Madhukumar, Neethu (2022) ***Decision support system for precision agriculture using deep learning.*** **PhD Thesis, James Cook University.**

Decision Support System for Precision Agriculture using Deep Learning

Neethu Madhukumar

Master's in Electronics and Communication Engineering

Bachelor's in Electronics and Communication Engineering

Neethu Madhukumar's ORCID

A thesis submitted for the degree of Doctor of Philosophy at James Cook University, Australia in 2022

College of Science and Engineering

Electronic Systems and Internet of Things Engineering

Acknowledgments

My Ph.D. research and thesis would not have been possible without the support, motivation, and assistance of my supervisors, colleagues, family, and friends.

I would like to thank my supervisor Dr. Eric Wang whose outstanding guidance and continuous assistance enabled me to complete this thesis. His valuable guidance in all aspects of this thesis has been instrumental in overall thesis upliftment.

I thank my secondary supervisor, Prof. Wei Xiang, for his valuable comments and expert advice in my thesis project. Working with Prof. Wei Xiang has been an opportunity to improve my research through his insightful and professional feedback.

Thank you, Dr. Yi-Fan Zhang and Prof. Clinton Fookes, for the expert technical insights provided in the weather and soil moisture research chapters.

I am grateful for all the support and encouragement provided at various stages of my Ph.D. research by my colleagues, including but not limited to Dr. Tao Huang, Assoc. Prof. Rabin Tuladhar, Dr. Bronson Philippa, Prof. Mohan Jacob, Ms. Christine Martyn, Ms. Shangyang Guo, Ms. Laura Blackmore, Ms. Bethany Keats, Ms. Melissa Norton, Ms. Debbie Berry, and Ms. Jodie Wilson.

Special thanks to Ms. Linden Woodward for helping in taking my thesis findings to a wider audience through JCU media and Assoc. Prof. Elizabeth Tynan for thesis editorial support as part of JCU's skills for international postgraduate program.

I am also thankful to all my friends in Australia and India for their support and prayers throughout this academic journey.

To my husband Ashwin Antony, thanks for always being my pillar of strength. Your support means the world. Finally, I dedicate this work to our family (mine and Ashwin's) for their endless love and support.

Statement of the Contribution of Others

I acknowledge the following contributions to this work with gratitude.

Financial assistance: The JCU Postgraduate Research Scholarship provided the final assistance, which included living assistance and tuition fee support. The annual funding and additional grants from the College of Science and Engineering and Graduate Research School enabled the publication of the content of this thesis.

Editorial assistance: My supervisors, Dr. Eric Wang, and Prof. Wei Xiang provided editorial assistance for my thesis. Chapters 3-6 are published or submitted for publication. These chapters received editorial support from my supervisors and co-authors, Dr. Yi-Fan Zhang (chapter 3) and Prof. Clinton Fookes (chapter 4).

Contributions to co-authored publications: In co-authored publications, the contributions from co-authors were editorial assistance and technical feedback. I am the first author of all publications related to this thesis. I developed all models, performed experiments, analyzed results, compared performance with existing state-of-the-art models, drew figures, graphs and tables, and wrote all text contents.

Data: The data used in chapter 3 is from the open-access data from the Bureau of Meteorology, Australia. In chapter 4, the data used is the open-access data from NASA earthdata. In chapters 5-6, open-access data from the City of Melbourne-open data portal is used.

Abstract

Precision agriculture is a highly efficient farming practice in which precision irrigation (PI) targets the optimal use of water. For implementing PI, an irrigation decision support system (IDSS) is required, which needs information from rainfall forecasts and soil moisture (SM). Although the spatial accuracy of official rain forecasts inferred from climate models (CMs) has been improving, the temporal accuracy is not good enough for PI scheduling. In the case of SM measurement, we may need to rely on satellites as deploying large-scale in-situ sensors is not feasible. Existing satellites require an overall improvement in spatial and depth aspects for PI. The advent of deep learning (DL) offers new opportunities for considerably improving rain and SM information. However, existing DL approaches require further development and adaptation in the spatio-temporal context for rain forecasts and satellite SM data. Therefore, the development of such DL models is pursued in this thesis.

In the first phase of this thesis, to improve CM's rainfall forecasting accuracy, a hybrid climate learning model (HCLM) was developed. Conventionally, from multiple CM forecasts, the one closest in time to the prediction target is assumed to be the best. However, the preliminary analysis conducted in this thesis revealed that the latest is not always the best forecast. Hence, the multiple CM forecasts are input to the HCLM for time-series refinement. The HCLM has a novel probabilistic network layer (considering the stochastic nature of rain forecasts) and a new sequential analysis layer (correlating the forecasts with corresponding CM observation data), making it a unique and highly suitable DL model for rain forecast refinement. The HCLM refined CM forecast highly accurately across six major climate zones in Australia.

Apart from rainfall, as SM information is required in PI decisions, improving SM was considered next. Downscaling, high-resolution spatially distributed SM data from satellites using existing methods generates empty spaces (spectral information loss) due to data corruptions, leading to the loss of farm SM information. Literature studies revealed that the SM data changes are spatio-temporal, and downscaled image pixels have bi-directional associations. Motivated by these factors, in the second phase of this thesis, a new DL downscaling method was developed which, unlike existing approaches, simultaneously associates bi-directionally the spectra-spatio-temporal

sequences for better downscaling. This considerably reduced information loss in the proposed downscaling compared to the state-of-the-art methods.

But, the downscaled satellite image can only reveal spatially distributed SM and not the moisture information across deeper soil layers. This depth information is directly measurable using large-scale root-zone soil moisture (RZSM) sensor deployment, but this is not economical, therefore, we need to explore in-direct means of RZSM estimation. For in-direct RZSM estimation, the auxiliary RZSM data (like weather, soil, and plant characteristics) associations within a location and multiple sensor sites with the target need to be found. Conventional transformer neural networks (TNN), the most suitable DL model for such sequential associations, cannot incorporate such simultaneous multi-associations. Hence in the third phase of this thesis, a new TNN model is developed called the hybrid TNN model, which can perform this action. From the analysis conducted in this thesis, the proposed model showed better RZSM estimation than other popular recent related works.

Finally, the models developed in phases one, two, and three of this thesis are integrated into an IDSS to provide refined rainfall forecast and SM information for irrigation scheduling. This integrated system considerably enhanced irrigation water and cost savings compared to existing state-of-the-art related works in literature.

List of Publications

[1] **Neethu Madhukumar**, Eric Wang, Yi-Fan Zhang, and Wei Xiang, "Consensus Forecast of Rainfall Using Hybrid Climate Learning Model," *IEEE Internet of Things Journal*, vol. 8, pp. 7270-7278, May 2021 (Impact Factor: 10.238, h-index: 144, Journal rank (Computing systems): 1).

[2] **Neethu Madhukumar**, Eric Wang, Clinton Fookes, and Wei Xiang,"Three-dimensional Bi-directional LSTM for Satellite Soil Moisture Downscaling," *IEEE Transactions on Geoscience and Remote Sensing*, Accepted for publication on 22 Nov. 2022 (Impact Factor: 8.125, h-index: 113, Journal rank (Remote sensing): 3).

[3] **Neethu Madhukumar**, Eric Wang, and Wei Xiang, "Hybrid Transformer Network for Root Zone Soil Moisture Estimation for Decision Support in Precision Irrigation," *Agricultural Water Management*, Under review (Impact Factor: 6.611, h-index: 65, Journal rank (Agronomy & crop science): 7).

Research Impact

1. Altmetrics Attention Score: Twenty-two, indicating a ranking in the top 25% of all research outputs scored by Altmetric.
2. News Articles: 40 plus news articles published worldwide, such as:
   a) American Society of Mechanical Engineers: IoT Bolsters Weather Forecasting for Farmers
   b) SBS Australia: Women's Involvement in Environmental Protection.
3. Social Media
   a) James Cook University (JCU) Twitter
   b) JCU Discover Facebook Page
   c) JCU LinkedIn
4. Mainstream Media

a) Australian Broadcasting Corporation (ABC) Radio: Mention in Far North Qld Local News

b) 7News Australia: Queensland researchers have been ’under the weather’, working on new AI climate forecasting

5. Symposiums and Three Minute Thesis

   a) Three Minute Thesis (3MT) Division Level Competition: Nominated to represent college of science and engineering (CSE), JCU, Aug. 2019.

   b) Post Graduate Symposium: 3rd Prize winner, JCU, Jun. 2019.

   c) Post Graduate Symposium: 3rd Prize winner, JCU, Nov. 2018.

6. Ph.D. Spotlights

   a) JCU International Student Stories: The international student experience at JCU, Australia.

   b) Women in Earth and Environmental Sciences in Australasia (WOMEESA) Spotlight: Ph.D. journey at JCU, Australia.

7. Real-world Implications: Taking the research to a wider audience through professional membership activities with:

   a) WOMEESA

   b) Women in Technology (WiT)

   c) e-Agriculture (part of Food and Agriculture Organization of the United Nations)

Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviations	Expansions
1D	One-Dimensional
2D	Two-Dimensional
3D	Three-Dimensional
3D-Bi-LSTM	Three-Dimensional Bi-directional Long Short-Term Memory
ACCESS	Australian Community Climate Earth-System Simulator
AMSR2	Advanced Microwave Scanning Radiometer-2
ANN	Artificial Neural Network
ASCAT	Advanced Scatterometer
Bi-LSTM	Bi-directional Long Short-Term Memory
BOM	Bureau of Meteorology
CM	Climate Model
CNN	Convolutional Neural Network
D-MLP	Dynamic Multi-Layer Perceptron
DI	Drought Indices
DL	Deep Learning
DOY	Date Of Year
DNN	Deep Neural Network
ESA	European Space Agency
ESM	Earth System Modeling
EVI	Enhanced Vegetation Index
GCM	General Circulation Model
GRU	Gated Recurrent Unit
HCLM	Hybrid Climate Learning Model
HD-LSTM	Hybrid Deep Long Short-Term Memory
IDSS	Irrigation Decision Support System
IDSS-PI	Irrigation Decision Support System for Precision Irrigation
LSTM	Long Short-Term Memory
ML	Machine Learning
MLP	Multi-Layer Perceptron

MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NMBE	Normalized Mean Bias Error
NSE	Nash Suttcliff coefficient of Efficiency
PA	Precision Agriculture
PI	Precision Irrigation
PMLP	Probabilistic Multi-Layer Perceptron
POAMA	Predictive Ocean Atmosphere Model for Australia
PP	Plant Parameters
R	Correlation Coefficient
R^2	Determination Coefficient
RF	Random Forest
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
RS	Remote Sensing
RZSM	Root Zone Soil Moisture
SHAP	SHapley Additive exPlanations
SLP	Single-Layer Perceptron
SM	Soil Moisture
SMAP	Soil Moisture Active Passive
SMI	Soil Moisture Indice
SMOS	Soil Moisture and Ocean Salinity
SSM	Surface Soil Moisture
SWB	Soil Water Balance
TNN	Transformer Neural Network
TVDI	Temperature-Vegetation Dryness Index
ubRMSE	unbiased Root Mean Square Error
VI	Vegetation Indices

Chapter 1

Introduction

1.1 Background and Motivation

Precision agriculture (PA) is a new agriculture practice that uses technology to gather, combine, process, and analyze temporal, spatial, and individual data to support management decisions for improved resource use according to farm requirements [4]-[6]. Among the resources, water is highly critical, as irrigation accounts for 70% of total freshwater withdrawals on average [7] and will further increase if improved practices are not adopted since irrigated land will further increase by 17% by 2050 [8], [9]. The PA approach, which targets the optimal usage of this crucial resource, is called precision irrigation (PI) [10] - [14].

The PI approach applies water to the plant to a selected location at the right time in small measured quantities to provide optimal growing conditions [15]. For practical application in any selected location, a PI system needs an irrigation decision support system (IDSS) to "know" when to irrigate, and how much water to apply [15], [16]. The IDSS deduces this information by analyzing and evaluating the farm rainfall and soil moisture (SM) data since, apart from irrigation, plants can also receive water from rain or use the water stored in the soil as moisture to meet their requirements [17] - [21].

Meteorological forecasts can provide information on the amount of future rainfall that may fall on a farm. Currently, official meteorological forecasts in most countries (including Australia) are produced using climate models (CMs) [22], [23]. Although the spatial prediction precision has improved significantly over the last decade, providing more regional forecasts for local farms, the temporal accuracy of CM forecasts needs further refinement to reduce IDSS irrigation schedule overlaps with rainfall [24]. Therefore, we should explore new methods to enhance the temporal forecasting accuracy of the CM to improve IDSS irrigation scheduling.

Next, to measure SM, in-situ sensors can be used, which can provide information at high temporal frequency but not the distributed farm SM details leading to a general dependency on satellites for this information [25]. However, existing SM satellite missions do not provide the spatial and depth resolutions required for PI application [26], [27]. Therefore, better spatial and depth resolution SM information has to be derived for plant water need assessments by IDSS.

For better temporal, spatial, and depth information extraction from real data sources (such as CMs, satellites, and in-situ sensors), compared to conventional approaches, deep learning (DL) methods have shown higher performance when a large amount of data is available [28] - [30]. Currently, a deluge of open-source data is available from CMs, satellites, and in-situ sensors, therefore, DL methods look very suitable to improve spatio-temporal information on rainfall and SM for the proposed application.

To summarise, this thesis will explore novel DL-based solutions for better information extraction of four key elements (a) rainfall forecast, (b) spatial SM distribution: surface soil moisture (SSM), (c) depth SM distribution: root zone soil moisture (RZSM), and (d) PI decisions.

1.2 Research Problems

This thesis focuses on improving the decisions taken for PI through better information extraction using novel DL methods. The research problems investigated include the development of a model for refining CM rainfall forecasts, a downscaling model for obtaining high spatial resolution daily satellite SSM, methods for estimating RZSM at required depths for target locations, and improving irrigation decisions using rainfall and SM information. These research problems are detailed as follows.

1.2.1 Rainfall Forecast Refining: To Enhance Temporal Forecasts Accuracy

The Bureau of Meteorology (BOM), Australia progressed from Predictive Ocean Atmosphere Model for Australia (POAMA) CM to the Australian Community Climate

Earth-System Simulator (ACCESS) CM, resulting in better spatial accuracy [24]. The current BOM model, ACCESS, produces multiple time-series rainfall forecasts for a target day. The latest rainfall forecast is assumed to be the best as it is closest in time to the actual event. However, the preliminary analysis conducted in this thesis revealed that the BOM's latest forecast is not always the best. In most cases, better rain forecasts were available in preceding time instants. Hence, finding the best forecast from multiple CM forecasts can enhance the accuracy of CM forecasts. A DL-based approach to refine CM forecasts is proposed, as existing studies reveal a comparative higher ability in DL methods to improve similar CM outputs [31], [24]. Since the target problem is time series, a sequential DL model will be suitable for refining the CM forecasts. However, existing state-of-the-art sequential DL models need to be made adaptable to the features of rainfall forecast data [18], [11] by incorporating additional layers for (a) capturing underlying minute functional relationships with related climate variables, such as temperature [32], (b) untangling the random forecasting structures [33], and (c) finding hidden forecasting patterns [34] essential for forecast refinement. Hence, a sequential DL-based climate learning model that integrates these essential data processing layers to refine the BOM's rainfall forecasts is proposed.

1.2.2 Surface Soil Moisture (SSM): To Obtain High-Resolution SM Spatial Distribution With Reduced Information Loss

The best daily SSM resolution provided by existing satellite missions is 9 km [25]. This spatial resolution is significantly low for agriculture applications [35]. Hence, we should downscale this SSM to a higher resolution. However, existing methods that target higher SSM spatial resolution, result in numerous empty spaces (loss of spectral information) in the downscaled image due to satellite data corruptions in selected days [36], [37]. Existing literature illustrates the changes in SSM have bi-directional spatio-temporal associations [38], [39]. Therefore, along with bi-directional spatial data analysis, incorporating an additional SSM image (2D) analysis over time (1D) bi-directionally and using this to predict the missing image pieces can help reduce the spectral information losses in downscaled output. But, this requires simultaneous bi-directional examination of the spectral sequence in the spatial (2D) and time (1D) domains. Unfortunately, existing downscaling models do not support this mode of analysis due to their structural limitations and the complex bi-directional 3D (2D plus 1D) data associations [40] - [43]. This thesis addresses these research problems by developing a DL model with the desired bi-directional 3D analysis structure for obtaining high-resolution SSM with reduced information loss.

1.2.3 Root Zone Soil Moisture (RZSM): To Determine SM Depth Distribution

Downscaled satellite SM provides information on SSM (0-5 cm). For knowledge of SM depth distribution ($>$ 5 cm), RZSM information is needed [44]. Direct RZSM measurement through large-scale in-situ sensor deployment at the subsurface will disturb the soil properties and is not economical [45]. Hence, in-direct RZSM estimation methods are required. The literature illustrates that RZSM at a location is related to changing weather, soil, and plant characteristics [46]. Therefore these characteristics can provide auxiliary information on the RZSM changes. These auxiliary RZSM variables can be associated with available sensor data, and relationships can be extrapolated to non-sensor locations to get the spatially distributed RZSM. Compared to existing methods, DL is most suitable for such data associations as they are autotuned to extract such relative relationships from such diverse big data [29], [30]. Among DL models, sequential models are apt for finding these relationships as all these variables are time-series sequences [47]- [56]. However, for in-direct RZSM estimation, the data associations within a location and multiple sensor sites with the target need to be found. Existing state-of-the-art DL models do not incorporate such simultaneous multi-associations [47] - [56]. Hence, in this thesis, such a DL model is developed, which combines ground and satellite-based information for in-direct RZSM estimation.

1.2.4 Irrigation Decision Support System (IDSS): To Determine When and How Much Irrigation Required

Existing scholarly literature highlight the importance of integrating rainfall forecasts and SM information into the IDSS to enhance the accuracy of its decisions on when and how much water to apply [17]. Due to inaccuracy in existing rainfall forecasting and SM measurement systems, current PI-based IDSS results in inaccurate irrigation scheduling [17], [50], [51]. Furthermore, existing IDSS shows RZSM error accumulation problems due to faulty SM inputs [52], [53] and over-irrigation issues due to overlaps with rainfall [17], [54], [55]. To address these problems, the DL-based forecast and SSM/RZSM models developed in this thesis are integrated to provide irrigation decision support.

1.3 Original Contributions

To address the research problems outlined in Section 1.2., the following original contributions to the literature are presented in this thesis:

1. Chapter 3: A new DL model that can learn from the past CM forecasts and refine them for accurate rainfall prediction is proposed. The model uses multiple CM forecasts and corresponding observations as inputs for rainfall prediction.

2. Chapter 4: A novel three-dimensional DL model is proposed for bi-directional spectral sequence analysis in both space-time domains simultaneously. The proposed model uses optical, thermal, and microwave remote sensed data as inputs for downscaling.

3. Chapter 5: A DL-based methodology is designed to combine satellite and ground-based data for daily in-direct RZSM estimation without requiring the large-scale installation of sensing stations.

4. Chapter 6: A DL-based IDSS is proposed for PI. The proposed model provides decision support for irrigation of each plant or selected area by processing improved rainfall and SM information obtained using original contributions 1-3.

1.4 Significance

The original contribution 1 in Section 1.3 is the first CM's multiple forecast refinement model in the literature. It significantly outperformed previous rainfall prediction works in literature over ten locations across six major climate zones in Australia, offering a better data-driven solution for climate model forecast refining.

Original contribution 2 in Section 1.3 is the first approach that bi-directionally associates the spectra-spatial-temporal relationships for SSM downscaling. It reduced the information loss in the downscaled surface soil moisture map compared to existing state-of-the-art downscaling models by 36.6667% on three globally used datasets for SSM downscaling validation.

Original contribution 3 in Section 1.3 provides a new method for in-direct RZSM estimation using ground and satellite data. It reduced RZSM prediction error compared to widely used DL models in similar applications by 21.341% for three test locations in Australia.

Finally, original contribution 4 in Section 1.3 provides a new integrated IDSS model which utilizes DL-processed rain and SM data rather than unrefined information. It showed 15.632% water and 15.125% cost-saving than competing relevant models from the literature on the selected sites. Overall, this thesis fills significant gaps in the rainfall, SM, and PI literature by developing novel DL models that outperform existing state-of-the-art models in related applications.

1.5 Document Organization

The structure of this thesis is as follows. Chapter 2 reviews related literature, with a focus on research in the areas of weather forecasting, downscaling models for SSM, relevant models for RZSM estimation, and problems in existing decision support systems in PI. The gaps in existing works are identified and lead to the four research chapters.

Chapters 3-6 are the research chapters that address the problems outlined in Section 1.2 by achieving the objectives outlined in Section 1.3. The following figure provides an overview of the structure of the research chapters and the relationships between them. As Fig. 1.1 illustrates, Chapter 2 provides the details of related works. Later, the existing state-of-the-art models from the literature relevant to each research problem, are compared for performance evaluation with the proposed models in Chapters 3-6.

Chapter 3 presents a DL-based climate learning model for refining rainfall forecasts.

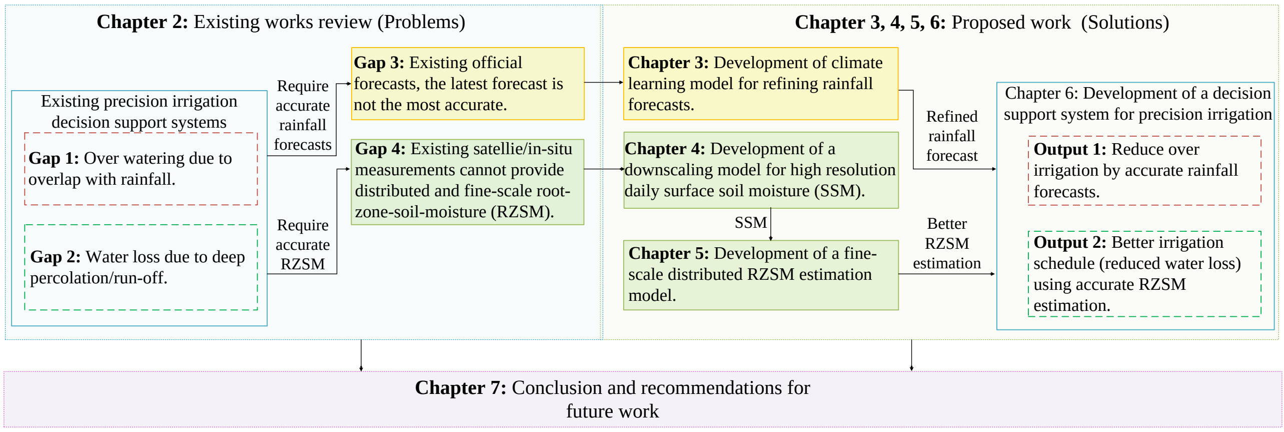


Figure 1.1: Illustration of the thesis organization.

The model uses CM forecasts and observations as inputs. Chapter 4, demonstrates a downscaling model for producing high-resolution daily SSM. Two new DL models have been designed and concatenated to form a downscaling model that can reveal the daily SSM distribution more clearly with an increased number of multi-spectral pixels. The spectrum of each pixel in the downscaled image will reveal the amount of SSM at a higher resolution spatially. Remote sensed high-resolution auxiliary SM indices (from MODIS satellite) are used as inputs to produce fine-scale SSM. Chapter 5 proposes a fine-scale distributed RZSM estimation model. The model works on a DL-based scheme to produce high-resolution RZSM by using downscaled SSM from Chapter 4 along with ground-based data. Using the rainfall forecasts from Chapter 3 and RZSM from Chapter 5 as inputs, a decision support system for PI is developed in Chapter 6. A conclusion is drawn in Chapter 7 based on the problems from Chapter 2 and solutions found from the research in Chapters 4-6. Possible future research directions are also discussed in Chapter 7.

Chapter 2

Background

This chapter reviews existing irrigation decision support systems (IDSSs) for precision irrigation (PI) and identifies their limitations and scope for improvement. Furthermore, this chapter examines existing novel models and methods for the estimation/prediction of rainfall, surface soil moisture (SSM), and root zone soil moisture (RZSM), which are identified as relevant for improving the irrigation decisions for PI.

2.1 Irrigation Decision Support System

PI is an agricultural approach that enables the application of water in small measured quantities to plants in any selected location at the correct time to provide optimal growing conditions by controlling an automatic irrigation system using an IDSS [16], [74]. The IDSS used in PI can be mainly classified as [50], [17] (1) soil water balance (SWB)-based models, (2) soil moisture (SM) based models, (3) plant parameters (PP)-based models, and (4) process-based models. Each of these models is detailed further.

The SWB IDSS model is based on the theoretical representation of a limited portion of the water cycle, the soil–plant–atmosphere interface [50]. The literature illustrates it to be a highly powerful IDSS tool under rain-fed irrigation schemes [76], [77], [19]. However, the SWB model's irrigation scheduling performance drops when the error of estimated daily RZSM accumulates [52], [53]. To overcome this SWB issue, an accurate estimation of SM in the plant root zone is essential.

In the case of the SM-based IDSS models, it compares monitored SM to the threshold moisture depletion tolerance to trigger irrigation [17]. It mainly focuses on determining when irrigation should be applied to maintain moisture in the plant root zone within an appropriate range to benefit plant growth, yield, and quality [79].

Next, the PP-based IDSS models focus on the relationship between plant water stress and soil water deficit [17]. It uses the plant-water-status-based indices to trigger irrigation [17], [78]. In conclusion, both the SM-based and PP-based models require soil water retention information. Although direct sensor measurements eliminate estimation errors, they require extensive installation of sensing systems (high cost) in the field [45]. Therefore spatially distributed accurate information without large-scale deployment of sensors is required for both SM-based and PP-based IDSS models.

Finally, the process-based models are based on the theoretical study of plants' growth processes and consider the effects of the soil-plant-atmosphere system in a holistic manner [51]. They simulate soil-plant-atmosphere dynamics (not direct measurements like SM-/PP-based models) to decide on irrigation requirements. However, some field-based calibrations are needed for practical implementation [81], [82]. For real-time irrigation scheduling, without linking to accurate weather forecasts, process-based models derived errored timing and quantity of irrigation. This error happens especially when a high rainfall event occurs soon after irrigation [17], [54], [55], indicating a need to input accurate rainfall forecasts to reduce water loss.

To summarise, IDSS models predict plant water requirements generally as a function of moisture and rain input information. Hence, the inaccuracies in these input measurements result in faulty irrigation scheduling [83], [84]. Therefore, accurate information extraction from these inputs is essential for PI.

2.2 Rainfall Forecast

For agricultural fields exposed to rainfall, irrigation decisions should be based on rainfall information [94]. For such agriculture farms, irrigation decisions taken by existing IDSS models resulted in water losses due to overlap with rain [17], [54], [55]. As rainfall has a highly random disposition, irrigation schedules should supplement rain to reduce water loss. More accurate rainfall forecasting will result in less irrigation water loss. Currently, there are two main ways to forecast rainfall: (a) climate models (CMs) and (b) DL approaches [95]. CMs simulate atmospheric operation, and DL approaches are based on the theories of various DL algorithms [95], [96]. In this section, existing CMs and DL models for rainfall forecasting are critically analyzed.

2.2.1 Climate Models

CMs are the primary tool available for investigating the response of the climate system to various kinds of atmospheric forcing (a dynamic process that forces the air to rise) for making rainfall predictions [98]. Atmospheric weather and ocean circulation form the backbone of CMs [99]. These phenomenons are governed by the fundamental laws of physics that describe the conservation of mass, energy, and momentum [95], [98].

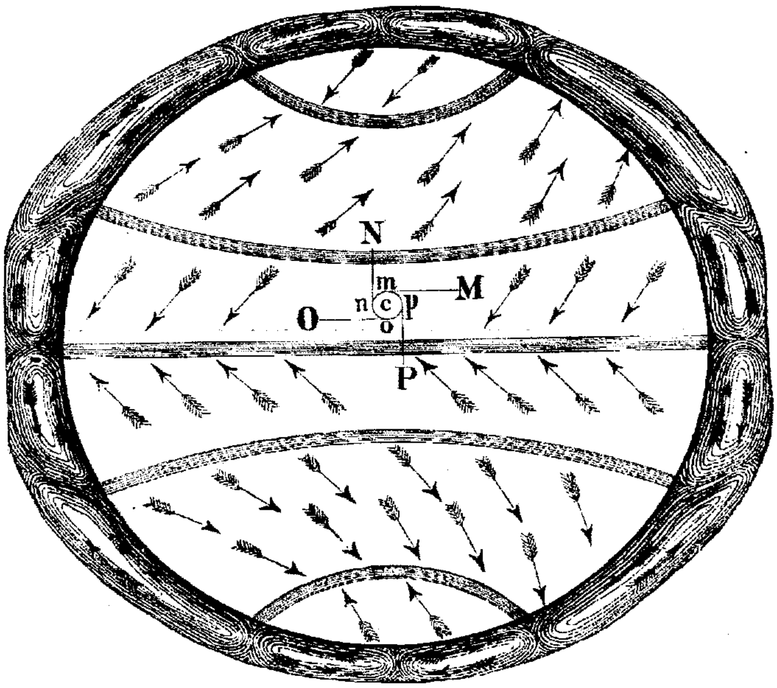


Figure 2.1: Ferrel's conceptual climate model illustrates the three-cell general circulation diagram, representing air circulations along the three axes (the poles and equator). The air circulation direction is evaluated with reference to the vectors N, M, P, and O, representing latitude, longitude, height, and depth [87].

Historically, climate modeling begins with conceptual models. These conceptual models explained the prevailing winds over the globe and captured the fundamental energy-transport function of the climate system. Among the many conceptual models, an important model is Ferrel's model for gross atmospheric circulation [87], as the atmospheric circulation concepts have not changed much from Ferrel's model. Fig. 2.1 illustrates Ferrel's model, which predicts climate based on the three-dimensional atmospheric circulation moments.

In the 19th century, mathematical models followed conceptual modeling [100], [22], [101]. Since the 1950s, the principal tools of climate science have been computer simulation models of the global general circulation. Fig. 2.2 illustrates the general circulation model (GCM). The GCM model produces weather forecasts based on the general circulation of a planetary atmosphere or ocean [23].

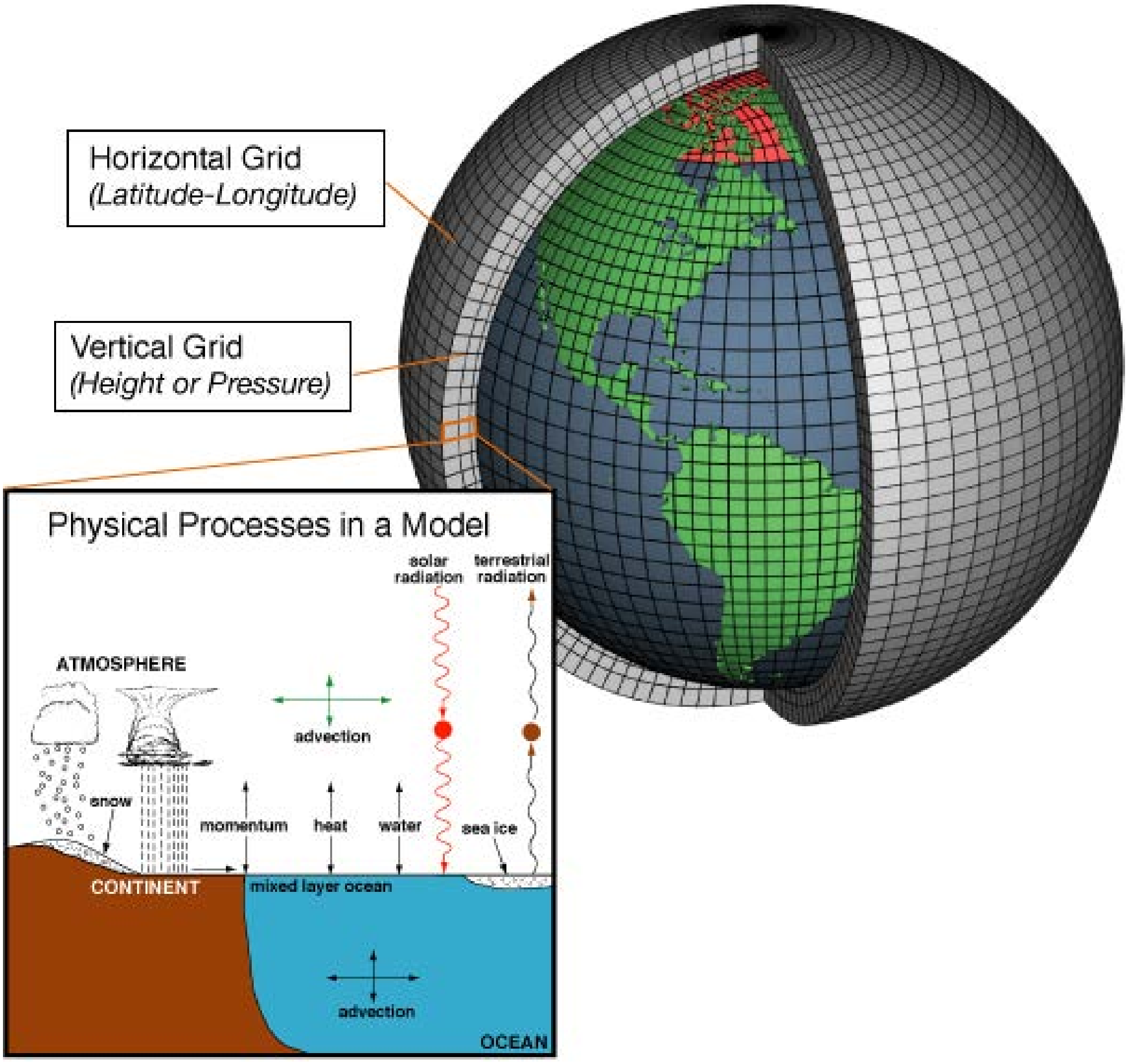


Figure 2.2: The general circulation model (GCM) is an atmospheric circulation computer simulation model [23]. The GCM computer model divides the planet into a three-dimensional grid, apply the basic equations, and evaluates the results for climate predictions. The horizontal grid represents the latitude and longitude, while the vertical grid height. Each cell (representing a particular latitude, longitude, and height) in the GCM evaluates the atmospheric changes (shown by the square illustrating the physical process in a model).

From the 1980s to the present, comprehensive coupled models of the entire climate system has dominated the field. Initially, coupled models were created at individual laboratories. However, the increase in component models led to an increasing number of mix-and-match models. In response to the rise in such models, in 2002, the National Center for Atmospheric Research (NCAR), National Oceanic and Atmospheric Administration (NOAA), National Aeronautics and Space Administration (NASA), and the Department of Defense constituencies group developed an Earth System Modeling (ESM) framework [88]. Fig. 2.3 illustrates the ESM, which couples additional information from the biogeochemical cycle systems to the GCM system for better predictions [24].

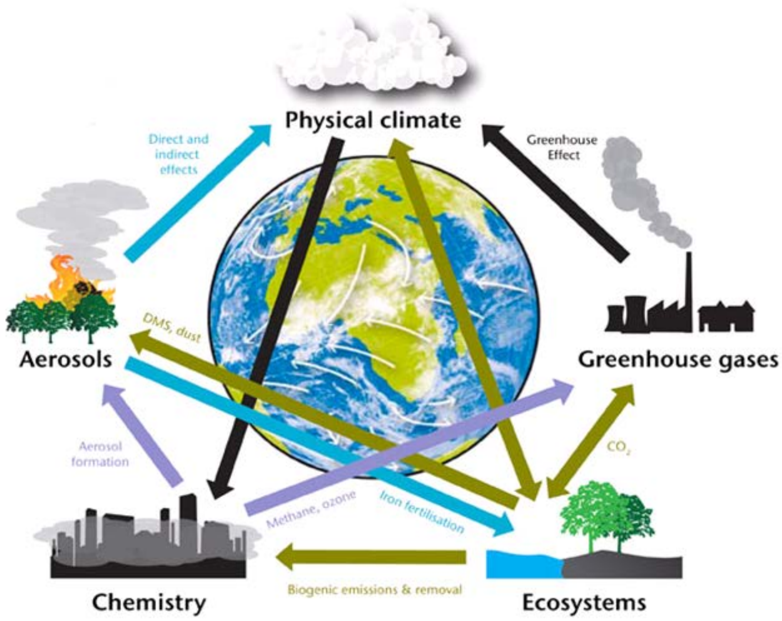


Figure 2.3: Schematic representation of the earth system model (ESM) [88]. The ESM couple the GCM to the biogeochemical cycle systems for more advanced weather predictions.

Currently, in Australia, an ESM model called the Australian Community Climate Earth-System Simulator (ACCESS) is used for weather forecasting by the Bureau of Meteorology (BOM). The GCM model called Predictive Ocean Atmosphere Model for Australia (POAMA) preceded ACCESS. ACCESS includes additional information from the biogeochemical cycle for weather forecasting, unlike POAMA, resulting in more regional forecast details (Fig. 2.4). BOM provides multiple forecasts using its CMs for the target day by updating its predictions over a period leading to the actual day. Usually, the latest forecast is believed to be the best, as it has the shortest lead time to the actual rainfall event [24]. Table 2.1 shows the monthly average performance of these multiple BOM forecasts, it can be observed that Forecast 7 (latest forecast) predictions were most accurate only for three months out of twelve in 2017. Hence, if the best forecast can be refined from multiple, it can further enhance the accuracy of CM's rain predictions.

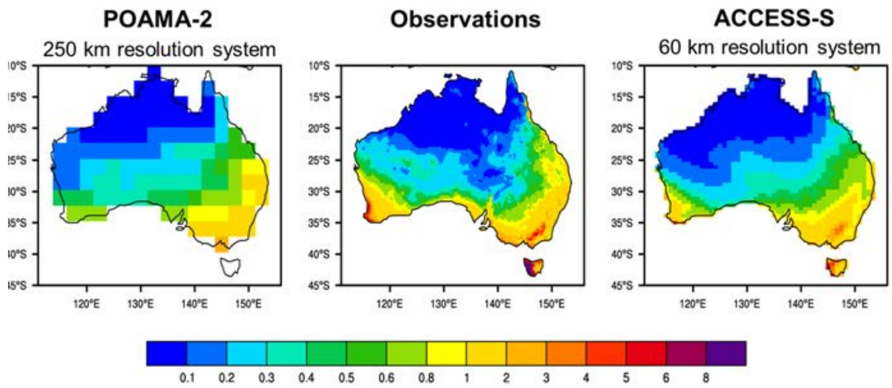


Figure 2.4: The rainfall forecasts of BOM Australia's previous model, the POAMA (GCM), and the current model ACCESS (ESM) [57]. Compared to POAMA, ACCESS has improved spatial resolution due to the addition of biogeochemical cycle information, resulting in forecasts with more regional detail.

Table 2.1: Comparison of BOM Climate Model Seven Day Forecasts using RMSE in 2017 [1]. F1 is the oldest forecast, and F7 is the latest forecast.

