utilise feature extraction and reveal the compounded association between predictors and targets [24]. Hence, soil moisture prediction with advanced algorithms is a highly practical tool for agricultural water management. DL methods, however, are yet to be explored in the present study region (i.e., Australian Murray Darling Basin). In this study, we adopt a gated recurrent unit (GRU) neural networks as a modified long-short term memory (LSTM) that has attracted good research attention [25]. There appear to be only a few studies on GRU-based models, especially in hydrology [26,27]. Convolutional Neural Networks (CNNs) is a useful feature extraction method to improve the overall predictive process [28]. Therefore, an integration of CNN and GRU can, in foreseeable possibilities, lead to a robust pre-processing of data providing a viable option to improve the model's forecasting skill. This has been evident in some studies that integrated CNN with LSTM for improved performance, with Ghimire et al. [28] showing the superior skill of the CNN-LSTM model in the problem of solar radiation. Integration of deep learning (i.e., CNN-GRU) for soil moisture forecasting is yet to be tested explicitly, with no studies previously using this method, the focus of this study.

Given the stochastic nature of hydrological variables, multi-resolution analysis (MRA) can enhance any model's performance as a tool to reveal the data features. Conventional MRA, for example, discrete wavelet transforms (DWT), have long been implemented [29–32]. However, DWT appears to have drawbacks, and this critical issue is resolved by the maximum-overlap discrete wavelet transform (MODWT), an advanced DWT method [11,33,34]. In this study, we adopt an improved version of EMD, i.e., complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) to implement a self-adaptive decomposition of the predictor variables [23]. In CEEMDAN-based decomposition, a coefficient representing Gaussian white noise with a unit variance is added consecutively at each stage to reduce the forecasting procedure's complexity, avoiding the time series' intricacy [35]. Previous studies used CEEMDAN in forecasting soil moisture [23,36] with an earlier version (i.e., EEMD) used in forecasting streamflow [37] and rainfall [38–40]. Moreover, The multivariate empirical mode decomposition (MEMD) is a self-adaptive algorithm that establishes multivariate inputs to perform a proper investigation [41]. The MEMD method has been successfully applied in time series forecasting [42,43]. The study incorporates the CEEMDAN method as neither the EEMD nor the CEEMDAN decomposition approach has been assimilated with any deep learning approach (i.e., GRU) to produce a soil moisture forecast system, as attempted in the present study.

Climate indices have long been recognised as a useful synoptic-scale indicator of teleconnections representing climate variability [9,44]. La Niña, represented by climate indices, is accountable for substantial rainfall in eastern Australia, whereas the El Niño phenomenon is related to drought [45]. However, El Niño Southern Oscillation (ENSO) has a potential impact on precipitation in northern and eastern Australia [46]. Considering the substantial effects of ENSO phenomena on Australia's climate variability, some

studies [9,47,48] have correlated ENSO effects with hydrological variables (e.g., streamflow, rainfall, and droughts). Rashid et al. [49] aimed to predict a drought index in Australian catchments by aggregating synoptic-scale climate mode indices. Considering these studies, the design of an artificial intelligence model utilising synoptic-scale climate indices, as done in this paper, can be of great practical value in developing sustainable river systems and drought management strategies.

In our paper, we rely on satellite (i.e., MODIS) sensors providing a flexible remote system to explore the nexus between physical, chemical, and biological parameters related to ground variables (i.e., observations) and how these affect future changes in daily soil moisture. However, the inclusion of three distinct datasets has a high potential to address the uncertainties in the predictor variables, especially the remote sensing data's errors. The variables from satellite sensors are associated with errors that propagate to the prediction of hydrologic variables [50–52]. To address this issue, it is preferable to integrate satellite and ground-based variables. Ghimire et al. [53] integrated GIOVANNI data with ECMWF Reanalysis to predict long-term solar radiation. However, the integration of satellite-based, ground-based SILO data, and climate indices for soil moisture forecasts, particularly with deep learning methods (e.g., LSTM), is yet to be implemented.

The objectives are, therefore, fourfold. (1) To build deep learning approaches to forecast surface soil moisture (SSM) at 2 cm depth, incorporating CEEMDAN (i.e., data splitting method) with CNN (i.e., feature extraction method) to generate a GRU-based predictive model. This predictive system, denoted as the CEEMDAN-CNN-GRU hybrid model, is improved with neighbourhood component analysis as a feature selection tenet on diverse predictors obtained from MODIS data, climate mode indices, and ground-based SILO product. (2) To adopt the hybrid CEEMDAN-CNN-GRU model for daily SSM forecasts at a multi-step horizon (i.e., 1st, 5th, 7th, 14th, 21st, and 30th day lead time). (3) To explore the contributory influence of climate indices on the accuracy of the CEEMDAN-CNN-GRU model. (4) To comprehensively benchmark the objective model against alternative tools such as the GRU standalone algorithm, CEEMDAN-GRU, and CNN-GRU hybrid model. This study's primary contribution is to generate a skilful deep learning method for soil moisture prediction, capitalising on remote sensing and ground data while capturing pertinent relationships between soil moisture and synoptic-scale drivers of climate variability in the Australian Murray Darling Basin.

2. Materials and Methods

2.1. Theoretical Frameworks

2.1.1. Convolutional Neural Network

To build the CEEMDAN-CNN-GRU hybrid model trained for daily SSM forecasts, this study purposely employs the Convolutional Neural Networks (CNN) for optimal feature extraction from the input dataset. CNN's have some similarities with conventional neural networks. They are, however, different in their connectivity between and within neuronal layers. In conventional neural networks, every neuron is wholly connected to all neurons in prior layers, whereas single layer neurons do not contribute to the model's network. CNN's are similar to Feed Forward Neural Networks [54], with its model architecture having three layers based on pooling, convolutional, and fully connected layer settings.

The connected layer is employed to estimate objective variables depending on the predictor variable's input features. CNN has proven to be a reliable modelling tool to extract hidden features in inputs and generating filters capturing data features in predictors [55]. To extract the pattern in an objective variable (i.e., SSM) and associated predictor variables, each convolutional layer is established as follows [56]:

$$h_{ij}^k = f \left(\left(W^k \times x \right)_{ij} + b_k \right) \quad (1)$$

Here, W^k is referred to as the weight of the kernel associated with k th feature map, f is the activation function, and the operator of the convolutional procedure is denoted by

multiplication sign (\times). The rectified linear unit (ReLU) is used as an activation function and the adaptive moment estimation (Adam) is selected as an optimisation algorithm using the grid search approach. The ReLU is described as:

$$f(x) = \max(0, x) \tag{2}$$

A one-dimensional convolutional operative was adopted to directly forecast the 1-Dimensional dataset, which eventually simplifies the modelling procedures for real-time forecasting execution.

2.1.2. Gated Recurrent Unit Network

The hybrid CEEMDAN-CNN-GRU model utilises Gated Recurrent Unit (GRU) neural network as the predictive tool after extracting features based on the CNN algorithm (Section 2.1.1). GRU is a distinct type of long short-term memory (LSTM) network presented by Cho et al. [57]. Along with similarities, GRU possesses different characteristics from the LSTM. For instance, the GRU owns two gates, namely the update gate and reset gate, whereas the LSTM has three gates (i.e., the input gate, forget gate, and output gate). Figure 1 provides a schematic of the hybrid CEEMDAN-CNN-GRU model with CEEMDAN data decomposition and model architecture. Moreover, Figure 1b shows the structure of the gated recurrent unit network.

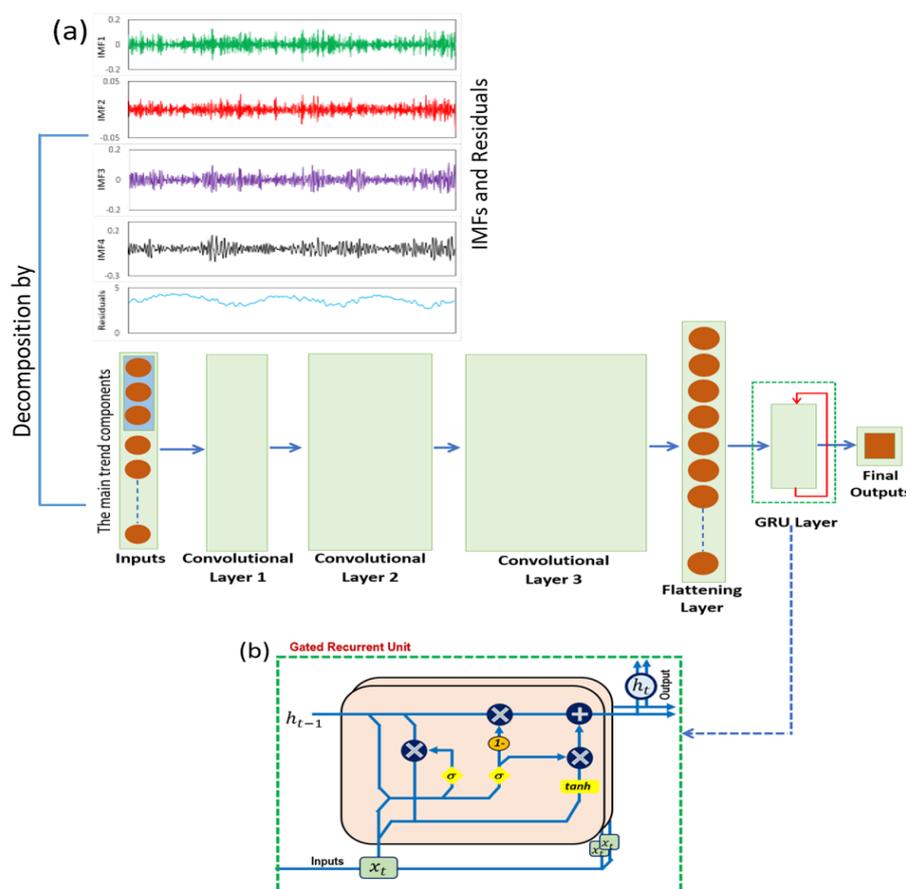


Figure 1. (a) Schematic of the hybrid CEEMDAN-CNN-GRU model with Complete Ensemble Empirical Model Decomposition (CEEMDAN), Convolutional Neural Networks (CNN), and Gated Recurrent Unit (GRU) Neural Network arrangement. The IMFs (Intrinsic Mode Functions) and residual series are generated in the CEEMDAN process, whereas the CNN algorithm represents the feature extraction stage. (b) 2-layered GRU model.

In a GRU Network, two input features, including the input vector $x(t)$ and output vector $h(t - 1)$, are present in each layer. The yield of each gate is achieved by logical operation and non-linear transformation of predictors. Moreover, the association between predictors and predictand can be defined as follows:

$$r(t) = \sigma_g(W_r x(t) + U_r h(t - 1) + b_r) \quad (3)$$

$$z(t) = \sigma_g(W_z x(t) + U_z h(t - 1) + b_z) \quad (4)$$

$$h(t) = (1 - z(t))o(t - 1) + z(t)o\hat{h}(t) \quad (5)$$

$$\hat{h}(t) = \sigma_h(W_h x(t) + U_h(r(t))o h(t - 1)) + b_h \quad (6)$$

where $r(t)$ is the reset gate vector, $z(t)$ is defined as the update gate vector, W and U are parameter metrics and vector. σ_h is referred to as a hyperbolic tangent, and σ_g is defined as a sigmoid function. Finally, given the architecture of GRU, a training approach is chosen, which includes backpropagation through time. Based on previous studies, *Adam* optimiser was implemented as it has enhanced expertise.

2.1.3. Hybrid CNN–GRU. Neural Network

In this paper, the hybrid modelling approach utilises a deep learning method built upon a feature extraction procedure under a forecast model framework. This research demonstrates how the CNN–GRU model comprised of three-layered CNN is used for feature extraction to generate future changes in the objective variable (i.e., SSM). In particular, the GRU layer is employed to integrate input features extracted by the CNN algorithm to finally forecast the target variable (i.e., SSM) with minimal training and testing error.

2.1.4. Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)

As elucidated in Section 1, CEEMDAN is adopted as an improved version of EMD and EEMD to perform a self-adaptive decomposition of model input signals [23] prior to modelling the target variable. The CEEMDAN decomposition process commences by discretising the n -length inputs of any model $\chi(t)$ into intrinsic mode functions (IMFs) and residues to comply with tolerability provision. Nevertheless, to ensure no leakage of information in the IMFs and residues from the training series into the future (i.e., testing and validation subset), the decomposition is performed separately for training, validation, and testing. The actual IMF is produced by taking the mean of the EMD-grounded I.M.F.s across a trial and the combination of white noise to model the predictor–target variables.

Assume that we have D -dimensional set, with n -length X_i matrix (i.e., inputs selected by two-phase decomposed sub-series achieved during the decomposition) and the 1-dimensional surface soil moisture as the target variable. The difference between CEEMDAN and EEMD is that in the CEEMDAN case, a restricted noise (ε_i) across $[0, 1]$ is included at every single decomposition stage, calculated to induce the IMF to take the lead to insignificant error. Considering $E_j(\cdot)$ as an operator producing J th modes obtained from EMD, we follow Torres et al. [58] to implement the CEEMDAN process as follows:

Step 1: The decomposition of p -realizations of $\chi[n] = \varepsilon_1 \omega^p[n]$ using EMD to develop their first intrinsic approach, as explained according to the equation:

$$IMF_1[n] = \frac{1}{p} \sum_{p=1}^p IMF_1^p[n] = \overline{IMF}_1[n] \quad (7)$$

Step 2: Putting $k = 1$, the 1st residue is computed following Equation (7).

$$Res_1[n] = \chi[n] - IMF_1[n] \quad (8)$$

Step 3: Putting $k = 2$, the 2nd residual is obtained as:

$$IM\hat{F}_2[n] = \frac{1}{p} \sum_{p=1}^P E_1(r_1[n] + \varepsilon_1 E_1(\omega^p[n])) \quad (9)$$

Step 4: Setting $k = 2 \dots K$ calculates the k th residue as:

$$Res_k[n] = Res_{k-1}[n] - IM\hat{F}_k[n] \quad (10)$$

Step 5: Now we decompose the realisations $Res_k[n] + \varepsilon_1 E_1(\omega^p[n])$, Here, $k = 1, \dots, K$ until their first model of EMD is reached; here, the $(k + 1)$ is:

$$IM\hat{F}_{(k+1)}[n] = \frac{1}{p} \sum_{p=1}^P E_1(r_k[n] + \varepsilon_k E_k(\omega^p[n])) \quad (11)$$

Step 6: Now the k value is incremented, and steps 4–6 are repeated. Consequently, the final residue is achieved:

$$RES_k[n] = \chi[n] - \sum_{k=1}^K IM\hat{F}_k \quad (12)$$

Here, K is defined as the limiting case (i.e., the highest number of modes). To comply with the replicability of the earliest input, $\chi[n]$, the following is performed for the CEEMDAN approach.

$$\chi[n] = \sum_{k=1}^K IM\hat{F}_k + RES_k[n] \quad (13)$$

The additive noise demonstrates that signal-to-noise ratio (ε) is operated at every phase [59,60] and must connect the low magnitude with high-frequency signals in the data [61,62]. Figure 1a provides the CEEMDAN decomposed IMFs and residuals and CNN architecture.