Month:	January	February	March	April	May	June	July	August	September	October	November	December
F1	10.485	27.955	12.086	13.916	1.575	1.040	0.862	6.334	**0.398**	3.171	2.313	**2.316**
F2	9.886	26.601	12.149	14.358	1.580	1.662	0.659	2.460	0.560	3.558	3.602	2.448
F3	8.693	25.544	11.174	**11.252**	1.554	**0.604**	0.509	2.239	1.093	4.039	**2.132**	2.437
F4	8.752	25.026	9.097	14.913	**1.391**	0.976	0.705	2.400	0.916	3.312	2.493	2.375
F5	8.369	24.992	8.710	15.864	1.563	0.706	0.723	**0.514**	0.456	**3.128**	2.753	2.454
F6	7.184	**24.266**	7.921	12.776	1.500	1.071	0.691	2.198	0.772	4.872	2.410	2.431
F7	**6.887**	25.388	**7.796**	11.368	1.855	0.610	**0.420**	2.236	1.100	4.522	2.500	2.381

Currently, the final forecast summary is usually provided by an expert using the CM, requiring long-term experience and subjective judgment, possibly causing the fluctuating accuracy of the CM's latest rainfall forecast [102]. As rainfall forecasts is the numerical relationship with the changing values of atmospheric and environmental parameters [95], [98], DL models can be applied to refine them since they can learn and execute complex analytical tasks better than existing conventional methods [103]. The details of existing DL-based models that can decode similar climate characteristics are thoroughly examined in the next section.

2.2.2 Deep Learning Models

Artificial neural networks (ANN) are traditional DL models with simple structures and shallow layers [104]. An ANN has a collection of connected units called artificial neurons, which mimics the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal to other neurons. And each connection is assigned a weight, which controls the strength of the signal transmitted. The ANN will randomize weight values before learning initially begins. The connection weights are adjusted by the ANN during training through back-propagation of errors to help reconcile the differences between the actual and predicted outcomes.

In the case of rainfall forecasting, studies have shown that ANN rainfall predictions have higher accuracy than the main prediction (final) from the CMs like GCM [104]. Previously, since the final forecast was assumed to be the best, this model was sufficient. However, as illustrated in Table 2.1, the final forecast is not the most accurate. ANNs might not be sufficient to decode the best forecast because they tend to fall into local minimum [105] and over-fitting [106] problems. Local minima problems can lead to the model choosing suboptimal points, while over-fitting can lead to the model over-fitting on training data, thus choosing a forecast that is not ideal. These issues in ANNs can be resolved using deep neural networks (DNNs) [107], which are multi-processing layered structures with outstanding learning capability [108]. DNNs have become the most effective DL models and have shown stronger learning ability compared with traditional ANN in both spatial and sequential learning classifications/prediction problems [109], [110].

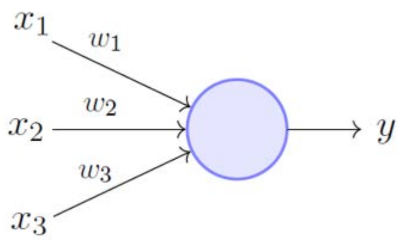


Figure 2.5: Single-layer perceptron network for binary classification. The output y is activated to 1 if the sum of the products of the weights (w_1, w_2, w_3) and the inputs (x_1, x_2, x_3) are above the threshold, otherwise deactivated.

DNNs can be classified as feedforward and recurrent neural networks [111]. In feedforward neural networks, the inputs are directly passed to the outputs through a series of connection weights. The simplest traditional feedforward neural network is a single-layer perceptron (SLP) network. It implements a learning rule for binary classification using a set of inputs and an artificial neuron, illustrated in Fig. 2.5. It consists of a single layer of output nodes. The inputs are fed directly to the outputs via a series of weights. The sum of the products of the weights and the inputs is calculated in each node, and if the value is above the threshold, the output is activated as one; otherwise, it is deactivated to zero.

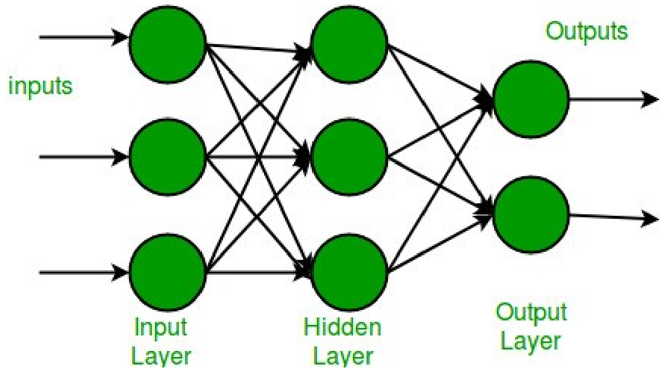


Figure 2.6: Multi-layer perceptron network [201]. In this network, perceptrons are organized into an input layer, a hidden layer, and an output layer. Depending on the type of activation function, it can perform as a classifier or regressor.

The multi-layer perceptron (MLP) network is a deeper version of the SLP network [201]. Fig. 2.6 illustrates an MLP network. It contains many perceptrons that are organized into layers. An MLP should have at least three layers of perceptrons: an input layer, a hidden layer, and an output layer. An MLP with more than two hidden layers is classified as a DNN. Also, MLP is a fully connected network, i.e., all the neurons in one layer are connected to the neurons in the next layer. While the SLP can only perform binary classification like heavy and light rainfall (as rainfall is non-linear and not linearly separable [195]), an MLP is free to either perform classification or regression, depending upon its activation function.

Convolutional neural network (CNN) are regularized versions of MLP [112], [113], [114]. Typical ways of regularization are penalizing parameters during training (such as weight decay) or trimming connectivity (such as dropout). CNNs regularize by the use of a convolution layer (not present in MLP) before the fully connected layer. This layer performs a mathematical operation called convolution in place of general weight matrix multiplication in MLP. Here it takes advantage of the hierarchical pattern in data and assembles patterns of increasing complexity using smaller and simpler patterns embossed in their filters (convolution kernels).

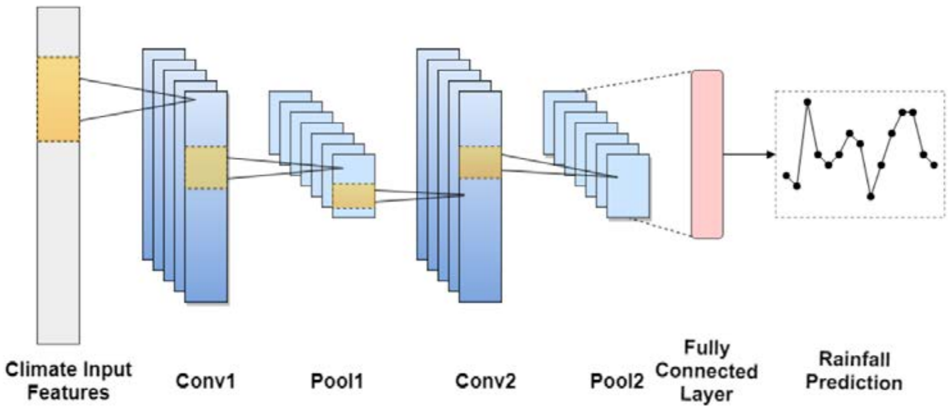


Figure 2.7: Convolutional neural network for rainfall prediction [114]. The Conv layers perform convolution operations on inputs, Pool layers reduce dimensionality through max pooling, and the fully connected layer implements the hyperbolic tangent activation function for rainfall prediction.

The literature illustrates that CNN exhibited better performance compared to the widely used feed-forward neural network, the Multilayer perceptron (MLP), and BOM's CM (ACCESS-S1) to predict rainfall [112], [113], [114]. Fig. 2.7 illustrates the CNN model for rainfall prediction. In this model, the convolutional layers convolve the input and pass its result to the pooling layer. Pooling layers reduce the dimensions of data by combining the outputs of neuron clusters at one layer into a single neuron in the next layer. This pooling layer uses the maximum value of each local cluster (max pooling). Finally, the fully connected layer connects every neuron in the previous layer to every neuron in the output layer. This final layer is the same as the traditional MLP. In this work, the CNN model analyzed the changes in various climate parameters and objectively extracted features to predict rainfall [18]. However, rainfall data is a time series sequence, and a dynamic neural network with memory units (recurrent neural network models) exhibited improvement in predicting rainfall compared with feedforward models [33], [199]. This indicates that sequential models such as recurrent neural networks might be better for predicting the best rainfall forecast as they are time series sequences with similar features as actual rainfall.

Fig. 2.8 illustrates a recurrent neural network (RNN). Unlike feedforward networks,

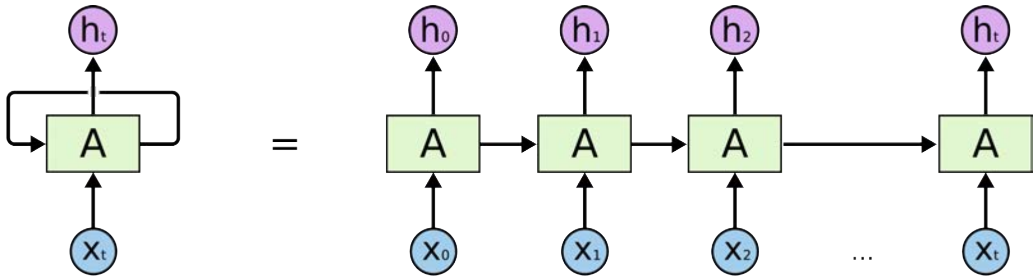


Figure 2.8: Structure of recurrent neural network [150]. The inputs at previous time instants (X_t) are provided as feedback for subsequent sequence predictions.

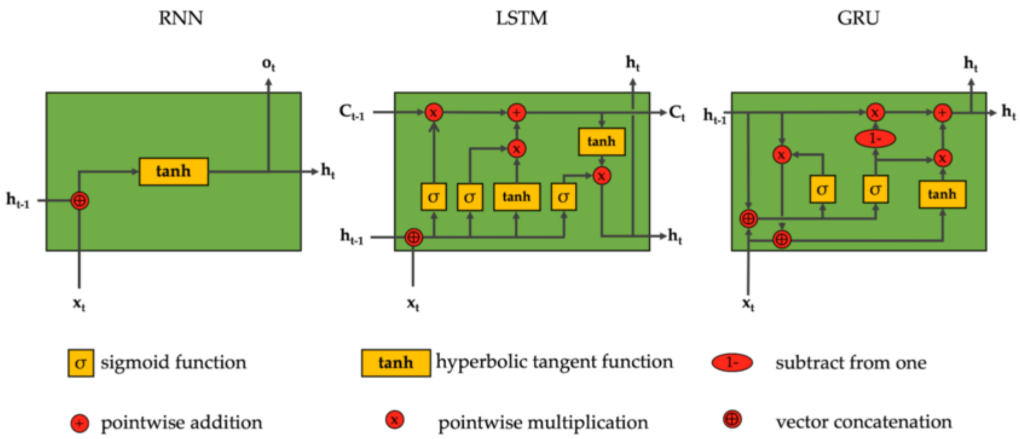


Figure 2.9: Comparison between an RNN, LSTM, and GRU cell [115]. The LSTM introduced gate control units (three gates: σ) to overcome the gradient disappearance/explosion problems in the RNN structure. GRU cell reduced the number of gate control units from 3 to 2 in the LSTM to reduce training time.

RNNs have a feedback loop [150]. Due to this feature, they can correlate contextual information effectively, which is suitable for modeling sequence data. They also fine-tuned the weights of the network based on the error obtained in the previous training iteration. However, the RNN structure is prone to gradient disappearance or gradient explosion problems when the sequence distance is long.

To overcome the problems caused by RNN in long-distance memory, the long short term memory (LSTM) neural network was proposed [115]. As can be observed from Fig. 2.9 LSTM introduces a gate mechanism to the RNN structure for updating information. While the LSTM gate mechanism helped address the problems in RNN, it made the overall structure complex, resulting in high training times. The gated recurrent unit (GRU), reduced the number of gate control units from 3 to 2 in the original LSTM [115]. Fig. 2.9 illustrates that GRU has less number of gates compared to LSTM. The reduced gate makes GRU faster compared to LSTM, however, it is less accurate on a larger dataset (such as meteorological data) compared to LSTM [11], [115]. Therefore, for more accurate rainfall predictions, the use of LSTM might be more appropriate.

More recent works have identified hybrid-sequential DNNs as candidates that can better decode many complex rainfall characteristics compared to recurrent-DNNs [11], [119]. Hybrid structures can incorporate diverse analyses in a single model for advanced feature extraction for classifications/predictions due to the concatenation of multiple model features. As climate forecasts are quite complex and patterns are difficult to find, more advanced analysis using hybrid-recurrent DNNs might be required to find the optimum forecast from CM. However, a major drawback in these models is that they focus on location-specific predictions, rather than predictions of entire precipitation fields, limiting their operational utility and usefulness across different spatial and temporal scales [11], [119]. Furthermore, the forecasts issued by these models show higher uncertainty with increasing lead times, which could be due to not including small-scale weather patterns, critical for improving forecast value [120].

Overall, DL methods look promising for the rainfall domain and hybrid-sequential DNN models seems the best fit to achieve this goal.

2.3 Surface Soil Moisture

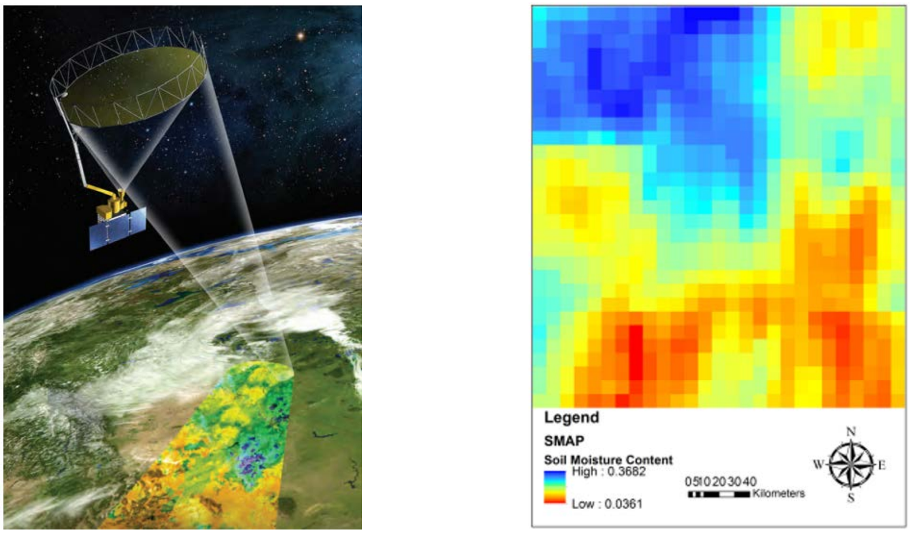
When overwatering occurs, part of the rainfall/irrigation water percolates below the root zone of the plants, and some flows away over the soil surface as run-off [59], [60]. Only the remaining part stored in the soil as moisture can be used by the plants [123]. Hence, the estimation of the water stored in soil is critical for improving agricultural irrigation efficiency. However, the direct measurement of spatially distributed SM using in-situ sensors is expensive. Satellites can measure the surface soil moisture (SSM) distribution information, but existing satellite missions have poor resolution [25], [65], [67], making them not suitable for irrigation applications. Hence, downscaling of satellite SSM is essential for use in irrigation applications.

2.3.1 Satellite SSM

In this section, the literature on existing satellite SM missions and downscaling methods for the current best-resolution satellite are thoroughly examined. Many global SSM products, based on remote sensing earth observations have been developed and have achieved much technical progress over the years [130]. Most of these SSM products are derived from microwave observations due to the direct connection between electromagnetic radiation and SSM. The popular shorter wavelength SM missions are Advanced Microwave Scanning Radiometer for the Earth-observing system (AMSR-E) [131], Advanced Scatterometer (ASCAT) [132], and Advanced Microwave Scanning Radiometer-2 (AMSR2) [133]. The AMSR-E sensor is a radiometer operating onboard the Aqua satellite since 2002 and was the first satellite sensor to incorporate SM as a standard product [131]. The ASCAT sensor is a C-band scatterometer operating onboard the Metop satellite since 2006 and uses the Vienna University of Technology (TUWIEN) change detection algorithm for SSM retrieval by scatterometer data [132]. The AMSR2 is onboard the GCOM-W1 satellite [133] and is a passive sensor that measures SM through C-band. Global SM products from these missions are of low spatial resolution (approximately 25–40 km), which greatly limits their use in regional agriculture.

In 2009, the European Space Agency (ESA) launched the first dedicated SM mission, the Soil Moisture and Ocean Salinity (SMOS) Earth Explorer [134]. Following suit, the National Aeronautics and Space Administration (NASA) launched the Soil Moisture Active Passive (SMAP) mission in 2015 [163]. Both SMOS and SMAP operate in L-band (higher wavelength) and can provide better-resolution SSM products than shorter wavelength missions. Currently, SMAP provides the best resolution daily SSM product (9 km) [135]. However, this resolution is not sufficient for irrigation applications [35], [64]. Existing research indicates that at least a resolution of 1 km is required for agriculture applications [35], [64]. Since SMAP provides the best resolution products, further downscaling of SMAP will result in better resolution products than the other satellite missions. SMAP and existing downscaling approaches for SMAP are reviewed further.

2.3.2 SMAP and SSM Downscaling



(a) Satellite (b) 9-km soil moisture

Figure 2.10: Illustration of Soil Moisture Active Passive (SMAP) mission's (a) satellite, and (b) Backus–Gilbert (BG) algorithm-based 9-km (best available resolution) soil moisture product.

Fig. 2.10 (a) illustrates the SMAP satellite. It is one of the latest L-band missions, launched on January 31, 2015, and aims to generate a global SSM map at multiple resolutions with an L-band radar (active) and radiometer (passive) [163]. It is operated in a near-polar orbit with the ascending and descending overpass times of 6:00 PM and 6:00 AM at local time. The satellite was designed to provide SM products with three main spatial resolutions: 36 km passive [164], 9 km active-passive [165], and 3 km active (radar). However, a radar failure on July 7, 2015, limited the generation of active and active-passive SM data. After the SMAP radar failure, SMAP adopted Sentinel-1 (the first Copernicus program satellite) instead of using radar data [26]. Sentinel-1 was used as a replacement due to its ability to provide both the co-polarization and cross-polarization measurements needed for producing active-passive SMAP products [26], [165]. Currently, SMAP can provide daily 9 km active-passive products using the Sentinel-1 data in the Backus-Gilbert (BG) algorithm [166], Fig. 2.10 (b) illustrates this SMAP product. However for regional agricultural applications 9-km resolution is not sufficient [35], [64]. Hence downscaling of the 9 km SMAP product is required.

Sentinel-1 has a narrow swath resulting in a temporal resolution of 6-12 days [36]. This drawback can be addressed by using the higher temporal resolution Moderate

Resolution Imaging Spectroradiometer (MODIS) data instead of Sentinel-1 data in downscaling algorithms [37]. MODIS is a key instrument aboard the Terra and Aqua satellites. Potential auxiliary SM information is derived by exploiting the absorptive and reflective characteristics of the canopy and bare soil in the visible to shortwave spectral bands. Based on the negative correlations between water and the red to shortwave infrared (SWIR) band reflectance, numerous soil moisture indices (SMIs) were proposed for SM mapping and drought status monitoring [40].

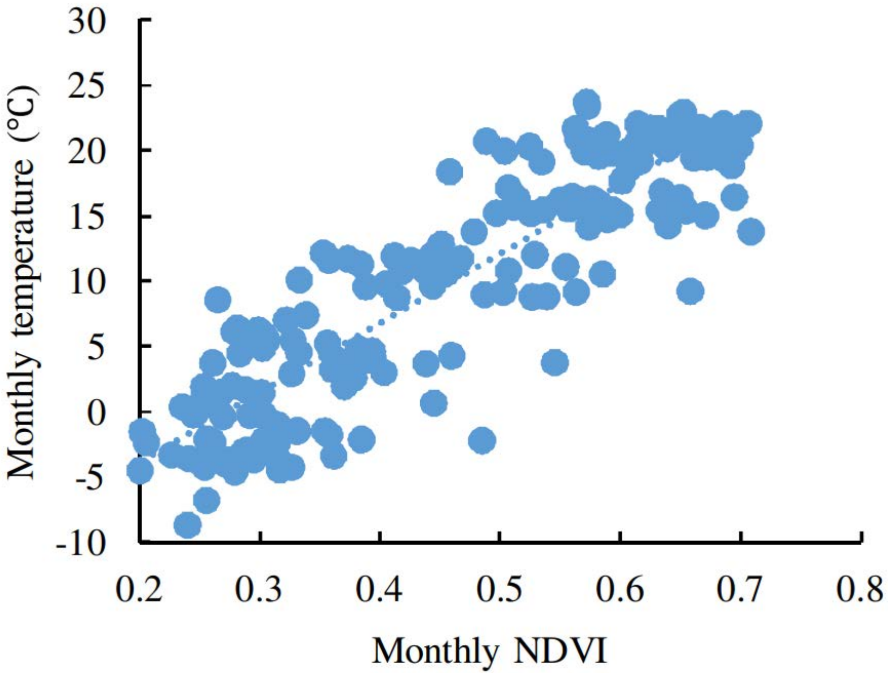


Figure 2.11: Scatter plot illustrating that NDVI has a positive correlation with soil moisture [140].

Normalized Difference Vegetation Index (NDVI) is the most typically used SMI [140]. It describes the difference between visible and near-infrared reflectance of vegetation cover. Existing literature shows that NDVI has a positive correlation with SM (Fig. 2.11). Another commonly used vegetation-based SMI is the Enhanced Vegetation Index (EVI) [141]. It is a vegetation index designed to enhance the vegetation signal with improved sensitivity in high biomass regions. Existing studies indicate that EVI has a positive correlation with SM (Fig. 2.12). Similarly, SWIR-NIR-based indices such as Temperature-Vegetation Dryness Index (TVDI) shows a negative correlation with SM (Fig. 2.13) [142]. Many such indices are available in the literature. Table 2.2 provides a list of existing SMIs which can be used for SMAP downscaling [40].

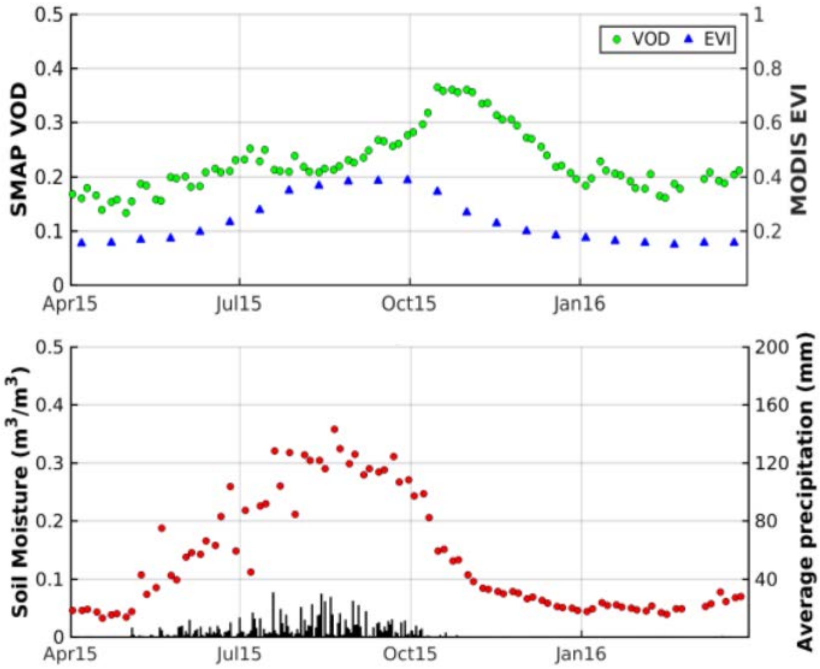


Figure 2.12: Time series comparison of SMAP VOD-vegetation optical depth (green dots), MODIS EVI (blue triangles), SMAP soil moisture (red dots), and GPM-global precipitation measurement (black vertical bars) [141] . The EVI is in phase with precipitation and soil moisture, and the VOD the peak comes after the end of the rainy season.

Next, existing SMAP downscaling models are examined. They can be broadly classified as spatial and sequential downscaling models [126], [143]. Spatial downscaling models only consider spectra-spatial relationships to produce SM maps [37], [126], [143]. Hence the daily downscaled SMAP SSM maps produced by these models resulted in empty spaces (Fig. 2.14) when MODIS data got corrupted for the selected date.

In the case of sequential downscaling models, they leverage spectral context from adjacent SMAP and MODIS slices either over time (within a feature spectrum) or space (between feature spectra) [144], [145], [146]. But if we can leverage inter-slice spectral sequence context simultaneously over time and space, it may further improve downscaling as the SSM changes are both time and space-dependent [38], [39], [43], [64]. However, existing downscaling models do not incorporate such spectral analyses (both spatial and sequential) simultaneously due to complexity and structural limitations [146]- [166]. Therefore, if a structure that can incorporate all these analyses (spatial/sequential) without significantly increasing the model complexity can be designed, it will substantially reduce the gaps and increase the accuracy of

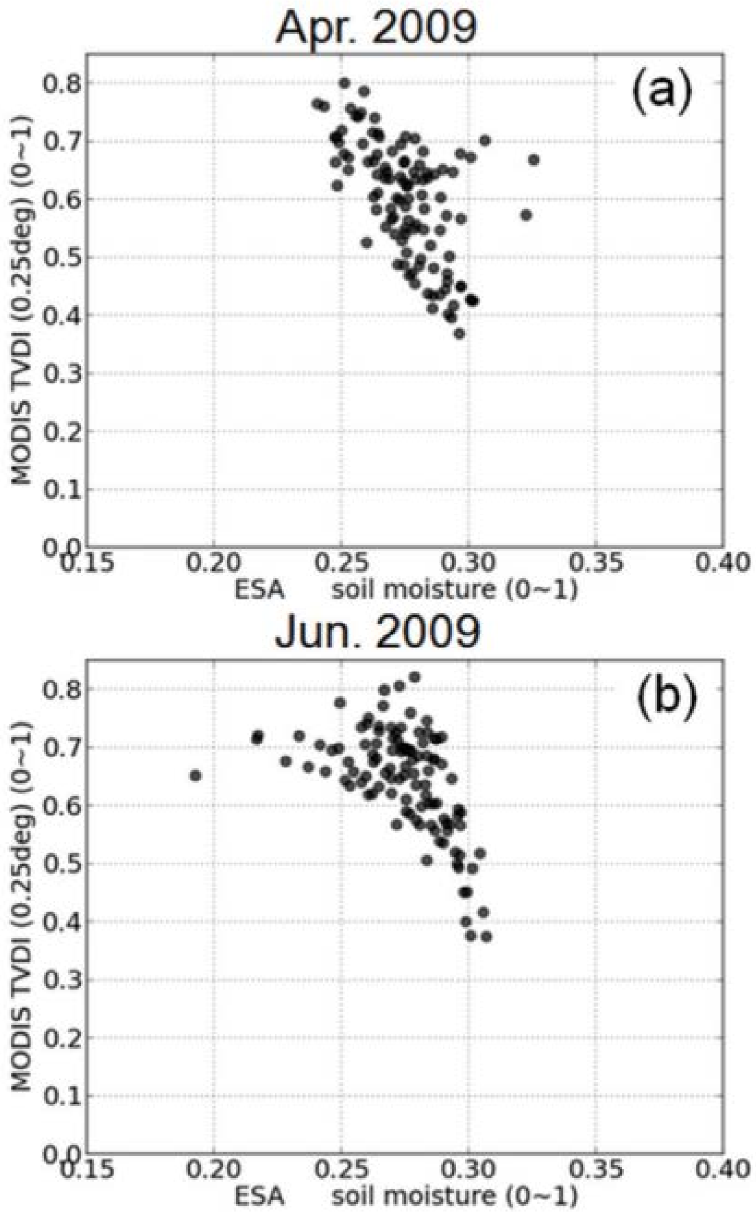


Figure 2.13: Temperature-Vegetation Dryness Index (TVDI) and soil moisture scatterplots show distinguishable negative correlations during (a) April 2009 and (b) June 2009 [142].

pixels in downscaled SMAP product.

2.4 Root Zone Soil Moisture

The soil can be classified into surface and root zone layers (Fig. 2.15). However, based on the plant type the root depths differs (Fig. 2.16) [75]. The SSM only provides information on the water content in the top few centimeters of the soil [124], and this information is not sufficient for irrigation decisions for deep-rooted plants. Hence RZSM estimation is required. RZSM represents the percentage of available

Table 2.2: Candidate MODIS Soil Moisture Indices (SMIs) for SMAP Downscaling [177]

SMIs	Formula	Symbol
NDVI	$(R_{b2}-R_{b1})/(R_{b2}+R_{b1})$	R_{b1}: Red band; R_{b2}: NIR1 band.
EVI	$2.5(R_{b2}-R_{b1})/(R_{b2}-6R_{b1}-7.5R_{b3}+1)$	R_{b3}: Blue band.
ANIR	α_{b2}	α_{b2}: Angle formed at the vertex of NIR by R–NIR–SWIR1 reflectances.
NDWI	$(R_{b2}-R_{b5})/(R_{b2}+R_{b5})$	R_{b5}: NIR2 band.
DDI	$(\sqrt{(R_{b1}^2)+(R_{b2}^2)}$)/(1+NDVI)	-
GVMI	$((R_{b2}+0.1)-(R_{b5}+0.02))/((R_{b2}+0.1)-(R_{b5}+0.02))$	-
PDI	$((R_{b1}+MR_{b2})/\sqrt{M^2+1}$	M: Slope of the soil line.
MPDI	$((R_{b1}+MR_{b2})-f_v(R_{v,b1}+MR_{v,b2})/(1-f_v)\sqrt{M^2+1}$	f_v: Vegetation fraction; $R_{v,b1}$ and $R_{v,b2}$: Vegetation reflectance in NIR band and Red band.
MPDI1	$\sqrt{PDI^2+PVI^2}$	PVI: Perpendicular vegetation index.
MSI	R_{b6}/R_{b2}	R_{b6}: SWIR1 band.
MSPSI	$((R_{b6}-R_{b1})+M(R_{b6}+R_{b1})/\sqrt{M^2+1}$	-
NDII6	$(R_{b2}-R_{b6})/(R_{b2}+R_{b6})$	-
NDII7	$(R_{b2}-R_{b7})/(R_{b2}+R_{b7})$	R_{b7}: SWIR2 band.
NDTI	$(R_{b6}-R_{b7})/(R_{b6}+R_{b7})$	-
NMDI	$(R_{b2}-(R_{b6}-R_{b7}))/(R_{b2}+(R_{b6}+R_{b7}))$	-
SANI	$\beta_{b6}(R_{b6}-R_{b2})/(R_{b6}+R_{b2})$	β_{b6}: Angle formed at vertex SWIR1 by NIR–SWIR1–SWIR2 reflectances.
SASI	$\beta_{b6}(R_{b2}-R_{b6})$	-
SPSI	$(R_{b6}+MR_{b2})/(\sqrt{M^2+1})$	-
SRWI	R_{b2}/R_{b5}	-
VSDI	$1-(R_{b6}+R_{b1}-2R_{b3})$	-
VSWI	$NDVI/T_c$	T_c: Surface radiometric temperature.
MSAVI	$(2R_{b2}+1-\sqrt{(2R_{b2}+1)^2-8(R_{b2}-R_{b1})})/2$	-
SIMI	$0.707\sqrt{(R_{b6})^2+(R_{b7})^2}$	-
SW	$1-(1-\beta EVI)\triangle T/((1-EVI)\triangle T_{max}+EVI\triangle T_e$	$\beta: Waterstressparameter; \triangle T, \triangle T_{max}, \triangle T_e$: All MODIS surface temperature data.
DDI	$\sqrt{(R_{b1})^2+(R_{b2})^2}/(NDVI+1)$	-
TWI	$log_{10}(\alpha/tan_{\beta})$	-
TVDI	$(T_s-T-smin)/(a+bNDVI-T_{smin})$	T_s: Surface temperature; a & b are parameters of the dry edge modelled as a linear fit to data.

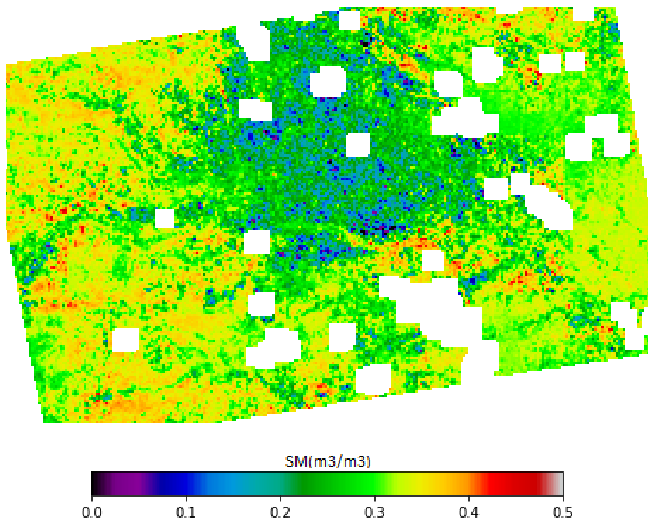


Figure 2.14: Information loss (gaps) in downscaled SMAP SM at Little Washita, Oklahoma due to corrupted MODIS data on day-of-the-year 193, 2016 [2].

water content in the top 1 m of the soil profile [159], [236]. From existing literature, RZSM estimation methods can be broadly classified as traditional and data-based methods [67]. In the further sub-sections, these methods are critically analyzed.

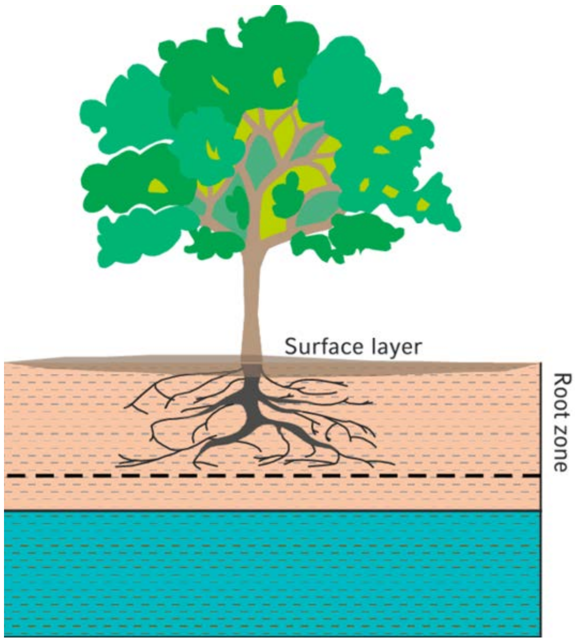


Figure 2.15: The surface and root zone soil layers.

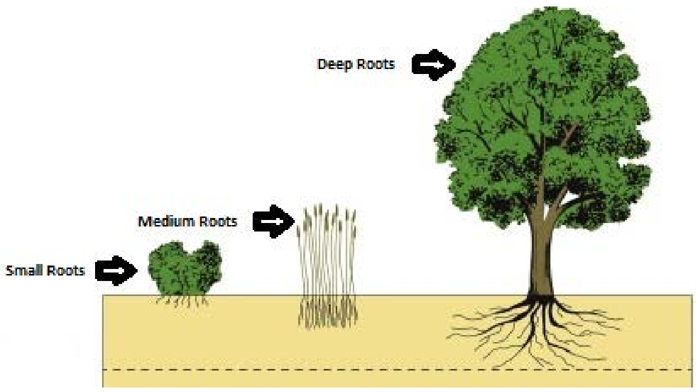


Figure 2.16: Plants with different root depths [75]. The small root plant extends only to the surface soil layer. The medium and deep root plants grow to the root zone soil layer.

2.4.1 Traditional Methods

In-situ devices offer direct measurement of SM in the root zone [160]. However, installing sensors to measure RZSM at a large spatial scale would be time-consuming and expensive [45]. Also, installing RZSM sensors at the subsurface is likely to disturb the soil properties [161], [162]. Moreover, in-situ SM data are generally representative of the SM condition in the measured sites or homogeneous to the surrounding environment, their validation is limited to available sites and may suffer from the representativeness error [168].

Model-based SM is usually representative of hundreds of square kilometers and is not limited by representativeness error [25], [169]. Also, RZSM estimated through land surface models (LSMs) or global hydrological models (GHMs) on a regional or global scale require site-specific initial conditions (e.g. land surface temperature) to be assimilated into these models to obtain synthesized estimates of regional RZSM [170], [171]. As SSM and RZSM are correlated to each other, SSM observations from satellite sensors are assimilated into these models to improve the RZSM estimation [173], [174]. Also, it is becoming relatively common practice to extract RZSM from SSM since SSM is relatively easier to obtain than RZSM [172].

Satellite measurements are not constrained by data limitations as in model-based or in-situ-based approaches [25]. However, current SM missions only offer the best resolution of 9 km daily measurements [163]. Also, the present sensors onboard these satellites can only detect 0-5 cm beneath the soil surface [25]. So one pixel of the satellite SM represents the 0-5 cm depth measurements of a 9 km $\times$ 9 km area. This 9 km $\times$ 9 km is too coarse a resolution to represent a farm. Before further deeper SM measurements can be obtained, the pixel resolution needs to be improved (downscaling) for farm RZSM derivation. Once SSM of the required resolution is obtained, RZSM needs to be derived.

2.4.2 Data-driven Methods

Data-driven methods to estimate RZSM include time series analysis (TSA) and machine learning (ML) techniques [67], [186], [187]. These methods aim to extract knowledge by first evaluating data patterns and then further take actions that are dictated by the data. In the context of RZSM estimation, data-driven methods implicitly incorporate and evaluate all the interacting features (precipitation, temperature, soil, and plant characteristics [46], [261]) that control the RZSM state. TSA methods, such as the application of an exponential filter [175], [176], a cumulative distribution function [72], [177], or transfer-functions [70] primarily utilize SSM data to derive a functional relation with RZSM. However, calibration of functional parameters is necessary when applied to a different study area (each time) in TSA methods to obtain high accuracy.

ML algorithms build mathematical models based on training sets and covariates to extract information from data [178]. Furthermore, they are tuned to handle diverse and large volumes of data sets, which may be relevant for large-scale studies or operational water management applications like PI. In many hydrological research domains, ML techniques have showcased high performance for prediction [179], [180], [181], sensitivity or optimization of model parameters [182], [183], and uncertainty estimation [184], [185]. However, the application of ML in soil hydrology has only started to gain attention in the past 20 years. For instance, only simple ML techniques have been so far applied to estimate model-derived SM, like ANNs [186] or satellite-derived SSM using support vector machines [187]. These works indicate the ability of ML models to both upscale or downscale SM obtained from satellite data [188], [189]. Research works that analysed ML models ability for forecasting SM using values at discrete soil depths [190], [191] at regional scales indicate their high performance. Interestingly, SSM has also been estimated from in-situ measurements of RZSM using ML for satellite SSM validation [192]. This work illustrates both the possibility of RZSM derivability from SSM as well as the capability of ML models for RZSM computations.

A comparison study showed that ML may provide a useful alternative to process-based models using limited input data for SM estimation [193]. DL is a sub-field of ML, that has shown exceptional performance in many hydrological applications and has exhibited higher accuracy compared to traditional ML models [43], [43], [61]. However, the application of DL in the RZSM field is still in the early stages compared to the SSM field [61]. Existing RZSM estimation research employs mainly traditional ANN models for RZSM estimation [19], [128], [244]. ANNs might not be sufficient to estimate the distributed spatial RZSM as they tend to fall into the local minimum [105] and over-fitting [106] problems. DNNs have a stronger learning ability compared to conventional DL models such as ANNs and can overcome the limitations in traditional ANNs [56], [245], [246]. Furthermore, DNNs can learn deep links among data, which has proved effective in dealing with SSM classification and prediction problems [248], [249]. Since SSM and RZSM are correlated and literature has illustrated the applicability of SSM estimation models for RZSM, DNN might be better suited for RZSM prediction compared to traditional ANNs. Overall DNN looks like a promising avenue for RZSM estimation.