2.1.5. Feature Selection: Neighbourhood Component Analysis

The selection of features within the inputs used to forecast soil moisture is vital in applying a predictive model. This is implemented to reduce the dimensionality of model inputs and computational cost, including the desired improvements in the forecasting accuracy and interpretation of the predictive model characteristics and nature of its predictors [59,63–65]. This study has adopted Neighbourhood Component Analysis (NCA) based on regressions applied to segregate the potential input variables from 52 predictor variables. Introduced by Yang et al., this method uses a competent, non-rectilinear, and non-parametric implanted approach. The MATLAB function called “*fsrnca*” performs NCA feature selection with regularisation to learn feature weights for the minimisation of an objective function that measures the average ‘leave-one-out’ regression loss over the training data. The NCA process’s *fsrnca* approach is adopted to train a variable set to better understand the importance of features through weight by minimising the objective function and calculating the regression loss of predictive model for soil moisture forecasts.

Consider training a dataset $T = \{(x_i, y_i) : i = 1, 2, 3, \dots, N\}$ where $x_i \in R^P$ is the feature vectors (i.e., predictor variables), $y_i \in R$ is the target (i.e., SSM), and N is the sample number for the training set. A function $g(x) : R^P \rightarrow R$ is absorbed by *fsrnca* algorithm to forecast the response y from several input variables, optimising their nearest spaces. The weighted distance (D_w) amongst any two samples is calculated as:

$$D_w(x_a, x_b) = \sum_{j=1}^J w_j^2 |x_a, x_b| \quad (14)$$

where x_a and x_b are the two samples used during training, and w_j is defined as the weight-related to the j th feature. Furthermore, the probability distribution ($p_{\alpha\beta}$) is employed to increase its leave-one-out forecasting correctness in the training phase. By contrast, the probability is that x_a chooses x_b as its reference argument. The algorithm acquires a weighting vector 'w' for gradient the ascent method to determine the feature subset with a regularisation factor to prevent overfitting.

3. Study Area and Data

3.1. Study Area and Description of Predictive Model Development Dataset

For the first time, this study aims to build a new forecast for daily surface soil moisture (SSM) with convolutional-gated recurrent unit neural networks within the Australian Murray Darling Basin (MDB). The MDB covers $\sim 1,042,730$ km² (or 14%) of mainland Australia [24,66] and $\sim 67\%$ of agricultural lands [67]. As illustrated (Figure 2), the sites are selected based on climate class and soil type diversity, namely Menindee, Deniliquin, Fairfield, and Gabo Island.

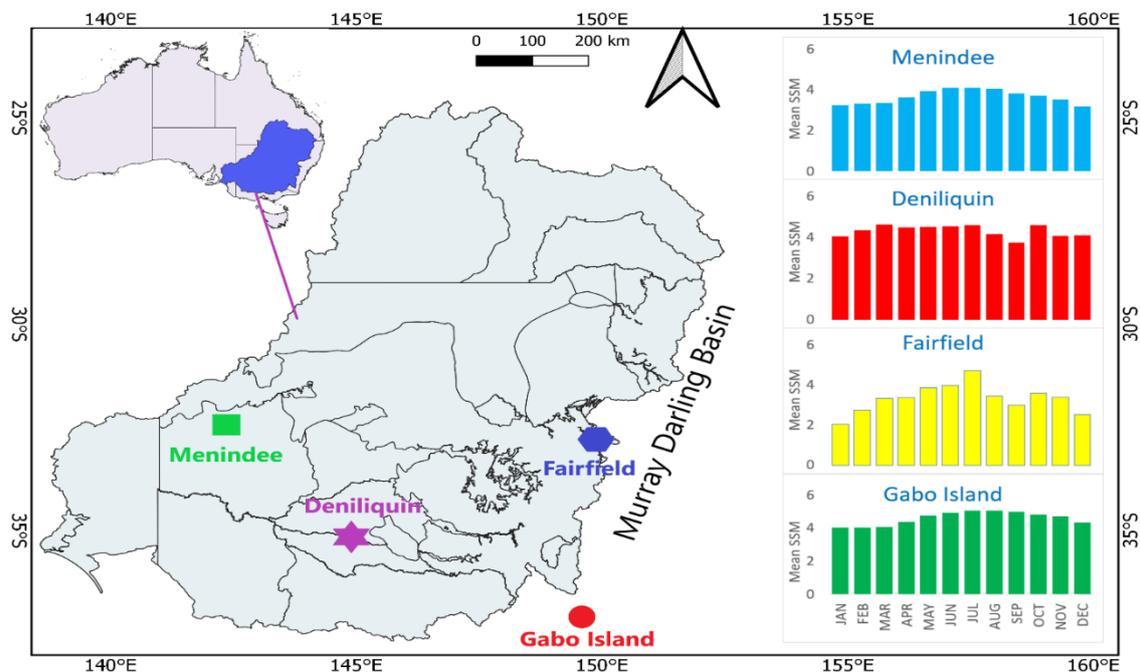


Figure 2. The Australian Murray Darling Basin with study sites and Surface Soil Moisture (SSM, kgm⁻²) where the hybrid CEEMDAN-CNN-GRU model at multi-step daily SSM forecasting.

The geographical locations and physical characteristics of the sites in Murray Darling Basin are tabulated in Table 1. It should be noted that the site Gabo Island is located at the border of the MDB region for comparison purposes with the other study stations, whereas ~ 20 lakes surround Menindee in a harsh desert environment. The site Fairfield lies within the savannah climate class with land-use patterns of dryland cropping [23]. Figure 2 also shows a histogram of monthly surface soil moisture patterns for the candidate sites.

Table 1. Geographic locations and physical characteristics of selected sites in the Murray Darling Basin.

Station Name	BOM Station ID	SILO Position (MODIS Grid Area)	Major Climate Class [68]	Soil Type [69]	Elevation [70]
Menindee	047019	32.39°S, 142.42°E (142.5°E, 32.5°S, 142.25°E, 32.25°S)	Desert	Calcarosol	61
Deniliquin	074128	35.53°S, 144.97° (145°E, 35.25°S, 144.75°E, 35°S)	Savannah	Calcarosol	94
Fairfield	066137	33.92°S, 150.98°E (149.75°E, 37.75°S, 150.0°E, 37.5°S)	Savannah	Vertosol	15
Gabo Island	084016	37.57°S, 149.92°E (150°E, 37.75°S, 149.75°E, 37.5°S)	Sub-Tropical	Sodosol	15

BOM = Bureau of Meteorology, Australia.

The appropriate selection of predictors related to the objective variable has a crucial role in predictive model design. To build a robust model, we adopt remotely sensed MODIS satellite-derived data identified as potential predictor variables in other studies, e.g., solar radiation prediction [24,71,72]. We consider different studies that demonstrate the potential utility of synoptic-scale climate indices that modulate Australian rainfall and crops [41,73,74]. This study integrates three unique data (i.e., satellite-derived data, climate indices, and ground-based variables) to capture a diverse suite of predictive features to forecast SSM, enabling the deep-learning approach a significant edge over the solely station-based models.

3.1.1. MODIS Satellite Dataset

Our hybrid deep learning model (i.e., CEEMDAN-CNN-GRU) is built upon NASA's Geospatial Online Interactive Visualization and Analysis Infrastructure (GIOVANNI) repository (1 February 2003 to 31 March 2020). GIOVANNI represents a powerful online visualization and analysis tool for geoscience datasets, capturing 2000 satellite variables [75,76]. In this study, MODIS-based predictor variables, presented in Table 2, are utilised to design and evaluate the hybrid CEEMDAN-CNN-GRU model for SSM forecasting. These are extracted from the GLDAS system representing the high-temporal resolution terrestrial modelling system consisting of the land surface state and several flux parameters with three temporal resolution products: hourly, daily, and monthly. Our study has used GLDAS 2.0 datasets extracted in daily temporal resolutions available publicly. The study utilised MODIS-based surface soil moisture (SSM) data as a target variable obtained from the GLDAS 2.0 model.

3.1.2. Scientific Information for Landowners (SILO) Dataset

To increase the pool of predictors, enabling effective feature engineering and increased performance of the DL model, this study selects nine meteorological variables from Scientific Information for Landowners (SILO): <https://www.longpaddock.qld.gov.au/silo/ppd/index.php> (accessed on 31 December 2020). SILO, managed by the Department of Environment and Science, Queensland Government [77], is popular for studying the Australian climate. Table 2 provides a list of SILO data.

3.1.3. Climate Indices

In previous studies, e.g., [9,29,59,74] on modelling precipitation, streamflow, and soil moisture, the role of synoptic-scale and climate indices were found significant in improving the overall model. In this study, twenty-one climate indices are thus obtained from many sources: National Climate Prediction Centre, Australian Bureau of Meteorology [70], and National Oceanic and Atmospheric Administration (NOAA) with daily sea

surface temperature (Nino1 + 2SST, Nino3SST, Nino3.4SST, Nino4SST) over 1 March 2003 to 31 March 2020 from KNMI Climate Explorer [78]. As the positive SOI is related to La-Nina and negative SOI concurs with El-Nino events [79,80], this study has used all of these indices due to strongly correlated rainfall with lagged SOI showing high predictability of rainfall from August–November [44,81]. To further enhance the predictive skill of the deep learning model, we consider Madden-Julian Oscillation (MJO) known to produce a substantial effect on tropical weather [70], which indeed entails a change in rainfall, wind, sea surface temperature (SST), and cloudiness [82]. Hence, eight daily MJO indices were adopted from KNMI. Climate Explorer [78], together with Interdecadal Pacific Oscillation (IPO), was introduced by Henley et al. [83], collected from NOAA National Climate Prediction Centre. Detailed information on climate indices and SSTs are in Table 2.

Table 2. Description of the global pool of 52 predictor variables used to design and evaluate hybrid CEEMDAN-CNN-GRU predictive model for daily surface soil moisture forecasting.

GLDAS 2.0: Modis Satellite Data from Giovanni Repository			
Predictor Variable	Notation	Description	Units
SurT	St	Average Surface Skin temperature	K
CSW	CW	Plant canopy surface water	Kg m ⁻²
CWE	CE	Canopy water evaporation	kg m ⁻² s ⁻¹
Esoil	Es	Direct Evaporation from Bare Soil	kg m ⁻² s ⁻¹
ET	ET	Evapotranspiration	kg m ⁻² s ⁻¹
Esnow	Es	Snow Evaporation	kg m ⁻² s ⁻¹
GWS	GW	Groundwater storage	mm
LWR.	LW	Net longwave radiation flux	W m ⁻²
Qg	Qg	Ground heat flux	W m ⁻²
Qh	Qh	Sensible heat net flux	W m ⁻²
Qle	Qle	Latent heat net flux	W m ⁻²
Qs	Qs	Storm surface runoff	Kg m ⁻² s ⁻¹
Qsb	Qb	Baseflow-groundwater runoff	Kg m ⁻² s ⁻¹
Qsm	Qm	Snow-melt	Kg m ⁻² s ⁻¹
Snd	Sn	Snow depth	m
Snt	Snt	Snow Surface temperature	m
SMp	Sp	Profile Soil moisture	Kg m ⁻²
SMrz	Sz	Root Zone Soil moisture	Kg m ⁻²
SSM	SSM	Surface Soil moisture	Kg m ⁻²
SWE	SW	Snow depth water equivalent	Kg m ⁻²
SWR	SR	Net short-wave radiation flux	W m ⁻²
Tra	Tr	Transpiration	Kg m ⁻² s ⁻¹
TWS	TW	Terrestrial water storage	mm
SILO (Ground-Based Observations)			
T.Max	Tx	Maximum Temperature	°C
T.Min	Tn	Minimum Temperature	°C
Rain	r	Rainfall	mm
Evap	Ep	Evaporation	mm
Radn	Rd	Radiation	MJ m ⁻²

Table 2. Cont.

GLDAS 2.0: Modis Satellite Data from Giovanni Repository			
Predictor Variable	Notation	Description	Units
VP	VP	Vapour Pressure	hPa
RHmaxT	Rx	Relative Humidity at Temperature T.Max	%
RHminT	Rn	Relative Humidity at Temperature T.Min	%
Mpot	Mp	Morton potential evapotranspiration overland	mm
SYNOPTIC-SCALE (Climate Mode Indices)			
Nino3.0	N3	Average SSTA over 150°–90°W and 5°N–5°S	
Nino3.4	N34	Average SSTA over 170°E–120°W and 5°N–5°S	
Nino4.0	N4	Average SSTA over 160°E–150°W and 5°N–5°S	
Nino1+2	N12	Average SSTA over 90°W–80°W and 0°–10°S	
AO	A	Arctic Oscillation	
AAO	AO	Antarctic Oscillation	
MJO1	MJ1	Madden Julian Oscillation-1	
MJO2	MJ2	Madden Julian Oscillation-2	
MJO4	MJ4	Madden Julian Oscillation-4	
MJO5	MJ5	Madden Julian Oscillation-5	
MJO6	MJ6	Madden Julian Oscillation-6	
MJO7	MJ7	Madden Julian Oscillation-7	
MJO8	MJ8	Madden Julian Oscillation-8	
MJO10	MJ10	Madden Julian Oscillation-10	
EPO	EP	East Pacific Oscillation	
GBI	G	Greenland Blocking Index (GBI)	
WPO	WP	Western Pacific Oscillation (WPO.)	
PNA	PN	Pacific North American Index	
NAO	N	North Atlantic Oscillation	
SAM	SM	Southern Annular Mode index	
SOI	SOI	Southern Oscillation Index, as per Troup [84]	

NONE

SSTA = Sea Surface Temperature anomalies (°C).

3.2. Predictive Model Development

To design a forecast model for SSM over multi-step periods of 1st, 5th, 7th, 14th, 21st, and 30th day lead time, three distinct datasets from satellites (i.e., GIOVANNI), climate indices, and ground source (SILO) for 17 years, 1 February 2003 to 31 March 2020 are used. Hybrid DL is implemented under Intel i7 @ 1.5 GHz and 16 GB memory. The proposed model algorithms were demonstrated using freely available DL libraries, namely the Keras [85,86] and TensorFlow [87] libraries. MATLAB 2020 software is used for Neighbourhood Component Analysis feature selection with packages matplotlib, and Minitab is used to visualise the forecasted SSM in the testing phase.

Data-driven models were built by normalising the input variables, transforming these predictors into a more consistent form [88]. To ensure the variable features were given proportional attention in network training, all were normalised [89] between (0, 1) [41,53,90].

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{15}$$

In Equation (15), x is the respective variable, x_{min} is the minimum value, x_{max} is the maximum and x_{norm} is the normalised value. After normalising the variables, the datasets are partitioned into training (February 2003–December 2013), validation (January 2014–December 2016), and testing (January 2017–March 2020) subsets. Figure 3 shows the methodological steps of the proposed CEEMDAN-CNN-GRU model. CEEMDAN is implemented in four stages.