2.5 Conclusion

This literature review first provides a systematic analysis of existing IDSS for PI, identifying their limitations in producing optimum water scheduling due to inaccurate rainfall and SM input information. Consequently, next, existing methods for the prediction/estimation of rainfall and SM were investigated.

Existing literature illustrates that current official rainfall forecasting methods are not accurate enough for PI application. Therefore, a significant research gap remains in developing a robust model for refining official rain forecasts for more accurate predictions for PI. Analysis of existing related works revealed DL methods to be suitable for such data refinement applications when big data is available. However, existing DL methods need to be made adaptable for rainfall data as they are highly stochastic.

From rainfall literature, next, the focus turns to methods for obtaining information on SM to improve PI decisions. For distributed SM information we turn to satellites. However, current SM missions can only provide the daily SM at a 9 km resolution. This resolution is not sufficient for PI applications. Hence, SM satellite data downscaling is required. Existing SM downscaling works suffer from spatially distributed spectral information loss, a gap that needs to be addressed for improved IDSS decisions.

Also, existing satellites can only detect SSM (SM at 0-5 cm depth) and do not provide RZSM (SM at $>$ 5 cm depth) details required for PI. Direct measurement using large-scale RZSM sensor deployment is not feasible. Hence, RZSM needs to be derived indirectly for distributed depth measurements. Sequential DNN looks like a promising avenue to investigate for this information extraction.

In conclusion, this thesis will address these four significant research gaps, providing a broad range of models designed for improving decisions for PI through better rain forecasting and spatial/depth estimation of SM.

Chapter 3

Consensus Forecast of Rainfall using Hybrid Climate Learning Model (HCLM)

This chapter contains materials published in the IEEE Internet of Things Journal:

[1] **Neethu Madhukumar**, Eric Wang, Yi-Fan Zhang, and Wei Xiang, "Consensus Forecast of Rainfall Using Hybrid Climate Learning Model," *IEEE Internet of Things Journal*, vol. 8, pp. 7270-7278, May 2021 (Impact Factor: 10.238, h-index: 144, Journal rank (Computing systems): 1).

3.1 Introduction

A reliable model for predicting weather events can provide useful information for irrigation planning and management [194] [121]. A good weather prediction model should seamlessly link together diverse atmospheric changes to provide improved weather predictions. Most countries around the world use climate models (CMs), such as the general circulation model or, the earth system model, to officially forecast weather patterns [22], [23], [24]. These models are built upon differential equations based on the laws of physics, fluid motion, and chemistry, which are applied to the input data. On that account, forecasts can be considered as the numerical relationship between various atmospheric and environmental changes that result in the occurrence of climate events.

The data inputted to the CM are obtained from the earth-observing satellites, the weather radars, and various climate sensors [195] [196]. These data then get

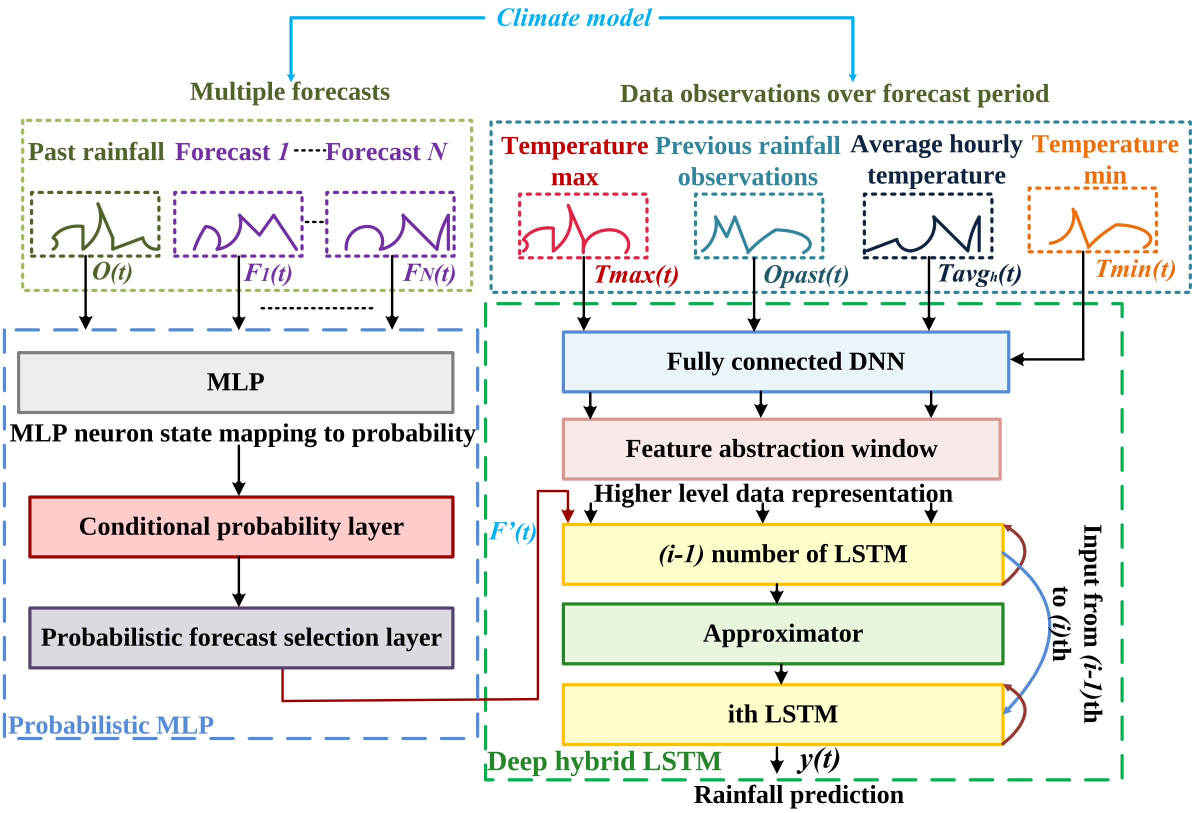


Figure 3.1: Illustration of the HCLM structure. Multiple forecasts from the CM are input to the probabilistic MLP (PMLP) network, which conducts a conditional probability evaluation of the forecasts for finding the best of the N forecasts. The output of PMLP is input to HD-LSTM along with input data observations over the forecast period. HD-LSTM analyse the PMLP forecast and the input data observed over the selected forecast period to predict rainfall.

processed using various scientific tools for climate impact, adaptation analysis, and weather forecasting thereby providing a higher representation of the input data. Over the years weather forecasting by the CM has greatly improved with the arrival of new technologies, integration of better theory, and increased computational power. Nevertheless, there is still a gap in finding the best forecast from multiple forecasts produced by the CM. The discovery of hidden weather patterns and relationships from the CM could significantly improve its performance. Deep learning techniques can capture underlying subtle functional relationships, unearth intricate structures from big climate data, and find the hidden patterns from complex CM forecast data. These features of climate and deep learning models make them ideal to form a teacher-student network for knowledge distillation from the CM to the deep learning model [197].

Neural networks have been used for predicting various weather attributes [122], [34]. Comparative analysis has indicated a higher performance accuracy for the neural network models than prominent data mining techniques for rainfall predictions [195]. A deep neural network (DNN) is a multi-layer neural network with outstanding learning ability [11]. The success of deep learning models is due to its deep architectures. DNN-based approaches have become the most effective techniques for rainfall prediction due to its ability to progressively build higher-level representation from the weather data.

The DNN-based rainfall prediction approaches that utilise feedforward neural networks provide only a static network, that lack memory capacity [196] [198]. They do not consider temporal relationships. Comparative analysis has shown that a dynamic neural network with memory units exhibited improvement in predicting the rainfall than static feedforward neural networks [199] [33]. So from all the above works, it can be concluded that both the CM and the DNN-based approaches for rainfall predictions have been successful in forecasting certain rainfall features but fails to completely emulate the actual rainfall pattern. Henceforth, a new prediction method that utilises both the CM and the DNN for the more accurate prognosis of rainfall events is proposed in this chapter.

The discovery of the best forecast pattern is difficult and consequently has not been investigated to improve the performance of climate models. Learning the best forecast pattern from the CM, and determining its relationship with other weather changes will improve the rainfall event prediction. I propose a hybrid climate learning model (HCLM) that learns from the CM using deep learning. The HCLM shown in Fig. 3.1 is designed using a probabilistic multi-layer perceptron (PMLP) and a hybrid deep long short-term memory (HD-LSTM) network for reprocessing the multiple forecasts produced by the CM for improved next-day rainfall prediction.

PMLP is used to rank and select the best forecast, and HD-LSTM finds the correlation between the PMLP-selected forecast and the corresponding data observations for predicting the next-day rainfall. The PMLP network evaluates the probability of each forecast and selects one forecast with the highest probability. The output from PMLP along with weather changes observed over the forecast period is analysed by the HD-LSTM to produce the next-day rainfall forecast. The rainfall forecasts and the observation data from ten sites from different weather zones across Australia are selected to evaluate the performance of the HCLM.

To summarise in this chapter a new HD-LSTM network is proposed for learning the relationship between the best forecasts from the CM and the corresponding data observations for producing the next-day rainfall forecast. The PMLP network analyses, compares and ranks multiple forecasts for different rainfall patterns and the hybrid learning model utilizes both the CM and the DNN to refine multiple forecasts from the CM.

The remainder of this chapter is organised as follows. Section 3.2 describes the proposed HCLM model. Section 3.3 gives the experimental results. Concluding remarks are drawn in Section 3.4.

3.2 Hybrid Climate Learning Model

In HCLM the PMLP learns the pattern of the best forecast from the CM over multiple data sub-samples and selects the best forecast from the evaluation data. The PMLP output is fed to HD-LSTM as part of knowledge distillation. The use of the PMLP output label for data selection reduces the LSTM search space for sequential data analysis.

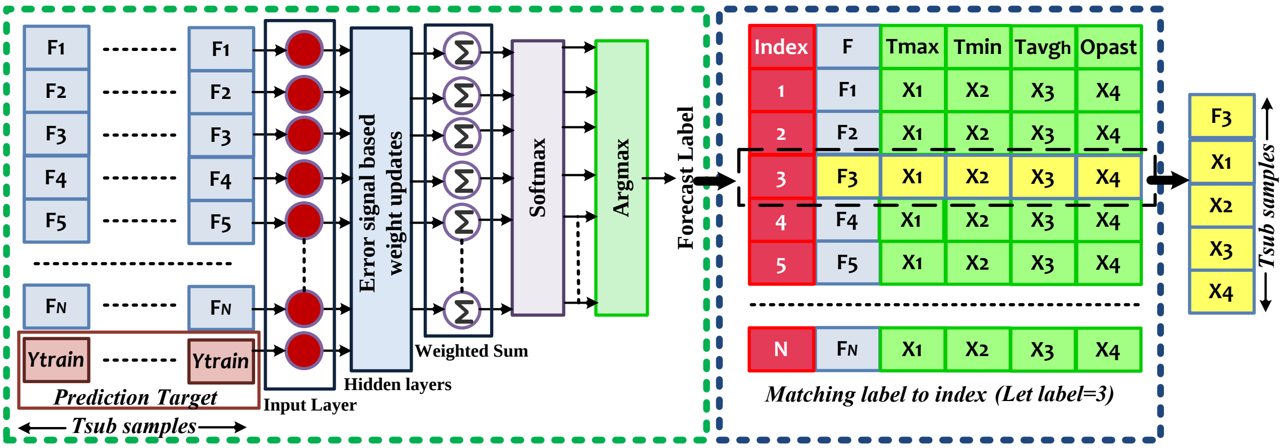


Figure 3.2: Illustration of the PMLP structure. All forecasts and actual rainfall (red box) pass from the input layer to the hidden layer, which generates error signals for a given set of input forecasts. Errors are used by the nodes to update the values for each connection weight. A weighted sum of the inputs goes to the Softmax layer, which provides a probabilistic estimate to the Argmax layer. The Argmax layer outputs the label of the best forecast, which is used to select the corresponding observation data for the best forecast.

3.2.1 PMLP

In the PMLP structure illustrated in Fig. 3.2, N input forecasts are passed to the input layer of the PMLP. The N forecasts produced for the rainfall event by the CM are denoted by $F_n(t)$ for $n \in \{1,2,...,N\}$ at $t \in \{1,2,...,T\}$ and are passed on to the input layer of the PMLP for further refinement. $O_{past}(t)$ is the past rainfall observations. The pattern of each forecast changes over different time blocks, so the PMLP learns the relationship between $F_n(t)$ and $O_{past}(t)$ for multiple T_{sub} data blocks. Table 3.1 lists the monthly Root Mean Square Error (RMSE) of the 7-day forecasts obtained from the Australian Bureau of Meteorology (BOM) for the Burdekin region in Queensland [200]. The dataset was prepared by the BOM's evidence-targeted automation team for the science to services program. In Table 3.1, Forecast 1 is the first forecast produced 7 days before the event, while Forecast 7 is the last forecast produced one day before the event. Forecast 7 is thought to be the best forecast as it has the shortest lead time to the actual rainfall event [24]. However, my analysis of the monthly forecast data (from 2015-2017 using [200]) shown in Table 3.1 indicated that this is not always true. For example, Forecast 7 predicted 6.4 mm of rain, and Forecast 1 predicted 0.9 mm of rain for 18-10-2017, but the actual rain was 0.6 mm. So to forecast rain on 18-10-2017, information from Forecast 1 is more useful.

Table 3.1 illustrates that for each period there exists a forecast which is more reliable than the other forecasts. So the rest of the forecasts can be dismissed for that period. Hence it is proposed to rank each forecast based on the patterns of each forecast and rainfall observations for multiple T_{sub} data blocks for selecting the best forecast. A probabilistic selection of any feature over different data blocks performs better than many benchmark methods [201] [202]. Therefore such a selection is adopted. For the data set used in this chapter, $T_{sub} = 21$ days is chosen using the random search algorithm [204].

The PMLP analyses all $F_n(t)$ for the rainfall patterns over multiple T_{sub} data blocks. After training, the PMLP for similar rainfall patterns and similar forecast inter-relationship selects one from $F_n(t)$ as its output, which it predicts to be the best forecast. Error signals for each hidden layer node are determined for the given inputs. The errors are used by the nodes to update the values for each connection weight.

For the PMLP, let W^m_{lv} denotes the values of the weights from the v^{th} neuron of layer $(m-1)$ to l^{th} neuron of layer m, and W be the collection of W^m_{lv}, $\forall\ l,v,m.$ So for the given set of inputs, the outputs can be represented as $S_n(F_n(t);W)$, a function of inputs and weights. The PMLP softmax layer estimates the conditional probability of all forecasts. Using [201] this can be expressed as

$$\begin{aligned} P_j(t) &= P(F_j(t)|F_n(t);W) \ \text{ for } \ j \in \{1,2,...,N\}, \forall \ j \neq n \\ &= \frac{e^{S_j(F_n(t);W)}}{e^{S_1(F_n(t);W)} + e^{S_2(F_n(t);W)} + ... + e^{S_N(F_n(t);W)}}, \end{aligned} \tag{3.1}$$

where $P(F_j(t)|F_n(t);W)$ is the conditional probability of forecast j for the given set of inputs $F_n(t)$ and the given set of weights W. The PMLP ranks and selects the best forecast using (1) over T_{sub} blocks. Let $F'(t)$ be the PMLP selected forecast, then

$$F'(t) = F_n(t), \ \text{ if } \ P_n(t) > P_j(t), \tag{3.2}$$

where $j \in \{1,2,...,N\}\ \forall\ j \neq n$. Using (3.2) the best forecast $F'(t)$, is selected from $F_n(t)$. The PMLP output $F'(t)$ is given as input to the HD-LSTM for temporal analysis with the temperature and rainfall observations during the forecast time. For the

Table 3.1: Monthly RMSE of 7 Days Forecasts of BOM Australia in 2017. Forecast i Produced $(8-i)$ Days before the Event.

Month	January	February	March	April	May	June	July	August	September	October	November	December
Forecast 1	10.485	27.955	12.086	13.916	1.575	1.040	0.862	6.334	**0.398**	3.171	2.313	**2.316**
Forecast 2	9.886	26.601	12.149	14.358	1.580	1.662	0.659	2.460	0.560	3.558	3.602	2.448
Forecast 3	8.693	25.544	11.174	**11.252**	1.554	**0.604**	0.509	2.239	1.093	4.039	**2.132**	2.437
Forecast 4	8.752	25.026	9.097	14.913	**1.391**	0.976	0.705	2.400	0.916	3.312	2.493	2.375
Forecast 5	8.369	24.992	8.710	15.864	1.563	0.706	0.723	**0.514**	0.456	**3.128**	2.753	2.454
Forecast 6	7.184	**24.266**	7.921	12.776	1.500	1.071	0.691	2.198	0.772	4.872	2.410	2.431
Forecast 7	**6.887**	25.388	**7.796**	11.368	1.855	0.610	**0.420**	2.236	1.100	4.522	2.500	2.381

given rainfall events, the temporal changes in observation data at the forecast period $t-z$, where z is the forecast lead time, affect its decisions. The PMLP argmax layer provides the label of the best forecast for matching with the index of the data set to select the observation data. The HD-LSTM analyses the relationship of the selected forecast with the observations. The HD-LSTM is detailed below.

3.2.2 HD-LSTM

In the HD-LSTM structure illustrated in Fig. 3.3, the output of PMLP is given to the HD-LSTM for further analysis. The HD-LSTM learns the relationship be-

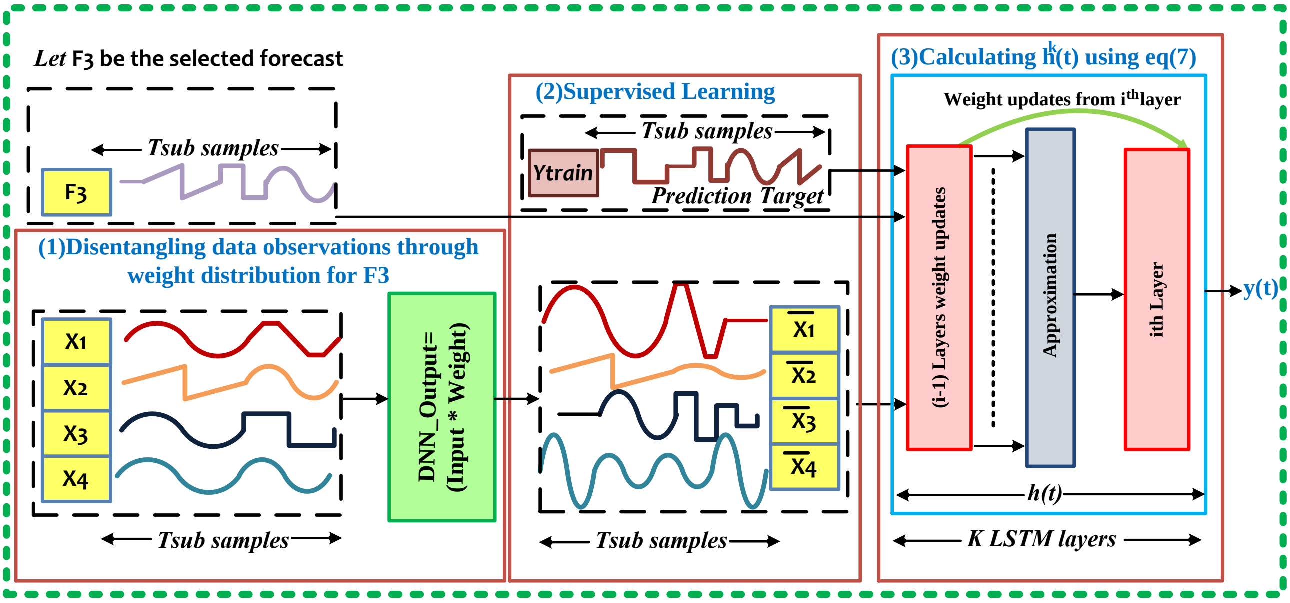


Figure 3.3: Illustration of the HD-LSTM structure. The learning process is implemented in three steps: (1) DNN disentangles data observations through weight distribution; (2) Supervised learning uses PMLP forecast and observation data with real rainfall as the prediction target; (3) Input state transition by the network through $h^k(t)$ from Eq. (7), and the network predicts output using Eq. (8).

tween the PMLP output and the observation data to predict the next-day rainfall. For the CM used, the data observations are the maximum temperature $T_{max}(t)$, the minimum temperature $T_{min}(t)$, the past rainfall $O_{past}(t)$, and the average hourly temperature $T_{avg_h}(t)$. For ease of derivation, $x(t)$ is denoted as the input data for the HD-LSTM and U types of data are denoted by $x_u(t)$. Here for derivation, $[T_{max}(t), T_{min}(t), O_{past}(t), T_{avg_h}(t)] = [x_1(t), x_2(t), x_3(t), x_4(t)]$ is used. For simplifying derivation, first derive for one type of observation, denoted by $x(t)$.

For the selected forecast $F'(t)$ by the PMLP, the HD-LSTM applies a fully-connected DNN at its input to select the data observations over the best forecast period. This data is passed on to the K LSTM layers. The network then analyses the temporal correlation between the forecast and changes in the observation data. To make the comparison further easier for the HD-LSTM hidden layers a summary of the changes in the $(i-1)$ hidden states are provided to the i^{th} hidden state. This implementation aids the network to look back at the summary of historical changes with current changes. The HD-LSTM within the HCLM progressively builds up a higher-level representation of input weather data. From [203], the equation of the

LSTM can be expressed as

$$h(t) = H\Bigg(W_{xh}x(t) + W_{hh}h(t-1) + b_h(t)\Bigg), \tag{3.3}$$

where the input sequence is $x(t)$, the hidden vector is $h(t)$, W_{xh} denotes the input-hidden weight matrix, W_{hh} denotes the hidden to hidden weight matrix, $b_h(t)$ is the hidden bias function, and H is the hidden layer function.

A fully-connected DNN at the input of the HD-LSTM implements a selection window. It selects the data observations x, which are the same as data observations over the best forecast period of z. The DNN matches the selected forecast label to the index of the observation dataset. This assists the network to establish relationships between the temporal changes in the PMLP forecast and observations. This is highly necessary for establishing the puzzling forecast relationships with other climatic changes $x(t)$ over the forecast period z. The DNN disentangling data observations through weight distribution for the selected forecast. So for the proposed network, $W_{xh}x(t)$ in (3.3) becomes $W_{\bar{x}h}\overline{x(t)}$

$$W_{\bar{x}h}\overline{x(t)} = \sum_{z=a}^{b} W_{h,z}x(t-z), \tag{3.4}$$

where $a \leq z \leq b$ defines the size of the selection window that ranges between 0 to 7, t corresponds to the time of the actual event and z corresponds to the forecast time of $F'(t)$. It is the window that the network looks back to learn the temporal correlation with $F'(t)$. $W_{h,z}$ is the input weight from the deep input transcription layer over the range z. The $W_{\bar{x}h}\overline{x(t)}$ and $F'(t)$ are analysed by the HD-LSTM to learn what changes in $x(t)$ over forecast interval led to the best forecast prediction.

For learning the relationship of the current dynamic changes with the past CM data, an approximator between $(i-1)$ hidden states and i^{th} hidden state is implemented in the HD-LSTM. This provides a useful summary of the past to the present hidden states along with current input. Thus enabling the HD-LSTM to establish the relationship between the two. This is implemented using a vanilla neural network. Hence for the HD-LSTM used in this chapter $h(t)$ in (3.3) can be expressed as

$$h(t) = \psi_h\Bigg[W_i\psi_{i-1}\Big[\ldots\big[\psi_1[W_{hh}h(t-1) + W_{\bar{x}h}\overline{x(t)} + W_{F'h}F'(t)]\big]\Big] + b_h(t), \tag{3.5}$$

where $F'(t)$ is obtained from the PMLP, $W_{F'h}$ denote the PMLP output to the HD-LSTM hidden weight matrix, $W_{\bar{x}h}\overline{x(t)}$ is obtained using (6.1), ψ_h is the non-linear function obtained using the vanilla neural network, W_i is the weight to the vanilla neural network between $(i-1)$ hidden states and i^{th} hidden state. For $k \in \{1,2,...,K\}$ layers, the hidden vector sequence $h(t)$ in (4.2) becomes $h^k(t)$. This can be expressed as

$$h^k(t) = \psi_h \Big[W_i \psi_{i-1} \Big[... \big[\psi_1 [W_{h^k h^k} h^k(t-1) + W_{\bar{x}h^k}\overline{x(t)} + W_{F'h^k} F'(t)] \big] \Big] \Big] + b_h^k(t). \quad (3.6)$$

Substituting for u types of data observations to (4.3), then

$$h^k(t) = \psi_h \Big[W_i \psi_{i-1} \Big[... \big[\psi_1 [W_{h^k h^k} h^k(t-1) + (W_{\overline{x_1}h} \overline{x_1(t)} + ... + W_{\overline{x_u}h}\overline{x_u(t)}) + W_{F'h^k} F'(t)] \big] \Big] \Big] + b_h^k(t). \quad (3.7)$$

The network computes all the K hidden states. Once the value of $h^k(t)$ at $k = K$ is obtained, the network output $y(t)$ can be calculated as

$$y(t) = W_{h^N y} h(t)^N + b_y(t), \quad (3.8)$$

where $b_y(t)$ is the output bias function. The process of building this structure is discussed in the following section.

3.2.3 HCLM Implementation

The HCLM was developed using Python 3.5 with TensorFlow as the backhand. A detailed explanation of the main steps illustrated in Fig. 3.4 for developing the HCLM is given below:

3.2.3.1 Data Pre-processing and Feature Extraction

The PMLP inputs are the multiple rainfall forecasts and past rainfall observations. They are inputted to the PMLP for selecting the best forecast $F'(t)$ from $F_n(t)$. The training set for the PMLP can be denoted as $Q = \{F_n(t), t = 1,2,...,T\}$ for T number of training data. The PMLP output $F'(t)$ and $x_u(t)$, the U types of observation data

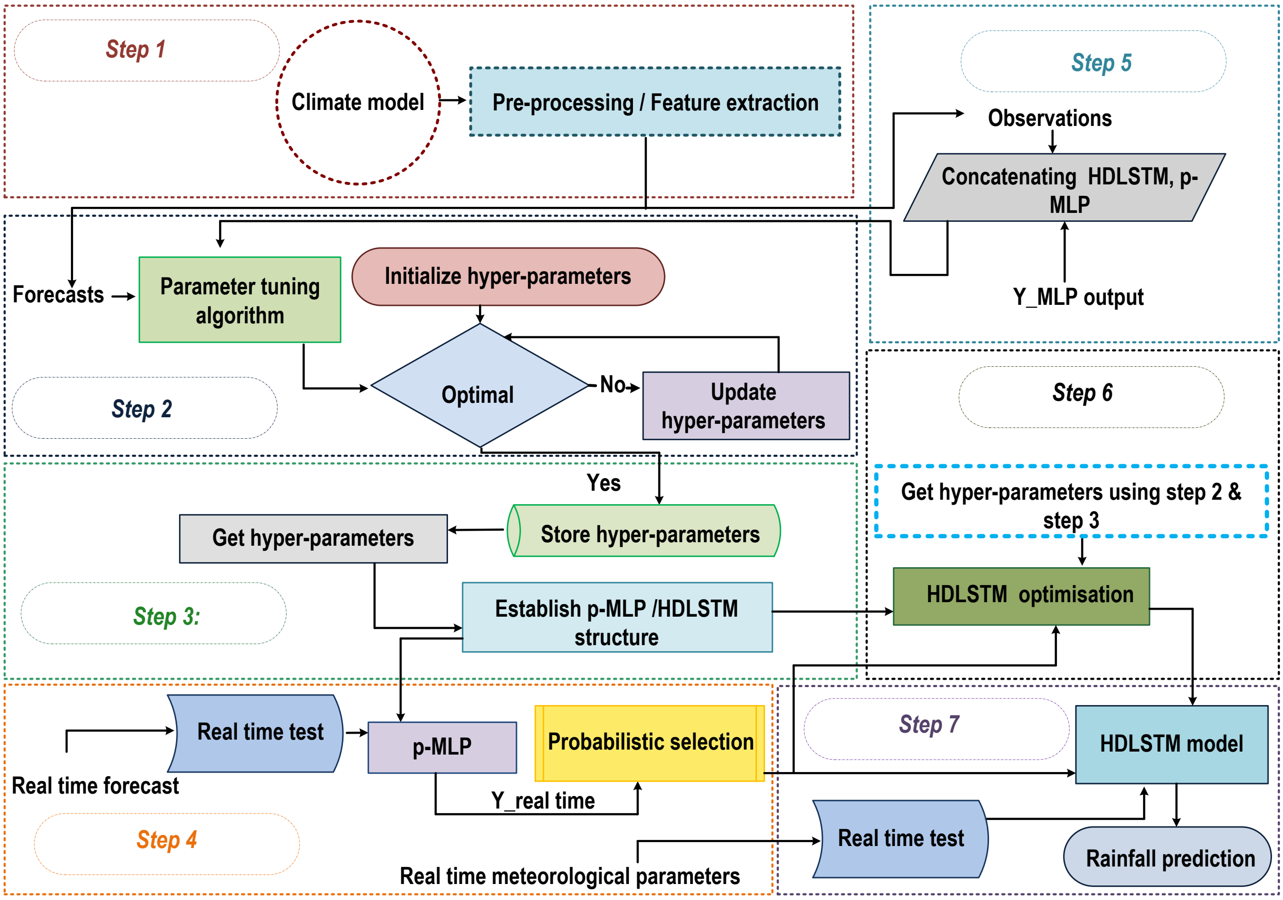


Figure 3.4: The process of establishing HCLM is done in seven steps: (1) Preprocessing & feature extraction; (2) PMLP structure selection & parameter tuning; (3) PMLP model optimisation; (4) Forecast selection using PMLP; (5) The PMLP is concatenated with the HD-LSTM for temporal analysis of forecasts with observations over forecast interval; (6) HD-LSTM model optimisation is done using step 2 & 3, and (7) Rainfall prediction is obtained.

are passed to the HD-LSTM. The training set for the HD-LSTM can be denoted as $R = \{(x_u(t), F'(t)), t = 1, \ldots, T\}$ for T number of data.

3.2.3.2 Structure Selection and Parameter Tuning

Parameter tuning is tedious and costly as the network goes through numerous iterations according to the performance of different sets of hyper-parameters. The random search method exhibits the same performance as many common hyper-parameter optimization methods with lower computation cost [204]. Hence I have chosen the random search method. The key hyper-parameters optimized are the length of the input sequence and the number of nodes in the hidden layer. A dropout of 20% is used to prevent over-fitting.

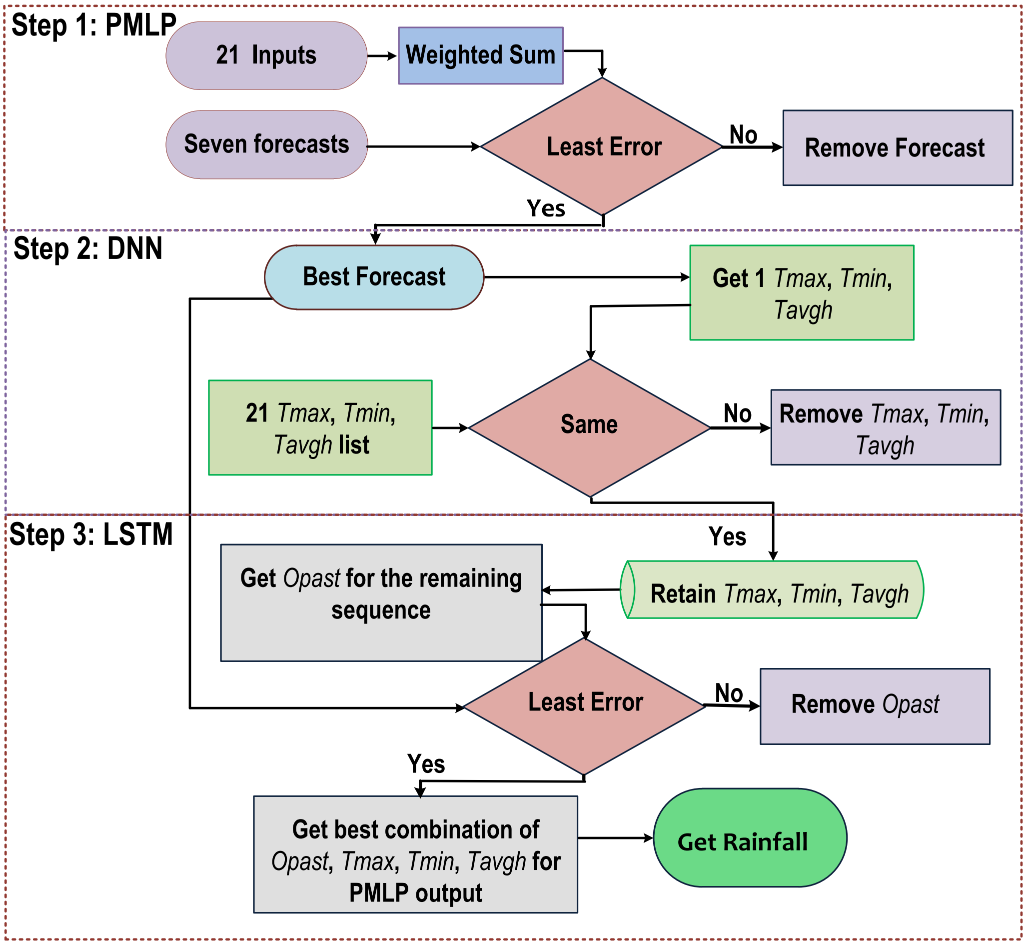


Figure 3.5: Illustration of the HCLM working mechanism. (1) The forecast having the least error with the weighted sum of the past 21 days is chosen as the best forecast by PMLP; (2) DNN removes all $T_{max}(t)$, $T_{min}(t)$, and $T_{avgh}(t)$ data from the list of 21 sequence that does not match with the observations corresponding to PMLP selection and retains $O_{past}(t)$ for the remaining values; and (3) HD-LSTM analyse the DNN input sequence and selects the one with minimal error with PMLP output. For the final selected combination of $T_{max}(t)$, $T_{min}(t)$, $T_{avgh}(t)$, $O_{past}(t)$, and PMLP best forecast the LSTM predicts the rainfall.

3.2.3.3 PMLP Model Optimisation

Once tuning is completed, optimal parameters are obtained. Next, the model undergoes training to find the best forecasts for the given inputs. PMLP learns through the backpropagation of errors by searching the data set containing all forecasts to find the best forecasts that enhance the performance. This provides the network with efficient learning updates with high-quality gradients.

3.2.3.4 Real-time Prediction

Once training is completed the PMLP will predict $F'(t)$; the best forecast by analyzing the given set of multiple forecasts $F_n(t)$. The PMLP uses Eq. (3.2) for selecting the best forecast as its output.

3.2.3.5 Model Concatenation

$F'(t)$ is the best forecast obtained using the PMLP. For U types of weather data changes observed over the forecast period, the observations $x_u(t)$ undergo temporal analysis with the best of forecasts $F'(t)$. For this purpose, the PMLP output is concatenated to the HD-LSTM.

3.2.3.6 HD-LSTM Model Optimisation

Steps two and three are applied to the HD-LSTM for model optimization. For rainfall events generally, it is thought that recent weather data is more important for prediction. However, older data can help the model in recognising general trends and movements. So the selection of sequence length is done for the rainfall data.

3.2.3.7 Rainfall Event Prediction

For the given data observations $x_u(t)$ and $F'(t)$ the best forecast selected by the PMLP, the HD-LSTM predicts the rainfall for the next day.

Fig. 3.5 illustrates how the PMLP selects the best forecast, and HD-LSTM correlates with the observations after completing the above seven steps of establishing the HCLM. The difference between the seven forecasts and the weighted sum of the past 21 days' PMLP inputs is calculated. The one with the least error is selected as the best forecast. For the selected forecast, the corresponding $T_{max}(t)$, $T_{min}(t)$, and $T_{avgh}(t)$ values are taken. From the list of 21 $T_{max}(t)$, $T_{min}(t)$, and $T_{avgh}(t)$ values, those that do not match with the output time observations selected by the PMLP are removed by the DNN. The $O_{past}(t)$ values corresponding to the DNN retained $T_{max}(t)$, $T_{min}(t)$, and $T_{avgh}(t)$ are taken for further examination. Next the remaining values of $T_{max}(t)$, $T_{min}(t)$, $T_{avgh}(t)$, and $O_{past}(t)$ are passed on to the LSTM for analysis. For the given sequence, the LSTM looks for a row with the minimum difference between $O_{past}(t)$ and the PMLP selected rainfall forecast, which is chosen as the output. The values of the selected row are weighted to scale the data to predict the rainfall for the target day. Hence to summarise, the HCLM model uses a PMLP network to select the best forecast. The best forecast produced by the PMLP is compared with the other observed weather changes over the forecast period by the HD-LSTM to predict the

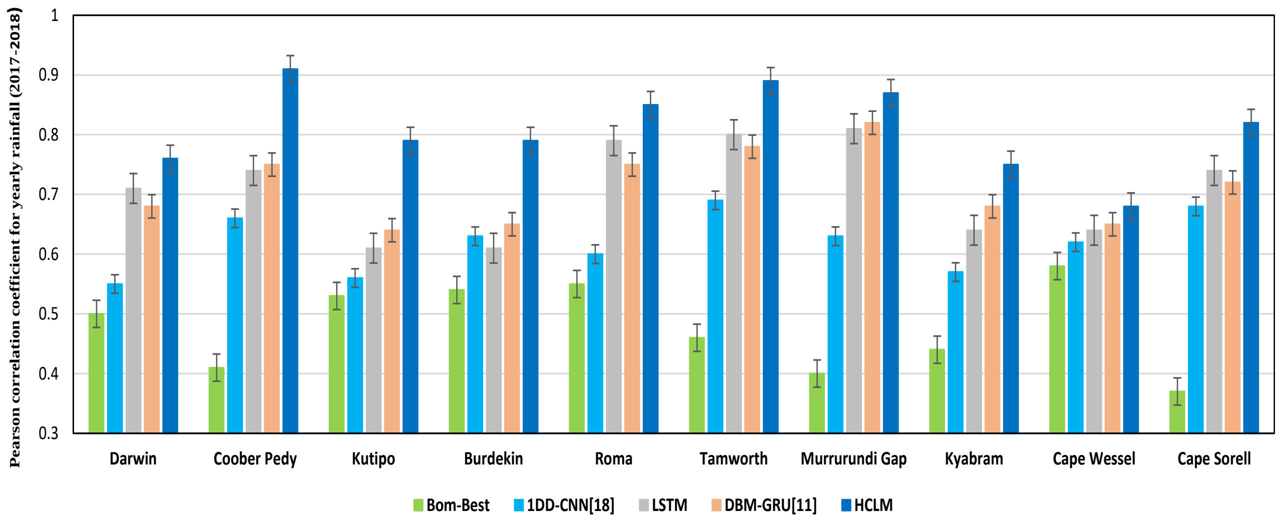


Figure 3.6: Pearson correlation values between the BOM-Best, LSTM, 1DD-CNN, DBM-GRU & HCLM predicted rainfalls. The HCLM, LSTM, 1DD-CNN [18], DBM-GRU [11] & BOM-Best exhibit average correlation of 0.81, 0.72, 0.61, 0.70, and 0.48, respectively.