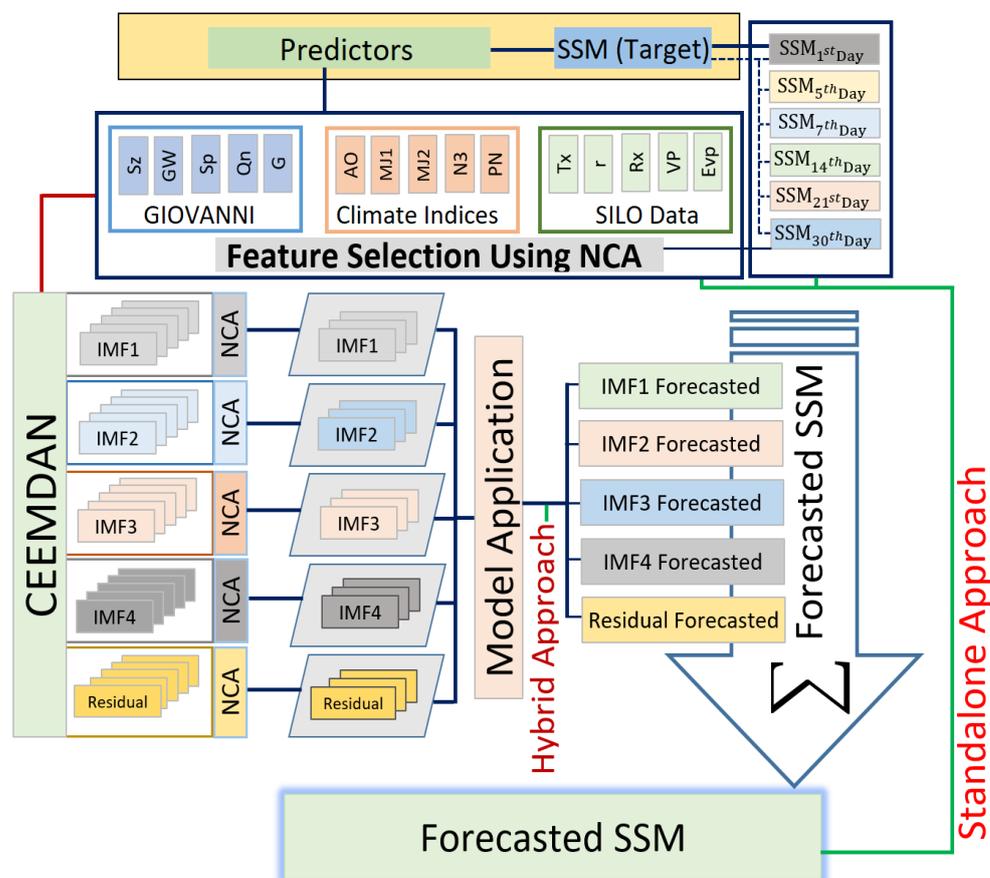


Figure 3. Workflow with the steps in model design for hybrid CEEMDAN-CNN-GRU predictive model. SSM = Surface Soil Moisture, NCA = neighbourhood component analysis for regression, IMF = Intrinsic Mode Function, CEEMDAN = Complete Ensemble Empirical Model Decomposition with adaptive noise, GRU = Gated Recurrent Units.

3.2.1. Feature Selection

By incorporating the MODIS satellite and ground and climate indices, this study has utilised 52 different predictors for SSM forecasting; hence, feature selection was crucial for data pre-processing. This is because irrelevant and redundant features increase the network size, congestion and cause a reduction in the algorithm’s speed, reducing the efficiency of the predictive model [91]. Therefore, our study has used the NCA algorithm to screen an optimal set of predictor variables out of the 52-variable set. In general, f_{srnca} calculates every predictor’s relative weight against a target (SSM), illustrated in Figure 4. Following this, the standalone GRU and hybrid CNN-GRU models were executed with

predictors added one by one from the highest feature to the lowest feature weight until an optimal performance was achieved. Figure 5 illustrates the the relative root mean squared error (RRMSE) value of different combinations prepared based on NCA. Tables A1–A6 shows the GRU and CNN-GRU model’s performance accordingly.

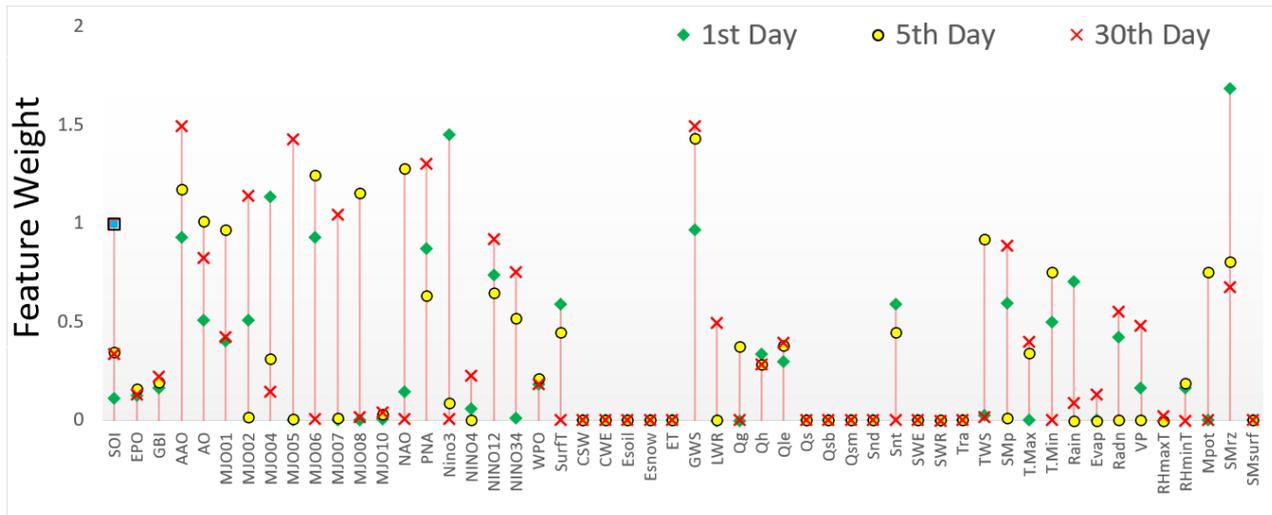


Figure 4. Feature weight matrix of predictor variables from a pool of 52 data sources using neighbourhood component analysis at the n th ($n = 1, 5,$ and 30) day lead time forecasting of surface soil moisture shown for the case of Menindee study station. Details of the variables are mentioned in Table 2.

Figure 4 illustrates the respective feature weights of predictor variables, using the Menindee station as an example. For the 1st day of SSM forecasting, the root zone soil moisture (kg m^{-2}) is found to generate the highest feature weight, whereas, for the 5th day, groundwater storage (mm) is found to be the most significant feature weight. Notably, the groundwater storage contributed to the largest feature weighted for the 7th, 14th, 21st, and 30th day SSM forecasting. This evaluation indicates that groundwater has a strong influence on SSM over inter-daily scales. Tables S1–S6 illustrates the input combination for SSM forecasting in the n th day lead period with their respective forecasting performance with CNN-GRU and GRU model. It is imperative to note that *fsrnca* algorithm is used in two distinct phases before applying the hybrid-deep learning (i.e., CEEMDAN-CNN-GRU) model. In the first phase, *fsrnca* attains the feature weights and acquires the optimal predictor variable list required for SSM forecasts. Subsequently, the second phase incorporates the data decomposition process utilising CEEMDAN to each variable selected from the feature weights. Finally, the feature weight is calculated for IMF (t) deduced for each predictor variable against the objective variable (i.e., SSM). Here, the term t refers to the number of IMFs for each variable, removing four to five least significant features from the hybrid CEEMDAN-CNN-GRU model.

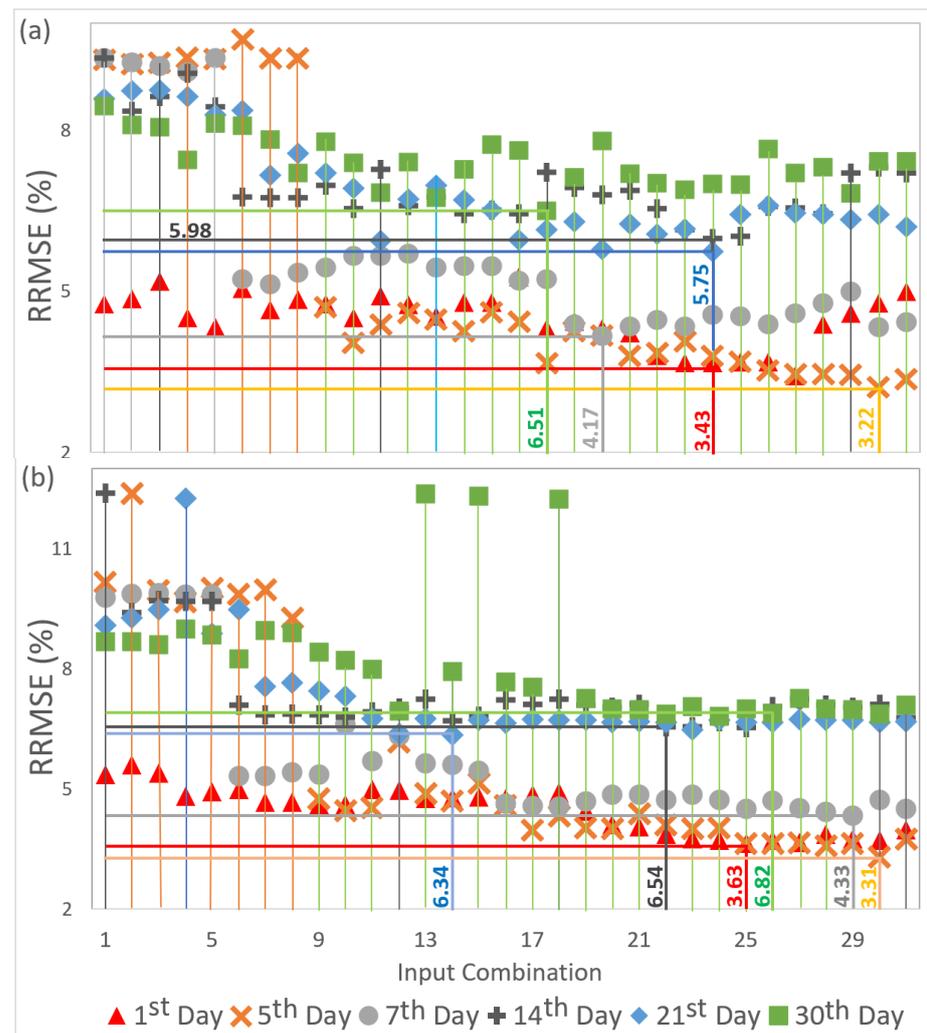


Figure 5. Stair plot showing the relative root mean squared error (RRMSE, %) for (a) CNN-GRU and (b) GRU applied at different input combinations for the Menindee station at the 1st, 5th, 7th, 14th, 21st and 30th day lead time.

3.2.2. Hybrid Deep Learning Algorithm Implementation

Before applying the CEEMDAN-CNN-GRU model in the 1st, 5th, 7th, 14th, 21st, and 30th day SSM forecasts, hyperparameter selection is undertaken through a grid search procedure whose theoretical descriptions are provided in Section 2. Table 3 shows the hyperparameters, optimal GRU architecture, and CNN-GRU with input combinations deduced from the feature weight matrix. Finally, the deep learning forecast model combining a data decomposition (i.e., CEEMDAN) stage with a three-layered feature extraction stage (i.e., CNN) and feature selection stage (i.e., *fsrnca*) is implemented to forecast SSM.

The proposed CEEMDAN-CNN-GRU model is implemented in four stages, as shown in Figure 3. Firstly, CEEMDAN is applied to decompose historical training data into IMFs and residual signals (Figure 1a) followed by segregation of each IMFs and residual, such as collecting all the IMF1 for predictor variables. The relative feature weights of respective IMFs related to IMF of the target variable (i.e., SSM) are determined. The optimal signal selection enables the algorithm to remove the least important feature-weighted IMFs, allowing the predictive model network to be noise-free. Finally, the forecasted SSM utilising the CEEMDAN-based model (i.e., the hybrid CEEMDAN-CNN-GRU) is obtained by aggregating the IMFs of the predictor variables. The robustness of the model is investigated by several evaluation criteria (Section 3.2.3).

Table 3. (a) Range of tested hyperparameters in designing hybrid CNN-GRU and GRU predictive models through trial and error method. (b) Optimally selected hyperparameters. ReLU stands for Rectified Linear Units, SGD stands for stochastic gradient descent optimiser.

(a) Tested Range of Model Hyper-Parameters		
Model	Model Hyper-parameter Names	Search Space for Optimal Hyper-Parameters
CNN-GRU	Filter 1	(70, 80, 100, 150)
	Filter 2	(70, 80, 100,150)
	Filter 3	(70, 80, 100, 150)
	GRU Cell Units	(40, 50, 70, 80, 100, 150)
	Epochs	(500, 800, 1000)
	Activation function	(ReLU)
	Optimiser	(Adam, SGD)
	Batch Size	(5, 10, 20, 50, 100)
GRU	GRU Cell 1	(70, 80, 100, 110)
	GRU Cell 2	(70, 80, 100,150, 200, 210)
	Epochs	(500, 800, 1000)
	Activation function	(ReLU)
	Optimiser	(Adam, SGD)
	Batch Size	(5, 10, 20, 50, 100)
(b) Optimally Selected Hyper-Parameters		
CNN-GRU	Convolution Layer 1 (C1)	80
	C1-Activation function	ReLU
	C1-Pooling Size	1
	Convolution Layer 2 (C2)	70
	C2-Activation function	ReLU
	C2-Pooling Size	1
	Convolution Layer 3 (C3)	80
	C3-Activation function	ReLU
	C3-Pooling Size	1
	GRU Layer 1 (L1)	200
	L1-Activation function	ReLU
	GRU Layer 2 (L2)	60
	L2-Activation function	ReLU
	Drop-out rate	0.2
	Optimiser	Adam
	Padding	Same
	Batch Size	5
	Epochs	400
GRU	GRU Cell 1 (G1)	110
	G1-Activation function	ReLU
	GRU Cell 2 (G2)	250
	G2-Activation function	ReLU
	Epochs	300

Table 3. Cont.

(a) Tested Range of Model Hyper-Parameters	
Optimiser	SGD
Drop-out rate	0.2
Batch Size	15
Epochs	1000

It is worth noting that climate indices (CIs) have a notable signature of climate variability in Australia, leading to substantial influence on rainfall and a potential effect on future surface soil moisture patterns. In the final task, climate indices' relative contribution to building the CEEMDAN-CNN-GRU model is assessed by Multivariate Adaptive Regression Splines (MARS) utilising the ARESLab toolbox. Following Friedman [86], MARS can determine each predictor variable's significance by evaluating its complex and non-linear interaction with the target (i.e., SSM) based on best regressors and provide the importance of each variable. The relative importance of any predictor variable is the square root of GCV (Generalised Cross-Validation) with all basic functions involving the respective variable minus the root square of the GCV score of that full model. However, this process is scaled in such a way that the relative importance has a value of 100, expressed:

$$GCV = \frac{MSE}{\left(1 - \frac{enp}{N}\right)^2} \quad (16)$$

Here, enp is the significant number of model parameters, $p = k + c(k - 1)/2$; k = basis function in MARS model; c = penalty (set to 2 or 3). However, if enp is greater or equal to N , GCV is an Inf, which indicates the model is flawed [92].

3.2.3. Predictive Model Evaluation

The efficacy of deep learning hybrid model is evaluated using different performance evaluation criteria e.g., Pearson's Correlation Coefficient (r), root mean square error (RMSE), Nash-Sutcliffe efficiency (NSE) [93], mean absolute error (MAE), and Kling-Gupta efficiency [94]. Due to geographic differences between the study stations, we employ relative error-based metrics: i.e., relative RMSE (denoted as RRMSE) and relative MAE (denoted as RMAE). The appraisal of a predictive model's efficacy depends on the exactness between the predicted and observed values. RMSE is an appropriate measure of model performance compared to MAE when the error distribution in the tested data is Gaussian [95] but for an improved model evaluation, the Willmott's Index (WI) and Legates-McCabe's (LM) Index are used as more sophisticated and compelling measures [96,97]. Mathematically, the metrics are as follows:

Correlation coefficient (r):

$$r = \left\{ \frac{\sum_{i=1}^N (SSM_{obs} - \overline{SSM}_{obs}) (SSM_{for} - \overline{SSM}_{for})}{\sqrt{\sum_{i=1}^N (SSM_{obs} - \overline{SSM}_{obs})^2 \sum_{i=1}^N (SSM_{for} - \overline{SSM}_{for})^2}} \right\}^2 \quad (17)$$

Mean absolute error (MAE: kg m^{-2}):

$$MAE = \frac{1}{N} \sum_{i=1}^N |SSM_{for} - SSM_{obs}| \quad (18)$$

Root mean squared error (RMSE: kg m^{-2}):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (SSM_{for} - SSM_{obs})^2} \quad (19)$$

Nash-Sutcliffe Efficiency (NSE):

$$NSE = 1 - \left[1 - \frac{\sum_{i=1}^N (SSM_{for})^2}{\sum_{i=1}^N (SSM_{obs} - \overline{SSM}_{for})^2} \right] \quad (20)$$

Kling-Gupta efficiency (KGE):

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\overline{SSM}_{for}}{\overline{SSM}_{obs}} - 1 \right)^2 + \left(\frac{CV_p}{CV_s} \right)^2} \quad (21)$$

Mean Absolute Percentage Error (MAPE, %):

$$MAPE = \frac{1}{N} \left(\sum_{i=1}^N \left| \frac{SSM_{for} - SSM_{obs}}{SSM_{obs}} \right| \right) \times 100, \quad (0\% \leq MAPE \leq 100\%) \quad (22)$$

Willmott's Index (WI):

$$WI = 1 - \left[\frac{\sum_{i=1}^N (SSM_{for} - SSM_{obs})^2}{\sum_{i=1}^N (|SSM_{for} - \overline{SSM}_{obs}| + |SSM_{obs} - \overline{SSM}_{obs}|)^2} \right] \quad (23)$$

Legates-McCabe's Index (LM):

$$LM = 1 - \left[\frac{\sum_{i=1}^N |SSM_{for} - SSM_{obs}|}{\sum_{i=1}^N |SSM_{obs} - \overline{SSM}_{obs}|} \right] \quad (24)$$

Relative Root Mean Squared Error (RRMSE, %):

$$RRMSE(\%) = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (SSM_{for} - SSM_{obs})^2}}{\frac{1}{N} \sum_{i=1}^N (SSM_{obs})} \times 100 \quad (25)$$

Relative Mean Absolute Error (RMAE, %):

$$RMAE(\%) = \frac{\frac{1}{N} \sum_{i=1}^N |SSM_{for} - SSM_{obs}|}{\frac{1}{N} \sum_{i=1}^N (SSM_{obs})} \times 100 \quad (26)$$

Absolute percentage bias (APB, %):

$$APB = \left[\frac{\sum_{i=1}^N |SSM_{obs} - SSM_{for}| \times 100}{\sum_{i=1}^N |SSM_{obs}|} \right] \quad (27)$$

In Equations (17)–(27), SSM_{obs} and SSM_{for} represents the observed and forecasted values for i th test value; \overline{SSM}_{obs} and \overline{SSM}_{for} refer to their averages, accordingly, and N is defined as the number of observations, while CV stands for the coefficient of variation. CV is a standardised measure of the dispersion of the frequency distribution.

4. Results

The practical utility of the hybrid DL (i.e., CEEMDAN-CNN-GRU) model is established by integrating diverse data in its training and model testing phase. Significant features from predictor variables are used by incorporating NCA, and the predictive model is evaluated using statistical metrics (Equations (17)–(27)), infographics, and visualisations to appraise the degree of agreements between simulated and observed soil moisture. By

several measures, the CEEMDAN-CNN-GRU model appears to outperform all the comparative models with superior r and NSE and low RMSE, MAE, and APB in the testing phase. An extensive analysis of tabulated results (Table 4) provides convincing arguments that the hybrid deep learning method is effective for surface soil moisture forecasts and can perhaps be a potential tool in agriculture water management. However, among all study sites, the CEEMDAN-CNN-GRU model for the Menindee station showed the best performance, considering r (0.996), NSE (0.995), and lowest RMSE (0.021), MAE (0.013), and APB (0.359) values for the 1st day of SSM forecasting. The performance of this model is followed by the CEEMDAN-GRU and CNN-GRU model.

For the 5th day of SSM forecasting, the results of the objective model for Menindee had the best performance ($r = 0.993$; $NSE = 0.991$; $RMSE = 0.040 \text{ kg m}^{-2}$) followed by Deniliquin ($r = 0.989$; $NSE = 0.975$; $RMSE = 0.091 \text{ kg m}^{-2}$). Likewise, for the 7th, 14th, 21st, and 30th days of SSM forecasting, the CEEMDAN-CNN-GRU model outperformed the other models by a notable margin for all the respective periods of SSM forecasting. However, a site-specific signature in the model accuracy was also evident, with the results for Menindee registering the lowest value of RMSE generated by the CEEMDAN-CNN-GRU model. In terms of MAE, the CEEMDAN-CNN-GRU model returned the lowest value for Menindee, suggesting that the CEEMDAN-CNN-GRU model was a potential forecasting tool SSM at the 1st, 5th, and 7th day ahead periods. Not surprisingly, in accordance with other studies, e.g., the present study indicates that as the length of the forecasting period was increased, the model's performance appear to reduce at a significant rate in such a way that the r -values reduced by 0.30%, 1.10%, 9.15%, 11% and 15% for the 1st to 5th, 7th, 14th, 21st and 30th day of SSM forecasting. The change of the performance metrics (i.e., NSE, MAE, and APB) for longer-term horizons relative to the shorter-term horizons also concurred with the respective changes in the r -values and is consistent with earlier studies [60,98]. For a longer-term horizon, the present r value was lower, and the MAE increased, suggesting that for the longer forecast horizon, the model appeared to lose the relevant data features in the predictor variables required to maintain precise SSM forecasting performance. The hybrid CEEMDAN-CNN-GRU model is further evaluated using a probability plot of errors at the 95th percentile, including those of the benchmark model (i.e., CNN-GRU, CEEMDAN-GRU) and the standalone model (i.e., GRU) with an illustration for Menindee at the different n th ($n = 1, 5, 7, 14, 21$ and 30) days (Figure 6). The CEEMDAN-CNN-GRU model results show that ~95% of SSM forecasting had the lowest error (<0.1) for the 1st and 5th days of SSM forecasting. Among all the predictive models and the forecast periods over n th days, the GRU-based model showed a more significant proportion of $|FE|$ values at a 95% confidence level. Notably, consistently good results were also achieved for the other stations (i.e., Deniliquin, Fairfield, and Gabo Island), which are shown in supplementary materials (Figure S1a–c). The lowest value of $|FE|$, with <0.063 with a 95% percentile, was evident for Fairfield compared to the other two study stations. The correlation between observed and forecasted daily surface soil moisture datasets generated by the proposed CEEMDAN-CNN-GRU model vs. the corresponding benchmark models (i.e., CNN-GRU and GRU), for the case of Menindee station, is illustrated in Figure 7. The correlations for the hybrid GRU model are positioned close to the observed SSM values up to the 7th day, revealing a high degree of forecasting accuracy. An improvement in the model's forecasting performance was attained by applying the CNN algorithm (i.e., soil moisture generated by the CNN-GRU model) and data decomposition (i.e., CEEMDAN-CNN-GRU) method on standalone GRU model. The disparity between the forecasted SSM and the reference SSM values was significantly higher for the 14th, 21st, and 30th days of SSM forecasting, which concurs with earlier metrics suggesting a potential inadequacy of the data features long time ahead periods [60].

Table 4. Evaluation of hybrid CEEMDAN-CNN-GRU vs. benchmark (CNN-GRU, CEEMDAN-GRU, GRU) models for the specific case of Menindee study site. The correlation coefficient (r), root mean square error (RMSE; Kg m^{-2}), mean absolute error (MAE; Kg m^{-2}), and Nash-Sutcliffe coefficient, (NS) is computed between forecasted and observed surface soil moisture for the 1st day, 5th day, 7th day, 14th day, 21st day, and 30th day ahead periods in the testing phase. The optimal model is boldfaced.

	Soil Moisture Forecasting Horizon, n th Day Lead Time																													
	1st Day					5th Day					7th Day					14th Day					21st Day					30th Day				
	r	NSE	RMSE	MAE	APB	r	NSE	RMSE	MAE	APB	r	NSE	RMSE	MAE	APB	r	NSE	RMSE	MAE	APB	r	NSE	RMSE	MAE	APB	r	NSE	RMSE	MAE	APB
Study Station 1: Menindee																														
CEEMDAN-CNN-GRU	0.996	0.995	0.021	0.013	0.359	0.993	0.991	0.040	0.030	0.823	0.985	0.967	0.075	0.057	1.559	0.906	0.896	0.226	0.185	5.079	0.895	0.787	0.230	0.186	5.098	0.869	0.714	0.255	0.201	5.493
CNN-GRU	0.967	0.892	0.135	0.112	3.061	0.966	0.918	0.117	0.094	2.569	0.945	0.861	0.152	0.121	3.330	0.892	0.770	0.235	0.193	5.285	0.899	0.788	0.210	0.168	4.594	0.851	0.765	0.238	0.181	4.945
CEEMDAN-GRU	0.976	0.937	0.116	0.094	2.234	0.970	0.933	0.120	0.095	2.265	0.957	0.909	0.140	0.110	2.613	0.882	0.738	0.237	0.186	4.424	0.864	0.781	0.262	0.206	4.918	0.866	0.742	0.275	0.217	5.163
GRU	0.962	0.893	0.134	0.110	3.020	0.962	0.933	0.121	0.094	2.589	0.940	0.851	0.158	0.126	3.452	0.882	0.745	0.244	0.197	5.390	0.887	0.748	0.243	0.196	5.360	0.863	0.726	0.251	0.197	5.386
Study Station 2: Deniliquin																														
CEEMDAN-CNN-GRU	0.990	0.899	0.048	0.034	0.778	0.989	0.975	0.091	0.065	1.489	0.959	0.917	0.165	0.113	2.611	0.801	0.607	0.355	0.247	5.716	0.768	0.573	0.374	0.266	6.130	0.703	0.465	0.415	0.295	6.807
CNN-GRU	0.979	0.955	0.098	0.075	1.799	0.945	0.866	0.169	0.137	3.270	0.929	0.846	0.181	0.143	3.405	0.866	0.624	0.283	0.224	5.333	0.873	0.749	0.231	0.181	4.298	0.848	0.687	0.258	0.202	4.806
CEEMDAN-GRU	0.987	0.958	0.106	0.081	1.930	0.968	0.929	0.123	0.096	2.279	0.969	0.920	0.131	0.106	2.524	0.872	0.730	0.240	0.189	4.505	0.859	0.712	0.249	0.197	4.701	0.869	0.671	0.264	0.207	4.926
GRU	0.967	0.927	0.125	0.099	2.350	0.947	0.889	0.154	0.121	2.874	0.918	0.822	0.195	0.153	3.655	0.867	0.722	0.244	0.191	4.560	0.868	0.695	0.256	0.201	4.787	0.850	0.659	0.269	0.217	5.152
Study Station 3: Fairfield																														
CEEMDAN-CNN-GRU	0.975	0.976	0.035	0.024	0.554	0.972	0.975	0.069	0.052	1.189	0.959	0.920	0.162	0.110	2.524	0.842	0.628	0.349	0.238	5.493	0.762	0.573	0.374	0.264	6.088	0.746	0.523	0.374	0.261	6.078
CNN-GRU	0.945	0.935	0.061	0.048	1.099	0.962	0.943	0.135	0.091	2.107	0.907	0.821	0.240	0.156	3.612	0.764	0.560	0.376	0.264	6.109	0.759	0.554	0.379	0.259	5.988	0.708	0.477	0.410	0.289	6.671
CEEMDAN-GRU	0.947	0.943	0.048	0.034	0.778	0.939	0.935	0.091	0.065	1.489	0.929	0.917	0.165	0.113	2.611	0.801	0.607	0.355	0.247	5.716	0.768	0.573	0.374	0.266	6.130	0.703	0.465	0.415	0.295	6.807
GRU	0.925	0.919	0.153	0.096	2.205	0.913	0.905	0.177	0.115	2.659	0.904	0.809	0.250	0.168	3.864	0.778	0.585	0.369	0.254	5.850	0.775	0.568	0.376	0.267	6.165	0.666	0.411	0.435	0.314	7.267
Study Station 4: Gabo Island																														
CEEMDAN-CNN-GRU	0.988	0.966	0.085	0.067	1.455	0.987	0.971	0.079	0.062	1.346	0.978	0.944	0.109	0.086	1.887	0.931	0.899	0.188	0.147	3.206	0.909	0.764	0.224	0.175	3.829	0.913	0.807	0.202	0.158	3.456
CNN-GRU	0.979	0.951	0.101	0.078	1.707	0.973	0.944	0.109	0.084	1.826	0.948	0.897	0.147	0.113	2.457	0.921	0.843	0.182	0.141	3.087	0.911	0.803	0.204	0.160	3.493	0.879	0.862	0.193	0.151	3.284
CEEMDAN-GRU	0.986	0.966	0.085	0.067	1.472	0.983	0.964	0.087	0.069	1.508	0.974	0.945	0.107	0.085	1.844	0.924	0.821	0.194	0.153	3.340	0.913	0.814	0.198	0.156	3.394	0.912	0.798	0.206	0.161	3.520
GRU	0.977	0.950	0.102	0.081	1.773	0.970	0.940	0.113	0.086	1.868	0.951	0.902	0.144	0.111	2.423	0.919	0.825	0.192	0.150	3.283	0.912	0.813	0.199	0.156	3.411	0.815	0.743	0.203	0.160	3.499