3.3 Experimental Results

To evaluate the HCLM, I test it on the dataset from the BOM of Australia [200]. According to the BOM, there exist six major climate zones, namely the equatorial, the tropical, the subtropical, the desert, the grassland, and the temperate zones. I compare HCLM model with the best BOM forecasts produced using the climate model [24], the LSTM which is the most widely used neural network for time series data analysis [33] [32] and two recent related work 1DD-CNN [18], DBN-GRU [11]. For real data analysis, ten stations that span across six climate zones are considered for model training and preliminary verification by all the models (123,640 data points). For testing, I used rainfall data from June 2017 to May 2018. For training the HCLM, LSTM, 1DD-CNN, and DBN-GRU, I used data from May 2015 to May 2017. I used seven forecasts for the daily rainfall and the observation data, i.e., $O_{past}(t)$, $T_{max}(t)$, $T_{min}(t)$, and $T_{avg_h}(t)$. HCLM, LSTM, 1DD-CNN, and DBN-GRU use this data to predict the next-day rainfall. Daily rainfall forecasts from the CM are produced over a 24-hour period back to back, where the last forecast is available only one day before the event. Therefore, this study aims at only the next-day prediction.

The 1DD-CNN [18] is used for comparison as it uses observational data from the BOM as input for rainfall prediction. The simulation results in [18] indicate a higher performance than BOM's forecasts. Therefore this provides a fair comparison to my approach. While [11] uses a hybrid gated recurrent unit (DBN-GRU) model to analyze time series meteorological data. This comparison is done since LSTM and GRU introduce additional gating components that handle vanishing gradients problems in traditional RNNs for time series data analysis. A comparison will help understand which of the two can better analyze time-series weather data. Hence I find these works very important for comparison.

3.3.1 Statistical Analysis

The accuracy of daily weather prediction for 12 months is compared using the Pearson correlation coefficient. The closer the value is to the $+1$, the better the forecasts. The bar graph in Fig. 3.6 illustrates the Pearson correlation coefficient values obtained for 10 stations over one year. The graph illustrates a higher temporal correlation of the HCLM than [18], [11], the LSTM, and the BOM. There is a significant improvement in the fitting of time series fluctuations of the HCLM predicted rainfall. For the average temporal correlation for 10 stations for one year calculated using the Pearson correlation, the HCLM exhibits an average correlation of 0.81, LSTM exhibits an average correlation of 0.72, [18] exhibits an average correlation of 0.61, [11] exhibits an average correlation of 0.70 and the BOM exhibits an average correlation of 0.48.

Table 3.2: Rainfall Prediction Error Comparison

SE	0.67	0.58	0.64	0.52	0.47
Location	**BOM**	**LSTM**	**1DD-CNN[18]**	**DBN-GRU[11]**	**HCLM**
Darwin	8.64	7.56	8.20	6.89	6.15
Coober Pedy	1.42	1.26	1.30	0.96	0.85
Kutipo	3.86	2.40	2.82	2.22	1.79
Burdekin	5.09	4.92	5.28	4.28	3.64
Roma	3.93	2.94	3.21	2.55	1.80
Tamworth	3.55	2.73	2.96	2.62	2.20
Murrurundi Gap	3.49	2.57	2.94	2.50	2.09
Kyabram	3.00	2.25	2.39	2.15	1.56
Cape Wessel	6.48	4.83	5.51	4.46	3.41
Cape Sorell	6.77	4.00	5.05	3.62	2.63
Average	4.62	3.55	4.00	3.23	**2.61**

Note: SE stands for standard error. The data reported in this table are for the 10 stations during 2017-2018.

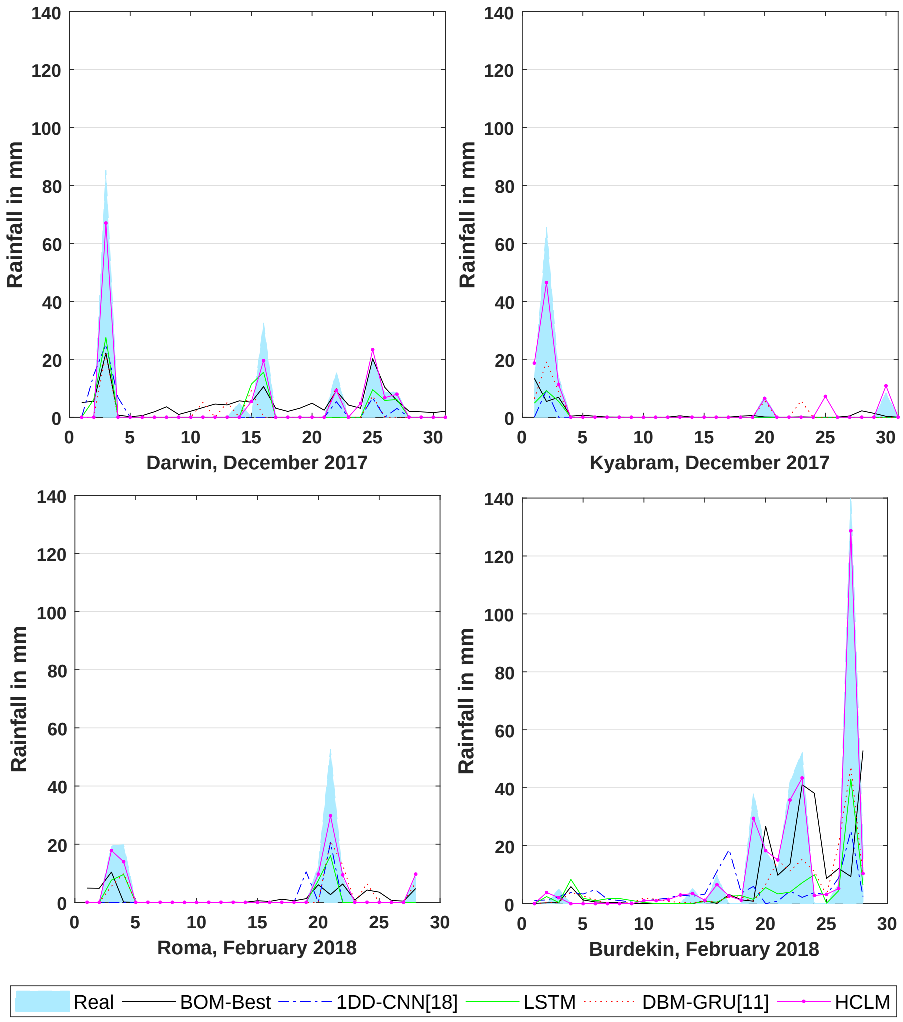


Figure 3.7: Rainfall forecast results of 4 stations that experienced the highest rainfall events between June 2017 to May 2018. Graphically compared HCLM with LSTM, 1DD-CNN [18], DBN-GRU [11], and BOM-Best.

In weather predictions, forecasting extreme rainfall events is a major challenge. Fig. 3.7 shows the extreme rainfall events for four locations located in high rainfall zones between June 2017 to May 2018. The comparison results show that the HCLM can outperform the BOM, the LSTM, [18], and [11]. Quantitative measurements are considered for synthesizing the effectiveness of the performances over 12 months for 10 stations. Here RMSE is used to calculate the accuracy of the HCLM. The closer the RMSE values are to zero, the better the forecasts. Fig. 3.8 shows the RMSE for rainfall predictions for 10 locations. The RMSE curves show that the proposed HCLM has a lesser error in predicting the rainfall amount in comparison to the BOM and the deep learning models. Table 3.2 gives the quantitative RMSE values for the 10 stations. It can be seen from this table that the average RMSE of the HCLM is the smallest among all approaches, indicating that the HCLM has higher rainfall prediction accuracy. The rainy and non-rainy days occurrence prediction accuracy test

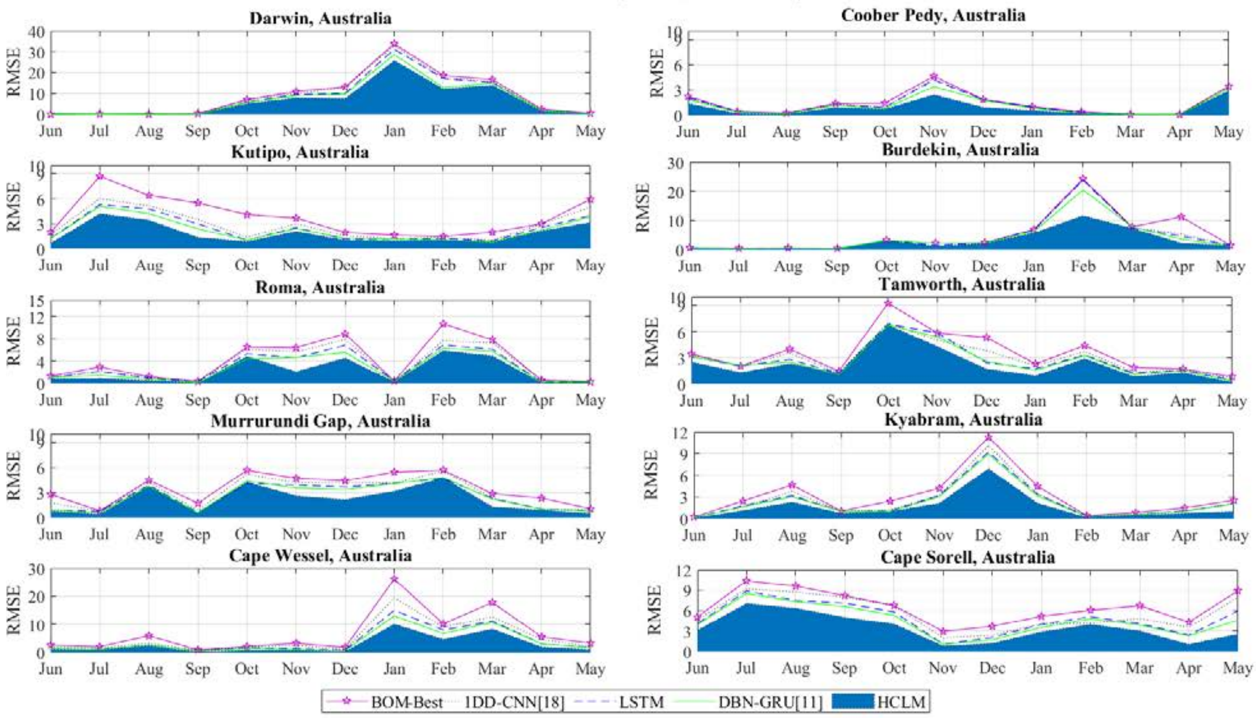


Figure 3.8: Comparison of the root mean square error for 10 locations across all major climate zones in Australia for 12 months (source: BOM).

was carried out for the 10 locations. Table 3.3 shows the prediction results. The results show that there is a notable improvement in the prediction accuracy of the HCLM in comparison with the BOM, the LSTM, the 1DD-CNN [18], and DBN-GRU [11]. There is an improvement of 18.65% for the HCLM over the best of BOM seven days forecasts. Also the HCLM is 8.64% more accurate than the LSTM. With the 1DD-CNN [18] there is an improvement of 10.57%. And to the DBN-GRU model in [11], there is an improvement of 7.10%. Hence from the results, it is apparent that the HCLM model is highly accurate in predicting the chance of occurrence of rainy days in comparison to the BOM, LSTM, 1DD-CNN, and DBN-GRU.

3.3.2 Ablation Studies of HCLM

In this section, the individual input and model components of the propose HCLM are assessed using seven forecasting approaches: 1) PMLP: performance of the PMLP module with the seven forecasts as input; 2) HD-LSTM+(All): HD-LSTM with all the forecasts and observations as input; 3) HD-LSTM+($Forecast$ 7+Obs): HD-LSTM with the last forecast and all observations as input; 4) HD-LSTM-(O_{past}): HD-LSTM

Table 3.3: Comparison of Probability of Rainfall Occurrence

SE	0.029	0.020	0.021	0.018	0.021
Location	**BOM**	**LSTM**	**1DD-CNN[18]**	**DBN-GRU[11]**	**HCLM**
Darwin	71.51%	74.25%	80.00%	78.63%	**90.68%**
Coober Pedy	86.84%	81.91%	89.31%	83.54%	**97.26%**
Kutipo	80.68%	84.18%	80.79%	85.93%	**90.57%**
Burdekin	78.15%	83.95%	79.34%	81.29%	**89.23%**
Roma	85.20%	91.23%	85.21%	90.14%	**92.33%**
Tamworth	69.31%	80.34%	77.15%	82.33%	**84.65%**
Murrurundi Gap	70.68%	81.23%	79.82%	80.95%	**86.57%**
Kyabram	68.76%	79.29%	76.87%	79.67%	**83.81%**
Cape Wessel	66.30%	77.83%	74.56%	78.31%	**80.43%**
Cape Sorell	56.16%	66.98%	64.10%	68.01%	**73.92%**
Average	73.36%	80.12%	78.72%	80.88%	**87.04%**

Note: SE stands for standard error. The data reported in this table are for the 10 stations during 2017-2018.

Table 3.4: Comparison of Ablation Studies

Approach	**HCLM**	**PMLP**	**HD-LSTM+(All)**	**HD-LSTM+(Forecast 7+Obs)**	**HD-LSTM-(O_{Past})**	**HD-LSTM-(T_{max})**	**HD-LSTM-(T_{min})**	**HD-LSTM-(T_{avg_h})**
RMSE	**2.61**	4.97	3.49	7.93	6.38	6.17	5.82	3.51
SE	**0.47**	0.43	0.62	0.89	0.71	0.76	0.63	0.48

without $O_{past}(t)$ as input; 5) HD-LSTM-(T_{max}): HD-LSTM without $T_{max}(t)$ as input; 6) HD-LSTM-(T_{min}): HD-LSTM without $T_{min}(t)$ as input and 7) HD-LSTM-(T_{avgh}): HD-LSTM without $T_{avgh}(t)$ as input. Approaches 4 to 7 use all the BOM predictions. Table 3.4 shows the ablation study comparison with the HCLM.

The higher RMSE of PMLP than HCLM suggests that PMLP alone does not achieve the same performance as that of HCLM. Approach 2 shows that the HD-LSTM module has a smaller RMSE than PMLP, signifying that linking changes in observation data with the CM forecasts is essential. However, the higher RMSE of Approach 2 than HCLM suggests that rather than more training data, better training data will improve network learning. So the use of PMLP to select the best forecast may have boosted the performance of the HCLM. The higher RMSE of approach 3 denotes that the final forecast alone cannot improve the performance of HD-LSTM.

Approaches 4 to 7 remove the observation data one by one from HD-LSTM, so as to identify the removal of which observations has the most effect. Table 3.4 shows that the removal of $O_{past}(t)$ has the most effect on HD-LSTM, probably because there was a learning gap with the HD-LSTM unable to relate the past rainfall trends with the forecast. The removal of $T_{avg_h}(t)$ seems to have the least effect. This could be because the model was able to learn the average hourly temperature trends from

$T_{max}(t)$ and $T_{min}(t)$ data. Also, approaches 4 to 7 do not outperform HCLM, further indicating that directly inputting all forecasts does not enhance model learning. This must be the case since the forecast errors add up and misguide the model. So these results demonstrate that the PMLP and HD-LSTM work collaboratively to enhance the performance of the HCLM.

3.4 Conclusion

In this chapter, a hybrid climate learning model (HCLM) is developed. The HCLM builds its future rainfall predictions on the multiple forecasts produced by the climate model (CM) for the rainfall events. The HCLM is designed using a PMLP network and an HD-LSTM network. The PMLP selects the best forecast from the multiple CM forecasts, and the HD-LSTM finds the relationship of the PMLP-selected forecast with the temperature and rainfall observations during the forecast time. The proposed model's performance evaluation was conducted in 10 different locations across six major climate zones in Australia. For average yearly rain prediction results in all the locations analyzed, the HCLM predicted rainfall is more similar to the actual rain than the CM and state-of-the-art related works. The potentiality of the model predictions has been also established in terms of ablation studies, RMSE, and Pearson correlation. Hence it can be concluded that there is a significant improvement by the HCLM to predict the rainfall over the climate and deep learning comparison models.

The proposed model addresses the first research problem of refining climate model forecasts (presented in Section 1.2.1) and leads to the first original contributions presented in Section 1.3. The high performance of the rainfall prediction model in this chapter leads to the investigation of the second research problem, improved methods for high-resolution SSM (the spatially distributed moisture distribution) information with less spectral information loss in the next chapter of this thesis.

Chapter 4

3D Bi-directional LSTM for Satellite Soil Moisture Downscaling

This chapter contains materials which has been accepted (with minor revisions) for publication in IEEE Transactions on Geoscience and Remote Sensing:

[2] **Neethu Madhukumar**, Eric Wang, Clinton Fookes, and Wei Xiang,"Three-dimensional Bi-directional LSTM for Satellite Soil Moisture Downscaling," *IEEE Transactions on Geoscience and Remote Sensing*, Accepted for publication on 22 Nov. 2022 (Impact Factor: 8.125, h-index: 113, Journal rank (Remote sensing): 3).

4.1 Introduction

Surface soil moisture (SSM) is a prime factor that controls plant water requirements [205]. Microwave remote sensing is widely acknowledged as the most promising technique to measure the spatial distribution of water content in the soil over time. This is due to its direct relationship with the soil dielectric constant, its ability to penetrate clouds, its reduced sensitivity to vegetation, and surface roughness [206]. Microwave observations from various missions are used in global SSM mapping [25]. However, global SSM products from these missions are of low spatial resolution (approximately 25–40 km), which greatly limits their use in regional agricultural applications. The Soil Moisture Active Passive (SMAP) satellite was launched to provide high-resolution SSM, but due to its radar failure, it can only provide global daily soil moisture products at a 9-km resolution [26], [27]. However, for

PI application this resolution is very low [64]. In an agricultural setting, the SSM map resolutions greater than 1-km cannot cover small and micro-farms. They are too coarse to inform on-farm irrigation and other management practices. As there remain significant challenges associated with the spatial resolution of satellite SSM, further downscaling to at least 1-km is required for localized applications [64].

According to literature, for a spatial resolution of $\leq$ 1-km, SSM changes at any target region are influenced by the SSM changes in the neighboring regions [64], [43]. Furthermore, as SSM is part of the hydrological cycle, its spatial distribution is influenced by the changes in the hydrological cycle over time [38], [39]. To sum up, for a spatial resolution of $\leq$ 1-km, the information from both the spatial and time axis is relevant for SSM downscaling. From a broader deep learning (DL) perspective, such spatial-temporal factors have been addressed by the combination of spatial and sequence learning models. A classical example of combining these two aspects of DL is video and motion prediction [207], which has striking similarities to many of the dynamic geoscience (GS) and remote sensing (RS) problems. Satellite images are time-evolving multi-dimensional data structures. Because of this fact, studies have begun to apply combined spatial and sequential (convolutional–recurrent) approaches to many similar geoscientific problems such as SSM downscaling [208], [209].

There is a renewed interest in blending the spatial (two-dimensional) and time (one-dimensional) aspects of DL for GS and RS applications [61]. By blending the spatio-temporal aspects, a three-dimensional (3D) structure is produced. Machine learning (ML)-based SSM downscaling schemes typically build on hand-crafted features and build a linking model to link SSM and input features [40]. The linking model applies the learned relationship between the SM and the input features at coarse to fine scales. This can be better achieved using DL-based models, which have achieved great results in the RS and GS fields [210], [1]. Therefore, rather than amending classical ML models, DL models that can automatically extract 3D features are ideal for building a 3D linking model for SSM downscaling. Among various DL models, convolutional neural network-long short-term memory network (CNN-LSTM) based DL models are highly popular for spatio-temporal analysis of RS images [211], [212].

A general approach while using the CNN-LSTM network combination is to use the CNN for feature extraction and the LSTM for sequence analysis. The traditional LSTM performs forward sequence analysis, while the bi-directional LSTM (Bi-LSTM) performs forward and reverse sequence analyses [43]. At any given time for an RS image sequence, all three domains must be simultaneously analyzed to obtain a full picture of the RS spectra relationships [42]. The RS image sequence analysis in the time domain is performed for all preceding time RS images for daily predictions. This approach is logical for time-series data as future time values beyond the present are unavailable. However, when considering the spatial domain, all neighboring pixels correlate with the target pixel. As such, in the spatial domain, sequence analysis should be performed for both the before and after pixels from the target pixel in an RS image. This bi-directional analysis should be performed for both the spatial sequence within a feature and the spectral sequence between features since they both have correlations with the prediction target. Hence, for RS image sequence relationship analysis, each point in the time domain (1D) should be further analyzed in a bi-directional manner in the 2D spatial domain (within and between spectra).

In the case of downscaling, it establishes an association between the coarse-resolution parameter and high-resolution auxiliary data [282]. If the spectra-spatio-temporal (3D) bi-directional associations can be included, it will enhance the downscaling output. Such a downscaling method is currently unavailable [41]. Compared to existing downscaling methods, deep learning approaches look most apt for implementing such a downscaling approach as they are auto-tuned to extract relationships from diverse and large volumes of data, such as RS [61], [41], [283]. However, the required type of downscaling relationships cannot be obtained using existing deep learning sequential models [43], [284]. They only perform forward and/or backward analysis of the subsequent sequences before the target time and do not simultaneously perform another bi-directional analysis of the 2D spatial domain over the different spectral inputs in each time instant. To overcome the limitations of achieving a 3D bi-directional SM downscaling method, a three-dimensional (3D) Bi-LSTM model is proposed.

For predicting SM, soil moisture indices (SMIs) obtained from optical and thermal remotely sensed MODIS observations are provided as input to the proposed model. In Fig. 4.1, to predict image 15, along with the preceding time images, in each time instant a bi-directional sequence analysis is performed for the 2D spatial domain within an image frame as well as between image frames. Since SM data is correlated in the time domain, the bi-directional analysis from previous time instants is considered as well. As such, a three-dimensional (two-dimensional spatial domain plus one-dimensional time domain) bi-directional (spatial pixel neighbors in the horizontal and vertical spatial axis in the RS image) LSTM for spatial-temporal SM prediction is proposed. Furthermore, in existing CNN-LSTM models, a spectral correlation that can offer important distinctive information is not fully exploited. This is critical for SM downscaling approaches that use multi-spectral MODIS data for obtaining fine-resolution auxiliary SM information. To address this problem, a covariance-adaptive CNN-based feature extraction method is proposed. The use of covariance helps exploit the correlation among differing SMIs and the SM for the target region, thus allowing for a more discriminative feature extraction process. The CNN kernel is made adaptive to the covariance matrix (CM) evaluated for the SM and SMIs. After feature extraction, time-series prediction is followed.

Fig. 4.2 illustrates the proposed satellite soil moisture downscaling methodology. First, with SM and SMIs as inputs, feature extraction is achieved using a covariance-adaptive CNN. Next, the proposed 3D-Bi-LSTM performs a bi-directional spectra-spatio-temporal relationship analysis. The proposed model is trained and validated on coarse-scale satellite data. After establishing the relationship between SM and SMIs, the model is given fine-scale input SMIs to produce 1-km SM outputs. To summarise in this chapter a new covariance adaptive CNN model for explicit region-specific MODIS SMI extraction for SMAP SSM downscaling is proposed. Next, a novel three-dimensional bi-directional LSTM (3D-Bi-LSTM) model is proposed for SSM downscaling using the covariance adaptive CNN model selected SMIs and SMAP SSM.

The rest of this chapter is organized as follows. Section 4.2 elaborates on the study area and dataset. Section 4.3 details the proposed three-dimensional bi-directional LSTM for satellite soil moisture downscaling. Section 4.4 presents the results and

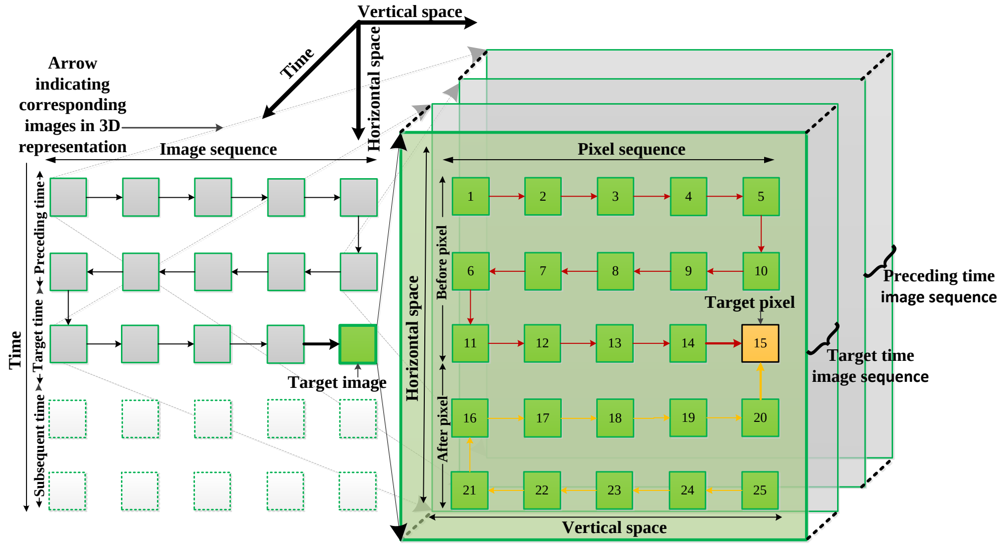


Figure 4.1: Illustration of three-dimensional RS image sequence analysis. In the time domain, the preceding time RS images (represented by grey blocks with a green-solid border) aid in the target image (represented by a single green block) prediction. The RS images in subsequent time instants from today (represented by white blocks with a green dashed border) are unavailable. For the 2D spatial domain, the target pixel represents the current pixel to be predicted in the target RS image. The red and yellow arrows inside the big green block represent the bi-directional 2D spatial domain sequence analysis for each input image at a given time. The pixel neighbors in the horizontal and vertical axes (represented by green blocks within the big green block) undergo the bi-directional analysis to predict the target pixel (represented by a yellow block with a black-solid border in the big green block). Similar bi-directional 2D spatial domain sequence analyses are conducted for all other images at a given time (represented by big grey blocks with a green-solid border) for the target pixel in the target image prediction.

discussions. Concluding remarks are drawn in Section 4.5.

4.2 Study Area and Dataset

The SMAP satellite provides daily soil moisture data at a 9-km resolution. The product name is SMAP Enhanced L3 Radiometer Global Daily 9-km EASE-Grid SSM V001 (SMAP L3-SM-P-E). To downscale soil moisture to a finer resolution, auxiliary information from MODIS observations is used. This approach is widely adopted in the SMOS and AMSR-E downscaling schemes. In this chapter, 27 indices are used as auxiliary soil moisture information. The 27 indices are NDVI, EVI, ANIR, NDWI, DDI, GVMI, PDI, MPDI, MPDI1, MSI, MSPSI, NDII6, NDII7, NDTI, NMDI, SANI, SASI, SPSI, SRWI, VSDI, VSWI, MSAVI, SIMI, LST, DDI, SW, and TVDI [40], [35]. Collectively they are referred to as SMIs in this chapter, which includes drought indices (DI) and vegetation indices (VI). Aqua MODIS and Tera MODIS are used in

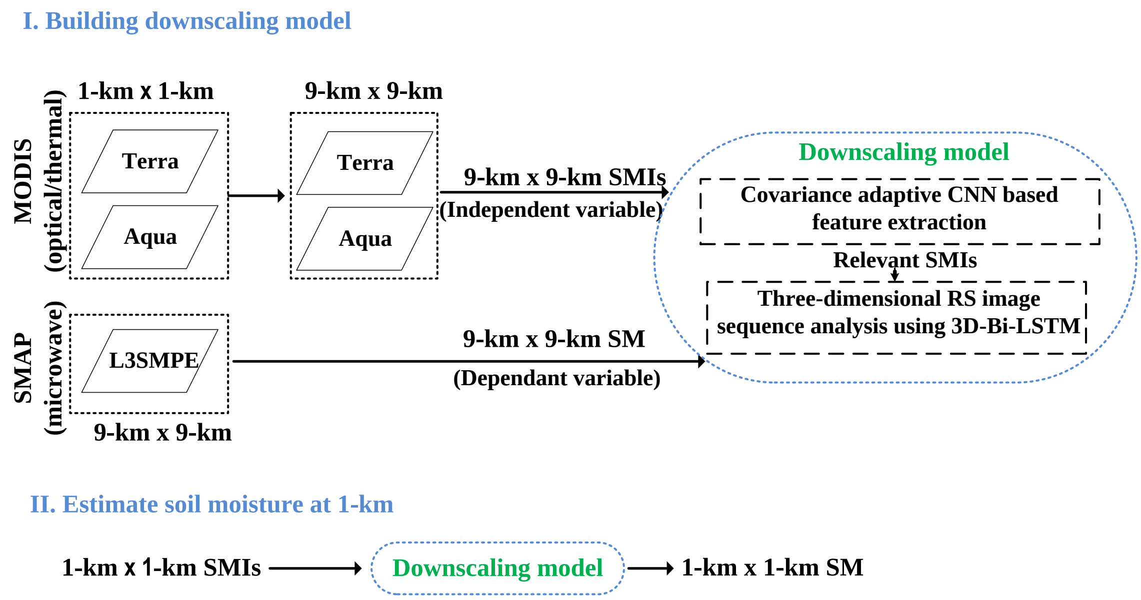


Figure 4.2: Illustration of the proposed satellite SSM downscaling method. The MODIS and SMAP remotely sensed observations are used as input for training the downscaling model. The downscaling model is established using 9-km × 9-km MODIS and SMAP observations. The model establishes a relationship between SSM and SMIs at 9-km and applies the learned relationship to the 1-km SMIs in an effort to predict 1-km SSM.

tandem with satellite local equatorial crossing times of the SMAP ascending and descending observations. As such, the Aqua MODIS is combined with ascending SMAP observations and Tera MODIS with SMAP descending observations for downscaling. The downscaled soil moisture is validated using ground in-situ observation data from the SMAP/IN-SITU CORE VALIDATION V001 [226], SOILSCAPE [227], SMAPVEX-16 [228], and SMAPVEX-19 [229] datasets. These datasets can be accessed from NASA earthdata (https://search.earthdata.nasa.gov/search). To assess the transferability of the proposed 3D-Bi-LSTM model, the study area is divided into geographically distinct locations, namely the Terra d'Oro, Carman, Petersham, Fort Cobb, Little Washita, Duero basin, and Walnut Gulch. The study sites and the SM dataset within each site are detailed further.

4.2.1 Terra d'Oro, California, United States

4.2.1.1 Site Characteristics

The 400 acres of Terra d'Oro vineyards centered at 38°30'36.0"N, 120°48' 00.0"W are located in the Amador County in California. The region's climate is classified as a tropical climate, with warmer, drier weather in summer and cooler, wetter weather in winter. For over three decades, Terra d'Oro's focused on zinfandel grapes cultivation. The vines are rooted in shallow topsoil above rocky granite hardpan on a series of gentle slopes between 1,300 and 1,600 feet in elevation.

4.2.1.2 SoilSCAPE Dataset

The soil moisture sensing controller and optimal estimator (SoilSCAPE) project provides in-situ measurements of soil moisture for the validation of spaceborne and airborne products. Terra d'Oro provides data captured by 27 SoilSCAPE sensor nodes between 2013-04-06 to 2015-10-26. From 2016-01-13 Terra d'Oro provides data from 24 SoilSCAPE sensor nodes. The dataset for 2016-01-14 file collections provides Terra d'Oro data from locations 38°30'36.0"N 120°48'00.0"W and 38°30'00.0"N 120°48'00.0"W, which are 2.9 km apart from each other.

4.2.2 Carman, Manitoba, Canada

4.2.2.1 Site Characteristics

The site is located Southwest of Winnipeg in the Carman-Elm Creek area. Clay and fine loamy soils account for approximately 76.5% of the study area. The region experiences a humid temperate climate. Soybeans, wheat, and canola account's approximately 70% of the crops grown in the study area.

4.2.2.2 SMAPVEX16 Dataset

The SMAPVEX16 is a temporary ground sensor data-based SMAP validation station, installed in May and removed in Jul-Aug. This data set contains data collected at 9 real-time in-situ soil monitoring for agricultural (RISMA) stations and 50 temporary soil stations. The proposed network covered the region bounded by latitude 49.756438° N to 49.384164° N, and longitude 97.756385° W to 98.098416° W.

4.2.3 Petersham, Massachusetts, United States

4.2.3.1 Site Characteristics

Petersham is a town in Worcester County, Massachusetts, United States. Paxton soil is found in this region. The climate is classified as a warm summer temperate climate.

Petersham mainly has a considerable amount of conservation land, including river reservations and forests.

4.2.3.2 SMAPVEX19 Dataset

The data consists of ground-based soil moisture measurements recorded by twenty-five temporary stations located in Petersham during the SMAPVEX19-21 campaign. The stations were installed across an area of 30-km by 40-km in May 2019 and operated through the middle of November 2019. The proposed network covered the region bounded by latitude 42.34° N to 42.69° N, and longitude 71.95° W to 72.29° W.

4.2.4 SMAP Core Validation sites

4.2.4.1 Site Characteristics

Due to the high spatial heterogeneity inherent in soil moisture, the point-scale measurements cannot be directly compared to the SMAP data. SMAP partnered with in-situ networks that have sufficiently dense soil moisture sensors to be reliably aggregated to a larger spatial scale. These sites are called core validation sites. In this chapter, the core validation sites studied are from the Contiguous United States (CONUS) and Duero basin. The locations within CONUS have diverse climates, ranging from cold to temperate climate types. The Duero basin climate is classified as temperate.

4.2.4.2 SMAP/IN-SITU CORE VALIDATION Dataset

Carman, Walnut Gulch, Fort Cobb, and Little Washita are distributed over CONUS and each of them contains multiple pixels matching the SMAP footprint. The major land cover type of Carman is croplands, Walnut Gulch is shrubs, Fort Cobb and Little Washita is grassland. REMEDHUS in-situ stations are located within an area of 1300-Km^2 (41.1°–41.5° N; 5.2°–5.6° W) in a central semiarid sector of the Duero basin. REMEDHUS network has 20 in-situ stations with major land cover types classified as croplands. The elevation of stations ranges between 700 and 900 m, with gentle slopes.

4.3 3D Bi-Directional LSTM for Satellite Soil Moisture Downscaling

The coarse-scale time-evolving two-dimensional SSM and SMIs are used to produce fine-scale SSM. Fig. 4.3 illustrates the proposed method for downscaling satellite SM. First, the MODIS bands are stacked together. Next, the SMIs are calculated using formulas from [40], [35]. The SMIs of 1-km $\times$ 1-km and 500-m $\times$ 500-m obtained from MODIS are aggregated and reprojected to the SMAP daily SSM resolution of 9-km $\times$ 9-km. The 9-km $\times$ 9-km SMIs and SSM are then stacked together, and the covariance matrix (CM) is calculated for the SSM and SMIs. This CM is used to adaptively update the CNN kernel weights for feature extraction. A two-layer CNN with an adaptive Gaussian kernel is adopted for FE. The important features for SSM prediction within the target region are extracted by the proposed CNN. Each 2D feature extracted is converted to a single 1D sequence using transposed ordering to enhance the spatial/spectral correlation within a time instant, and for $t \in \{T_0, T_1, ..., T_N\}$ time instants, 'T_N' number of 1D sequences are stacked for improving the spatial/spectral correlation between multiple time instants. The proposed 3D-bi-directional LSTM analyzes the data for SSM prediction and the model is calibrated using feedback errors (difference between the predicted and the actual). After this process, the downscaling model is established and 1-km $\times$ 1-km SMIs are given as input to produce 1-km $\times$ 1-km SSM. The proposed CNN-based feature extraction strategy and the 3D-Bi-LSTM analysis for spatio-temporal SSM prediction are detailed below.

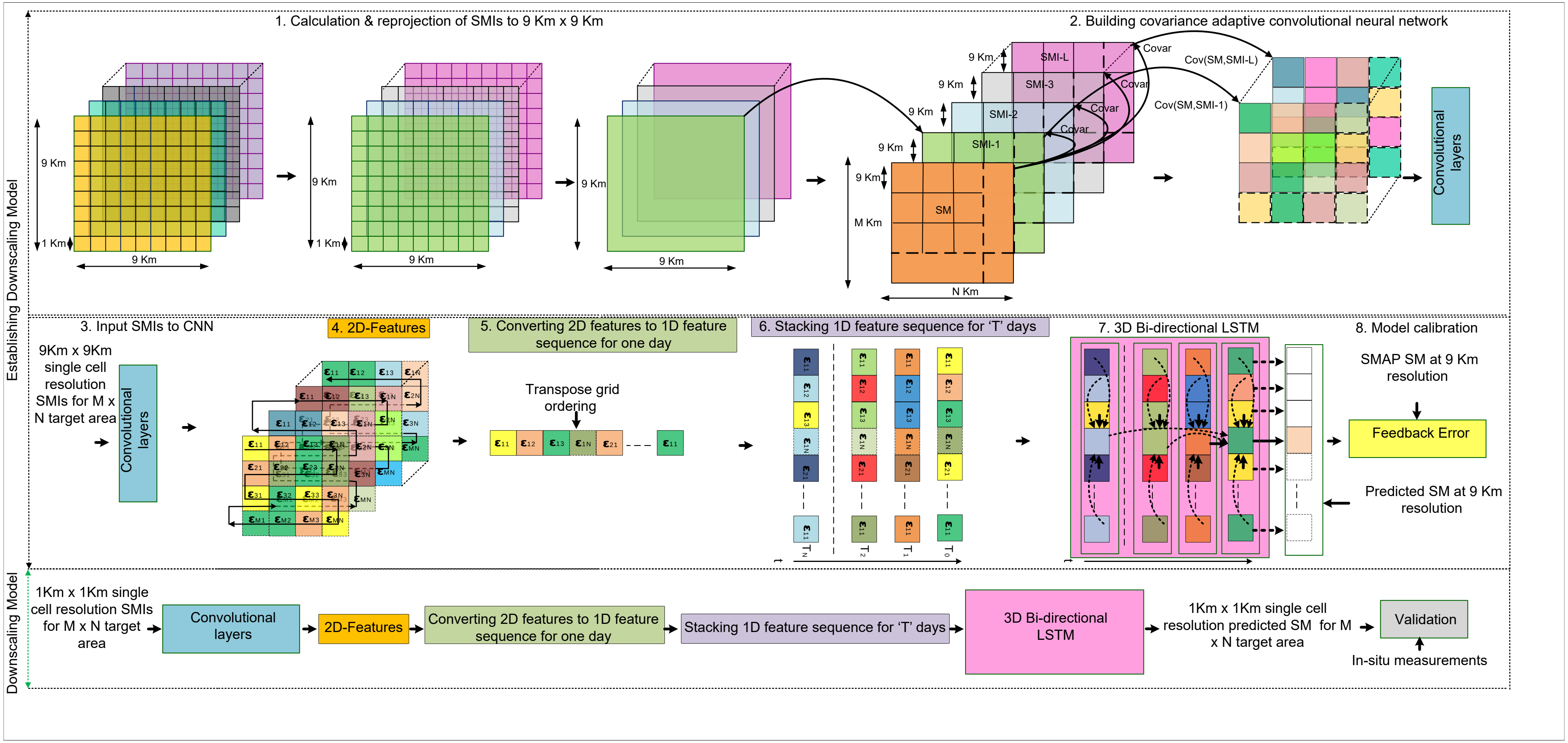


Figure 4.3: Method for achieving the bi-directional spectra-spatio-temporal downscaling. The downscaling is established in 8 steps: (1) SMIs are calculated using formulas on MODIS bands and the 1-km $\times$ 1-km (and/or 500-m $\times$ 500-m) SMIs are aggregated and reprojected to SMAP daily SM resolution (9-km $\times$ 9-km singles raster grid resolution); (2) 9-km $\times$ 9-km SMIs and SM are stacked together, the covariance matrix is calculated for the SM and SMIs, and the CNN kernel weights are updated according to the covariance matrix; (3) the CNN is established for feature extraction; (4) important features for SM prediction within the target region are extracted; (5) 2D features are converted to a 1D sequence using transposed ordering for enhancing spatial correlation within a time instant; (6) For $t \in \{T_0, T_1, \ldots, T_N\}$, T_N 1D sequences are stacked for analysis; (7) 3D-bi-directional LSTM analysis for SM downscaling; and (8) model calibration using feedback errors. After the above 8 steps, the proposed downscaling is established and 1-km $\times$ 1-km SMIs are given as input to produce 1-km $\times$ 1-km SM.