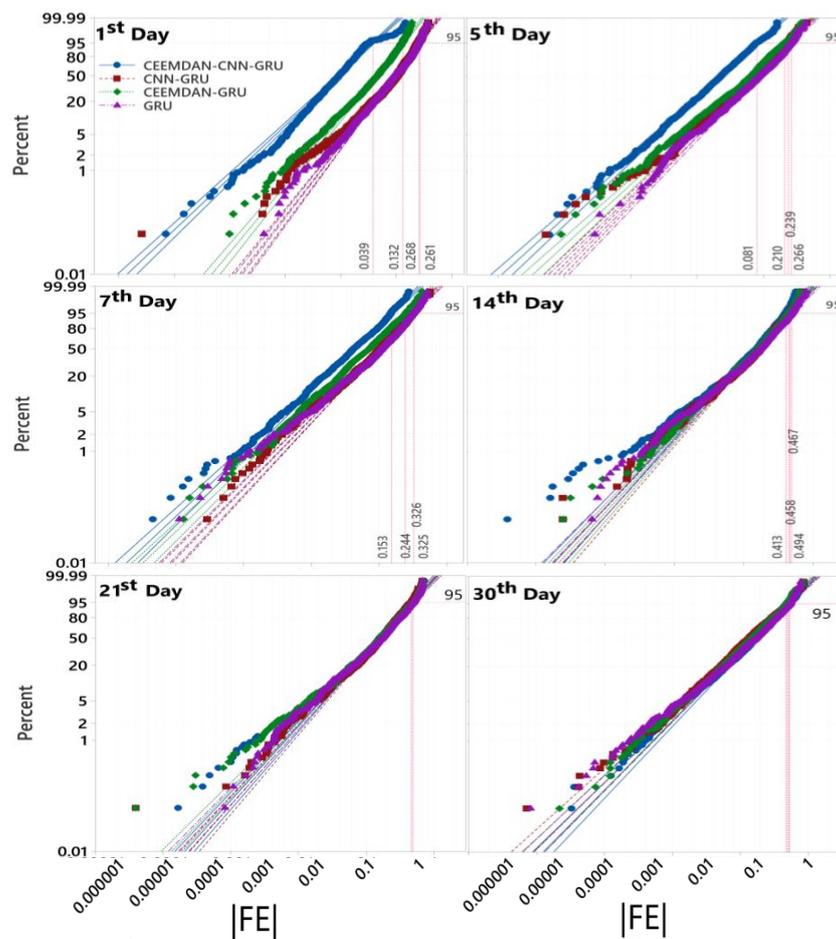


Figure 6. Probability plot (95 percentiles) for hybrid CEEMDAN-CNN-GRU, CNN-GRU, CEEMDAN-GRU, and GRU model for Menindee at different n th ($n = 1, 5, 7, 14, 21$ and 30) day lead time.

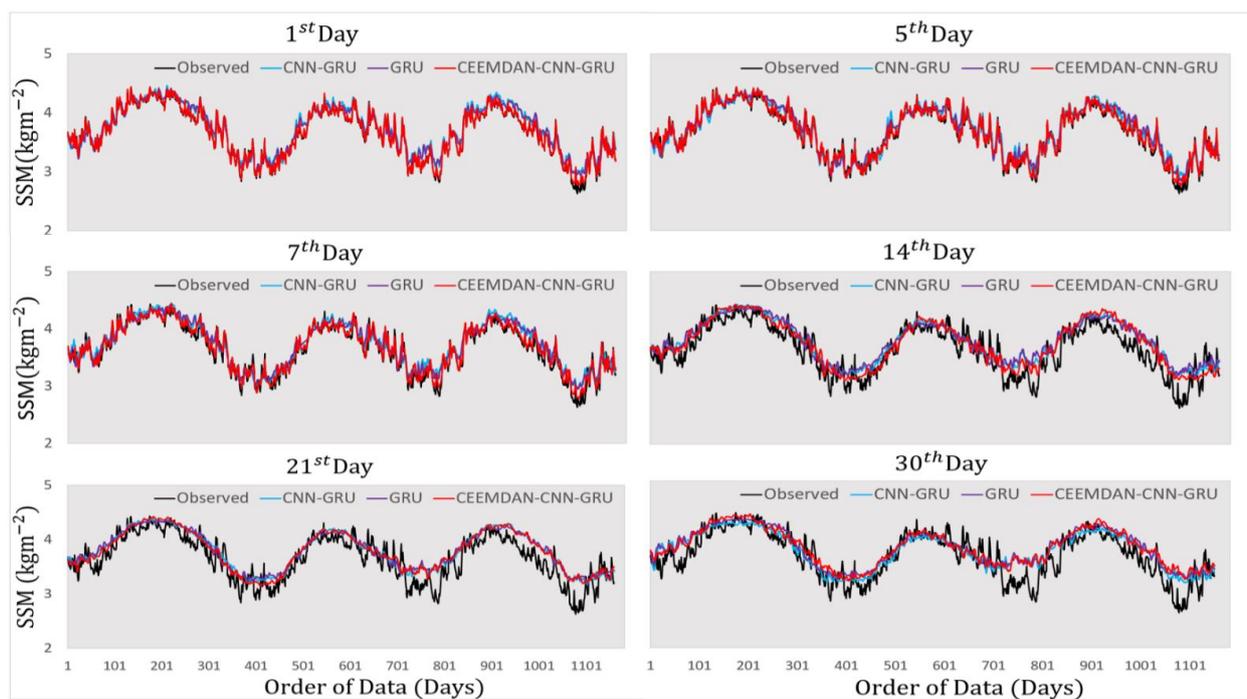


Figure 7. Time series of daily surface soil moisture (SSM, kg m^{-2}) for observed SSM (Gray) and Figure 1. 5, 7, 14, 21, and 30) day lead times.

Figure 8 shows a scatter plot of forecasted and observed SSM for the 1st and 7th days across the Murray Darling Basin with a least square regression line, $y = mx + C$, and the coefficient of determination in each sub-panel. Notably, the objective model (i.e., CEEMDAN-CNN-GRU) is seen to attain more accurate results with considerably larger r^2 values. The SSM forecast with a hybrid deep learning model for Menindee station performed significantly better than the comparative model (i.e., CNN-GRU). In the case of Menindee, for example, the values for m and r^2 are in reasonably good agreement against the 1:1 line representing the forecasted and observed SSM values in such a way that ($m \mid r^2$) is 0.994 | 0.995 for the hybrid CEEMDAN-CNN-GRU model relative to (0.931 | 0.933) for CNN-GRU for the 1st day ahead of SSM forecasting. Moreover, for the 1st day of SSM forecasting, the CEEMDAN-CNN-GRU model provided results in significant proximity to the other three stations, such as Deniliquin: 0.962 | 0.966, Fairfield: 0.928 | 0.964, and the Gabo Island: 0.958 | 0.976). Alternatively, the y-intercept of the regression line was close to trivial, i.e., 0.002 (Menindee: 1st day), 0.193 (Deniliquin: 1st day), 0.05 (Fairfield: 1st day), and 0.303 (Gabo Island: 1st day), revealing the efficacy of the deep learning hybrid method for surface soil moisture forecasting. For the 14th, 21st, and 30th day ahead of SSM forecasting, the y-intercept, as expected, deviated slightly from the ideal value of 0, caused by more outliers between simulated and reference values in the testing phase.

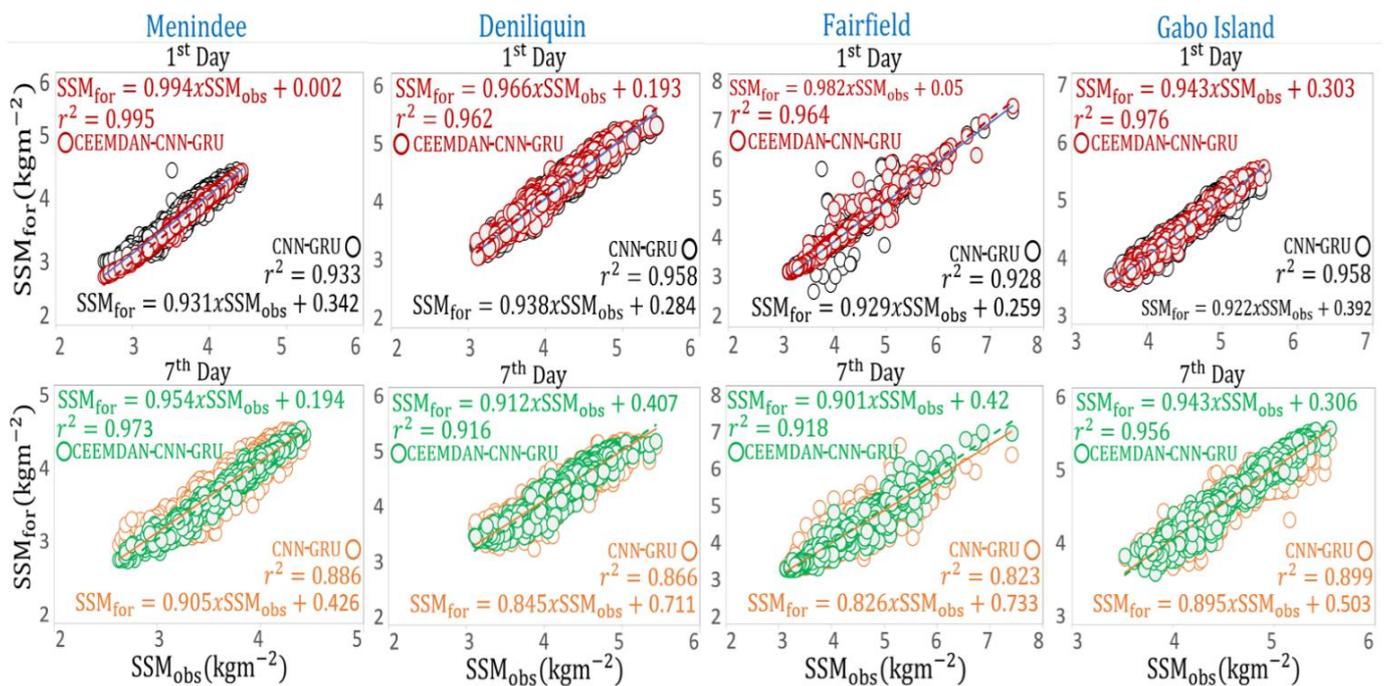


Figure 8. Scatter plot of the forecasted and observed SSM for Menindee, Deniliquin, Fairfield, and Gabo Island stations at different n th ($n = 1$ and 7) day ahead. A least square regression line, $y = mx + C$, and coefficient of determination (R^2) is shown in each sub-panel.

To further analyse the tested models' performances, we adopt the Legates and McCabe's Index [99] as a cross-validation metric for simulated data. This metric has a better model penalisation skill when high SSM values are expected in the testing set [41]. This is illustrated in Figure 9 in terms of a polar plot of the LM values for the hybrid deep learning approach (i.e., CEEMDAN-CNN-GRU) and other models for the different day ahead forecasting. The LM values accumulated across all stations in the case of CEEMDAN-CNN-GRU have a superior result with the highest LM ≈ 0.962 for Menindee and the lowest LM for the case of Gabo Island (LM ≈ 0.846) in the 1st Day ahead SSM forecasting. In agreement with earlier results, the LM values for the 14th, 21st, and 30th day ahead for other models were comparatively smaller. Figure 10a,b is a contour plot of KGE and MAPE for the hybrid DL approach (i.e., CEEMDAN-CNN-GRU) along with its benchmark (i.e.,

CNN-GRU) and standalone (i.e., GRU) methods for all four stations in MDB at different n th ($n = 1, 5, 7, 14, 21$ and 30) days in forecasting SSM. This infographic verifies the robustness of the proposed objective model that attains the highest KGE values and the lowest MAPE values for 1st and 5th day of SSM forecasting.

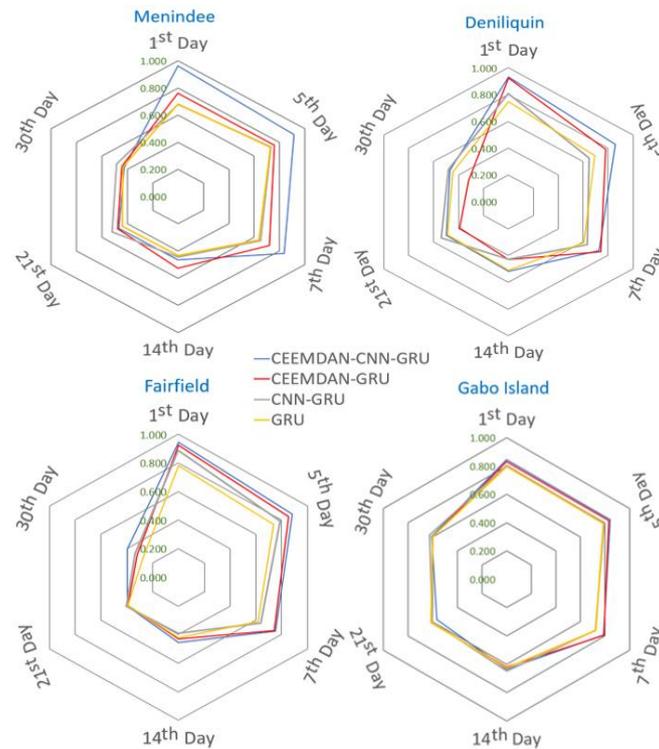


Figure 9. Polar plot showing the Legates and McCabe's Index (LM) in the testing period computed for the hybrid CEEMDAN-CNN-GRU against comparative models at different n th ($n = 1, 5, 7, 14, 21$, and 30) day ahead forecasting of SSM.

However, for the 14th, 21st, and 30th day of SSM forecasting, the KGE values range between 0.40 and 0.80, and the MAPE values range from 4–11%, demonstrating a slightly lower forecast accuracy relative to the 1st and 5th day of SSM forecasting. Figure 11 illustrates the absolute forecasted error ($|FE|$) using all the four candidate study sites' implemented models. The box plot demonstrates the data dispersal in terms of the forecasted (SSM_{for}) SSM. Figure 11 provides a clear visualisation of the closed distribution of error values for Menindee and Fairfield stations in the hybrid CEEMDAN-CNN-GRU model for 1st day ahead SSM forecasting. The lower end of the plot for $|FE|$ is situated within the lower quartile (25th) and upper quartile (75th). Moreover, the GRU and CNN-GRU models for these stations show an increased distribution of $|FE|$, except for the Fairfield station. Moreover, the forecasting of SSM for the 14th, 21st, and 30th day periods have a comparatively higher value of the absolute forecasting error for all tested models. A more comprehensive inspection of the absolute forecasting error ($|FE|$) in the case of the hybrid GRU model for the four study stations further cements the suitability of the CEEMDAN-CNN-GRU model in forecasting SSM for the 1st, 5th, and 7th day ahead periods in Australian Murray Darling Basin, evidenced by the narrowest error distribution in comparison with the other models.

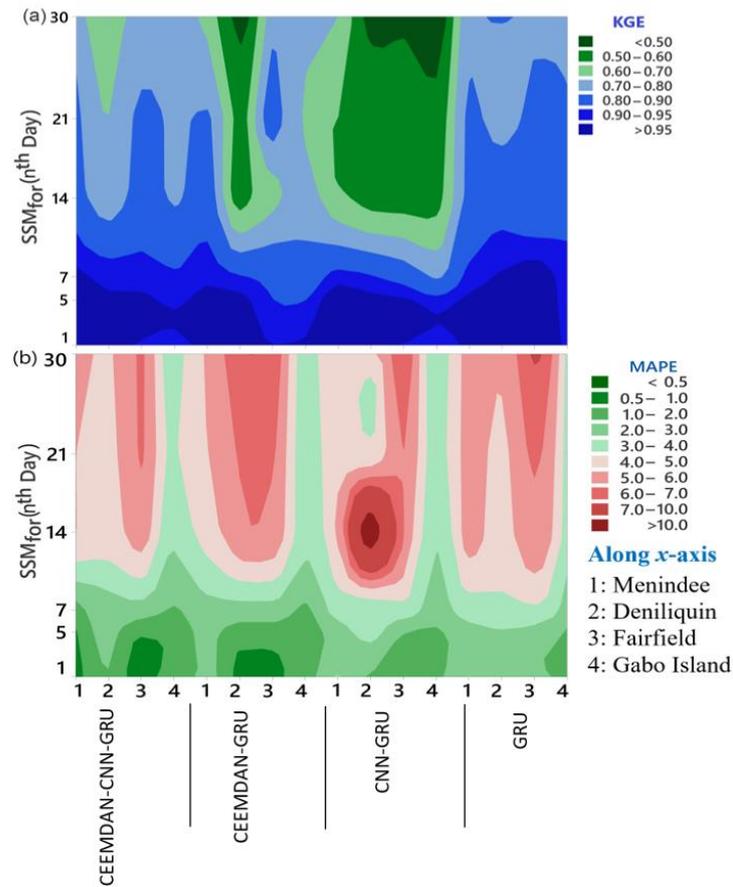


Figure 10. Contour plot of (a) KGE, (b) MAPE for hybrid CEEMDAN-CNN-GRU model against comparative models for different n th ($n = 1, 5, 7, 14, 21,$ and 30) day ahead forecasting of SSM.

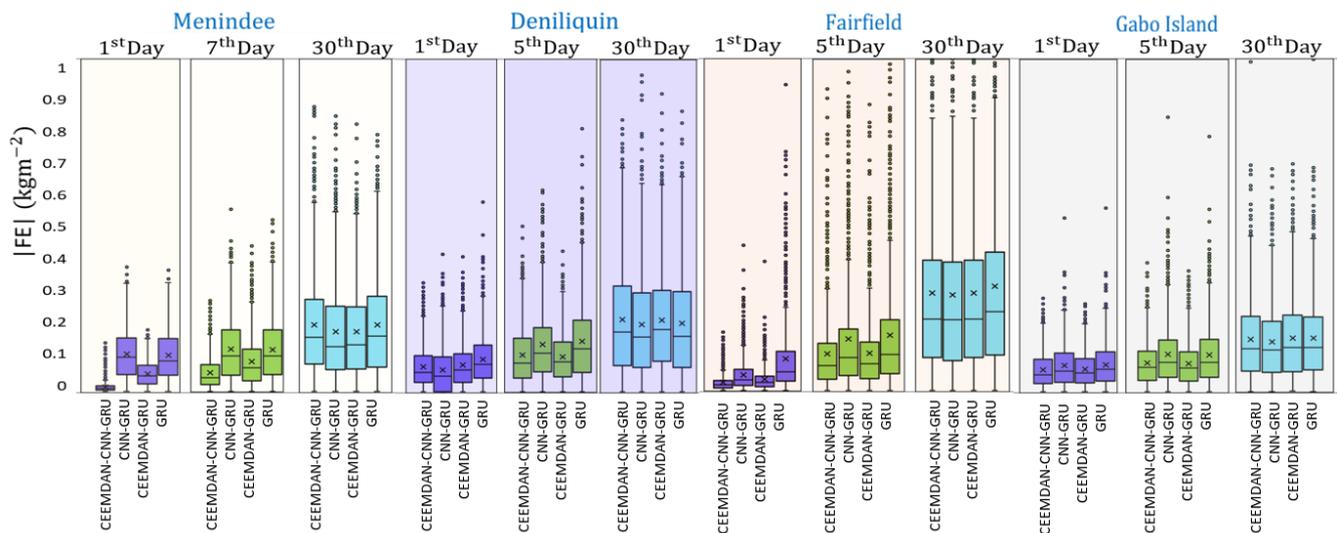


Figure 11. Box plot of errors in the testing phase for hybrid CEEMDAN-CNN-GRU against comparative models at different n th ($n = 1, 7,$ and 30) day ahead lead time forecasting SSM. (Note: CEEMDAN-CNN-GRU = Hybrid Model integrating the CEEMDAN and CNN algorithm with GRU; CEEMDAN-GRU = Hybrid Model integrating the CEEMDAN algorithm with GRU; CNN-GRU = Hybrid Model integrating the CNN algorithm with GRU).

It is noteworthy that in this study, two distinct algorithms, namely the CEEMDAN and CNN, are used to improve the GRU-based predictive model. Therefore Figure 12

shows the effect of applying CEEMDAN and CNN as data pre-processing and feature extraction methods incrementally, respectively, on the per cent change in RMAE values within the testing SSM values. In terms of 1st, 5th, and 7th day of Menindee station, the RMAE (%) values of CEEMDAN-CNN-GRU model (where both CEEMDAN and CNN are integrated) appeared to decrease by ~87%, 68%, and 54%, respectively. Similarly, for the 1st-day forecasting taking the example of Fairfield station, the CNN feature-extraction skill reduced the error of ~55%, whereas an additional decrease in RMAE of ~18% was noted integration of the CEEMDAN selected variables (CEEMDAN-CNN-GRU). Additionally, for Deniliquin and Gabo Island study sites, the SSM forecasting for the 1st day ahead evaluated through RMAE values decreased by slightly less than 20%. It is worth mentioning that the per cent increase in RMAE was ~5% for Menindee for the 30th day ahead SSM forecasting with similar deductions for the other stations.

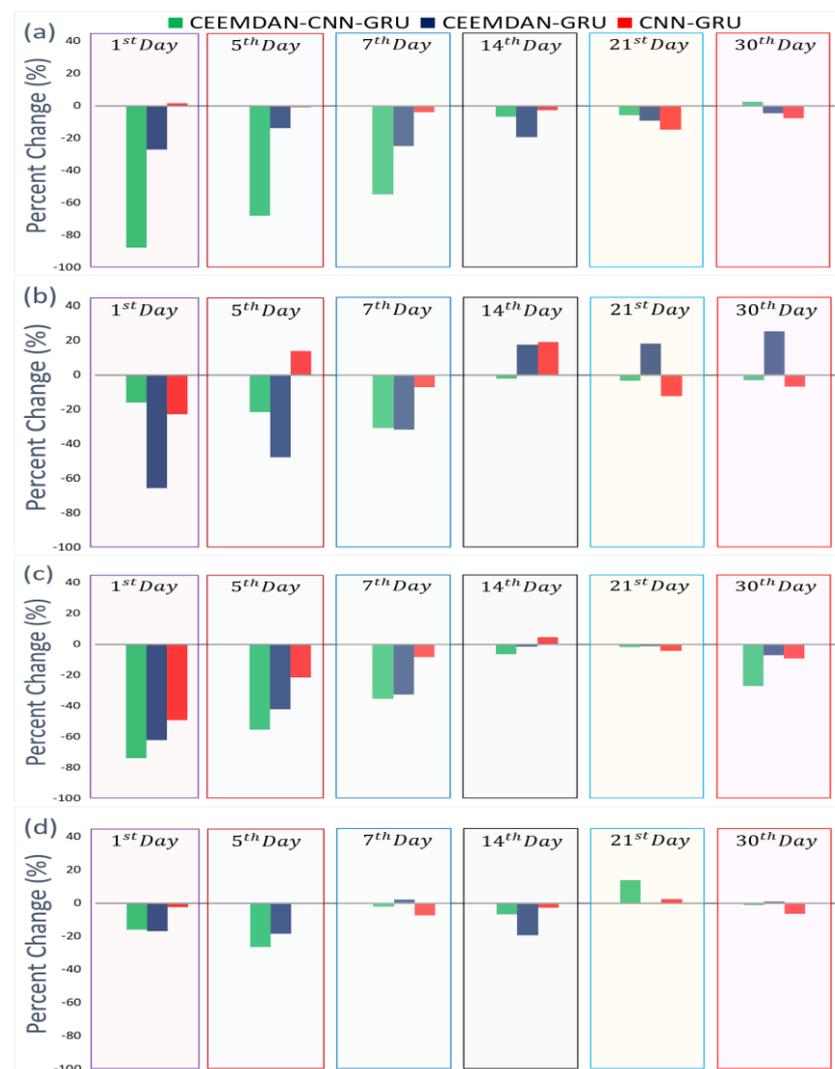


Figure 12. The percentage change in RMAE generated by the objective, and benchmark models using CEEMDAN and CNN methods (as data decomposition and feature extraction methods) adopted in forecasting SSM at four study sites: Murray Darling Basin. (a) Menindee, (b) Deniliquin, (c) Fairfield, (d) Gabo Island at different n th ($n = 1, 5, 7, 14, 21,$ and 30) day ahead forecasting SSM.

We further show the CEEMDAN-CNN-GRU hybrid model's skill for seasonal forecasting for the different day ahead periods to better understand the seasonal effects of models used in SSM prediction. Figure 13 displays the average observed vs forecasted SSM on a seasonal basis (i.e., austral summer, autumn, winter, and spring) generated by CEEMDAN-

CNN-GRU model in case of Menindee study site. The forecast error across these seasons is relatively insignificant, occupying values of (0, 0.16) kg m^{-2} to demonstrate the exceptional skill of the objective model. Notably, the 1st and 5th day ahead of observed and forecasted SSM for austral summer, spring, winter, and autumn appear to match with the forecast error ($|FE| < 0.04 \text{ kg m}^{-2}$), whereas, for winter, the $|FE|$ values are slightly higher for the 5th day ahead SSM forecasting. Not surprisingly, the CNN-GRU model possesses a larger error, ranging from 0.04 to 0.18 kg m^{-2} , establishing the CNN-GRU model's relatively poor performance compared with the hybrid CEEMDAN-CNN-GRU model. For the case of the 30th day ahead SSM forecasting, the study site Menindee registered a higher uncertainty for austral summer ($0.18 < |FE| < 0.18 \text{ kg m}^{-2}$) compared with winter and spring ($0.14 < |FE| < 0.15 \text{ kg m}^{-2}$). This indicates that the hybrid CEEMDAN-CNN-GRU model developed with NCA and CEEMDAN algorithms employing MODIS-derived satellite data, ground-based observations, and climate indices can be considered ideal in multi-step SSM forecasting.

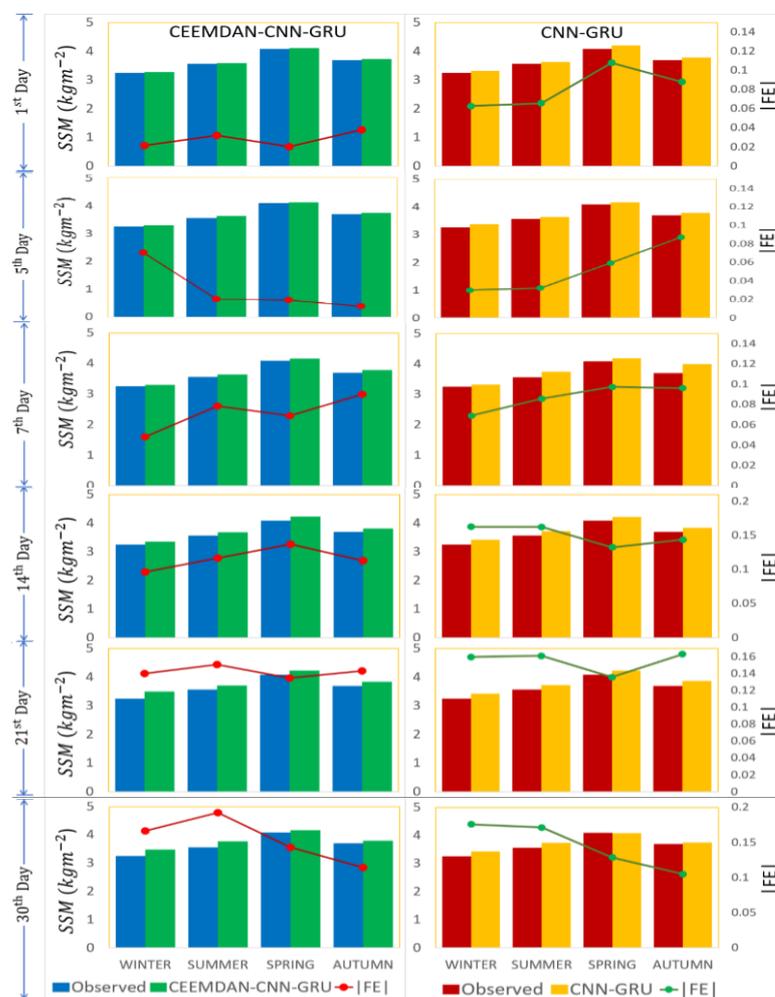


Figure 13. The average forecasted SSM vs. observed SSM on a seasonal basis using hybrid CEEMDAN-CNN-GRU and CNN-GRU models for Menindee at different n th ($n = 1, 5, 7, 14, 21,$ and 30) day ahead periods. The forecast error ($|FE|$) in each model is plotted on a secondary axis as a line chart.

5. Discussions

Based on the results, we note the effects of climate indices on surface soil moisture as non-negligible. In this paper, analysing this impact is undertaken using two ways. Firstly, the NCA algorithm provides key information about how climate indices affect SSM. For

example, for SSM forecasting, climate indices based on SOI, EPO, MJOs and SST were found to significantly affect the SSM. Secondly, GCV values based on a MARS model were calculated following Friedman [100] approach to deduce the importance of input features. The contributory influence appeared to be between 12% and 53% according to GCV for Menindee station, and similarly notable effect for the other study sites. Specifically, the lowest percentage of ~12% of GCV was found for the 14th, and the highest percentage (~53%) was found for the 28th Day ahead SSM forecasting. In a nutshell, we note that climate indices make a moderate to high contribution in forecasting surface soil moisture within the Murray Darling Basin.

Neighbourhood Component Analysis (NCA) was utilised to examine significant features from a relatively large pool (or 52 different) data related to soil moisture. In data-driven modelling, selecting predictor variables is crucial, as improper variables with weak relationships against SSM can lead to undesirable uncertainties in the model. As per evaluations in Tables S1–S6, a combination of predictor variables deduced by NCA at six different lead time SSM forecasting was significant, and this result concurred with previous studies [59,101].