4.3.1 Covariance Adaptive CNN Feature Extraction

The SMIs derived from MODIS are aggregated to the same spatial resolution as SMAP SSM. Let Γ^l_{MN} be the l^{th} SMI raster cell value at a coarse resolution and $\Gamma^{l\prime}_{pq}$ be the SMI raster cell values at a fine resolution within Γ^l_{MN}. For training the model, all $\Gamma^{l\prime}_{pq}$ are aggregated to the resolution of Γ^l_{MN}. This can be expressed as,

$$\Gamma^l_{mn} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \frac{\alpha_{pq}}{\beta} * \Gamma^{l\prime}_{pq} \, \forall \, m > p \text{ and } n > q, \tag{4.1}$$

where there are $m \times n$ raster grids of Γ^l_{mn} for $m \in \{1,2,...,M\}$, $n \in \{1,2,...,N\}$ in the target area. Each Γ^l_{mn} grid consists of $p \times q$ cells of rows and columns for $p \in \{1,2,...,P\}$ and $q \in \{1,2,...,Q\}$. Furthermore, α_{pq} is the total area covered by raster cell $\Gamma^{l\prime}_{pq}$, and β is the total area covered by the target coarse-resolution raster cell Γ^l_{mn}. The target region has M rows and N columns of SMAP raster cells. Let Γ be the SMI derived for the target region with $M \times N$ rows and columns of SMAP raster cells. Using the values from (6.1), the matrix Γ^l can be expressed as,

$$\Gamma^l = \begin{bmatrix} \Gamma^l_{11} & \Gamma^l_{12} & ... & \Gamma^l_{1N} \\ ... & ... & ... & ... \\ \Gamma^l_{M1} & \Gamma^l_{M2} & ... & \Gamma^l_{MN} \end{bmatrix}, \tag{4.2}$$

where $l \in \{1,2,...,27\}$ and Γ^l denotes the l^{th} SMI. That is, Γ^1 is the first SMI, Γ^2 is the second SMI, and so on. The new cell values of Γ^l calculated are re-projected over the target area using functions in the geopandas open-source library [224]. Denote by Υ the SSM matrix at coarse resolution, which can be expressed as,

$$\Upsilon = \begin{bmatrix} \Upsilon_{11} & \Upsilon_{12} & ... & \Upsilon_{1N} \\ ... & ... & ... & ... \\ \Upsilon_{M1} & \Upsilon_{M2} & ... & \Upsilon_{MN} \end{bmatrix}, \tag{4.3}$$

where each Υ_{mn} represents a 9-km $\times$ 9-km region. To derive an adaptive kernel for the changes in SSM with respect to SMI, a covariance matrix is first calculated. Covariance gives the joint variability of the two random variables Υ and Γ. The covariance is calculated using each cell values in (4.2) and (4.3) as,

$$C^l_{mn} = \frac{1}{(mn-1)} \sum_{m=1}^{M} \sum_{n=1}^{N} \Upsilon_{mn} \Gamma^l_{mn} - \frac{1}{(mn-1)} \sum_{m=1}^{M} \sum_{n=1}^{N} \Upsilon_{mn} \times \frac{1}{(mn-1)} \sum_{m=1}^{M} \sum_{n=1}^{N} \Gamma^l_{mm} \, \forall \, l \in \{1,2,...,27\}, \tag{4.4}$$

where C^l_{mn} checks the variance of Γ^l values for each l with respect to Υ values above and below. This analysis is conducted as soil moisture distribution occurs both upstream and downstream geographically. C^l_{mn} represents the m-th row and n-th column of the covariance matrix C^l. Matrix C is updated l times for l number of SMIs, denoted by C^l. The l denotes to which SMI the covariance matrix belongs, i.e., C^1 SSM covariance with respect to the first SMI, C^2 SSM covariance with respect to the second SMI, and so on. Let C^l be the covariance matrix derived from SSM and the l^{th} SMI. It follows from (4.4) that the covariance matrix can be expressed as,

$$C^l = \begin{bmatrix} C^l_{11} & C^l_{12} & \cdots & C^l_{1N} \\ \cdots & \cdots & \cdots & \cdots \\ C^l_{M1} & C^l_{M2} & \cdots & C^l_{MN} \end{bmatrix} \forall \, l \in \{1,2,...,27\}. \tag{4.5}$$

The diagonal elements of the matrix are the covariance between Υ and Γ^l in the same spatial area. The non-diagonal elements represent the variance in SSM with respect to the SMI over the total target area. The covariance matrix C^l is a good measure to track changes in the dependent variable Υ with respect to the independent variables Γ^l. So if the convolutional kernel is adaptive to C^l, this will result in satisfactory feature extraction. For any l^{th} SMI the convolutional kernel can be expressed as,

$$K_{mn} \propto C^l_{mn} \forall \, l \in \{1,2,...,27\}. \tag{4.6}$$

Eq. (4.6) can be satisfied using a Gaussian kernel [225]. The Gaussian kernel can be built adaptive to the values in (4.5) as,

$$K_{mn} = \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{1}{(2\pi)^{\frac{mn}{2}} |C^l_{mn}|^{\frac{1}{2}}} e^{-1/2(\Gamma^l_{mn}-\mu)^T (C^l_{mn})^{-1} (\Gamma^l_{mn}-\mu)}, \tag{4.7}$$

where μ denotes the mean of Γ^l_{mn}. The kernel matrix K is constructed using K_{mn}.

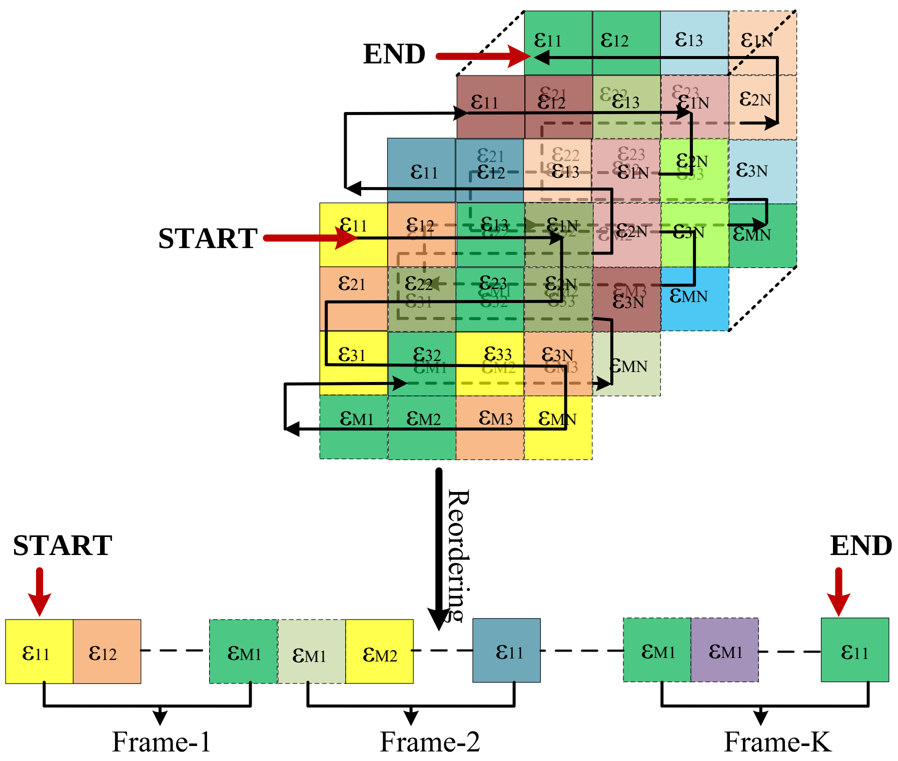


Figure 4.4: A re-ordering strategy for the 2D RS spatial data in each time instant. Each frame represents a 2D spatial image sequence of a feature and ε represents the cells in each image frame.

The CNN feature extraction using K on the l^{th} SMI can be expressed as,

$$\xi^l = K * \Gamma^l, \tag{4.8}$$

where $*$ indicates the convolution operation.

Let ξ^l be the CNN output for $l \in \{1, 2, ..., 27\}$ SMIs denoted by Γ^l, and then ξ^l for the l^{th} SMI can be expressed as,

$$\xi^l = \begin{bmatrix} \xi^l_{11} & \xi^l_{12} & \cdots & \xi^l_{1N} \\ \cdots & \cdots & \cdots & \cdots \\ \xi^l_{M1} & \xi^l_{M2} & \cdots & \xi^l_{MN} \end{bmatrix}. \tag{4.9}$$

The CNN output ξ^l for $l \in \{1, 2, ..., 27\}$ at $t \in \{T_0, T_1, ..., T_N\}$ is subject to a further spatio-temporal analysis by the proposed 3D-Bi-LSTM. The 3D-Bi-LSTM analysis of the CNN output is detailed below.

4.3.2 Transposed 3D-Bi-LSTM Time Series Prediction

The input to 3D-Bi-LSTM is $\xi^l(t)$, and $\Upsilon(t)$ is the prediction target during training. For time-series analysis both $\xi^l(t)$ and $\Upsilon(t)$ should be transformed from 3D to 2D

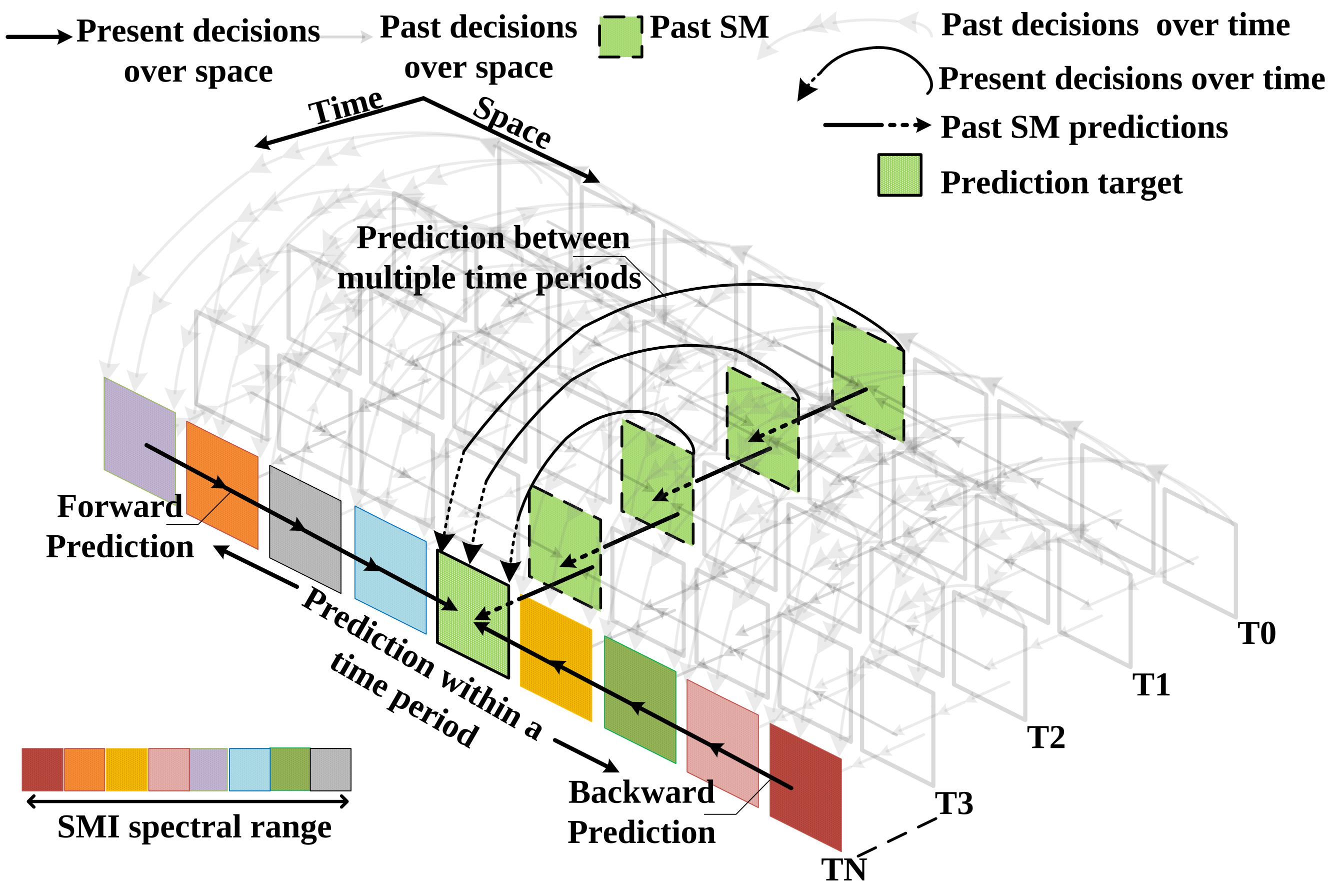


Figure 4.5: Three-dimensional bi-directional LSTM soil moisture prediction structure. The middle green cell with a solid border represents the predicted SSM cell at present time $t = T_N$, where $t \in \{T_0, T_1, ..., T_N\}$, while the middle green cells with a dashed border are the predicted SSM in the previous time instants. The input cells that are at locations before the middle green cell with a solid border undergo the forward prediction analysis, whereas the pixel points after the middle green cell with a solid border undergo the backward prediction analysis for SSM prediction. The bi-directional analysis over time instants T_0 to T_{N-1} along with the bi-directional analysis of the cells at $t = T_N$ to predict the middle green cell with a solid border, i.e., the SSM value at T_N.

sequence by transposing both $\xi^l(t)$ and $\Upsilon(t)$. However, when all the SMIs are transposed and arranged, a single time instant itself will have a long sequence. For $t \in \{T_0, T_1, ..., T_N\}$ the sequence will be very long and many of the important sequences may be subject to the vanishing gradient problem. Hence the 3D to 2D conversion is performed through transposed ordering to enhance the spatial and spectral correlation within a time instant so that important information is not lost. This is illustrated in Fig. 4.4. For $t \in \{T_0, T_1, ..., T_N\}$, there are a total of T_N 2D sequences. Fig. 4.5 illustrates the proposed 3D-bi-directional satellite SSM prediction structure.

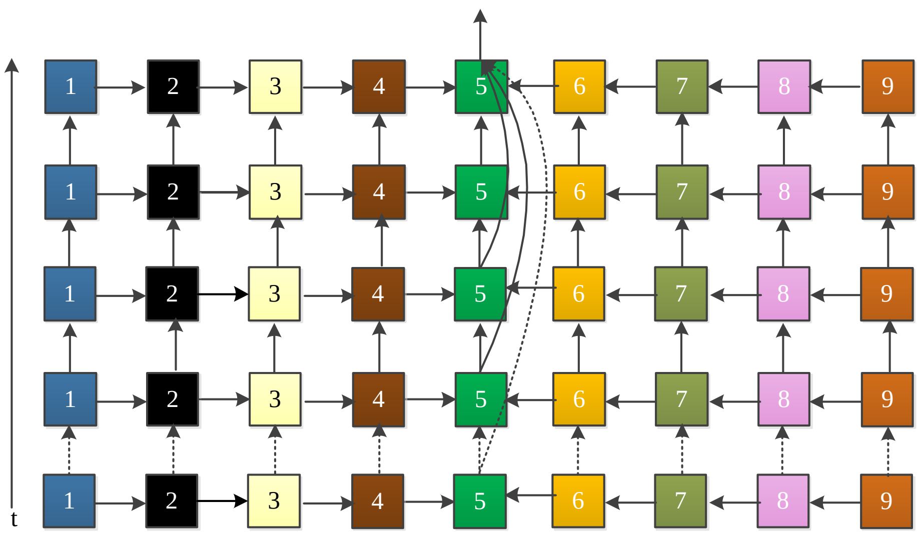
The conventional way to transform the CNN output is,

$$\xi^l(t) = \{\xi^l_{11}(T_0), ..., \xi^l_{MN}(T_0), \xi^l_{11}(T_1), ...,$$
$$\xi^l_{MN}(T_N - 1), \xi^l_{11}(T_N), ..., \xi^l_{MN}(T_N)\}. \tag{4.10}$$

To analyze the spatial correlations within and between the time instants as in Fig. 4.5, $\xi(t)$ in (4.10) changes to,

$$\hat{\xi}^l(t) = \{\xi^l_{11}(T_0), ..., \xi^l_{MN}(T_0), \xi^l_{MN}(T_1), ...,$$
$$\xi^l_{MN}(T_N - 1), \xi^l_{MN}(T_N), ..., \xi^l_{11}(T_N)\}. \tag{4.11}$$

The LSTM can only obtain the previous information in the sequence data. Different from conventional LSTM, Bi-LSTM adds an additional layer of reverse LSTM and simultaneously synthesizes the forward and reverse information of the before pixel points. The reverse layer LSTM is calculated in a similar manner to its forward counterpart, except that the direction is reversed to obtain the subsequent time information. The Bi-LSTM forward hidden vector $\overrightarrow{h}(t)$ and backward hidden vector $\overleftarrow{h}(t)$ are calculated as,



(c) 3D-Bi-LSTM: SM prediction structure

Figure 4.6: Prediction structure for predicting cell 5 in a raster of 9 cells using the 3D-Bi-LSTM. The raster grid is predicted using not just the preceding grid cells (cells 1 to 4) information but future grid cells after cell 5 (cells 6 to 9) which are next in the spatial location from the preceding target cell location along with multiple time instants from both directions.

$$\begin{aligned}
\overrightarrow{h}(t) &= H\left(W_{\xi\overrightarrow{h}}\xi(t) + W_{\overrightarrow{h}\overrightarrow{h}}\overrightarrow{h}(t-1)\right), \\
\overleftarrow{h}(t) &= H\left(W_{\xi\overleftarrow{h}}\xi(t) + W_{\overleftarrow{h}\overleftarrow{h}}\overleftarrow{h}(t+1)\right),
\end{aligned} \tag{4.12}$$

where $h(t)$ denotes the hidden vector, $W_{\xi h}$ is the input-hidden weight matrix, $W_{\overrightarrow{h}\overrightarrow{h}}$ symbolize the forward hidden-hidden weight matrix, $W_{\overleftarrow{h}\overleftarrow{h}}$ represent the reverse hidden-hidden weight matrix, and H is the hidden layer function. For the proposed 3D-Bi-LSTM, the backward hidden vector $\overleftarrow{h}(t)$ in (12) changes to,

$$\begin{aligned}
\hat{\overleftarrow{h}}(t) = H\left(W_{\hat{\xi}^l\overleftarrow{h}}\hat{\xi^l}(t) + W_{\overleftarrow{h}\overleftarrow{h}}\overleftarrow{h}(t+1)\right) \\
\forall\ l \in \{1,2,...,27\},
\end{aligned} \tag{4.13}$$

where $\hat{\xi(t)}$ is given by (4.11). For the proposed 3D-Bi-LSTM, $\overrightarrow{h}(t-1)$ and $\overleftarrow{h}(t+1)$ are rearranged as shown in Fig. 4.6. The output of the proposed Bi-LSTM for SSM prediction can be expressed as,

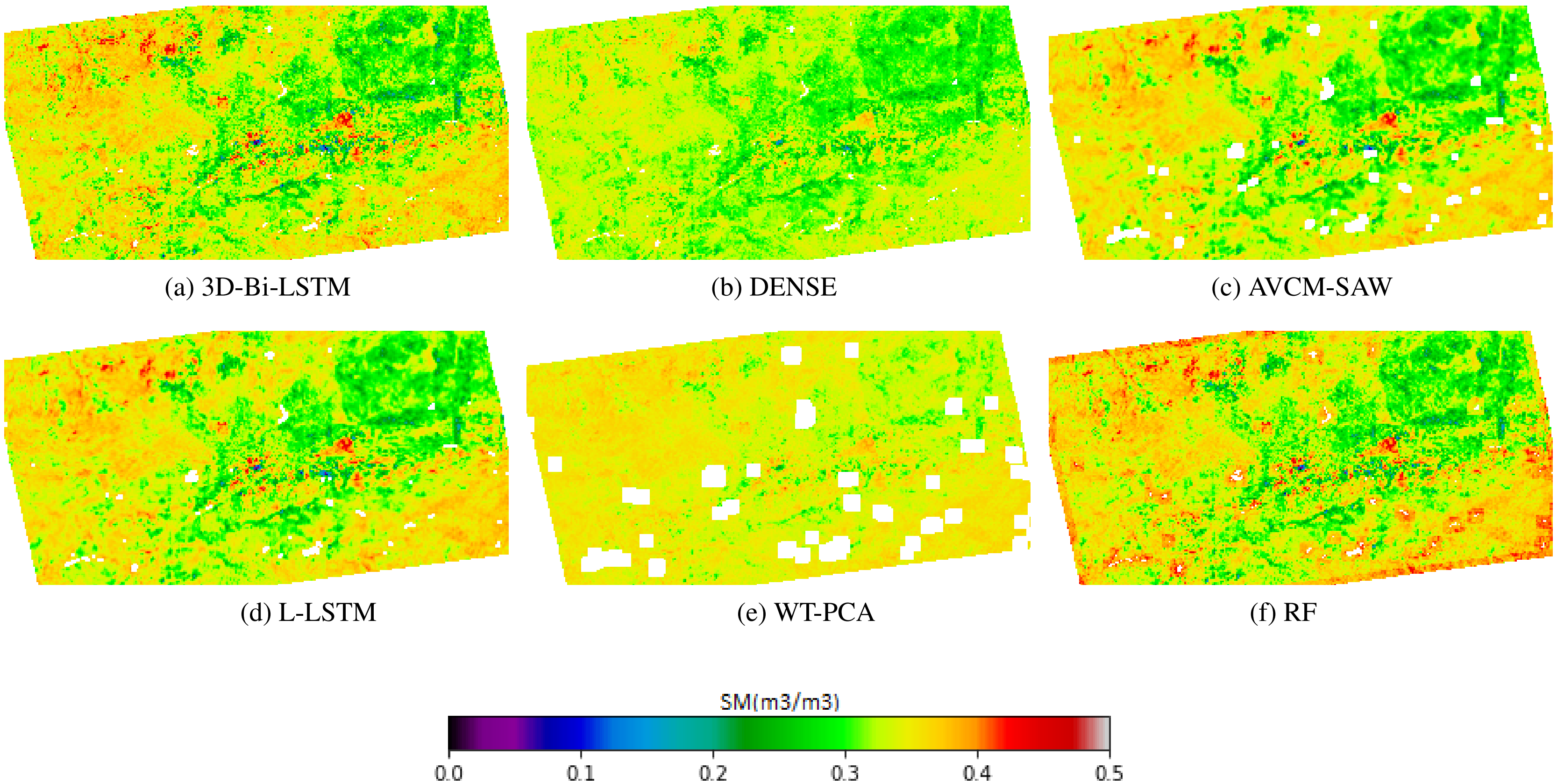
$$\hat{\Gamma(t)} = G\left(W_{\overrightarrow{h}\hat{\Gamma}}\overrightarrow{h}(t) + W_{\hat{\overleftarrow{h}}\hat{\Gamma}}\hat{\overleftarrow{h}}(t) + b_{\hat{\Gamma}}(t)\right), \tag{4.14}$$

where $\hat{\Gamma}(t)$ is the predicted output, $b_{\hat{\Gamma}}$ is the output bias function, G is the output transfer function, $\overrightarrow{h(t)}$ is the forward hidden vector, and $\overleftarrow{h(t)}$ is the backward hidden vector.

Once training is completed and the model is validated, the input at fine resolution (1-km) is fed to the proposed downscaling model to yield output at fine resolution (1-km). The downscaled SSM is validated using ground sensor data. The proposed model is tested in Section IV.

4.4 Results and Discussions

To evaluate the proposed 3D-Bi-LSTM, test is conducted on the SMAP/IN-SITU CORE VALIDATION V001 [226], SOILSCAPE [227], SMAPVEX-16 [228], and SMAPVEX-19 [229] datasets. The downscaling results were compared with five state-of-the-art related works the DENSE [40], AVCM-SAW [41], Long term LSTM (L-LSTM) [43], WT-PCA [42], and random forest (RF). The simulation



(a) 3D-Bi-LSTM (b) DENSE (c) AVCM-SAW

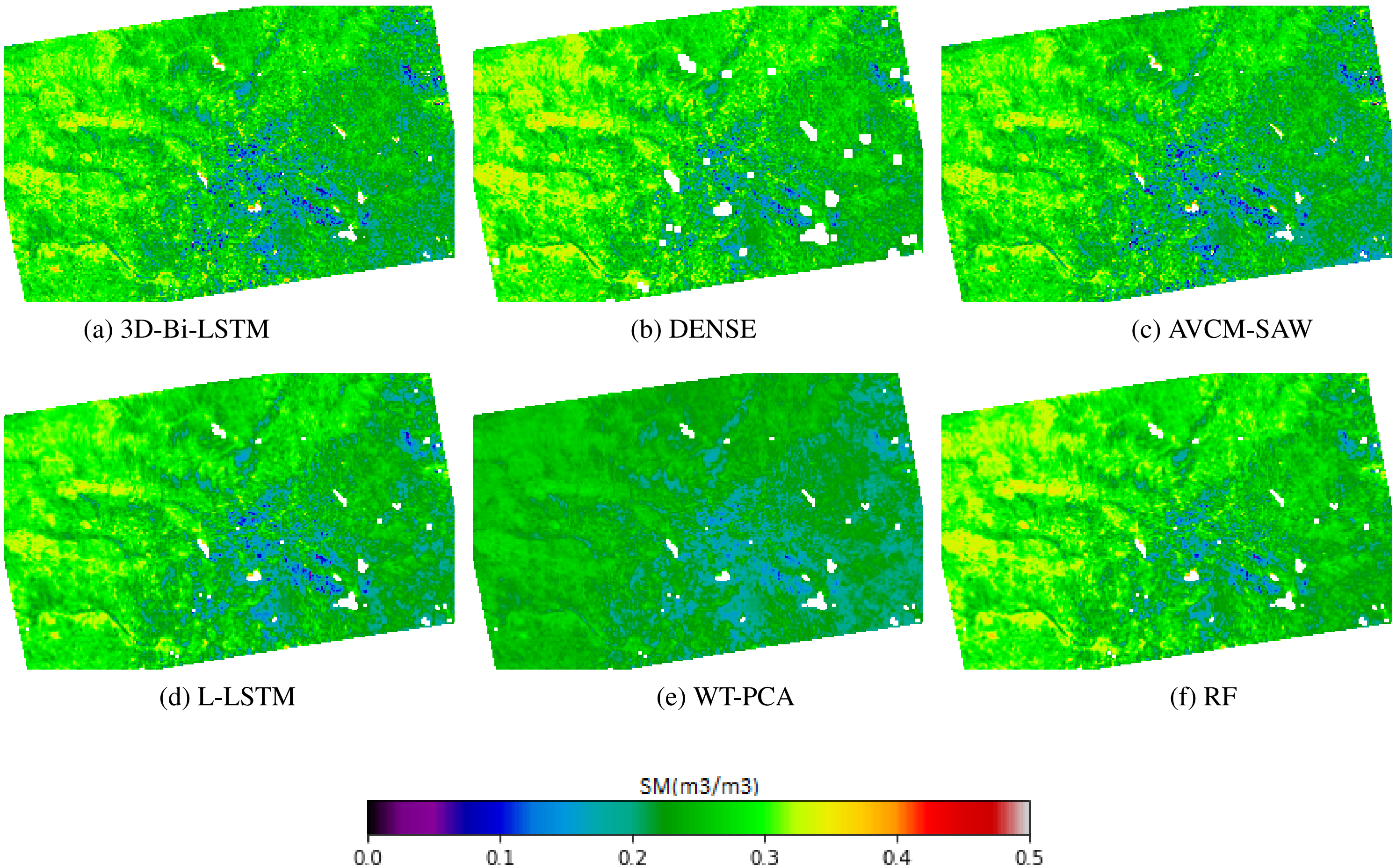
(d) L-LSTM (e) WT-PCA (f) RF

Figure 4.7: Comparison of downscaled soil moisture maps produced by: (a) 3D-Bi-LSTM, (b) DENSE, (c) AVCM-SAW, (d) L-LSTM, (e) WT-PCA, and (f) RF at Remedhus, Spain on DOY of 193, 2016.

results of DENSE and AVCM-SAW validate the use of ML/DL sequential models for downscaling SSM to a finer resolution using the knowledge learned at the coarse resolution, but both DENSE and AVCM-SAW do not provide any comparisons with other ML- or DL-based models. Similarly, WT-PCA uses the wavelet transform, which is a traditional signal processing technique, and the principal component analysis is a traditional ML technique for downscaling SMAP soil moisture. The proposed framework uses CNN and Bi-LSTM for spatio-temporal data analysis for SMAP downscaling. The comparison will provide a benchmark for accuracy comparisons with traditional approaches. Furthermore, to ensure a fair comparison, the same SMIs are provided as input to all competing models.

4.4.1 Visual Inspection of Downscaled Results

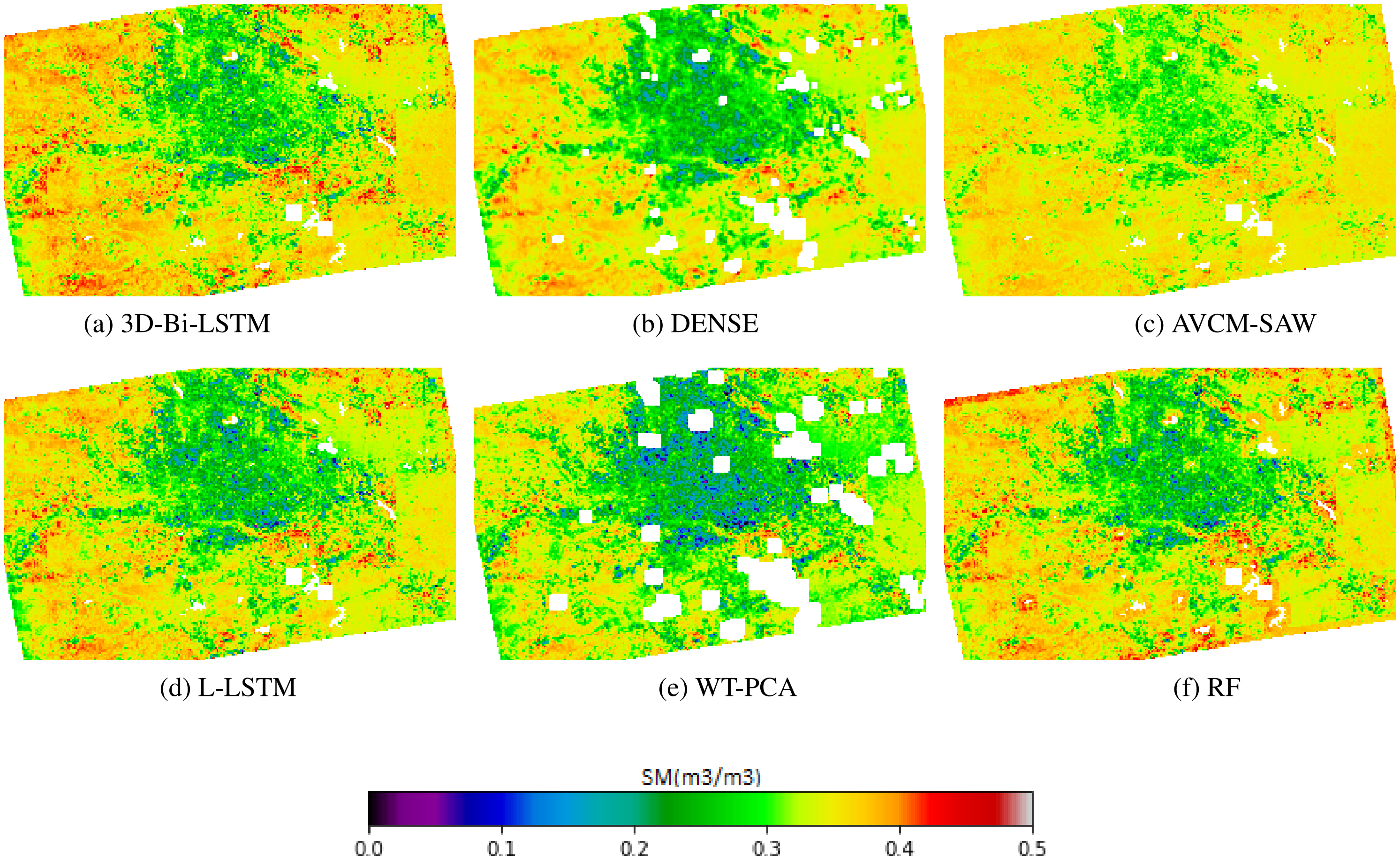
To present the downscaling effect, the results for the image acquired from the proposed 3D-Bi-LSTM, DENSE [40], AVCM-SAW [41], L-LSTM [43], WT-PCA [42] and RF (random forest) are illustrated in Fig. 4.7-4.9. Fig. 4.7 shows the comparison for a very hot and humid day with 30.5^0 C temperature and 41.62% humidity (day-of-the-year (DOY) of 193, 2016) at Remedhus, Spain, while Fig. 4.8 illustrates the cold and dry day with 7^0 C temperature and 18% humidity (DOY of 39, 2016) at Fort Cobb, Oklahoma. Over different climate conditions (Fig. 4.7-4.8), the spectral variability is better represented with less distortion by the proposed 3D-Bi-LSTM than all other models. It is clear that the proposed model is able to better capture the seasonal soil moisture variability than DENSE, AVCM-SAW, L-LSTM, and WT-PCA. Fig. 4.9 illustrates the satellite coverage for a different location (Little Washita, Oklahoma) for the same DOY 193, 2016. Fig. 4.7 and Fig. 4.9 illustrate that more accurate downscaling which is robust to the image contamination introduced by clouds is possible using 3D-Bi-LSTM compared to other methods, which could be due to the inclusion of the third axis (time-axis) information on past SSM and SMI values for SSM prediction. The performance of WT-PCA is not as strong in comparison to the other methods. This could be due to the previously discussed inherent defects of using PCA for FE.



(a) 3D-Bi-LSTM (b) DENSE (c) AVCM-SAW

(d) L-LSTM (e) WT-PCA (f) RF

Figure 4.8: Comparison of downscaled soil moisture maps produced by: (a) 3D-Bi-LSTM, (b) DENSE, (c) AVCM-SAW, (d) L-LSTM, (e) WT-PCA, and (f) RF at Fort Cobb, Oklahoma on DOY of 39, 2016.



(a) 3D-Bi-LSTM

(b) DENSE

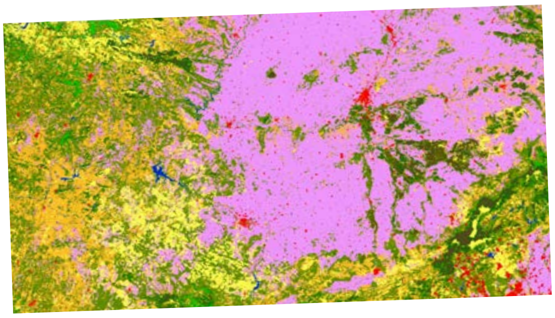
(c) AVCM-SAW

(d) L-LSTM

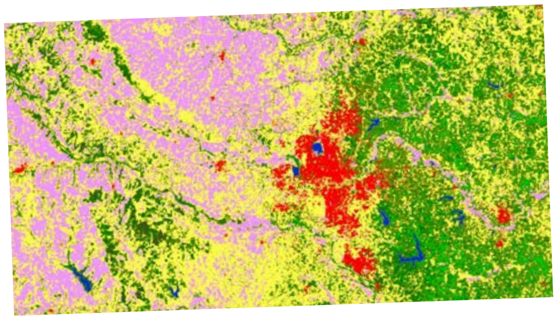
(e) WT-PCA

(f) RF

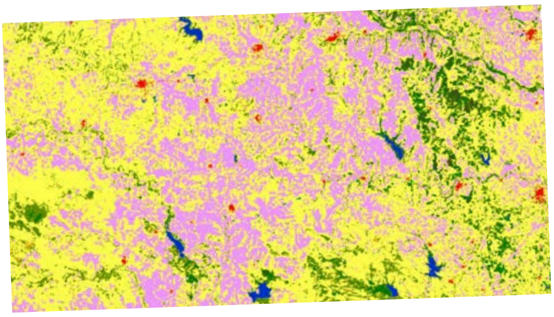
Figure 4.9: Comparison of downscaled soil moisture maps produced by: (a) 3D-Bi-LSTM, (b) DENSE, (c) AVCM-SAW, (d) L-LSTM, (e) WT-PCA, and (f) RF at Little Washita, Oklahoma on DOY of 193, 2016.



(a) Remedhus



(b) Fort Cobb



(c) Little Washita

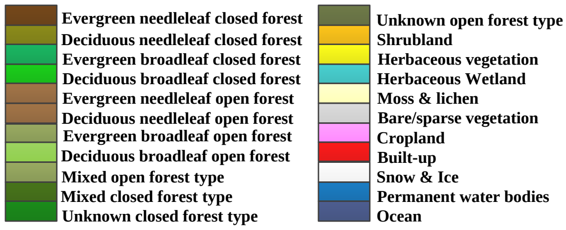
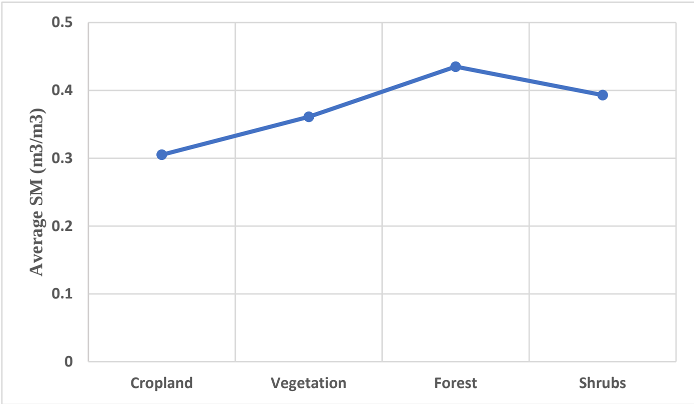


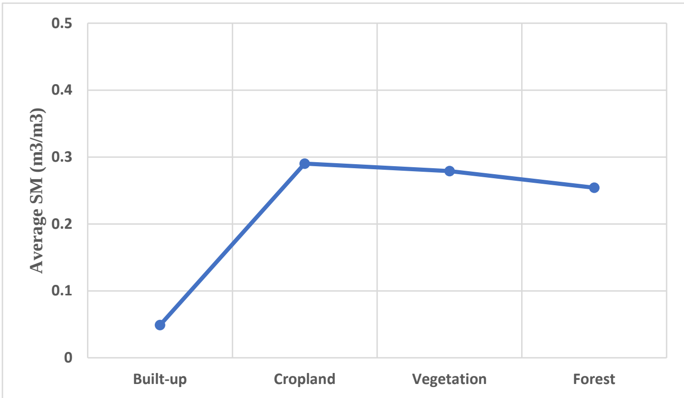
Figure 4.10: Land cover maps of: (a) Remedhus, Spain (b) Fort Cobb, Oklahoma, and (c) Little Washita, Oklahoma.

The spatial variation of SSM is dramatic for the selected locations. As can be visually observed from Fig. 4.7-4.9, L-LSTM and DENSE are not able to capture the spatial variation well, which may be because both methods only implement time-series prediction and the analysis over the spatial domain is not included in the final prediction. In contrast, time-series prediction has reduced the number of blank values in comparison to WT-PCA for L-LSTM and DENSE. Hence, the results emphasize the importance of considering the spatial-time correlations of all SMI spectra and SSM during the downscaling process. The proposed 3D-Bi-LSTM is more robust to contamination introduced by clouds than DENSE, AVCM-SAW, L-LSTM, and WT-PCA.

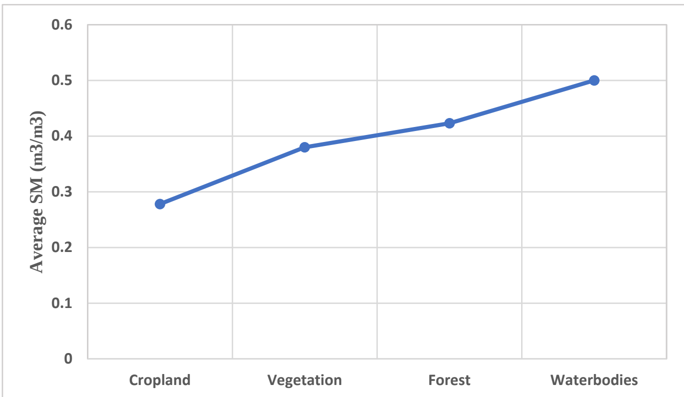
Fig. 4.10 illustrates the land cover maps corresponding to the downscaled SSM maps. The visual comparison of Fig. 4.7-4.9 and Fig. 4.10 show that land cover types such as built-ups are represented by lower intensity moisture spectrum and land cover types like permanent water bodies are represented by higher intensity moisture spectrum. The moisture changes for different land cover types are further analyzed in Fig. 4.11. In Fig. 4.11 the average SSM values at different land covers are plotted.



(a) Remedhus



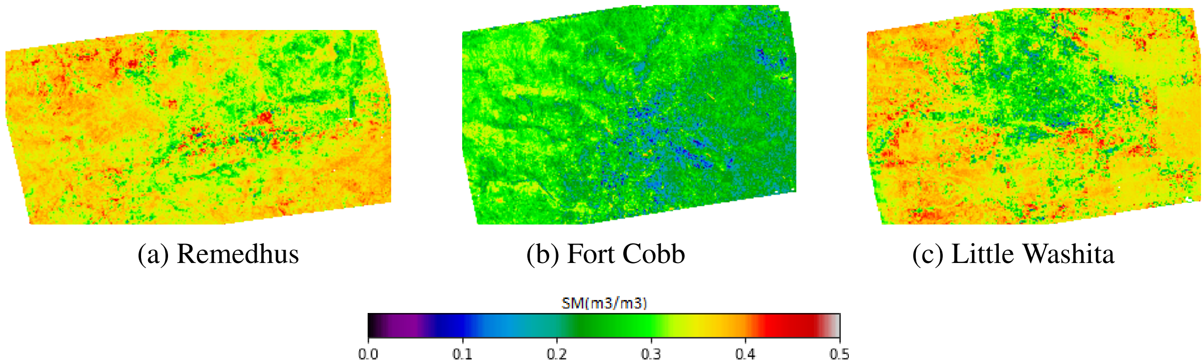
(b) Fort Cobb



(c) Little Washita

Figure 4.11: Average SSM values at different land covers by 3D-Bi-LSTM in: (a) Remedhus, Spain (b) Fort Cobb, Oklahoma, and (c) Little Washita, Oklahoma.

By comparing Fig. 4.11 (a) and Fig. 4.11 (c) with Fig. 4.11 (b) we can observe that except for the croplands, other land cover types reflect the dry season lower moisture values. Croplands normally receive irrigation for proper crop cultivation. This must



(a) Remedhus (b) Fort Cobb (c) Little Washita

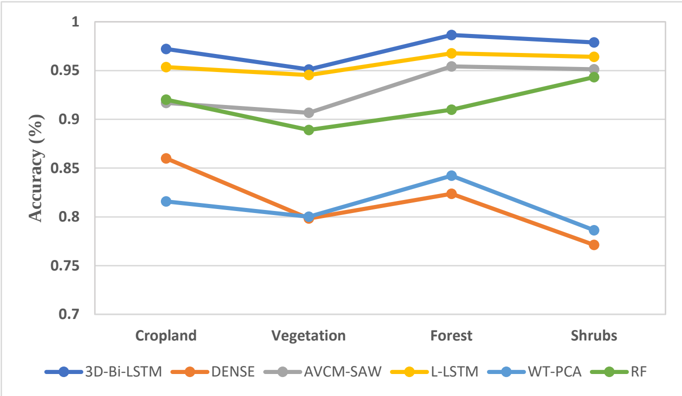
Figure 4.12: Upscaled soil moisture maps at: (a) Remedhus, Spain (b) Fort Cobb, Oklahoma, and (c) Little Washita, Oklahoma.

be the reason that in Fig. 4.11, the croplands have not experienced a fall in soil moisture.

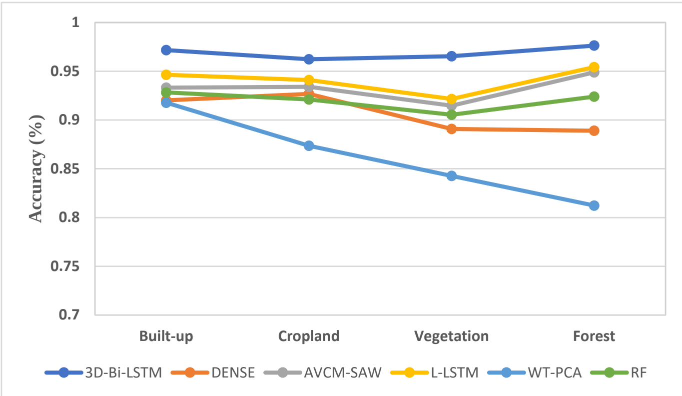
The upscaled SSM maps are illustrated in Fig. 4.12. To analyze the accuracy of similarity of spatial patterns with the 3D-Bi-LSTM downscaled SSM maps, the percentage accuracy in pixel matching is done with the upscaled maps. Fig. 4.13 illustrates the pixel matching percentage accuracy graphs. The average pixel matching accuracy at different land covers in each location is plotted. The average pixel matching accuracy of 3D-Bi-LSTM from the three locations for different land covers illustrate a 0.96 percent accuracy for land cover type built-up, 0.95 percent accuracy for land cover type cropland, 0.92 percent accuracy for land cover type vegetation, 0.94 percent accuracy for land cover type forest, 0.97 percent accuracy for land cover type vegetation shrubs and 0.93 percent accuracy for land cover type vegetation water bodies. Fig. 4.13 clearly illustrates a higher prediction accuracy for 3D-Bi-LSTM across different land cover types in comparison to all other competing models. Hence to summarise the 3D-Bi-LSTM is able to capture the variations in soil moisture well for different land cover types. Furthermore, for each land cover type, the proposed 3D-Bi-LSTM illustrates a higher prediction accuracy in comparison to all competing models.

4.4.2 Validation Results

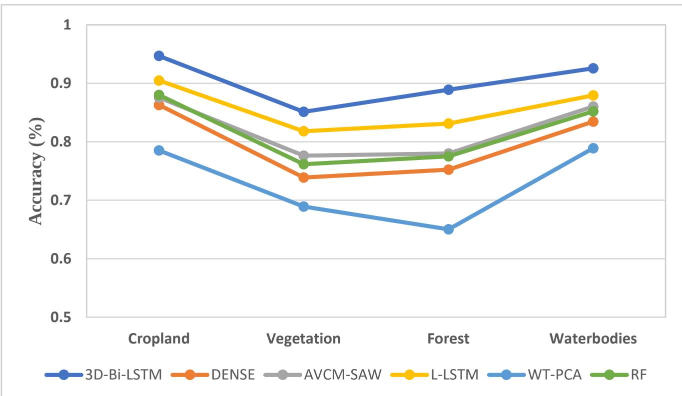
The downscaled SSM is first validated with the area average of the entire target region for the Soilscape network dataset. Fig. 4.14 shows the scatter plots for the downscaled SSM using (a) proposed 3D-Bi-LSTM, (b) DENSE, (c) AVCM-SAW,



(a) Remedhus



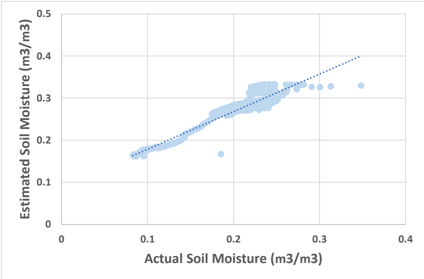
(b) Fort Cobb



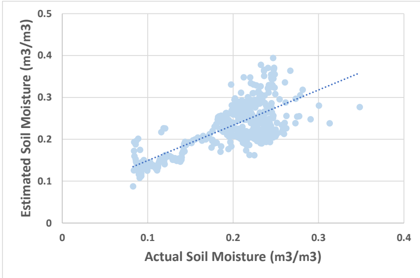
(c) Little Washita

Figure 4.13: Pixel matching percentage accuracy between the 3D-Bi-LSTM downscaled and upscaled soil moisture maps at different land covers in: (a) Remedhus, Spain (b) Fort Cobb, Oklahoma, and (c) Little Washita, Oklahoma.

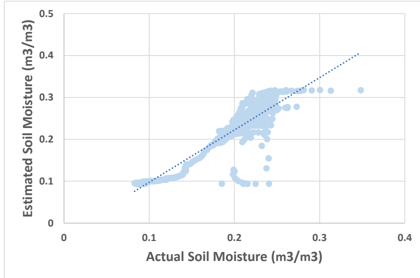
(d) L-LSTM, (e) WT-PCA, and (f) RF against the in-situ SSM measurements on the Soilscape network dataset. As can be observed from these scatter plots, the SSM values downscaled by the proposed 3D-Bi-LSTM agree well with their in-situ



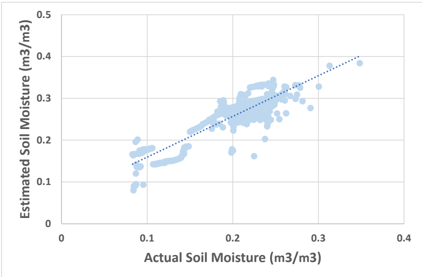
(a) 3D-Bi-LSTM



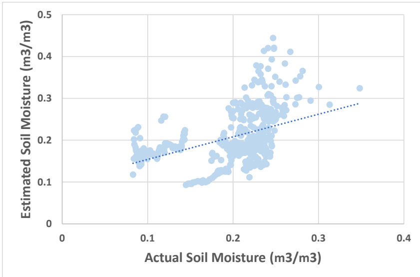
(b) DENSE



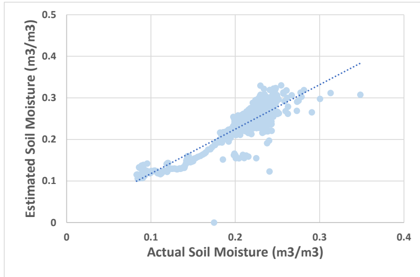
(c) AVCM-SAW



(d) L-LSTM



(e) WT-PCA

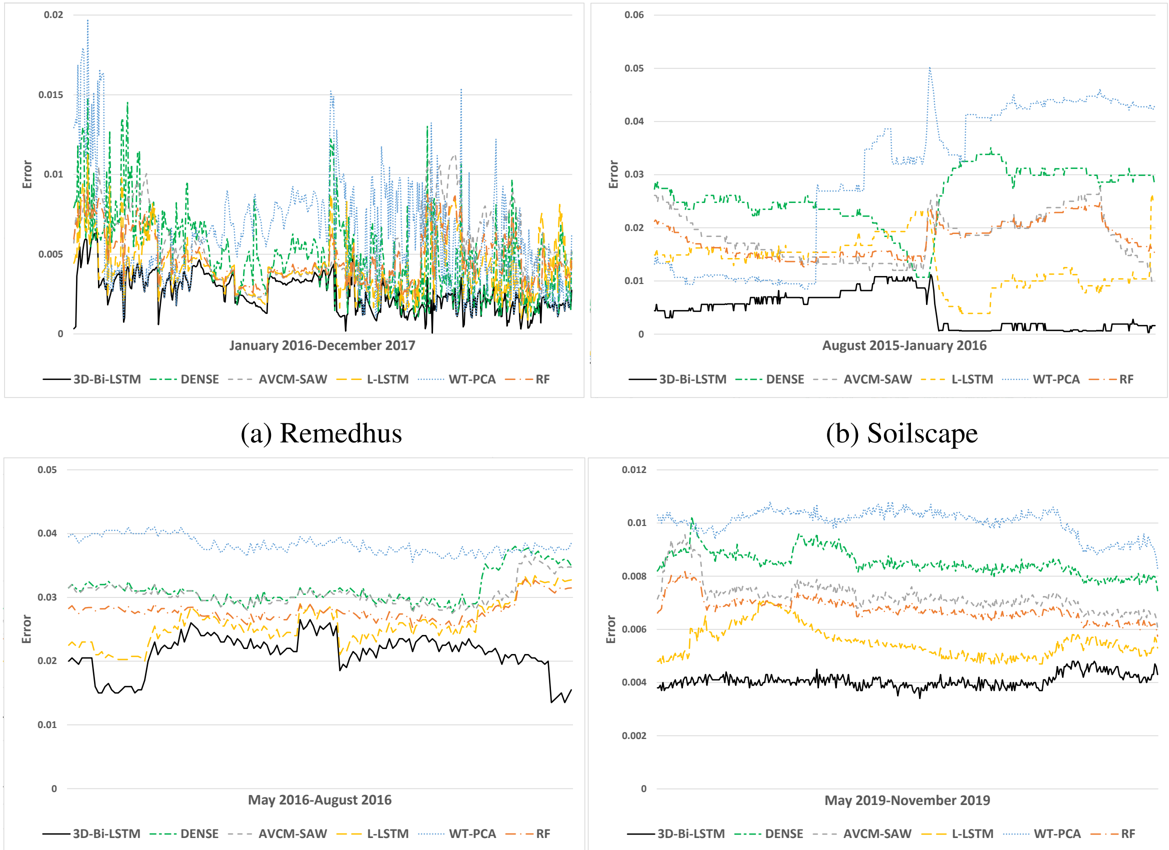


(f) RF

Figure 4.14: Scatter plots of downscaled area average SSM values against in-situ values of the SOILSCAPE dataset using: (a) 3D-Bi-LSTM, (b) DENSE [40], (c) AVCM-SAW [41], (d) L-LSTM [43], (e) WT-PCA [42], and (f) RF.

counterparts when compared to those downscaled by AVCM-SAW, RF, WT-PCA, DENSE, and L-LSTM. For the general performance, it seems evident that the 3D-Bi-LSTM downscaled SSM accurately captures the temporal variations of SSM over the Soilscape network dataset in the target region and maintains a reliable accuracy in most circumstances when compared to the other baseline methods. We can observe a generally higher correlation between the predicted and field-measured SSM values for 3D-Bi-LSTM than all the other comparison approaches. The accuracy of the daily SSM prediction is graphically compared using average error graphs. Fig. 4.15 graphically compares the average daily downscaling prediction errors achieved by 3D-Bi-LSTM as opposed to DENSE, AVCM-SAW, L-LSTM, WT-PCA, and RF. These results indicate a higher overall prediction accuracy for 3D-Bi-LSTM.

Quantitative measurements are considered to further demonstrate the effectiveness of the proposed 3D-Bi-LSTM on the target datasets. The correlation coefficient (R), root-mean-squared error (RMSE), unbiased RMSE (ubRMSE), and bias are calculated as the statistical metrics to analyze the reliability of downscaled results. The RMSE



(a) Remedhus

(b) Soilscape

(c) SMAPVEX16

(d) SMAPVEX19

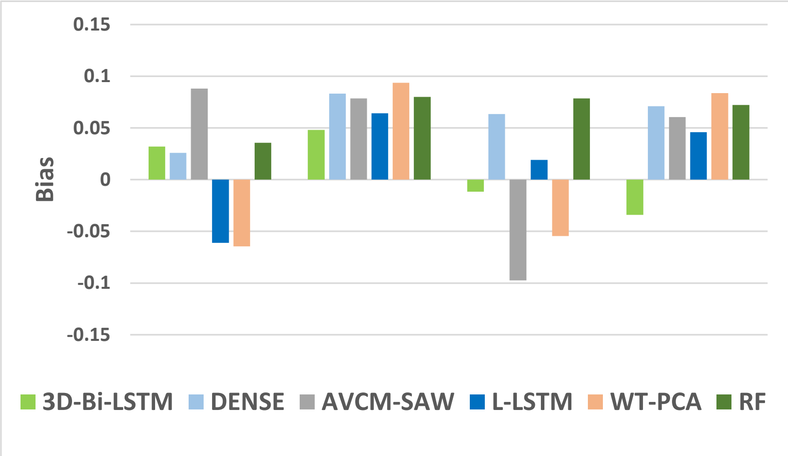
Figure 4.15: Time series comparison of the average error between the in-situ SSM and the SSM predicted by 3D-Bi-LSTM, DENSE, AVCM-SAW, L-LSTM, WT-PCA [19] and RF on: (a) the in-situ core validation set (Remedhus), (b) Soilscape, (c) SMAPVEX16, and (d) SMAPVEX19 datasets.

Table 4.1: Comparison of the Average Prediction Accuracy of Each Model

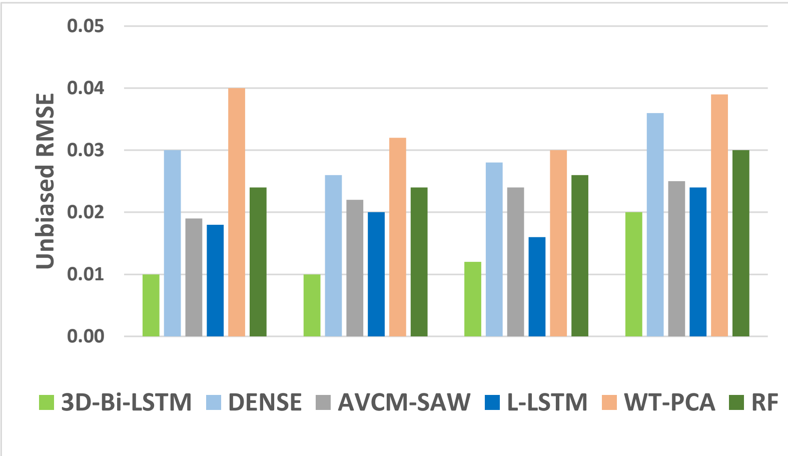
SMAP/IN-SITU CORE VALIDATION Dataset [35]						
	3D-Bi-LSTM	DENSE	AVCM-SAW	L-LSTM	WT-PCA	RF
RMSE	**0.041**	0.092	0.072	0.063	0.099	0.075
SOILSCAPE Dataset [36]						
	3D-Bi-LSTM	DENSE	AVCM-SAW	L-LSTM	WT-PCA	RF
RMSE	**0.026**	0.059	0.040	0.035	0.073	0.043
SMAPVEX16 Dataset [37]						
	3D-Bi-LSTM	DENSE	AVCM-SAW	L-LSTM	WT-PCA	RF
RMSE	**0.016**	0.054	0.044	0.031	0.063	0.047
SMAPVEX19-21 Dataset [38]						
	3D-Bi-LSTM	DENSE	AVCM-SAW	L-LSTM	WT-PCA	RF
RMSE	**0.034**	0.084	0.061	0.042	0.092	0.064
Avg	**0.029**	0.072	0.054	.047	0.082	0.057
SD	0.002	0.003	0.002	0.002	0.004	0.002

Note: SD is the standard deviation.

evaluation is conducted for all six models. Table 4.1 provides the RMSE results for all six comparison models on all four datasets. Fig. 4.16 provides the bias and



(a) Bais



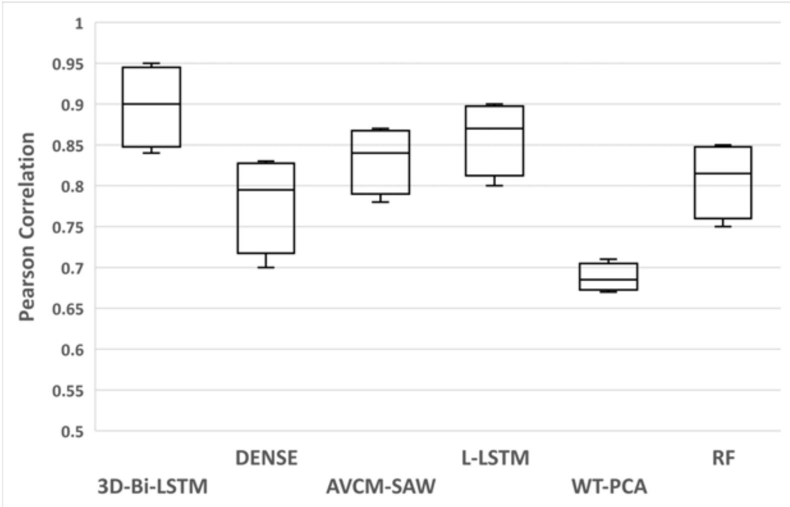
(b) Unbiased RMSE

Figure 4.16: Bias and unbiased RMSE values between: (a) 3D-Bi-LSTM, (b) DENSE [40], (c) AVCM-SAW [41], (d) L-LSTM [43], and (e) WT-PCA [42] and RF.

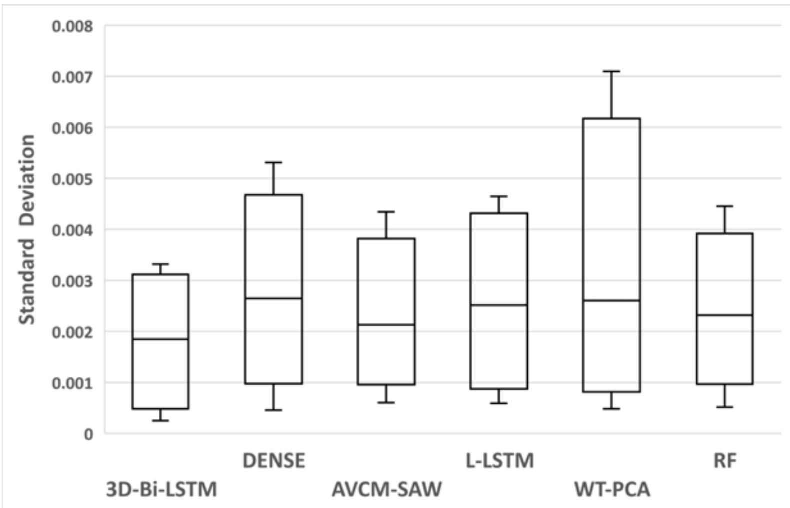
Table 4.2: Time Complexity Comparison of the Models

Time	**3D-Bi-LSTM**	**DENSE**	**AVCM-SAW**	**L-LSTM**	**WT-PCA**	RF
Train (hr)	70.4	31.5	22.3	47.3	19.8	27.9
Test (s)	1.64	1.57	1.12	1.15	1.41	1.06

unbiased RMSE for the six models over all four datasets. As can be seen from Table 4.1 and Fig. 4.16, the average RMSE, bias, and unbiased RMSE values of the The 3D-Bi-LSTM model is the smallest among all the comparison approaches, indicating the proposed 3D-Bi-LSTM model has higher SSM prediction accuracy. Table 4.2 provides the training and testing times of the six comparison models. Generally, deep



(a) Pearson correlation



(b) Standard deviation

Figure 4.17: Comparison of Pearson correlation and standard deviation among the SSM values predicted by 3D-Bi-LSTM, DENSE [40], AVCM-SAW [41], L-LSTM [43], WT-PCA [42] and RF.

learning models take longer to train than machine learning models. This is reflected in the table. However, the testing times of all six models are close to each other. Although the proposed 3D-Bi-LSTM requires a longer training time, it achieves a better prediction accuracy than the other five competing models. Further comparison of the models using the Pearson correlation and standard deviation is conducted. The closer the Pearson correlation is to $+1$, the better the accuracy. A low standard deviation indicates that the values tend to be close to the mean (the expected value) of the dataset, while a high standard deviation indicates that the values are spread out over a wider range. Fig. 4.17 gives the Pearson correlation and standard deviation of all six models on the four datasets. For the average temporal correlation over all the four datasets calculated using the Pearson correlation, the 3D-Bi-LSTM has the highest average correlation. Moreover, the lower standard deviation of 3D-Bi-LSTM in comparison to the competing five models indicates that the 3D-Bi-LSTM predicted SSM values are consistently closer to the actual SSM values than the L-LSTM, DENSE, AVCM-SAW, WT-PCA, and RF.

4.4.3 Explanatory Methods to Interpret 3D-Bi-LSTM

4.4.3.1 Ablation Studies of 3D-Bi-LSTM

In this section, the individual input and model components of the proposed 3D-Bi-LSTM are assessed. The 3D-Bi-LSTM was tested first by removing a set of inputs. The theoretical SSM prediction is mainly dependent on the drought index, and vegetation index. To test this, the DI and VI are removed, and check their effects on the proposed 3D-Bi-LSTM model outputs. In addition, the comparison model AVCM-SAW uses only LST and NDVI from the SMIs in Table 4.3. To check their effects LST and NDVI were removed from the set of inputs. Next, to validate the performance of individual components of the network model, the CNN module for feature selection is removed and directly feeds all the SMIs to 3D-Bi-LSTM.

Table 4.3: Comparison of Ablation Studies

Approach	1	2	3	4	5	6	7
	3D-Bi-LSTM	3D-Bi-LSTM-DI	3D-Bi-LSTM-VI	3D-Bi-LSTM-(NDVI+LST)	3D-Bi-LSTM-(DI+NDVI)	3D-Bi-LSTM-(VI+LST)	3D-Bi-LSTM-FE
RMSE	**0.029**	0.040	0.037	0.033	0.038	0.043	0.057
SD	0.002	0.002	0.002	0.002	0.002	0.002	0.003

Note: SD is the standard deviation.

For testing individual inputs and modules of proposed 3D-Bi-LSTM, seven prediction approaches in Table 4.3 are used: 1) 3D-Bi-LSTM: performance of the 3D-Bi-LSTM model with all the inputs; 2) 3D-Bi-LSTM-DI: performance of 3D-Bi-LSTM without DI; 3) 3D-Bi-LSTM-VI: performance of 3D-Bi-LSTM without VI; 4) 3D-Bi-LSTM-(NDVI+LST): performance of 3D-Bi-LSTM without NDVI and LST; 5) 3D-Bi-LSTM-(DI+NDVI): performance of 3D-Bi-LSTM without DI and NDVI; 6) 3D-Bi-LSTM-(VI+LST): performance of 3D-Bi-LSTM without VI and LST; 7) 3D-Bi-LSTM-FE: performance of 3D-Bi-LSTM without feature selection.

Approach 4 removes NDVI and LST from the set of inputs. This analysis is conducted because it is suggested in AVCM [41] that NDVI and LST have a strong relationship with SSM. The higher RMSE achieved by 3D-Bi-LSTM-(NDVI+LST) over the 3D-Bi-LSTM with all the inputs suggests that the use of LST and NDVI as SMIs does enhance the performance of SSM prediction. In approach 6, all drought indices including LST are removed and only one VI (NDVI) is removed. In approach 5, all vegetation indices are removed and only one drought index (LST) is removed. Approach 6 has a higher RMSE than approach 5, while approach 2 has a higher RMSE than approach 3. The results of approaches 2, 3, 5, 6 imply that the drought indices play a significant role in SSM prediction.

In approach 7, area-specific feature extraction based on the correlation between SMI and SSM is removed. The higher RMSE achieved by approach 7 than approach 1 suggests that feature extraction based on the spectral correlation within a spatial area enhances the performance of 3D-Bi-LSTM. So these results demonstrate that the NDVI and LST are important for SSM prediction and feature extraction through the covariance-adaptive CNN enhances the performance of the proposed 3D-Bi-LSTM model.

4.4.3.2 Transferability Assessment of 3D-Bi-LSTM

Table 4.4: Overview of the Eight Test Sites

Dataset	REMEDHUS	Carman	Walnut Gulch	Little Washita	Fort Cobb	Soilscape	SMAPVEX16	SMAPVEX19
Land cover	Croplands	Croplands	Shrub open	Grasslands	Grasslands	Croplands	Croplands	Forest
Climate	Temperate	Cold	Arid	Temperate	Temperate	Tropical	Temperate	Temperate
Location	Spain	Canada	USA (Arizona)	USA (Oklahoma)	USA (Oklahoma)	USA (California)	Canada	USA (New York)

The model's transferability is evaluated on data from eight geographically distinct locations. Table 4.4 gives an overview of these eight locations. Twelve distinct groups of the training and testing datasets from the NASA earthdata (https://search.earthdata.nasa.gov/search) by skipping one location (eight cases) and two locations (four cases) in training. The test cases for skipping one location in training are: Case-1) SoilScape removed from training, and used for testing transferability; Case-2) SMAPVEX16 removed from training, and used for testing transferability; Case-3) REMEDHUS removed from training and used for testing transferability; Case-4)

Table 4.5: Model Transferability Assessment using RMSE

Model \Test sites	Soilscape	SMAPVEX16	REMEDHUS	Carman	Walnut Gulch	Little Washita	Fort Cobb	SMAPVEX19
3D-Bi-LSTM	0.0263	0.0341	0.0114	0.0417	0.0568	0.0453	0.0477	0.0339

Skipping one location								
Case name	**Soilscape**	**SMAPVEX16**	**REMEDHUS**	**Carman**	**Walnut Gulch**	**Little Washita**	**Fort Cobb**	**SMAPVEX19**
3D-Bi-LSTM-(Soilscape)	0.0314	0.0361	0.0163	0.0470	0.0604	0.0482	0.0496	0.0374
3D-Bi-LSTM-(SMAPVEX16)	0.0286	0.0379	0.0144	0.0512	0.0590	0.0504	0.0529	0.0407
3D-Bi-LSTM-(REMEDHUS)	0.0268	0.0345	0.0197	0.0561	0.0580	0.0554	0.0540	0.0376
3D-Bi-LSTM-(Carman)	0.0296	0.0372	0.0179	0.0657	0.0632	0.0505	0.0562	0.0403
3D-Bi-LSTM-(Walnut Gulch)	0.0271	0.0364	0.0134	0.0429	0.0975	0.0470	0.0493	0.0357
3D-Bi-LSTM-(Little Washita)	0.0298	0.0352	0.0154	0.0450	0.0594	0.0630	0.0575	0.0352
3D-Bi-LSTM-(Fort Cobb)	0.0268	0.0350	0.0185	0.0455	0.0580	0.0602	0.0673	0.0374
3D-Bi-LSTM-(SMAPVEX19)	0.0280	0.0368	0.0169	0.0442	0.0613	0.0461	0.0505	0.0551
Max. percentage performance drop	**19.39%**	**11.14%**	**72.81%**	**57.55%**	**71.66%**	**39.07%**	**41.09%**	**62.54%**
Location overall model transferability rank	**2**	**1**	**8**	**5**	**7**	**3**	**4**	**6**

Skipping one location			
Overall model transferability	Good: 81.3%	Moderate: 15.6%	Poor: 3.1%

Skipping two locations								
Case name	**Soilscape**	**SMAPVEX16**	**REMEDHUS**	**Carman**	**Walnut Gulch**	**Little Washita**	**Fort Cobb**	**SMAPVEX19**
3D-Bi-LSTM-(REMEDHUS+Soilsacpe)	0.0322	0.0366	0.0208	0.0565	0.0617	0.0570	0.0558	0.0391
3D-Bi-LSTM-(SMAPVEX16+ Carman)	0.0298	0.0407	0.0186	0.0687	0.0641	0.0519	0.0566	0.0415
3D-Bi-LSTM-(Little Washita+ Fort Cobb)	0.0305	0.0357	0.0187	0.0462	0.0602	0.0719	0.0763	0.0377
3D-Bi-LSTM-(Walnut Gulch+SMAPVEX19)	0.0287	0.0376	0.0170	0.0445	0.0987	0.0479	0.0518	0.0572
Max. percentage performance drop	**22.43%**	**19.35%**	**82.46%**	**64.75%**	**73.76%**	**58.71%**	**59.96%**	**68.73%**
Location overall model transferability rank	**2**	**1**	**8**	**5**	**7**	**3**	**4**	**6**

Skipping two locations			
Overall model transferability	Good: 68.7%	Moderate: 21.9%	Poor: 9.4%

Note: The circled RMSE values are the maximum value for each site, and the corresponding maximum percentage performance drop is in bold font in the table. Also, the 3D-Bi-LSTM's transferability for eight locations is ranked based on the overall transferability test outcome. To categorize transferability as good (green), moderate (yellow), and poor (red), the increase in the root-mean-square error (RMSE) in the range of 0% to 100% is divided into three equal ranges. Transferability is said to be good, moderate, and poor if the RMSE increase from the actual 3D-Bi-LSTM is $< 33.33\%$, between $\geq 33.33\%$ and $< 66.67\%$, and $\geq 66.67\%$. The 3D-Bi-LSTM's overall transferability for skipping one/two locations is illustrated at the bottom of the table.

Carman removed from training, and used for testing transferability; Case-5) Walnut Gulch removed from training, and used for testing transferability; Case-6) Little Washita removed from training and used for testing transferability; Case-7) Fort Cobb removed from training, and used for testing transferability; Case-8) SMAPVEX19 removed from training, and used for testing transferability. For skipping two locations in training, the test cases are classified as: Case-1) REMEDHUS and SoilScape removed from training, and used for testing transferability; Case-2) SMAPVEX16 and Carman removed from training and used for testing transferability; Case-3) Little Washita and Fort Cobb removed from training, and used for testing transferability; Case-4) Walnut Gulch and SMAPVEX19 removed from training and used for testing transferability. In each of the above cases, 20% of training data from each location dataset is held back from the training for testing transferability in time.

For the transferability assessment by skipping one location from training in Table 4.5, generally, 3D-Bi-LSTM shows good transferability on the 20% of training locations data which is unused for training from fully unskipped locations (green cells). In the fully skipped one training location, the 3D-Bi-LSTM mainly shows moderate transferability (yellow cells). The highest transferability from single skipped location cases is observed for SMAPVEX16 and Soilscape sites. Table 4.4 shows that SMAPVEX16 and Soilscape have similarities with the other sites. The lowest transferability in REMEDHUS and Walnut Gulch for the 3D-Bi-LSTM is noted among the one skipped location cases. Table 4.4 illustrates that REMEDHUS and Walnut Gulch are very distinct from the rest of the sites. Hence the skipping of the REMEDHUS and Walnut Gulch site resulted in poor predictions by the model at these locations. For the skipping of one location, the proposed model shows 81.3% good transferability, 15.6% moderate transferability, and 3.1% poor transferability. Table 4.5 also displays the transferability assessment of the proposed model by skipping two locations for testing transferability. For skipping two locations, the skipping of Little Washita and Fort Cobb showed an interesting observation. Skipping them together resulted in a significant performance drop than skipping them individually ($\approx$ 20%). Table 4.5 reveals that Little Washita and Fort Cobb are quite similar sites. Hence the model might have transferred learning from Little Washita or Fort Cobb when only one of these two locations was skipped from training. Although REMEDHUS and Walnut Gulch are still the worst cases under the two locations skipped category, their performance decreased by only 9.65% and 2.1% respectively. This is because both the sites are quite distinct, and removing other sites does not have much effect on the predictions at these locations. For the overall transferability tests conducted, the proposed model shows 68.7% good transferability, 21.9% moderate transferability, and 9.4% poor transferability. For the overall transferability tests conducted, the model's transferability for the site SMAPVEX16 is ranked 1st, Soilscape is ranked 2nd, Little Washita is ranked 3rd, Fort Cobb is ranked 4th, Carman is ranked 5th, SMAOVEX19 is ranked 6th, Walnut Gulch is ranked 7th, and REMEDHUS is ranked 8th.

4.4.3.3 Feature Importance

The feature importance study is conducted using SHAP (SHapley Additive exPlanations) [235]. The interpretation of deep learning models is arduous due to their complex "black box" architecture. However, the shapley additive explanations are highly robust in interpreting deep neural networks. SHAP values calculate the importance of a feature by comparing what a model predicts with and without this feature. Since the increasing order in which a model sees each feature can affect its predictions, feature importance analysis is done in every possible order, so that the features are fairly compared. The SHAP summary plots in Fig. 4.18 are used to interpret the importance of each feature. The y-axis indicates the variable name in the increasing order of importance from top to bottom. The color represents the feature value (red high, blue low). All features are continuous and are vertically sorted by their average impact on the predictions.

Fig. 4.18 indicates the impact of each input feature on the model output. From Fig. 4.18 it is clear that LST is the most prominent feature, with low values of LST associated with high SSM predictions. This interpretation is compatible with the scientific knowledge that LST and SSM have a negative correlation [230]. Additionally, in Fig. 4.18 NDVI and EVI show a positive correlation with SSM. And NDVI shows a higher correlation with soil moisture than EVI. This result matches existing studies that NDVI shows a higher positive correlation with SSM than EVI [123]. As indicated in Fig. 4.18, both LST and NDVI play significant contributions to the model output. Furthermore, Fig. 4.18 illustrates that the high value of NDTI is associated with low values of SSM. This outcome is too in tandem with existing research that

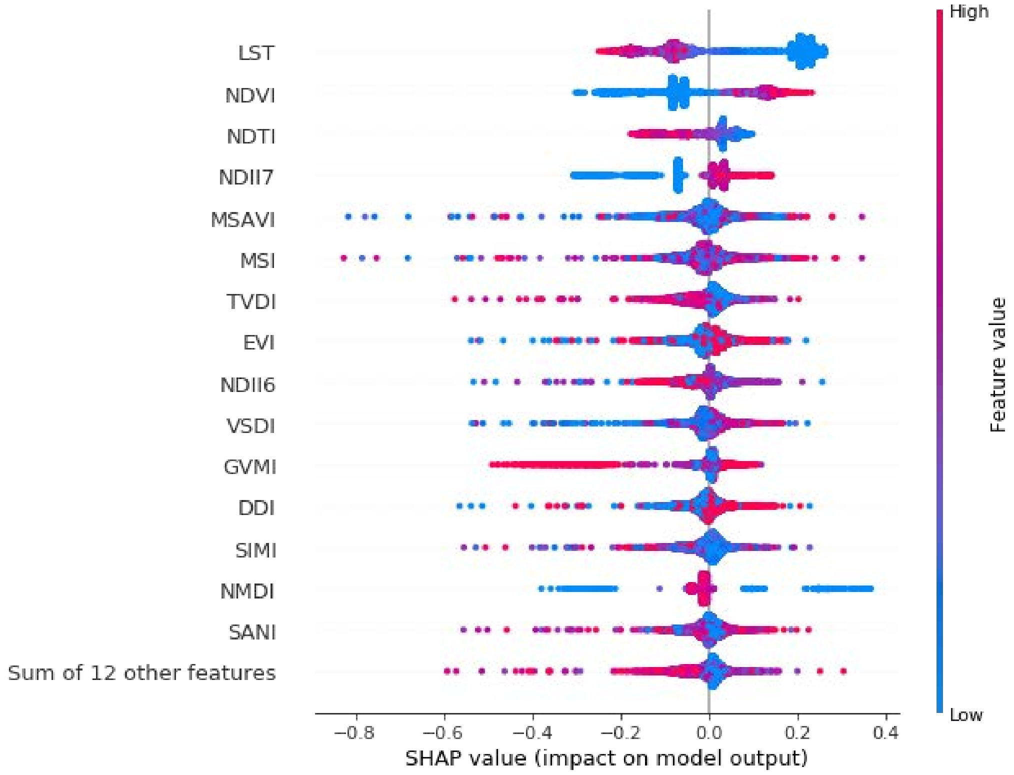


Figure 4.18: SHAP summary plots for the proposed 3D-Bi-LSTM network.

high tillage on soils increases undesirable changes in soil structure, and lowers water infiltration and moisture-holding capacity [231]. The MSI, on the other hand, shows a mixed pattern. This result aligns with existing research-based experimental evaluation studies, which indicate that the MSI values showed a mixed correlation with the measured soil moisture [232]. GVMI is shown as a lesser important feature. This model calculation might be because GVMI is suitable for retrieving vegetation water content when the LAI is equal to or greater than 2 [233]. Moreover, GVMI high values are associated both with high and low SSM predictions. This prediction could be because soil moisture retrievals through GVMI are related to vegetation water content, and vegetation water content is crop-specific [234]. Hence, the type of relationship it has with soil moisture may vary based on the vegetation types in the target area. It is interesting to note that the proposed 3D-Bi-LSTM model has been able to understand these real-world relationships between the input features and soil moisture through the spatio-temporal changes of different spectra. And SHAP has managed to very clearly illustrate how such changes in inputs affect the model output.