The objective approach based on NCA yielded good accuracy (i.e., CEEMDAN-CNN-GRU), demonstrating that best predictors were attained through a careful variable selection stage (by NCA) and feature extraction stage (by CNN and CEEMDAN methods). Accordingly, the proposed forecast model for SSM was sufficiently robust in daily and seasonal tests, as well as through the inclusion of synoptic-scale features, i.e., those captured from patterns in the SST and MJO series. The probability of absolute error placing within the 95th percentile and the substantial seasonal forecasting of SSM indicates that the model can handle satellite-derived variables' error. Our study also suggests that groundwater recharge, deep percolation, and plant uptake, which are essential factors to concentrate soil moisture in different layers [57], can be ideal variables to better understand SSM characteristics while also assisting in the prediction of future changes.

The present model's performance revealed that a shorter period forecast (i.e., 1st, 5th, or 7th) was more precise, whereas a longer forecast horizon (i.e., 14th, 21st, and 30th) registered a lower accuracy than that of the shorter span of SSM forecasting. One plausible reason for this is that our predictive model appeared to struggle to capture enough input features from the dataset for a more extended time-step forecast (i.e., 30th day against 7th day). Considering the reduction in feature capturing capability of the model, we can say that as the time series data approached close to the 7th-day boundary, the model would capture it with good forecast accuracy. Undoubtedly, this occurs due to a loss of data features in the predictor-target matrix. This indeed concurs with earlier studies (e.g., [60,98], where models for the 1- and 2-day ahead modelling horizon was more accurate than the 30-day horizon for river flow forecasting, and the 1- and 3-month runoff model was more accurate than the 6-month runoff model predicting 1-, 3-, and 6-month ahead runoff in the Yingluoxia watershed, Northwestern China. The hybrid deep learning approach (i.e., CEEMDAN-CNN-GRU) incorporated with MODIS satellite-derived data, ground-based SILO data, and climate mode indices (representing synoptic-scale climate features) can be a good modelling tool to predict soil moisture or other hydrological variables at multi-step lead times, including its future use in water resource management and sustainable agriculture.

6. Conclusions

This study reports the performance efficacy of a DL data-driven (CEEMDAN-CNN-GRU) model based on the Gated Recurrent Unit (GRU) for daily surface soil moisture forecasting at multi-step horizons. The hybrid CEEMDAN-CNN-GRU model was built by integrating MODIS sensors (satellite-derived data), ground-based observations, and climate indices tested at important stations in the Australian Murray Darling Basin. To attain an accurate and reliable model for soil moisture, a feature extraction (i.e., CNN) and feature (or variable) selection algorithm (i.e., NCA) was used, with tests at 1st, 5th, 7th, 14th,

21st, and 30th day ahead period. The input variables, comprised initially of 52 different predictors, were extracted from March 2003 to March 2020 and screened accordingly, using the NCA algorithm through a feature selection stage, to select the most relevant input variables required to forecast daily-scale soil moisture. Three other benchmarking models (i.e., CEEMDAN-GRU, CNN-GRU, and GRU) were built and evaluated against statistical score metrics and visual analysis to ascertain the predictive skill of the objective model of observed and forecasted datasets in the testing phase. The results revealed that NCA was a practical approach to acquire the best features from an optimal set of predictor variables. The hybrid CEEMDAN-CNN-GRU model has significantly improved the decomposition of input variables to provide more defined soil moisture prediction features. Thus, the proposed CEEMDAN-CNN-GRU model yielded an acceptable level of accuracy when applied at the 1st, 5th, and 7th day ahead SSM forecasting against standalone GRU model registering a comparatively higher forecast error at all these periods. This superior performance was also endorsed with low MAE values, ranging from 0.013 kg m^{-2} to 0.067 kg m^{-2} , 0.030 kg m^{-2} to 0.075 kg m^{-2} , and 0.057 kg m^{-2} to 0.113 kg m^{-2} for the 1st, 5th, and 7th day ahead period. Other results also supported the practical utility of the CEEMDAN-CNN-GRU model. For example, the probability plot of absolute error for Menindee station has 95% of SSM forecasting with the lowest error bracket (<0.1) at the 1st, and 5th day SSM prediction, and these results were better than earlier studies on forecasting soil moisture prediction, e.g., [23,36,59,102]. As the present study has focused on daily scale prediction, in a future study, researchers may also adopt the CEEMDAN-CNN-GRU model to utilise the global climate model (GCM) model-simulated variables to estimate future SSM under global warming scenarios.

Supplementary Materials: The following are available online at <https://www.mdpi.com/2072-4292/13/4/554/s1>, Table S1: Performance of CNN-GRU and GRU model to forecast the 1st day Surface Soil moisture of Minindee Station with the optimum forecasting results based on the Nash-Sutcliffe coefficient (NS) and mean absolute error (MAE; Kg m^{-2}) for the testing phase, Table S2: Caption identical to Table S1, except for the 5th day. Table S3: Caption identical to Table S1, except for the 7th day, Table S4: Caption identical to Table S1, except for the 14th day, Table S5: Caption identical to Table S1, except for the 21st day, Table S6: Caption identical to Table S1, except for the 30th day, Figure S1: Probability plot for the objective model (i.e., CEEMDAN-CNN-GRU), benchmark model (i.e., CNN-GRU, CEEMDAN-GRU), and the standalone model (i.e., GRU) for (a) Deniliquin, (b) Fairfield and (c) Gabo Island stations at different nth ($n = 1, 5, 7, 14, 21, \text{ and } 30$) days. The reference points are provided at 95 percentiles on the probability of comparing the model.

Author Contributions: Conceptualisation, A.A.M.A. and R.C.D.; methodology, A.A.M.A. and R.C.D.; software, A.A.M.A.; model development, A.A.M.A.; validation, A.A.M.A.; formal analysis, A.A.M.A.; investigation, A.A.M.A.; resources, A.A.M.A.; data curation, A.A.M.A.; writing—original draft preparation, A.A.M.A.; writing—review and editing, A.A.M.A., R.C.D., N.R., A.G., Q.F., Z.Y. and L.Y.; visualisation, A.A.M.A.; supervision, R.C.D.; funding acquisition, R.C.D. All authors have read and agreed to the published version of the manuscript.

Funding: The study was supported by the Chinese Academy of Science (CAS), and University of Southern Queensland (USQ) under the USQ-CAS Postgraduate Research Scholarship (2019–2021) awarded to the first author, managed by Graduate Research School under the leadership of Professor Feng Qi (CAS) and Associate Professor Ravinesh Deo (USQ).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analysed in this study. The data can be found here: [<https://giovanni.gsfc.nasa.gov/giovanni/>; <https://www.longpaddock.qld.gov.au/silo/>].

Acknowledgments: We thank the Editor and Reviewers for their insightful comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Van Loon, A.F.; Laaha, G. Hydrological drought severity explained by climate and catchment characteristics. *J. Hydrol.* **2015**, *526*, 3–14. [[CrossRef](#)]
2. Brocca, L.; Melone, F.; Moramarco, T.; Morbidelli, R. Spatial—temporal variability of soil moisture and its estimation across scales. *Water Resour. Res.* **2010**, *46*, W02516. [[CrossRef](#)]
3. Brocca, L.; Ciabatta, L.; Massari, C.; Camici, S.; Tarpanelli, A. Soil Moisture for Hydrological Applications: Open Questions and New Opportunities. *Water* **2017**, *9*, 140. [[CrossRef](#)]
4. Chang, X.; Zhao, W.; Zeng, F. Crop evapotranspiration-based irrigation management during the growing season in the arid region of northwestern China. *Environ. Monit. Assess.* **2015**, *187*, 699. [[CrossRef](#)]
5. Gill, M.K.; Asefa, T.; Kembrowski, M.W.; McKee, M. Soil moisture prediction using support vector machines 1. *JAWRA J. Am. Water Resour. Assoc.* **2006**, *42*, 1033–1046. [[CrossRef](#)]
6. Akbari Asanjan, A.; Yang, T.; Hsu, K.; Sorooshian, S.; Lin, J.; Peng, Q. Short-Term Precipitation Forecast Based on the PERSIANN System and LSTM Recurrent Neural Networks. *J. Geophys. Res. Atmos.* **2018**, *123*. [[CrossRef](#)]
7. Tripathi, S.; Srinivas, V.V.; Nanjundiah, R.S. Downscaling of precipitation for climate change scenarios: A support vector machine approach. *J. Hydrol.* **2006**, *330*, 621–640. [[CrossRef](#)]
8. Yang, L.; Feng, Q.; Yin, Z.; Wen, X.; Deo, R.C.; Si, J.; Li, C. Application of multivariate recursive nesting bias correction, multiscale wavelet entropy and AI-based models to improve future precipitation projection in upstream of the Heihe River, Northwest China. *Theor. Appl. Climatol.* **2018**, *137*, 323–339. [[CrossRef](#)]
9. Nguyen-Huy, T.; Deo, R.C.; An-Vo, D.-A.; Mushtaq, S.; Khan, S. Copula-statistical precipitation forecasting model in Australia's agro-ecological zones. *Agric. Water Manag.* **2017**, *191*, 153–172. [[CrossRef](#)]
10. Deo, R.C.; Şahin, M. Application of the extreme learning machine algorithm for the prediction of monthly Effective Drought Index in eastern Australia. *Atmos. Res.* **2015**, *153*, 512–525. [[CrossRef](#)]
11. Prasad, R.; Deo, R.C.; Li, Y.; Maraseni, T. Input selection and performance optimisation of ANN-based streamflow forecasts in the drought-prone Murray Darling Basin region using IIS and MODWT algorithm. *Atmos. Res.* **2017**, *197*, 42–63. [[CrossRef](#)]
12. Ahmed, A.M.; Shah, S.M.A. Application of artificial neural networks to predict peak flow of Surma River in Sylhet Zone of Bangladesh. *Int. J. Water* **2017**, *11*, 363–375. [[CrossRef](#)]
13. Hu, C.; Wu, Q.; Li, H.; Jian, S.; Li, N.; Lou, Z. Deep Learning with a Long Short-Term Memory Networks Approach for Rainfall-Runoff Simulation. *Water* **2018**, *10*, 1543. [[CrossRef](#)]
14. Kratzert, F.; Klotz, D.; Brenner, C.; Schulz, K.; Herrnegger, M. Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 6005–6022. [[CrossRef](#)]
15. Arto, I.; Garcia-Muros, X.; Cazcarro, I.; Gonzalez-Eguino, M.; Markandya, A.; Hazra, S. The socioeconomic future of deltas in a changing environment. *Sci. Total Environ.* **2019**, *648*, 1284–1296. [[CrossRef](#)] [[PubMed](#)]
16. Le, H.; Lee, J. Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. *Water* **2019**, *11*, 1387. [[CrossRef](#)]
17. Ahmed, A.M.; Deo, R.C.; Ghahramani, A.; Raj, N.; Feng, Q.; Yin, Z.; Yang, L. LSTM integrated with Boruta-random forest optimiser for soil moisture estimation under RCP4. 5 and RCP8. 5 global warming scenarios. *Stoch. Environ. Res. Risk Assess.* **2021**, 1–31. [[CrossRef](#)]
18. Gedefaw, M.; Hao, W.; Denghua, Y.; Girma, A. Variable selection methods for water demand forecasting in Ethiopia: Case study Gondar town. *Cogent Environ. Sci.* **2018**, *4*, 1537067. [[CrossRef](#)]
19. Mouatadid, S.; Adamowski, J. Using extreme learning machines for short-term urban water demand forecasting. *Urban Water J.* **2017**, *14*, 630–638. [[CrossRef](#)]
20. Ahmed, A.A.M. Prediction of dissolved oxygen in Surma River by biochemical oxygen demand and chemical oxygen demand using the artificial neural networks (ANNs). *J. King Saud Univ. Eng. Sci.* **2017**, *29*, 151–158. [[CrossRef](#)]
21. Ahmed, A.A.M.; Shah, S.M.A. Application of adaptive neuro-fuzzy inference system (ANFIS) to estimate the biochemical oxygen demand (BOD) of Surma River. *J. King Saud Univ. Eng. Sci.* **2017**, *29*, 237–243. [[CrossRef](#)]
22. Huang, C.; Li, L.; Ren, S.; Zhou, Z. Research of soil moisture content forecast model based on genetic algorithm BP neural network. In Proceedings of the International Conference on Computer and Computing Technologies in Agriculture, Nanchang, China, 22–25 October 2010; pp. 309–316.
23. Prasad, R.; Deo, R.C.; Li, Y.; Maraseni, T. Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition. *Geoderma* **2018**, *330*, 136–161. [[CrossRef](#)]
24. Ghimire, S.; Deo, R.C.; Raj, N.; Mi, J. Deep Learning Neural Networks Trained with MODIS Satellite-Derived Predictors for Long-Term Global Solar Radiation Prediction. *Energies* **2019**, *12*, 2407. [[CrossRef](#)]
25. Zhang, J.; Zhu, Y.; Zhang, X.; Ye, M.; Yang, J. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. *J. Hydrol.* **2018**, *561*, 918–929. [[CrossRef](#)]
26. Le, X.-H.; Ho, H.V.; Lee, G. Application of gated recurrent unit (GRU) network for forecasting river water levels affected by tides. In Proceedings of the International Conference on Asian and Pacific Coasts, Hanoi, Vietnam, 25–28 September 2019; pp. 673–680.
27. Gao, S.; Huang, Y.; Zhang, S.; Han, J.; Wang, G.; Zhang, M.; Lin, Q. Short-term runoff prediction with GRU and LSTM networks without requiring time step optimisation during sample generation. *J. Hydrol.* **2020**, *589*, 125188. [[CrossRef](#)]