4.5 Conclusion

In this chapter, a 3D bi-directional LSTM (3D-Bi-LSTM) model for surface soil moisture (SSM) downscaling was proposed for spatially distributed SM information. The daily MODIS and SMAP satellite data are used as input to the model. The proposed 3D-Bi-LSTM model predicts fine-scale (1-km) SSM based on the satellite SSM and SMIs acquired at a coarse scale. A downscaling framework was designed using a covariance-adaptive CNN and the 3D-Bi-LSTM model. The covariance-adaptive CNN is used for adaptive spatial feature extraction for selected locations from each daily SMIs. The 3D-Bi-LSTM performs spatio-temporal spectral sequence analysis for SSM prediction. The evaluation was conducted on four open-source SSM datasets. Results showed that the proposed 3D-Bi-LSTM model is capable of predicting SSM closer to the ground SSM measurements and has less spectral information loss than the other state-of-the-art ML/DL SSM downscaling models tested. The efficacy of the proposed model in predicting SSM was further validated in terms of ablation studies, transferability assessment, feature studies, Pearson correlation, standard deviation, RMSE, and unbiased RMSE. Hence it can be concluded that there is a significant improvement achieved by the proposed 3D-Bi-LSTM model to downscale SSM in comparison to recent state-of-the-art ML/DL downscaling models in the literature.

In conclusion, this chapter addresses the second research problem of SSM downscaling that was presented in Section 1.2 and provides the second original contribution listed in Section 1.3. In the next chapter, I derive RZSM for SM depth distribution information, another critical input required for taking decisions for precision irrigation.

Chapter 5

Hybrid Transformer Network for Root Zone Soil Moisture Estimation for Decision Support in Precision Irrigation

This chapter contains materials included in the following manuscript, which has been submitted to Agricultural Water Management Journal:

[3] **Neethu Madhukumar**, Eric Wang, and Wei Xiang,"Hybrid Transformer Network for Root Zone Soil Moisture Estimation for Decision Support in Precision Irrigation", *Agricultural Water Management*, under review (Impact Factor: 6.611, h-index: 65, Journal rank (Agronomy & crop science): 7).

5.1 Introduction

The Root zone soil moisture (RZSM) is an indicator of the vegetation drought stress and the crop water demand, which are valuable information for precision irrigation (PI) [67]. In-situ sensors installed at root zone-specific depths can provide direct field RZSM measurements. However, the vast installation of sensors across the field at the subsurface for RZSM measurements is not economically viable. Current satellite missions can provide kilometers of distributed horizontal spatial details, but only centimeters of vertical (depth) SM information [25], [238]. This vertical resolution can only provide knowledge of the surface soil moisture (SSM) layer, not the RZSM layer.

The limitations in direct RZSM measurements are predominantly addressed by indirect estimation through analytical methods using theoretical or empirical models to obtain fine-scale RZSM [239], [240], [241]. These analytical methods use the relationship between environmental variables, such as precipitation, temperature, soil, and plant characteristics (PTSPc) which determines the RZSM state for indirect estimations [46]. However, these methods require a target location-based optimization to improve the model's RZSM simulation accuracy [242]. Adoption of data-based methods can help remove these location-based optimization constraints as they are adaptive to data [243], [72], [73]. Among data-based methods, deep learning (DL) is becoming increasingly popular due to its deep architectures, which can extract accurate information from big environmental data [29], [30].

The application of DL in the RZSM field is still in the early stages [19], [244], [128]. However, RZSM is SSM diffused to lower soil layers [67], [65]. Therefore, the more advanced DL-based techniques for detailed SSM information such as performing downscaling (increasing spatial/temporal details) by correlating the SSM with auxiliary SSM information available at target location [43], [37], [35], [250] can be applied to RZSM. The environmental variables controlling the RZSM state can provide auxiliary information to these DL models when the RZSM source (in-situ/satellite) information is limited or unavailable. The literature illustrates that a sequential DL model will be apt for performing such correlation analysis [262], [47].

The transformer neural network (TNN) is the state-of-the-art DL model to process sequential data [49]. It is a new cognitive model eschewing recurrence implementations in the previous benchmark sequential DL models, such as Long short-term memory (LSTM) [287] and Gated recurrent unit (GRU) [286]. LSTM and GRU receive the input information indirectly as a set of hidden states passed on through multiple cells via series processing. While TNN can pay attention to every single input directly through parallel processing and attention mechanism to draw global dependencies between input and output. The attention mechanism enhances critical parts of the input while diminishing others so that the network can devote more focus to the essential information [251]. The indirect RZSM estimation requires data from multiple sensor locations to be perceived simultaneously to draw critical information from target prediction. Hence I find TNN more suitable for RZSM estimation.

Furthermore, the LSTM and GRU-based RZSM works mainly focus on estimating RZSM using meteorological variables alone [56], [47]. However, other environmental variables also affect the RZSM state and should be considered for estimation. Unlike LSTM and GRU, more variables will enhance TNN learning due to parallel processing and attention mechanism. Also, existing RZSM works aim at future predictions of a target location where ground truth is available [56], [47]. However, the main problem in the RZSM domain is the limited information on depth at locations without ground truth due to the shortcomings of direct measurement by in-situ and satellite devices. Such measurement constraints were addressed in the SSM domain using auxiliary variables correlations with the prediction target [43], [37], [35], [250]. This correlation approach is extrapolatable to the RZSM domain due to its similarities with the SSM. In order to extrapolate the correlations from sensor location RZSM and auxiliary data for prediction at a site without direct measurement, I introduce a hybrid TNN model. The proposed model uses auxiliary RZSM variables identified in the literature for RZSM estimation: SSM, soil type, surrounding RZSM, target distance, plant type, RZSM depth, root type, plant daily water requirement, humidity, temperature, rainfall, soil salinity, and soil temperature [67], [46], [65], [47], [56].

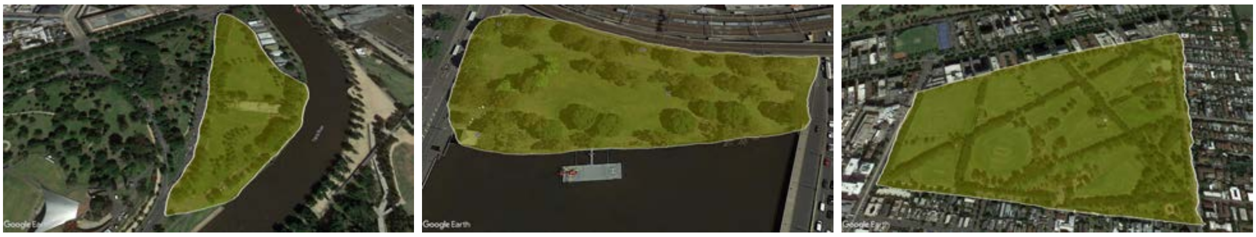
In the proposed model first, as a high-level input abstraction, sensor locations with similar satellite SSM values as the target are selected for further analysis. To distinguish the SSM information from different test locations for similarity assessment, I use the downscaled 1 km SM from Chapter 4 [2]. Next, a dynamic-MLP (D-MLP) finds the prominent PTSPc predictor variables from chosen sensor locations by determining the relative importance of these variables as a function of the neural network synaptic weights. The network assigns higher weights if the variables contribute more to the predictions. The selected sensor RZSM and auxiliary variables from the sensor and target locations are input to the attention layer of the proposed hybrid TNN model for further processing.

The attention block first analyses the relative relationship between the auxiliary variables at sensor and target locations. Next, the model performs a similar relationship analysis between sensor location auxiliary variables and RZSM. To achieve these proposed dual relative relationship analyses I concatenate an additional attention block to the basic TNN architecture. Inside the additional block, I implement a new multi-cross attention mechanism, which simultaneously analyses cross-relative relationships between features from multiple sensor sites with the target. Unlike the conventional attention block, which transforms the output using its inputs, this block transforms the input using the inputs from the other concatenated block. Such an arrangement helps correlate the similarities of RZSM-related environmental changes between target and selected sensor locations. Based on these relationships, probabilities were generated by the proposed hybrid TNN using the Bayes theorem in an additional layer implemented as a preceding layer to the traditional TNN output layer for each selected sensor RZSM. Finally, the probabilistic weighted sum of the selected RZSM sensors provides the required indirect estimates. A probabilistic weight assignment can reduce the uncertainties in multi-source ensembling [265] and can enhance the proposed model prediction accuracy as it combines multi-sensor data.

The proposed hybrid TNN model is compared to two recent related works using LSTM and GRU [56], [47] for performance evaluation. These evaluations and comparisons will provide new insights for improvements in the RZSM estimation. The remaining chapter is organized as follows: the study area and proposed model are elaborated in Section 5.2 and Section 5.3 respectively. The simulation results are shown in Section 5.4. Finally, a conclusion is drawn in Section 5.5.

5.2 Study Area and Dataset

To demonstrate the advantage of the proposed model I require a dataset having multiple RZSM sensing station readings installed over a vast spatially distributed area having similarities in the soil-plant-atmosphere domain. Hence I use the City of Melbourne Soil Sensor Readings (CoMSSR [252]) dataset. This dataset can help demonstrate the advantage of the proposed model. The CoMSSR dataset contains historical readings for RZSM sensors and related auxiliary RZSM variables within parks across the city of Melbourne. The units and readings from 78 soil sensors installed locations are included within the CoMSSR dataset.



(a) Alexandra Garden, Melbourne (b) Batman Park, Melbourne (c) Fawkner Park, Melbourne

Figure 5.1: The layout of the three test sites: (a) Alexandra Garden (5.2 hectares) is located at 37.82037°S 144.971938°E with clay loam soil. It has an avenue of oak trees around the garden; (b) Batman Park (1.47 hectares) is located at 37.821766°S 144.956348°E with medium to heavily textured clay plus some sand. The park consists of sparse maturing eucalyptus trees with no understorey or saplings, and (c) Fawkner Park (41 hectares) is located at 37.838434°S 144.981016°E with sandy loam soil. The park has elm, oak, and fig trees.

Next, the Soil Sensor Locations (SSL) dataset can be used to get the locations (latitude and longitude) where soil sensors have been deployed across the city. The CoMSSR dataset is joined to the SSL dataset [255] using the site-id column. The SSL dataset information can be further used to link with the soil type by area maps (STBAM) dataset to get the soil type information [256]. The CoMSSR, SSL, and STBAM datasets can be accessed from the city of Melbourne-open data portal (https://data.melbourne.vic.gov.au/).

The latitude and longitude information from the SSL dataset can be used to obtain the weather data from the nearest weather station accessible through the Bureau of Meteorology (BOM), Australia (http://www.bom.gov.au/). The weather station number from BOM can be used to join with the BOM forecast dataset [200]. The latitude and longitude information from the SSL dataset is further used to generate the fine-resolution SSM for the 78 sensor locations. The fine resolution SM is generated using the 9 km Soil Moisture Active Passive (SMAP) and 1 km Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data, which can be accessed from NASA earthdata (https://search.earthdata.nasa.gov/search). The sensor locations can be mainly classified as dry and wet regions. The Alexandra Gardens and Batman Park are wet regions as they are located near the Yarra River. Fawkner Park represents the case of a dry location far from any water bodies. Furthermore, these three test locations have the most widely available soil types in the world [257] and will be tested for wet and dry seasons to illustrate the model generalization. Hence my model is tested at these three locations illustrated in Fig. 5.1 to assess its performance using all these datasets. The proposed model is elaborated further.

5.3 Methodology

TNN is the state-of-the-art DL model for analyzing sequences. However, for in-direct RZSM estimation, the data associations within a location and multiple sensor sites with the target need to be found. Conventional TNN does not incorporate such simultaneous multi-associations, hence to suit the in-direct RZSM estimation application, we develop a new TNN model called the hybrid TNN model. First, sensor locations with similar downscaled 1-km satellite SM as the target are selected. Next, a dynamic multilayer perceptron (D-MLP) network layer selects the highly correlated auxiliary RZSM information using ground SM and downscaled 1-km satellite SM. D-MLP output is provided to a dual attention module along with the sensor RZSM. The dual attention module finds the required multi-associations using sensor and target region data. Finally, the Bayesian layer averages multi-location RZSM using the conditional probability generated based on relative relationships for getting the target location RZSM estimate. The proposed TNN is optimized using the random search method, which exhibits the same performance as many common optimization methods but with lower computation costs [204]. Table 5.1 illustrates the optimized architecture of the proposed hybrid TNN. With the chosen model configuration, the optimization results suggest a learning rate of 0.075 for good model performance. The proposed hybrid TNN-based RZSM estimation is detailed below.

5.3.1 SSM Similarity Assessment

Satellite imagery is made up of pixels and it represents the relative reflected light energy recorded for that part of the image. Each pixel represents a square area on an image and is a measure of the sensor's ability to resolve ground objects. The reflected energy from a light spectrum will be the same for objects that are similar in composition. Currently, satellites can only provide SSM information. Since SSM and RZSM are correlated, a similarity assessment of satellite pixels can help identify similar pixel-valued sensor sites with the target location without a sensor to further derive RZSM information.

A minimum of 1 km resolution is required to distinguish the SM pixels from the sensor sites in the CoMSSR dataset [252]. Currently, SMAP provides the highest resolution global daily SM of 9 km. Since this resolution is not sufficient for the PI application, I downscale the 9 km SMAP SM to 1 km using the 3D-Bi-LSTM model [2] from my previous work in Chapter 4. If the SMAP can be downscaled to $>$ 1 km (the minimum requirement), the overall RZSM estimation accuracy can further increase. However, the resolution of open-source MODIS data (essential for SMAP downscaling) limits further enhancement of the resolution [157], [158]. The comparison results (illustrated in the previous Chapter) with existing state-of-the-art downscaling methods show that the model developed in Chapter 4 is currently best for SMAP downscaling. Therefore, this model is adopted for downscaling in the current study.

Based on the downscaled SSM similarity the sensor locations are selected. The target location 1 km downscaled SSM is evaluated for similarity with all the sensor location's 1 km SSM (as all sensor locations may not be equally important for target prediction). The literature illustrates that the removal of irrelevant inputs results in higher predictive accuracy [201], [202]. Also, as the input data grows, the number of training instances needed grows exponentially, so in practical situations, it is necessary to reduce the input data using dimensionality reduction techniques [253]. Existing literature illustrates an MLP-based relative contribution technique most suitable for this task [253], [254]. For the proposed work, this technique recommends a selection of $\geq$ fifty-percent similarity. Hence, to identify the relevant pixels that can contribute to the target prediction and reduce the complexity of analysis, I select pixels with at least fifty-percent similarity to the target. Next, the selected features are passed on to hybrid TNN for further analysis.

5.3.2 RZSM Estimation

Fig. 5.2 illustrates the proposed RZSM estimation using the hybrid TNN model. The D-MLP network forms the first processing layer of the proposed neural network. In the D-MLP the number of neurons is dynamically set at each cycle based on the similarity assessment in the current iteration. This design helps to analyze all locations with high similarity scores for any selected time rather than limiting the analysis to a fixed number of sites. Also, for locations with similar similarity scores, the closer in the range locations are given higher weights. This weight allocation is because closer regions to the target will have more similar environmental conditions. Hence the D-MLP hidden layer weights are further tuned based on the range value from the target location.

The calculation of whether a sensor is close or far in the distance (d) is performed using the longitude and latitude of the target point and $g \in \{1,2,3,...,G\}$ selected sensor locations. The closest regions will have similar hydrological cycles compared to the farther located sensors. Hence the D-MLP assigns more weight to the decisions made from the information from closer sensor locations ($W_p > W_q \ \forall \ d_p < d_q$ if $p \neq q$).

Table 5.1: Hybrid Transformer Neural Network Architecture

Architecture		
Model Components	Parameter Name	Parameter Value
MLP	Number of hidden layers	Five
	Number of neurons in the hidden layer	Twenty-eight
	Activation function	ReLU
Transformer	Head size	Two-hundred-fifty-six
	Number of heads	Four
	Number of blocks	Four

Note: The model parameters are obtained using the random search optimisation.

Once the D-MLP weights are optimized, I auxiliary RZSM information $f_g^i \in \{f_g^1, f_g^2, ..., f_g^I\}$ from the g selected similar locations are passed on to the D-MLP for finding the prominent auxiliary RZSM information features. The D-MLP determines the relative prominence [258] of the predictor variables as,

$$V_{ij} = \sum_{k=1}^{K} W_{ik} \bullet W_{kj}, \tag{5.1}$$

where V_{ij} is the relative importance of any input variable f_g^i with respect to $j \in \{1,2,...,J\}$ output neuron (for J number of output neurons), K is the number of neurons in the hidden layer, W_{ik} is the synaptic connection weight between the input neuron i and the hidden neuron k, and W_{kj} is the synaptic weight between the hidden neuron k and the output neuron j. The D-MLP outputs the $I/2$ top-ranked features from the relevant list of I total predictor variables from selected sensor locations

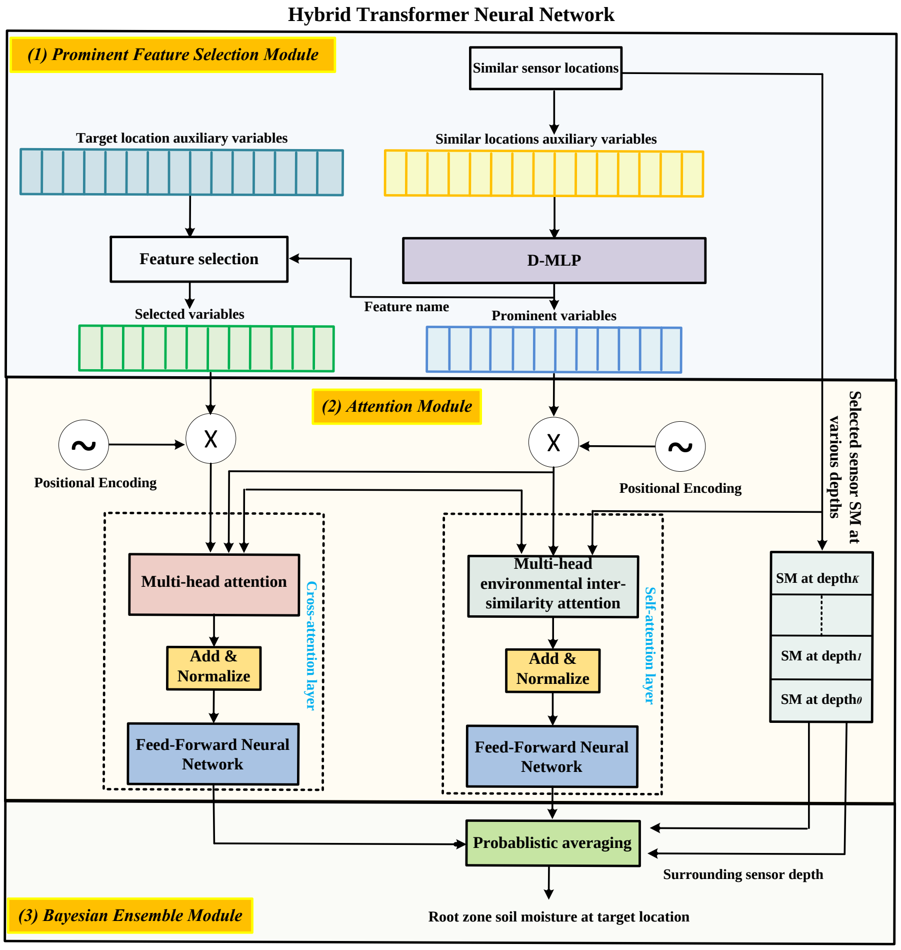


Figure 5.2: Root zone soil moisture estimation using the hybrid transformer neural network (hybrid TNN). The hybrid TNN has three modules: (1) Prominent feature selection module: from similar SSM sensor locations, the D-MLP (violet box) selects prominent predictor variables from auxiliary RZSM variables. The selected feature name is used to select the same feature from the target (white box); (2) Attention module: the selected variables undergo position encoding to track the order of the selected variable sequence. The position-encoded selected features from target and sensor sites are provided to the multi-head attention module (brown box). The module finds the relative relationship between the selected sensor and target location data. The multi-head environmental inter-similarity attention module (light blue box) finds the relative relationship between the multi-location RZSM and prominent predictor variables from sensor sites; (3) Bayesian ensemble module: performs a Bayesian average of multi-sensor RZSM using the conditional probability (green box) generated based on the attention module output to estimate target RZSM.

to the attention module for RZSM prediction. Using the attention module outputs proposed model performs a Bayesian average of selected sensor location moisture data to predict the target RZSM.

To calculate the probability values to perform the Bayesian average of the RZSM from selected sensors, first, the cross attention layer of the hybrid TNN evaluates the percentage similarity of features at the target location and selected sensor locations. Let $f_g^s \in \{f_g^1, f_g^2, \ldots, f_g^S\}$ be the D-MLP selected prominent features, where $S = I/2$ is the number of chosen features. The f_g^s is encoded as c_g^s through position encoding [259]. At a g^{th} location, the SM can be expressed as a function of $s \in \{1,2,...,S\}$ prominent auxiliary RZSM variables c_g^s,

$$y_g(t) = z_{P_g}\big(\sum_{s=1}^{S} c_g^s(t)\big) \,\forall\, t \in \{1,2,3,...,T\}, \tag{5.2}$$

where z_{P_g} is the mapping function for the given hybrid TNN predictor neuron P_g. Eq. (5.2) gives the RZSM at a g^{th} sensor location. However, I aim to find the RZSM at location $h \neq g$ without a sensor. Hence I estimate $y_h(t)$ through Bayesian ensemble of RZSM from g sensor locations. To estimate this, the relative relationship between the independent variable $c_g^s(t)$ at each g^{th} location and $c_h^s(t)$ at target location 'h'needs to be derived. The proposed hybrid TNN's inter-similarity attention layer queries how related the encoded feature c_h^s at the target location is to the same features at G locations. The relative relationship between c_h^s and $\{c_1^s,...,c_G^s\}$ is $\{r_{h\to1},...,r_{h\to G}\}$. For any location $u \in g$, this can be expressed as,

$$r_{h\to u}^s(t) = \frac{e^{\frac{min(c_h^s(t),c_u^s(t))}{max(c_h^s(t),c_u^s(t))}}}{\sum_{g=1}^{G} e^{\frac{min(c_h^s(t),c_g^s(t))}{max(c_h^s(t),c_g^s(t))}}} \,\forall\, s \in \{1,...,S\}, \quad t \in \{1,...,T\}. \tag{5.3}$$

The relative relationship in Eq. (5.3) is used to get the transformed input features from g locations '$\tilde{c_g^s}(t)$' as,

$$\tilde{c_g^s}(t) = \sum_{g=1}^{G} r_{h\to g}^s(t) c_g^s(t) \forall\, s \in \{1,...,S\}, t \in \{1,...,T\}. \tag{5.4}$$

Transforming the RZSM at g locations in Eq. (5.2) using the transformed auxiliary RZSM variables in Eq. (5.4) as,

$$\hat{y_g}(t) = fp_g(\tilde{c_g^s}(t)) \forall\, P_g, s \in \{1,...,S\}, t \in \{1,...,T\}, \tag{5.5}$$

where $\hat{y_g}(t)$ is the predicted RZSM. The objective of each hybrid TNN predictor neuron, denoted as P_g, is to find out the relationship between $\tilde{c_g^s}(t)$ and $\hat{y_g}(t)$. To ensemble the relative SM at g locations for getting the SM at the target location, the

Bayesian average is used. The ensemble SM prediction, given the data $\tilde{c_g^s}(t)$, is

$$y^{en}(t) = \sum_{g=1}^{G} p(\hat{y}|P_g, \tilde{c_g^s})(t)p(P_g|\tilde{c_g^s})(t)y_g(t)\forall\, t \in \{1,2,3,...,T\}, \tag{5.6}$$

where $y^{en}(t)$ represent the RZSM at the location $h \neq g$ for $t \in \{1,...,T\}$ time instants. $y_g(t)$ is the actual RZSM measured by the sensors at g locations. $p(\hat{y}|P_g, \tilde{c_g^s})$ is the SM PDF based on P_g alone, estimated from training data, and $p(P_g|\tilde{c_g^s})$ is the posterior probability of predictor P_g being correct for the given training data, $\tilde{c_g^s}(t)$.

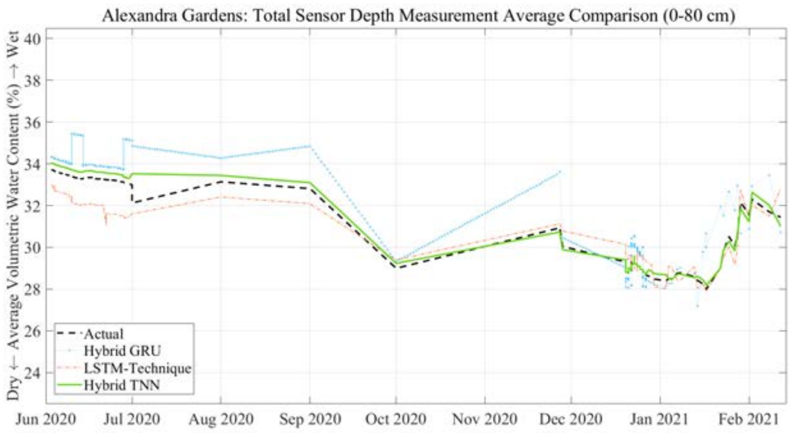
5.4 Results and Discussion

In this section, simulations are performed to evaluate the effectiveness and potentialities of the proposed hybrid TNN for RZSM at three different test sites in Melbourne, (1) Alexandra Gardens, (2) Batman Park, and (3) Fawkner Park. Comparisons are conducted with the hybrid GRU [47] and LSTM-Technique [56]. The simulation of RZSM prediction is implemented using Python 3.7.6. The NumPy and pandas libraries are used for data preprocessing and data management. The model is implemented using TensorFlow 2.8.0.

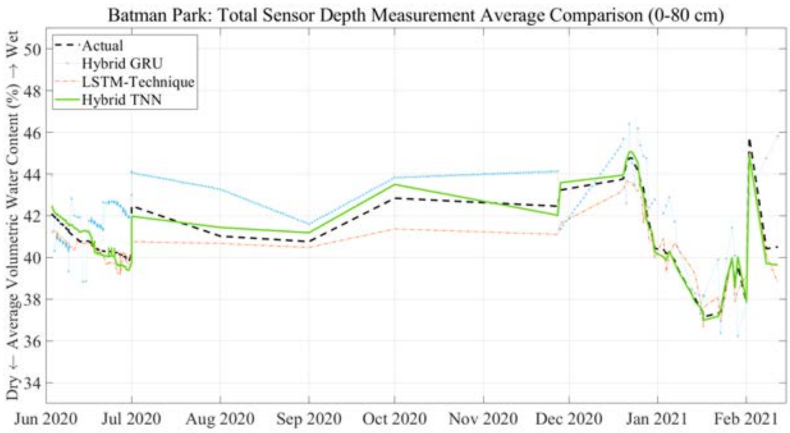
5.4.1 Qualitative Performance Evaluation

To demonstrate the effectiveness of the hybrid TNN, I compare the predicted and actual values at different soil depths over the three test locations. Figs. 5.3 (a)-(c) illustrate the comparison curve between hybrid GRU [47], LSTM-Technique [56] and the hybrid TNN predicted value, and the observed value of soil water content over eight soil layers. The time series predictions in Figs. 5.3 (a)-(c) by hybrid TNN from June 2020 to February 2021 for each test site show good temporal consistency. This result suggests that compared to the other two models, the hybrid TNN estimated RZSM mostly maintains the temporal changing features of the actual RZSM over the eight soil layers across the three test sites. Also, the hybrid TNN predicted RZSM closely following the actual RZSM than hybrid GRU [47], and LSTM-Technique [56], indicates the proposed model to be the more reliable sequential model.

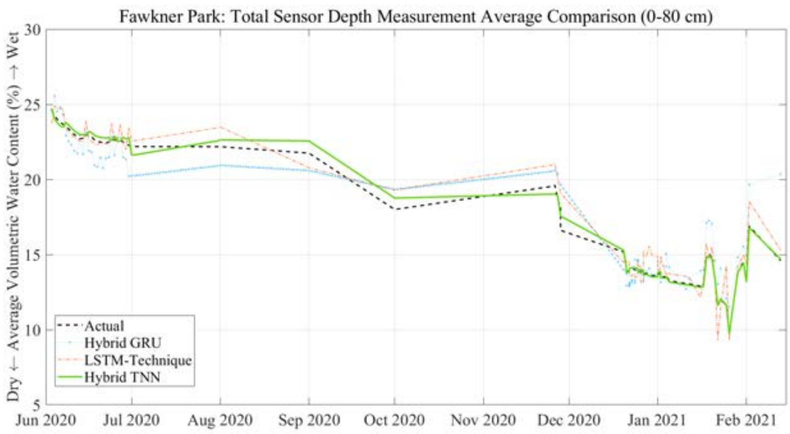
Further qualitative measurements are considered to synthesize the model's performance effectiveness over eight months for three stations. Table 5.2 shows the



(a) Alexandra Garden, Melbourne



(b) Batman Park, Melbourne



(c) Fawkner Park, Melbourne

Figure 5.3: Comparison of the predicted and the actual average soil moisture in the eight layers by Hybrid GRU [47] (blue dotted line with x marker), LSTM-Technique (red semicolon line) [56], and Hybrid TNN from June 2020 to February 2021 at: (a) Alexandra Gardens, (b) Batman Park, and (c) Fawkner Park. The Hybrid TNN prediction (green line) is closest to the actual value (black dashed line).

performance comparison of the hybrid GRU [47], LSTM-Technique [56], and hybrid TNN across eight different soil layers on the RZSM dataset with the three test

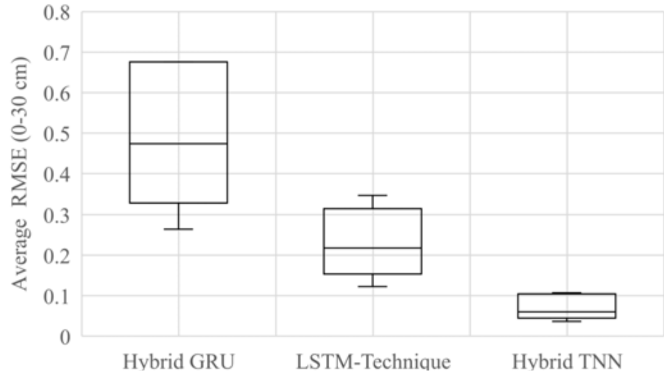
Table 5.2: RZSM prediction accuracy comparison of the proposed Hybrid TNN, Hybrid GRU [47], and LSTM-Technique [56] at eight soil layers on the RZSM dataset with three locations data.

Root zone depth	**Model**	**Performance Indices**		
		NSE	R^2	NMBE
0-10 cm	Hybrid GRU	0.6945	0.9633	-0.1502
	LSTM-Technique	0.7658	0.9721	-0.0620
	Hybrid TNN	0.8527	0.9913	-0.0252
10-20 cm	Hybrid GRU	0.7357	0.9783	-0.0873
	LSTM-Technique	0.8493	0.9818	-0.0397
	Hybrid TNN	0.9012	0.9879	-0.0156
20-30 cm	Hybrid GRU	0.6580	0.9368	-0.1490
	LSTM-Technique	0.7716	0.9475	-0.0812
	Hybrid TNN	0.8963	0.9817	-0.0320
30-40 cm	Hybrid GRU	0.5708	0.9679	-0.0926
	LSTM-Technique	0.9017	0.9715	-0.0577
	Hybrid TNN	0.9486	0.9981	-0.0109
40-50 cm	Hybrid GRU	0.7016	0.9528	-0.0532
	LSTM-Technique	0.9098	0.9643	-0.0156
	Hybrid TNN	0.9627	0.9826	-0.0073
50-60 cm	Hybrid GRU	0.7764	0.9283	-0.0299
	LSTM-Technique	0.9389	0.9542	-0.0207
	Hybrid TNN	0.9950	0.9973	-0.0079
60-70 cm	Hybrid GRU	0.6998	0.9017	-0.0229
	LSTM-Technique	0.8997	0.9218	-0.0125
	Hybrid TNN	0.8967	0.9640	-0.0089
70-80 cm	Hybrid GRU	0.7789	0.9463	-0.6512
	LSTM-Technique	0.8901	0.9623	-0.4531
	Hybrid TNN	0.9121	0.9729	-0.0895
Average	Hybrid GRU	0.7020	0.9469	-0.1545
	LSTM-Technique	0.8659	0.9594	-0.0928
	Hybrid TNN	**0.9207**	**0.9845**	**-0.0247**

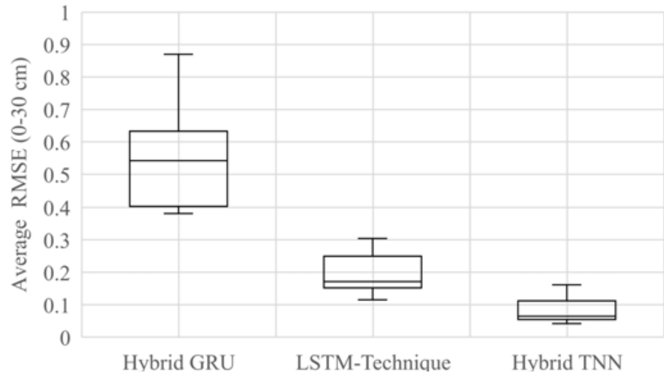
Note: The qualitative performance of these models is evaluated using Nash Suttcliff coefficient of efficiency (NSE), determination coefficient (R^2), and normalized mean bias error (NMBE). The NSE is commonly used to assess the performance of hydrological predictions. NSE values closer to one suggest the prediction model is a better predictor than the mean of the actual values. The R^2 values nearer to 1 indicate a highly reliable model for future predictions. NMBE captures the average bias in the RZSM prediction. Positive NMBE represents that the data from datasets is overestimated.

locations data using Nash Suttcliff coefficient of efficiency (NSE), determination coefficient (R^2), and normalized mean bias error (NMBE). The relative RZSM prediction accuracy improvement of the proposed B-TTN against hybrid GRU and LSTM is 23.7537% and 18.9283%, respectively. Specifically, the proposed hybrid TNN has R^2 values closer to 1, indicating a highly reliable prediction model. The negative NMBE values indicate that all the three models ([47], [56], and proposed hybrid TNN) do not overfit the dataset.

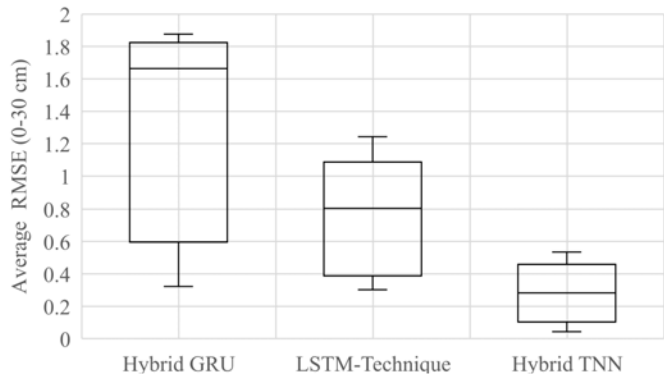
Figs. 5.4 (a)-(c) compares the average root mean square error (RMSE) over the active root zone depths from June 2020 to February 2021 at three test sites. Lower



(a) Oak Tree, Alexandra Garden



(b) Eucalyptus Tree, Batman Park



(c) Oak Tree, Fawkner Park

Figure 5.4: Comparison of RMSE average over the active root zone depths of target tree by the Hybrid GRU [47], the LSTM-Technique [56], and the proposed hybrid TNN for: (a) Oak Tree (0-50 cm), Alexandra Garden, Melbourne, (b) Eucalyptus Tree (0-30 cm), Batman Park, Melbourne, and (c) Oak Tree (0-50 cm), Fawkner Park, Melbourne from June 2020 to February 2021.

RMSE implies higher prediction accuracy. Fig. 5.4 (a) illustrates the RZSM prediction error for an oak tree at Alexandra Gardens, Melbourne. The oak tree's active root zone lies in the 0-50 cm layer [263]. Fig. 5.4 (b) illustrates the RZSM prediction error for a eucalyptus tree at Batman Park, Melbourne. Eucalyptus trees have shallow roots and grow up to 20-30 cm soil depth [263]. Fig. 5.4 (c) illustrates the RZSM prediction error for an oak tree at Fawkner Park, Melbourne. In Figs. 5.4 (a)-(c), the lowest prediction error is indicated by the proposed hybrid TNN in all three locations.

In addition to the RMSE, the vertical error bars in Figs. 5.4 (a)-(c) illustrate the standard deviation of prediction accuracy by each model. The error bars indicate a less standard deviation of prediction for the proposed hybrid TNN compared to hybrid GRU [47], and LSTM-Technique [56].

The lower RMSE of hybrid TNN also indicates its usability in addressing the performance drop in existing PI decision support systems ([239]) due to the RZSM estimation error accumulations than hybrid GRU [47] and LSTM-Technique [56].

5.5 Conclusion

In this chapter, a sequential DNN model, called hybrid-TNN, that can determine distributed RZSM indirectly using auxiliary information from satellite and ground data is developed. First, a dynamic MLP (D-MLP) network layer is designed for selecting the nearby soil moisture (SM) sensor locations data with the target location data. The selected location data is passed onto the proposed dual attention block and then to a Bayesian layer to estimate the RZSM. The model presents a way to derive the target location RZSM in-directly, thereby eliminating the requirement for a denser cluster of sensor installment. The test results illustrated the proposed model improved predictions compared to prominent sequential DNN models like GRU and LSTM by 23.7537% and 18.9283% respectively. Also, the NSE, R^2, RMSE, and NMBE performance analysis indicated better performance by the proposed hybrid TNN in terms of model performance, reliability, error reduction, and less overfitting. Overall, the performance of the proposed model indicates that a TNN-based sequential DNN model is highly suitable for RZSM estimation.

To conclude, the proposed model addresses the third research problem presented in Section 1.2.3 and leads to the third original contribution presented in Section 1.3. The high performance of hybrid TNN for RZSM estimation in this chapter leads to the further exploration of the final research question, ‘can incorporating better RZSM estimates’ help take better irrigation decisions in the final chapter of this thesis.