28. Ghimire, S.; Deo, R.C.; Raj, N.; Mi, J. Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. *Appl. Energy* **2019**, *253*, 113541. [[CrossRef](#)]
29. Deo, R.C.; Sahin, M. An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern Queensland. *Environ. Monit. Assess.* **2016**, *188*, 90. [[CrossRef](#)] [[PubMed](#)]
30. Nourani, V.; Komasi, M.; Mano, A. A multivariate ANN-wavelet approach for rainfall–runoff modeling. *Water. Resour. Manag.* **2009**, *23*, 2877–2894. [[CrossRef](#)]
31. Nourani, V.; Baghanam, A.H.; Adamowski, J.; Kisi, O. Applications of hybrid wavelet–Artificial Intelligence models in hydrology: A review. *J. Hydrol.* **2014**, *514*, 358–377. [[CrossRef](#)]
32. Deo, R.C.; Wen, X.; Qi, F. A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset. *Appl. Energy* **2016**, *168*, 568–593. [[CrossRef](#)]
33. Cornish, C.R.; Bretherton, C.S.; Percival, D.B. Maximal overlap wavelet statistical analysis with application to atmospheric turbulence. *Bound.-Layer Meteorol.* **2006**, *119*, 339–374. [[CrossRef](#)]
34. Rathinasamy, M.; Khosa, R.; Adamowski, J.; Ch, S.; Partheepan, G.; Anand, J.; Narsimlu, B. Wavelet-based multiscale performance analysis: An approach to assess and improve hydrological models. *Water Resour. Res.* **2014**, *50*, 9721–9737. [[CrossRef](#)]
35. Di, C.; Yang, X.; Wang, X. A four-stage hybrid model for hydrological time series forecasting. *PLoS ONE* **2014**, *9*, e104663. [[CrossRef](#)] [[PubMed](#)]
36. Prasad, R.; Deo, R.C.; Li, Y.; Maraseni, T. Weekly soil moisture forecasting with multivariate sequential, ensemble empirical mode decomposition and Boruta-random forest hybridiser algorithm approach. *Catena* **2019**, *177*, 149–166. [[CrossRef](#)]
37. Seo, Y.; Kim, S. Hydrological Forecasting Using Hybrid Data-Driven Approach. *Am. J. Appl. Sci.* **2016**, *13*, 891–899. [[CrossRef](#)]
38. Beltrán-Castro, J.; Valencia-Aguirre, J.; Orozco-Alzate, M.; Castellanos-Domínguez, G.; Travieso-González, C.M. Rainfall forecasting based on ensemble empirical mode decomposition and neural networks. In Proceedings of the International Work-Conference on Artificial Neural Networks, Tenerife, Spain, 12–14 June 2013; pp. 471–480.
39. Jiao, G.; Guo, T.; Ding, Y. A new hybrid forecasting approach applied to hydrological data: A case study on precipitation in Northwestern China. *Water* **2016**, *8*, 367. [[CrossRef](#)]
40. Ouyang, Q.; Lu, W.; Xin, X.; Zhang, Y.; Cheng, W.; Yu, T. Monthly rainfall forecasting using EEMD-SVR based on phase-space reconstruction. *Water Resour. Manag.* **2016**, *30*, 2311–2325. [[CrossRef](#)]
41. Ali, M.; Deo, R.C.; Maraseni, T.; Downs, N.J. Improving SPI-derived drought forecasts incorporating synoptic-scale climate indices in multi-phase multivariate empirical mode decomposition model hybridized with simulated annealing and kernel ridge regression algorithms. *J. Hydrol.* **2019**, *576*, 164–184. [[CrossRef](#)]
42. Adarsh, S.; Sanah, S.; Murshida, K.; Nooramol, P. Scale dependent prediction of reference evapotranspiration based on Multi-Variate Empirical mode decomposition. *Ain Shams Eng. J.* **2018**, *9*, 1839–1848. [[CrossRef](#)]
43. Hu, W.; Si, B.C. Soil water prediction based on its scale-specific control using multivariate empirical mode decomposition. *Geoderma* **2013**, *193*, 180–188. [[CrossRef](#)]
44. Schepen, A.; Wang, Q.J.; Robertson, D. Evidence for Using Lagged Climate Indices to Forecast Australian Seasonal Rainfall. *J. Clim.* **2012**, *25*, 1230–1246. [[CrossRef](#)]
45. Yuan, C.; Yamagata, T. Impacts of IOD, ENSO and ENSO Modoki on the Australian winter wheat yields in recent decades. *Sci. Rep.* **2015**, *5*, 1–8. [[CrossRef](#)]
46. Risbey, J.S.; Pook, M.J.; McIntosh, P.C.; Wheeler, M.C.; Hendon, H.H. On the remote drivers of rainfall variability in Australia. *Mon. Weather Rev.* **2009**, *137*, 3233–3253. [[CrossRef](#)]
47. Royce, F.S.; Fraisse, C.W.; Baigorria, G.A. ENSO classification indices and summer crop yields in the Southeastern USA. *Agric. For. Meteorol.* **2011**, *151*, 817–826. [[CrossRef](#)]
48. Shuai, J.; Zhang, Z.; Sun, D.-Z.; Tao, F.; Shi, P. ENSO, climate variability and crop yields in China. *Clim. Res.* **2013**, *58*, 133–148. [[CrossRef](#)]
49. Rashid, M.M.; Sharma, A.; Johnson, F. Multi-model drought predictions using temporally aggregated climate indicators. *J. Hydrol.* **2020**, *581*. [[CrossRef](#)]
50. Nikolopoulos, E.I.; Anagnostou, E.N.; Borga, M. Using high-resolution satellite rainfall products to simulate a major flash flood event in northern Italy. *J. Hydrometeorol.* **2013**, *14*, 171–185. [[CrossRef](#)]
51. Nikolopoulos, E.I.; Anagnostou, E.N.; Hossain, F.; Gebremichael, M.; Borga, M. Understanding the scale relationships of uncertainty propagation of satellite rainfall through a distributed hydrologic model. *J. Hydrometeorol.* **2010**, *11*, 520–532. [[CrossRef](#)]
52. Yong, B.; Hong, Y.; Ren, L.L.; Gourley, J.J.; Huffman, G.J.; Chen, X.; Wang, W.; Khan, S.I. Assessment of evolving TRMM-based multisatellite real-time precipitation estimation methods and their impacts on hydrologic prediction in a high latitude basin. *J. Geophys. Res. Atmos.* **2012**, *117*. [[CrossRef](#)]
53. Ghimire, S.; Deo, R.C.; Downs, N.J.; Raj, N. Self-adaptive differential evolutionary extreme learning machines for long-term solar radiation prediction with remotely-sensed MODIS satellite and Reanalysis atmospheric products in solar-rich cities. *Remote Sens. Environ.* **2018**, *212*, 176–198. [[CrossRef](#)]
54. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436. [[CrossRef](#)] [[PubMed](#)]
55. Oehmcke, S.; Zielinski, O.; Kramer, O. Input quality aware convolutional LSTM networks for virtual marine sensors. *Neurocomputing* **2018**, *275*, 2603–2615. [[CrossRef](#)]

56. Nunez, J.C.; Cabido, R.; Pantrigo, J.J.; Montemayor, A.S.; Velez, J.F. Convolutional neural networks and long short-term memory for skeleton-based human activity and hand gesture recognition. *Pattern Recognit.* **2018**, *76*, 80–94. [CrossRef]
57. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
58. Torres, M.E.; Colominas, M.A.; Schlotthauer, G.; Flandrin, P. A complete ensemble empirical mode decomposition with adaptive noise. In Proceedings of the 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP), Prague, Czech Republic, 22–27 May 2011; pp. 4144–4147.
59. Prasad, R.; Deo, R.C.; Li, Y.; Maraseni, T. Ensemble committee-based data intelligent approach for generating soil moisture forecasts with multivariate hydro-meteorological predictors. *Soil Tillage Res.* **2018**, *181*, 63–81. [CrossRef]
60. Wen, X.; Feng, Q.; Deo, R.C.; Wu, M.; Yin, Z.; Yang, L.; Singh, V.P. Two-phase extreme learning machines integrated with the complete ensemble empirical mode decomposition with adaptive noise algorithm for multi-scale runoff prediction problems. *J. Hydrol.* **2019**, *570*, 167–184. [CrossRef]
61. Wu, Z.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41. [CrossRef]
62. Wu, Z.; Huang, N.E.; Chen, X. The multi-dimensional ensemble empirical mode decomposition method. *Adv. Adapt. Data Anal.* **2009**, *1*, 339–372. [CrossRef]
63. Bowden, G.J.; Dandy, G.C.; Maier, H.R. Input determination for neural network models in water resources applications. Part 1—background and methodology. *J. Hydrol.* **2005**, *301*, 75–92. [CrossRef]
64. Maier, H.R.; Jain, A.; Dandy, G.C.; Sudheer, K.P. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environ. Model. Softw.* **2010**, *25*, 891–909. [CrossRef]
65. Yang, W.; Wang, K.; Zuo, W. Neighborhood Component Feature Selection for High-Dimensional Data. *JCP* **2012**, *7*, 161–168. [CrossRef]
66. Murray–Darling Basin Authority. *Guide to the Proposed Basin Plan*; Murray–Darling Basin Auth.: Canberra, Australia, 2010.
67. Australian Bureau of Statistics. *Household Use of Information Technology*; Australia Bureau of Statistics: Canberra, Australia, 2010.
68. Hijmans, R.J.; Cameron, S.E.; Parra, J.L.; Jones, P.G.; Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol. J. R. Meteorol. Soc.* **2005**, *25*, 1965–1978. [CrossRef]
69. ASRIS. *The Australian Soil Resource Information System*; Department of Agriculture, Fisheries and Forestry: Canberra, Australia, 2014. Available online: <https://www.asris.csiro.au/> (accessed on 12 December 2020).
70. BOM. Bureau of Meteorology. 2020. Available online: <http://www.bom.gov.au/> (accessed on 31 December 2020).
71. Deo, R.C.; Sahin, M. Forecasting long-term global solar radiation with an ANN algorithm coupled with satellite-derived (MODIS) land surface temperature (LST) for regional locations in Queensland. *Renew. Sustain. Energy Rev.* **2017**, *72*, 828–848. [CrossRef]
72. Deo, R.C.; Şahin, M.; Adamowski, J.F.; Mi, J. Universally deployable extreme learning machines integrated with remotely sensed MODIS satellite predictors over Australia to forecast global solar radiation: A new approach. *Renew. Sustain. Energy Rev.* **2019**, *104*, 235–261. [CrossRef]
73. Deo, R.C.; Syktus, J.I.; McAlpine, C.A.; Lawrence, P.J.; McGowan, H.A.; Phinn, S.R. Impact of historical land cover change on daily indices of climate extremes including droughts in eastern Australia. *Geophys. Res. Lett.* **2009**, *36*. [CrossRef]
74. Nguyen-Huy, T.; Deo, R.C.; Mushtaq, S.; An-Vo, D.-A.; Khan, S. Modeling the joint influence of multiple synoptic-scale, climate mode indices on Australian wheat yield using a vine copula-based approach. *Eur. J. Agron.* **2018**, *98*, 65–81. [CrossRef]
75. Berrick, S.W.; Leptoukh, G.; Farley, J.D.; Rui, H. Giovanni: A web service workflow-based data visualization and analysis system. *IEEE Trans. Geosci. Remote Sens.* **2008**, *47*, 106–113. [CrossRef]
76. Chen, C.; Jiang, H.; Zhang, Y.; Wang, Y. Investigating spatial and temporal characteristics of harmful Algal Bloom areas in the East China Sea using a fast and flexible method. In Proceedings of the 2010 18th International Conference on Geoinformatics, Beijing, China, 18–20 June 2010; pp. 1–4.
77. Morshed, A.; Aryal, J.; Dutta, R. Environmental spatio-temporal ontology for the Linked open data cloud. In Proceedings of the 2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, Melbourne, VIC, Australia, 16–18 July 2013; pp. 1907–1912.
78. Trouet, V.; Van Oldenborgh, G.J. KNMI Climate Explorer: A web-based research tool for high-resolution paleoclimatology. *Tree-Ring Res.* **2013**, *69*, 3–13. [CrossRef]
79. Adnan, M.; Rehman, N.; Sheikh, M.; Khan, A.; Mir, K.; Khan, M. Influence of natural forcing phenomena on precipitation of Pakistan. *Pak. J. Meteorol.* **2016**, *12*, 23–35.
80. Philander, S.G.H. El Nino southern oscillation phenomena. *Nature* **1983**, *302*, 295–301. [CrossRef]
81. Chiew, F.H.; Piechota, T.C.; Dracup, J.A.; McMahon, T.A. El Nino/Southern Oscillation and Australian rainfall, streamflow and drought: Links and potential for forecasting. *J. Hydrol.* **1998**, *204*, 138–149. [CrossRef]
82. Madden, R.A.; Julian, P.R. Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.* **1971**, *28*, 702–708. [CrossRef]
83. Henley, B.J.; Gergis, J.; Karoly, D.J.; Power, S.; Kennedy, J.; Folland, C.K. A tripole index for the interdecadal Pacific oscillation. *Clim. Dyn.* **2015**, *45*, 3077–3090. [CrossRef]
84. Troup, A. The ‘southern oscillation’. *Q. J. R. Meteorol. Soc.* **1965**, *91*, 490–506. [CrossRef]
85. Ketkar, N. Introduction to keras. In *Deep Learning with Python*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 97–111.

86. Brownlee, J. *Deep Learning with Python: Develop Deep Learning Models on Theano and TensorFlow Using Keras*; Machine Learning Mastery: Vermont, VIC, Australia, 2016.
87. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M. Tensorflow: A system for large-scale machine learning. In Proceedings of the 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16), Savannah, GA, USA, 2–4 November 2016; pp. 265–283.
88. Jayalakshmi, T.; Santhakumaran, A. Statistical normalization and back propagation for classification. *Int. J. Comput. Theory Eng.* **2011**, *3*, 1793–8201.
89. Maier, H.R.; Dandy, G.C. Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications. *Environ. Model. Softw.* **2000**, *15*, 101–124. [[CrossRef](#)]
90. Deo, R.C.; Downs, N.; Parisi, A.V.; Adamowski, J.F.; Quilty, J.M. Very short-term reactive forecasting of the solar ultraviolet index using an extreme learning machine integrated with the solar zenith angle. *Environ. Res.* **2017**, *155*, 141–166. [[CrossRef](#)]
91. Arhami, M.; Kamali, N.; Rajabi, M.M. Predicting hourly air pollutant levels using artificial neural networks coupled with uncertainty analysis by Monte Carlo simulations. *Environ. Sci. Pollut. Res.* **2013**, *20*, 4777–4789. [[CrossRef](#)]
92. Jekabsons, G. ARESLab: Adaptive Regression Splines Toolbox for Matlab/Octave. 2011. Available online: <http://www.cs.rtu.lv/jekabsons> (accessed on 18 January 2021).
93. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [[CrossRef](#)]
94. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* **2009**, *377*, 80–91. [[CrossRef](#)]
95. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [[CrossRef](#)]
96. Willmott, C.J.; Robeson, S.M.; Matsuura, K. A refined index of model performance. *Int. J. Climatol.* **2012**, *32*, 2088–2094. [[CrossRef](#)]
97. Legates, D.R.; McCabe, G.J. A refined index of model performance: A rejoinder. *Int. J. Climatol.* **2013**, *33*, 1053–1056. [[CrossRef](#)]
98. Yin, Z.; Feng, Q.; Wen, X.; Deo, R.C.; Yang, L.; Si, J.; He, Z. Design and evaluation of SVR, MARS and M5Tree models for 1, 2 and 3-day lead time forecasting of river flow data in a semiarid mountainous catchment. *Stoch. Environ. Res. Risk Assess.* **2018**, *32*, 2457–2476. [[CrossRef](#)]
99. Legates, D.R.; McCabe Jr, G.J. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, *35*, 233–241. [[CrossRef](#)]
100. Friedman, J.H. *Estimating Functions of Mixed Ordinal and Categorical Variables Using Adaptive Splines*; Stanford University CA Lab for Computational Statistics: Stanford, CA, USA, 1991.
101. Ghimire, S.; Deo, R.C.; Downs, N.J.; Raj, N. Global solar radiation prediction by ANN integrated with European Centre for medium range weather forecast fields in solar rich cities of Queensland Australia. *J. Clean. Prod.* **2019**, *216*, 288–310. [[CrossRef](#)]
102. Cai, Y.; Zheng, W.; Zhang, X.; Zhangzhong, L.; Xue, X. Research on soil moisture prediction model based on deep learning. *PLoS ONE* **2019**, *14*, e0214508. [[CrossRef](#)]