Chapter 6

Decision support system for precision irrigation

In this final chapter, the individual rainfall and SM models developed in chapters 3-5 are integrated to provide decision support in precision irrigation scheduling. Using this integrated model, the advantage of including more accurate rainfall and SM information for irrigation scheduling is quantified in terms of the amount of water and cost savings.

This chapter contains materials included in the following manuscript, which has been submitted to Agricultural Water Management Journal:

[3] **Neethu Madhukumar**, Eric Wang, and Wei Xiang,"Hybrid Transformer Network for Root Zone Soil Moisture Estimation for Decision Support in Precision Irrigation", *Agricultural Water Management*, under review (Impact Factor: 6.611, h-index: 65, Journal rank (Agronomy & crop science): 7).

6.1 Introduction

Irrigation decision support systems (IDSSs) are integrated solutions combining and interpreting meteorological, soil moisture (SM), and/or crop water stress (in-directly indicated by SM) data to help growers make correct irrigation decisions [285]. Literature reveals faulty IDSS timing and quantity of irrigation decisions under inaccurate rainfall forecasts and SM data interpretation circumstances [17], [54], [55], [52], [53]. Hence, I developed the rainfall forecast refinement model in Chapter 3 and SM downscaling/estimation models in Chapters 4-5 for improved irrigation scheduling through better information extraction. The experimental evaluations in Chapters 3-5 reveal that the models developed in this thesis for rainfall forecast refinement, SM downscaling, and SM estimation at the plant root zone perform better than existing state-of-the-art models in each of these applications. Therefore, in this chapter, I integrate the individual rainfall and SM components developed in Chapters 3-5 to form an IDSS.

In the proposed IDSS for precision irrigation (IDSS-PI), first using the models from Chapters 4-5, I provide the information about the level of SM to a central analysis hub that evaluates whether irrigation is required or not. Irrigation is not scheduled by the IDSS-PI when SM information indicates that it has met the daily plant water requirements. When irrigation is scheduled, to optimize irrigation, the IDSS-PI only provides the remaining plant water requirement, which is unmet through rainfall. To evaluate the efficiency of the proposed irrigation scheduling, I compare two recent related works, gated recurrent unit (GRU)-based [47], and long short-term memory (LSTM)-based [56] irrigations with the IDSS-PI. I find these comparisons important as GRU and LSTM are two widely adopted sequential DL models in soil-hydrological applications. This comparison will illustrate if, like other sequential analysis problems, TNN-based models can exhibit higher performance than GRU [47] and LSTM [56] for irrigation scheduling.

The remaining chapter is organized as follows: the study area is detailed in Section 6.2, the proposed IDSS-PI model is presented in Section 6.3, and the simulation results and conclusion are provided in Section 6.4 and Section 6.5 respectively.

6.2 Study Area and Dataset

The irrigation scheduling is performed by the proposed IDSS-PI based on the RZSM predicted in chapter 5. Hence, the irrigation is scheduled for the same test locations as in chapter 5.

The first test site is Alexandra Garden, Melbourne. It has an avenue of oak trees around the garden. The garden has a total area of 5.2 hectares. Clay loam soil type is found in this area.

The second test site is Batman Park, Melbourne. It has a total area of 1.47 hectares. Current vegetation in the park consists of sparse maturing eucalyptus trees with no understorey or saplings. Medium to heavily textured clay with some sand is found in this area.

The third test site is Fawkner Park, Melbourne. It is trapezoidal with an area of 41 hectares. The park is edged with elm, oak, and figs. Sandy loam soil type is found in Fawkner Park.

The proposed IDSS-PI extracts information required for irrigation using the models from chapters 3-5, hence, all the datasets used in chapters 3-5 are combined. The details of these datasets and how to combine them are elaborated further.

The City of Melbourne Soil Sensor Readings (CoMSSR [252]) dataset was used in chapter 5 to estimate RZSM. The CoMSSR dataset contains historical readings for RZSM sensors and related auxiliary RZSM variables within parks across the city of Melbourne. The units and readings from 78 soil sensors installed locations are included within the CoMSSR dataset.

The Soil Sensor Locations (SSL) dataset can be used to get the locations (latitude and longitude) where soil sensors have been deployed across the city. The CoMSSR dataset is joined to the SSL dataset [255] using the site-id column.

The SSL dataset information can be further used to link with the soil type by area maps (STBAM) dataset to get the soil type information [256]. The CoMSSR, SSL, and STBAM datasets can be accessed from the city of Melbourne-open-data-portal (MODP) (https://data.melbourne.vic.gov.au/).

The latitude and longitude information from the SSL dataset can be used to obtain the weather data from the nearest weather station accessible through the Bureau of Meteorology (BOM), Australia (http://www.bom.gov.au/). The weather station number from BOM can be used to join with the BOM forecast dataset [200].

The latitude and longitude information from the SSL dataset is further used to generate the fine-resolution SSM for the 78 sensor locations. The fine resolution SM is generated using the 9 km Soil Moisture Active Passive (SMAP) and 1 km Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data, which can be accessed from NASA earthdata (https://search.earth data.nasa.gov/search). The proposed IDSS-PI is elaborated further.

6.3 Model Design and Overview

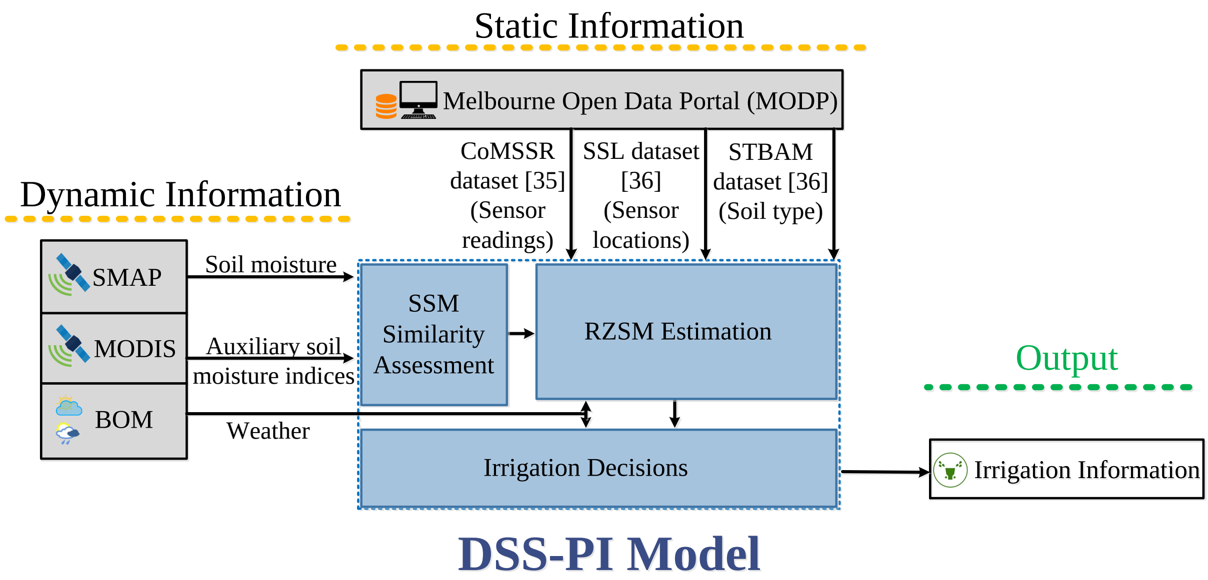


Figure 6.1: The framework of the decision support system for precision irrigation (IDSS-PI). The dynamic information (SMAP, MODIS, and BOM) and static information (CoMSSR, SSL, and STBAM) are given as input to the proposed IDSS-PI to get irrigation information. First, the IDSS-PI performs SSM similarity assessment using SMAP and MODIS data for selecting sensor locations for further analysis. Next, the RZSM estimation module outputs RZSM information using CoMSSR, SSL, STBAM, and BOM data. Finally, the irrigation decisions module provides the irrigation information using predicted RZSM and BOM rainfall data.

Fig. 6.1 illustrates the framework of the proposed IDSS for precision irrigation (IDSS-PI). The dynamic information (SMAP, MODIS, and BOM) and static information (CoMSSR, SSL, and STBAM) are given as input to the proposed IDSS-PI to get RZSM and irrigation information. The proposed IDSS-PI is made of three parts: 1) the SSM similarity assessment, 2) the RZSM estimation, and 3) the irrigation decision module.

Fig. 6.2 illustrate the proposed IDSS. The irrigation frequency module in the IDSS decides whether irrigation is required or not based on the water depletion tolerance rate ($R(t)$) for each irrigation type. The $R(t)$ value is calculated using [264] as follows,

$$R(t) = P_{wr}(t) - \left(\frac{P_{wr}(t) \times DT}{100} \right) \forall t \in \{1, 2, 3, ..., T\}, \tag{6.1}$$

where $P_{wr}(t)$ is the daily water requirement of the plant and DT is the soil water depletion tolerance by the PI system. The $P_{wr}(t)$ should not drop below $R(t)$ to maintain water levels above wilting point. The IDSS used for PI knows the irrigation

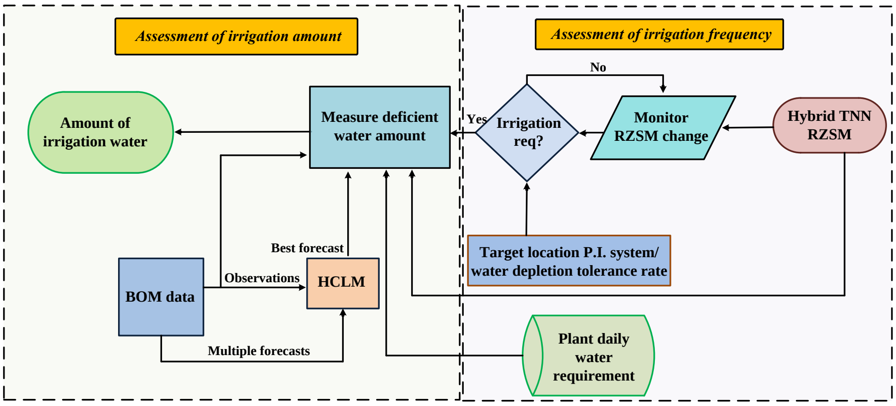


Figure 6.2: The IDSS with two main parts: (a) Irrigation frequency assessment module (light violet square with dashed black outline) and (b) Irrigation amount assessment module (light green square with dashed black outline). The irrigation frequency assessment module examines if irrigation is required using hybrid TNN predicted RZSM (brown ellipse) and water depletion tolerance (dark blue rectangle) data. The irrigation amount assessment module subtracts the water available from BOM rainfall data (dark blue square) and hybrid TNN RZSM from plant water requirement (green cylinder) data.

requirements in the farm using Eq. (6.1) and uses this information for controlling an automatic irrigation system for the precise application of water to plants. The main type of irrigation system used for PI is drip irrigation. Ideally, for drip irrigation systems, the DT should be in the range of 20 to 25 [264].

The proposed IDSS decides on when irrigation should occur as follows,

$$\begin{aligned} \forall\, y^{en}(t) \le R(t):&\ \text{Irrigation required} \\ \forall\, y^{en}(t) > R(t):&\ \text{Irrigation not required.} \end{aligned} \tag{6.2}$$

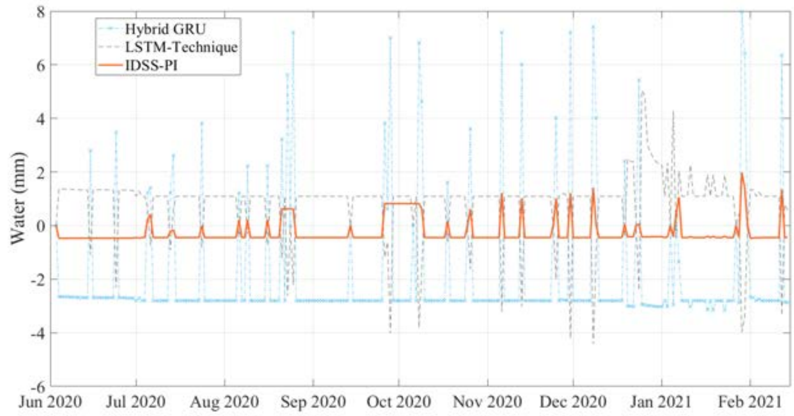
If irrigation is required, the amount of water to be given through irrigation is evaluated next. The amount of water to be provided through irrigation is calculated as

$$A(t) = P_{wr}(t) - (y^{en}(t) + W(t))\ \forall\, t \in \{1, 2, 3, ..., T\}, \tag{6.3}$$

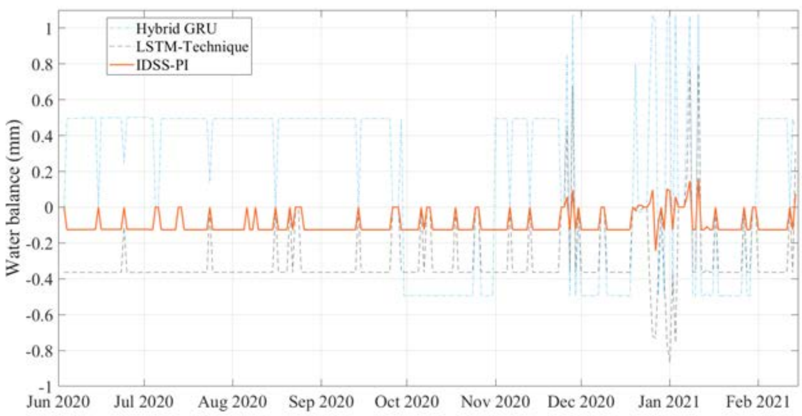
where $A(t)$ is the amount of irrigation water, $P_{wr}(t)$ is the daily water requirement of the plant, $y^{en}(t)$ is the hybrid TNN estimated RZSM, and $W(t)$ is the estimated water available from rainfall for the selected day using [1]. The proposed model is tested in Section 3.

6.4 Results and Discussion

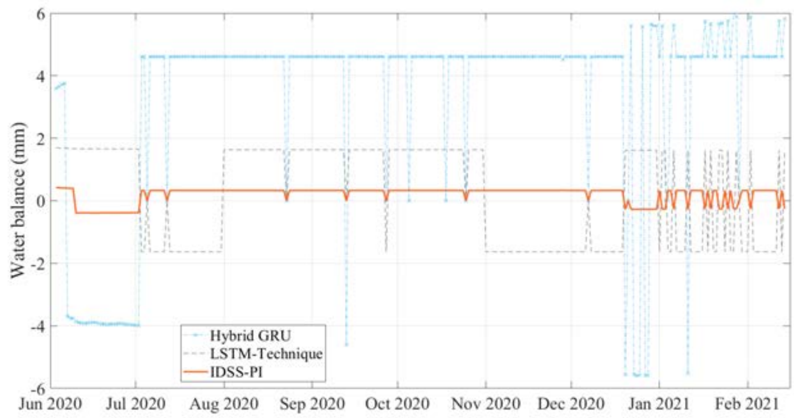
In this section, simulations are performed to evaluate the effectiveness and potentialities of the proposed model for irrigation predictions at three different test sites in Melbourne, (1) Alexandra Gardens, (2) Batman Park, and (3) Fawkner Park. Com-



(a) Oak Tree, Alexandra Garden, Melbourne



(b) Eucalyptus Tree, Batman Park, Melbourne



(c) Oak Tree, Fawkner Park, Melbourne

Figure 6.3: Comparison of irrigation water balance by Hybrid GRU [47] (blue semicolon line with x marker), LSTM-Technique [56] (grey dashed line), and (3) IDSS-PI (red line) at (a) Alexandra Garden, (b) Batman Park, and (c) Fawkner Park. The IDSS-PI line is closer to zero (most optimum irrigation) than the other two models.

parisons are conducted with the hybrid GRU [47] and LSTM-Technique [56]. The simulation of irrigation is implemented using Python 3.7.6. The NumPy and pandas libraries are used for data preprocessing and data management. The model is implemented using TensorFlow 2.8.0.

In this section, the drip irrigation simulation results of the hybrid GRU [47], LSTM-Technique [56], and proposed IDSS-PI are presented. As discussed before drip irrigation is a widely adopted irrigation system for PI. The irrigation is simulated from June 2020 to February 2021 for three test locations. The water balance, water savings, and cost savings are evaluated for quantitative performance comparisons between the hybrid GRU [47], LSTM-Technique [56], and proposed IDSS-PI.

Fig. 6.3 illustrates the water balance of the drip irrigation schedules produced using the three RZSM predictions. Drip irrigation systems are usually designed to keep root zone moisture close to the optimum level by daily moisture evaluation. Irrigation will be applied when the root zone water depletes below 20% of the requirement [264]. The daily water depletion tolerance for the trees on the three sites is 13.52 mm, 9 mm, and 21.03 mm. The values are calculated by substituting the daily water requirement of the tree species from literature ([263], [252]) in Eq. (6.1). All three models, schedule irrigation to meet the plant water requirement post subtracting the daily water obtained from rainfall. The Melbourne (Olympic Park), BOM, Australia weather station data is used to schedule irrigation at Alexandra Gardens and Batman Park for supplementing rainfall water. The Hawthorn (Scotch College), BOM, Australia weather station data is used for Fawkner Park irrigation.

Table 6.1: Total comparative water savings by IDSS-PI based irrigation against Hybrid GRU [47] and LSTM-Technique [56] across all three test locations during 2020-2021 using recycled water.

Saving against:	Water saving/tree/year (KL)	Water saving/total area/year (KL)
Hybrid GRU	8.24	60,325.04
LSTM-Technique	8.04	58,860.84
Average	8.14	59,592.94

Note: The water cost calculation is based on the 2020/2021 class B and C recycled water tariff rate from Western Waters, Victoria, Australia. https://www.westernwater.com.au/our-services/recycled-water/class-b-c-recycled-water. The energy cost calculation is based on the 2020/2021 public unmetered supplies energy tariff rate from CitiPower, Victoria, Australia. https://www.aer.gov.au/networks-pipelines/determinations-access-arrangements/pricing-proposals-tariffs/citipower-annual-pricing-2020. It is assumed that the pump capacity is 10 kW, under a constant rate of 65 L/s.

The irrigation water balance by the hybrid GRU [47], LSTM-Technique [56], and IDSS-PI irrigation scheduling systems is shown in Fig. 6.3 (a)-(c). The plant irrigation

Table 6.2: Total comparative cost savings by IDSS-PI based irrigation against Hybrid GRU [47] and LSTM-Technique [56] across all three test locations during 2020-2021 using recycled water.

Saving against:	Cost saving/tree/year (AUD)	Cost saving/total area/year (AUD)
Hybrid GRU	2.91	21,304.11
LSTM-Technique	2.69	19,693.49
Average	2.80	20,498.80

Note: The water cost calculation is based on the 2020/2021 class B and C recycled water tariff rate from Western Waters, Victoria, Australia. https://www.westernwater.com.au/our-services/recycled-water/class-b-c-recycled-water. The energy cost calculation is based on the 2020/2021 public unmetered supplies energy tariff rate from CitiPower, Victoria, Australia. https://www.aer.gov.au/networks-pipelines/determinations-access-arrangements/pricing-proposals-tariffs/citipower-annual-pricing-2020. It is assumed that the pump capacity is 10 kW, under a constant rate of 65 L/s.

water requirement at each test site is closely met by the IDSS-PI compared to the hybrid GRU and LSTM-Technique. Fig. 6.3 (a)-(c) illustrate that more irrigation water is wasted by hybrid GRU [47] and LSTM-Technique [56] compared to the proposed IDSS-PI.

To further quantify the water and cost saved by the proposed IDSS-PI, the irrigation water amount by the three competing models is compared for the three test locations. All three test locations use recycled water for irrigation. Table 6.1 shows that the proposed IDSS-PI conserved 60,325.04 KL and 58,860.84 KL of water compared to hybrid GRU [47] and LSTM-Technique [56] respectively in total over the three test locations. Furthermore, Table 6.2 illustrate that the IDSS-PI saved 21,304.11 AUD compared to hybrid GRU [47] and 19,693.49 AUD compared to LSTM-Technique [56] in total cost savings over the three test locations for a duration of one year. The average savings by IDSS-PI is 15.632% water and 15.125% cost compared to the two comparison models. From the results in this section, it can be concluded that the proposed IDSS-PI attains a more precise water balance than hybrid GRU [47] and LSTM-Technique [56], thus reducing water loss and increasing irrigation profitability.

6.5 Conclusion

In this chapter, an IDSS for precision irrigation (IDSS-PI) was developed by integrating the DL models developed in chapters 3-5 into an irrigation decision support framework. The proposed model scheduled irrigation based on moisture variations in the plant root zone. Furthermore, irrigation was provided as a supplement to rainfall to reduce water loss. The experimental results indicate that the integration of the DL models from chapters 3-5 into an IDSS improved the irrigation decisions through better RZSM and rainfall input information. Also, the proposed model demonstrated higher water and cost savings compared to existing DL-based irrigation scheduling models. Additionally, the proposed irrigation indicates better water balance and lesser over-irrigation compared to recent related works in the literature. Overall, this chapter addresses the fourth research problem discussed in Section 1.2. It also provides the fourth and final original contribution of this thesis.

Chapter 7

Conclusion

7.1 Summary

This thesis developed a novel deep learning (DL) based irrigation decision support system (IDSS) for precision irrigation (PI). The proposed IDSS overcame the irrigation scheduling errors shown in the existing models due to inaccurate rainfall, surface soil moisture (SSM), and root zone soil moisture (RZSM) inputs through better information extraction methods. Each of the contributions in this thesis is detailed further.

The first research problem presented in Section 1.2.1 was addressed in Chapter 3. Conventionally, the final climate model (CM) forecast is assumed to be the best as it has the least lead time with the prediction target. However, the preliminary analysis conducted in this thesis revealed that better forecasts for the target day were available in the preceding times. Therefore, if the multiple CM forecast can be refined, rain forecasting will further improve. To achieve the required CM forecast refinement, I developed a new DL model called the hybrid climate learning model (HCLM). The HCLM contained two parts, the probabilistic Multilayer perceptron (P-MLP) and the hybrid deep long short-term memory (HD-LSTM). The PMLP learned the pattern of the best forecast from the CM over multiple data sub-samples and select the best forecast. The PMLP output was fed to HD-LSTM as part of knowledge distillation. The use of the PMLP output label for data selection reduced the HD-LSTM search space for sequential data analysis, which enabled HD-LSTM to decode better rain predictions by correlating P-MLP output with CMs' input observations.

The HCLM significantly improved CM's rain predictions over six major climate zones in Australia.

Next, the second research problem, presented in Section 1.2.2, was considered. To solve this research problem a new downscaling method was developed for obtaining SSM information in Chapter 4. First, a covariance-adaptive convolutional neural network (CNN) was developed to select the most correlated soil moisture indices (SMIs) to reduce the data association complexity of the spectra-spatial-temporal (3D) structure and to create a more discriminative feature extraction process. Next, a strategy was developed to transpose the 3D spatial and temporal SMI time serials into a 1D sequence representation to enhance spatial correlation within any given time instant. Finally, a novel 3D-Bidirectional-LSTM model is developed to adapt to the 3D downscaling structure to predict fine-scale soil moisture maps from coarse inputs. The proposed model is the first downscaling model that bi-directionally associates the 3D relationship between coarse-resolution SSM and fine-resolution SMIs. This type of data association considerably reduced the empty spaces (spectral information loss) in the proposed downscaled images compared to widely used downscaling models.

Downscaled satellite SM provides information on SSM (spatially distributed SM). For knowledge of SM depth distribution, RZSM information is needed. Direct RZSM measurement through large-scale in-situ sensor deployment is not economical. Hence, in-direct RZSM estimation methods are required, this research problem was described in Section 1.2.3. To solve this research problem an in-direct RZSM estimation method using a novel hybrid transformer neural network was developed in Chapter 5. The proposed model provided a unique framework for integrating information on weather, soil, and plant characteristics from ground and satellite sources for estimation as RZSM changes are controlled by them. First, sensor locations with similar downscaled SSM as the target was selected. Next, a dynamic multilayer perceptron (D-MLP) network layer selected the highly correlated auxiliary RZSM information using ground SM and downscaled SSM. D-MLP output was provided to a dual attention module along with the sensor RZSM. The dual attention module found the required multi-associations using sensor and target region data. Finally, the Bayesian layer averaged multi-location RZSM using the conditional probability generated based on relative relationships and obtained the target location RZSM estimate. My proposed model showed significantly better RZSM estimation compared to the popular related works.

Finally, in Chapter 6, all the models developed in Chapters 3-5 were integrated into an IDSS to address research problem 4, presented in Section. 1.2.4. First using the models from Chapters 4-5, the IDSS checked if the water available in the soil is below the depletion tolerance of the PI system for precise irrigation assessment. Next, if water depletion is below the tolerance threshold, irrigation was supplied after considering the water than can be available from the rain using the model from Chapter 3, reducing overirrigation. This Chapter illustrated how the models developed in Chapters 3-6 can be combined to generate improved PI schedules. The better rainfall, SSM, and RZSM input information helped significantly reduce water loss and enhanced cost savings compared to state-of-the-art related works.

Overall, this thesis has addressed all the research problems outlined in Section 1.2 by developing innovative models/methods for better rain and SM (SSM/RZSM) information extraction to aid in making more precise decisions on irrigation scheduling.

7.2 Recommendations for Future Work

Precision agriculture (PA) is a broad field, while this thesis has addressed the key research gaps in the most critical PA approach (precision irrigation), there are certainly still opportunities for future work. Suggested areas for future directions include:

1. Farm implementation trails for the developed IDSS- Conducting trials of the proposed IDSS-based irrigation scheduling in farm settings is a vital and significant step in moving towards real-world implementation. The proposed system can be uploaded to a cloud and hooked to an automatic irrigation system on the farm for doing PI. Apart from the BOM and NASA open-access data, for practical utilization of the proposed IDSS, the installation of RZSM sensing stations and information collection on farm soil/crops is required.

2. Application of drones and cameras for RZSM information- The moisture in the plant root zone is not directly measurable from existing satellites, and large-scale

in-situ sensor installation is impractical. Therefore, I developed an in-direct estimation method. However, if we are able to use nearer sources than satellites like drones and cameras it can provide more spectral details from which RZSM may be better derivable. Hence, the applicability of these sources needs to be explored.

3. Investigation of the applicability to precision fertilization- Along with water, the application of an optimum amount of required nutrients is essential for ensuring a higher quantity and quality of yield. Similar to SM, vegetation indices are also related to soil fertilization. Hence, they can be used as auxiliary information and correlate with multispectral camera/drone images for distributed fertilization measurement using the SM models developed in this thesis. However, for pursuing this research path, extensive plant nutrient data acquisition is required.

4. Examine ML techniques to reduce training time- The proposed decision support system for precision irrigation has demonstrated high accuracy, but its long training times is a limitation that needs to be addressed. To overcome this limitation and improve the proposed model, some ML areas which can be explored are: (1) transfer learning, (2) distributed training, (3) parallel processing techniques to reduce the current training time. To perform this investigation, the purchase of advanced computational resources such as cloud computing, GPUs, or other technologies may be required.

5. Investigate explainable AI techniques- To ensure widespread adoption and successful implementation of this system, researchers and practitioners need to explore ways to further improve the model's interpretability and explainability. By developing and incorporating more advanced explainable AI techniques, we can create a more transparent decision support system for precision irrigation, instilling trust and increasing the model's usability in real-world scenarios. However, there is currently no standardized framework for explainable AI, which makes it difficult to compare different approaches and results. To address this issue, researchers and practitioners should work towards developing a set of shared principles, best practices, and evaluation metrics for explainable AI.

Bibliography

[1] N. Madhukumar *et al.*, "Consensus forecast of rainfall using hybrid climate learning model," *IEEE Internet of Things Journal*, vol. 8, pp. 7270 - 7278, May 2021.

[2] N. Madhukumar *et al.*, "Three-dimensional bi-directional LSTM for satellite soil moisture downscaling," *IEEE Transactions on Geoscience and Remote Sensing*, Accepted for publication, Nov. 2022.

[3] N. Madhukumar *et al.*, "Hybrid Transformer Network for Root Zone Soil Moisture Estimation for Decision Support in Precision Irrigation," *Agricultural Water Management*. (under review)

[4] S. Humphries, "The intensification of traditional agriculture among Yucatec Maya farmers: Facing up to the dilemma of livelihood sustainability," *Hum. Ecol.*, vol. 21, pp. 87-102, Mar. 1993.

[5] V. C. Schreiner *et al.*, "Paradise lost? Pesticide pollution in a European region with considerable amount of traditional agriculture," *Water Research*, vol. 188, pp. 1-10 , Jan. 2021.

[6] A. McBratney *et al.*, "Future directions of precision agriculture," *Precision Agriculture*, vol. 6, pp.7-23, Feb. 2005.

[7] F. A. O. Solaw, "The state of the world's land and water resources for food and agriculture," *Food and Agriculture Organization of the United Nations*, Rome, Italy, Final Rep., 2011.

[8] J. R. Holden and P. J. McGuire, "Irrigation of sugarcane manual," *BSES*, vol. 1, pp 1-64, Jan. 2013.

[9] B. R. Hanson *et al.*, "A comparison of furrow, surface drip, and subsurface drip irrigation on lettuce yield and applied water," *Agricultural Water Management*, vol. 33, pp. 139-157, Jun. 1996.

[10] Department of agriculture and fisheries: Queensland Government, "Sugarcane improved practices guide: Burdekin", 2016.

[11] Y. Cheng *et al.*, "Deep belief network for meteorological time series prediction in the internet of things," *IEEE Internet of Things J.*, vol. 6, pp. 4369-4376, Jun. 2019.

[12] J. W. Kijne *et al.*, "A water-productivity framework for understanding and action," *Water productivity in agriculture: limits and opportunities for improvement*, vol. 1, pp. 1-18, Jan. 2003.

[13] E. Fereres and M. A. Soriano, "Deficit irrigation for reducing agricultural water use," *Journal of experimental botany*, vol. 58, pp.147-159, Jan. 2007.

[14] S. H. Ahmadi *et al.*, "Effects of irrigation strategies and soils on field grown potatoes: yield and water productivity," *Agricultural Water Management*, vol. 97, pp. 1923-1930, Nov. 2010.

[15] R. J. Smith *et al.*, "Review of precision irrigation technologies and their application,"*National Centre for Engineering in Agriculture (NCEA)*, University of Southern Queensland, Toowoomba, NCEA Publication 1003017/1, Nov. 2010.

[16] E. Bwambale *et al.*, "Smart irrigation monitoring and control strategies for improving water use efficiency in precision agriculture: A review,"*Agricultural Water Management*, vol. 260, pp. 1-12, Feb. 2022.

[17] Z. Gu *et al.*, "Irrigation scheduling approaches and applications: A review," *J. Irrig. Drain. Eng.*, vol. 146, pp. 1-15, Jun. 2020.

[18] A. Haidar and B. Verma, "Monthly rainfall forecasting using one-dimensional deep convolutional neural network," *IEEE Access*, vol. 6, pp. 69053-69063, Dec. 2018.

[19] Z. Gu *et al.*, "Neural network soil moisture model for irrigation scheduling," *Computers and Electronics in Agriculture*, vol. 180, pp. 1-10, Jan. 2021.

[20] M. Smith, "The application of climatic data for planning and management of sustainable rainfed and irrigated crop production," *Agricultural and Forest Meteorology*, vol. 103, pp. 99-108, Jun. 2000.

[21] P. Fox and J. Rockström, "Supplemental irrigation for dry-spell mitigation of rainfed agriculture in the Sahel," *Agricultural Water Management*, vol. 61, pp. 29-50, Jun. 2000.

[22] A. H. Sodergren *et al.*, "An energy balance model exploration of the impacts of interactions between surface albedo, cloud cover and water vapour on polar amplification," *Climate Dynamics*, vol. 51, pp. 1639-1658, Sep. 2018.

[23] E. Price *et al.*, "GPU-accelerated long-wave radiation scheme of the rapid radiative transfer model for general circulation models," *IEEE J. Selected Topics Appld. Earth Observ. & Remote Sensing*, vol. 7, pp. 3660-3667, Aug. 2014.

[24] D. Hudson *et al.*, "ACCESS-S1: The new bureau of meteorology multi-week to seasonal prediction system," *J. Southern Hemisphere Earth Syst. Sci.*, vol. 67, pp. 132-159, Dec. 2017.

[25] M. Link *et al.*, "Soil moisture information content in SMOS, SMAP, AMSR2, and ASCAT Level-1 data over selected in-situ sites," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, pp. 1213 - 1217, Jul. 2020.

[26] N. N. Das *et al.*, "Combining SMAP and Sentinel data for high-resolution soil moisture product," in *Proc. IEEE International Geoscience and Remote Sensing Symposium* (IGARSS), Beijing, China, Jul. 2016, pp. 129-131.

[27] T. Jagdhuber *et al.*, "Microwave covariation modeling and retrieval for the dual-frequency active-passive combination of sentinel-1 and SMAP, " in *Proc. IEEE International Geoscience and Remote Sensing Symposium* (IGARSS), Fort Worth, Texas, Jul. 2017, pp. 3996-3999.

[28] E. Racah *et al.*, "Extreme weather: a large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events," *Adv. Neural Inform. Process. Syst.*, vol. 30, pp. 3405–3416, Dec. 2017.

[29] S. Chatterjee *et al.*, "Characterization of field-scale soil variation using a stepwise multi-sensor fusion approach and a cost-benefit analysis," *CATENA*, vol. 201, pp. 1-18, Jun. 2021.

[30] E. Santi *et al.*, "Neural network integration of SMAP and Sentinel-1 for estimating soil moisture at high spatial resolution," in *Proc. IEEE International Geoscience and Remote Sensing Symposium IGARSS*, Brussels, Belgium, Jul. 2021, pp. 6327-6330.

[31] D. Jacob *et al.*, "EURO-CORDEX: new high-resolution climate change projections for European impact research," *Regional environmental change*, vol. 14, pp. 563-578, Apr. 2014.

[32] Y. Wang *et al.*, "LASSO and LSTM integrated temporal model for short-term solar intensity forecasting," *IEEE Internet of Things J.*, vol. 6, pp. 2933-2944, Apr. 2019.

[33] Y. Yu *et al.*, "An LSTM short-term solar irradiance forecasting under complicated weather conditions," *IEEE Access*, vol. 7, pp. 145651-145666, Oct. 2019.

[34] M. Chen *et al.*, "A two-layer nonlinear combination method for short-term wind speed prediction based on ELM, ENN, and LSTM," *IEEE Internet of Things J.*, vol. 6, pp. 6997-7010, Aug. 2019.

[35] L. He *et al.*, "An improved method for soil moisture monitoring with ensemble learning methods over the Tibetan Plateau," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 2833-2844, Feb. 2021.

[36] M. Piles *et al.*, "Downscaling SMOS-derived soil moisture using MODIS visible/infrared data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, pp. 3156-3166, Sep. 2011.

[37] P. W. Liu *et al.*, “Assessing disaggregated SMAP soil moisture products in the United States, ” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 2577-2592, Feb. 2021.

[38] D. Liu and A. K. Mishra, “Performance of AMSRE soil moisture data assimilation in CLM4.5 model for monitoring hydrologic fluxes at global scale,” *Journal of Hydrology*, vol. 547, pp. 67-79, Apr. 2017.

[39] Y. Song *et al.*, “Soil Moisture Memory and Its Effect on the Surface Water and Heat Fluxes on Seasonal and Inter-annual Time Scales,” *J. Geophys. Res. Atmos.*, vol. 124, pp. 10730–10741, Oct. 2019.

[40] Z. Wei *et al.*, “Downscaling SMAP soil moisture estimation with gradient boosting decision tree regression over the Tibetan Plateau,” *Remote Sensing of Environment*, vol. 225, pp. 30-44, Mar. 2019.

[41] F. Wen *et al.*, “A value-consistent method for downscaling SMAP passive soil moisture with MODIS products using self-adaptive window,” *IEEE Trans. Geosci. Remote Sens.*, vol. 58, pp. 913-924, Feb. 2020.

[42] G. Gabriel and V. Virginia, “Downscaling SMAP soil moisture images with wavelets transform and PCA method,” in *Proc. IEEE International Symposium of Geoscience and Remote Sensing*, Valdivia, Chile, Jun. 2017, pp. 1-4.

[43] K. Fang *et al.*, “The value of SMAP for long-term soil moisture estimation with the help of deep learning,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, pp. 2221-2233, Apr. 2019.

[44] A. Pardossi *et al.*, “Root zone sensors for irrigation management in intensive agriculture,” *Sensors*, vol. 9, pp. 2809-2835, Apr. 2009.

[45] D. Li *et al.*, “High-throughput plant phenotyping platform (HT3P) as a novel tool for estimating agronomic traits from the lab to the field,” *Front Bioeng Biotechnol*, vol. 8, pp. 1-24, Jan. 2021.

[46] W. Tiejun *et al.*, “Evaluating controls of soil properties and climatic conditions on the use of an exponential filter for converting near surface to root zone soil moisture contents,” *Journal of hydrology*, vol. 548, pp. 683-696, May 2017.

[47] Y. Jingxin *et al.*, “A hybrid CNN-GRU model for predicting soil moisture in maize root zone,” *Agricultural Water Management*, vol. 245, pp. 1-10, Feb. 2021.

[48] C. D. Salles *et al.*, “Extreme precipitation seasonal forecast using a transformer neural network,” *arXiv:2107.06846 [cs]*, Jul. 2021.

[49] P. Pokhrel *et al.*, “A transformer-based regression scheme for forecasting significant wave heights in oceans,” *IEEE J. of Oceanic Engineering*, pp. 1-14, Jul. 2022.

[50] I. Campos *et al.*, “Combining a water balance model with evapotranspiration measurements to estimate total available soil water in irrigated and rainfed vineyards, ” *Agricultural Water Management*, vol. 165, pp. 141-152, Feb. 2016.

[51] L. R. Ahuja *et al.*, “Root zone water quality model sensitivity analysis using Monte Carlo simulation,” *Transactions of the ASAE*, vol. 43, pp. 883-895, Jan. 2000.

[52] M. D. Dukes *et al.*, “Water conservation potential of landscape irrigation smart controllers,” *Trans. ASABE*, vol. 55, pp. 563–569, Dec. 2011.

[53] Z. Gu *et al.*, “Real-time precise irrigation scheduling and control system in solar greenhouse based on ET and water balance,” *Trans. Chinese Soc. Agric. Eng.*, vol. 34, pp. 101–118, Dec. 2018.

[54] X. Chen *et al.*, “Responses of cotton photosynthesis and growth to a new irrigation control method under deficit irrigation,” *Field Crops Research*, vol. 275, pp. 1-10, Jan. 2022.

[55] H. Zhang *et al.*, “LoRaWAN based internet of things (IoT) system for precision irrigation in plasticulture fresh-market Tomato,” *Smart Agricultural Technology*, vol. 2, pp. 1-9, Apr. 2022.

[56] K. Wasayangkool *et al.*, “Moisture prediction system with LSTM technique for cactus farm,” in *Proc. IEEE 18th International Conference on Electrical*

Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Mai, Thailand, Jun. 2021, pp. 144-147.

[57] ACCESS-S, Bureau of Meteorology, Australia 2021-2022. [Online]. Available: http://poama.bom.gov.au/general/access-s.html.

[58] Y. Liu *et al.*, "Application of deep convolutional neural networks for detecting extreme weather in climate datasets," in *Proc. ABDA'16-International Conference on Advances in Big Data Analytics*, pp. 81–88, May 2016.

[59] K. Ostad-Ali-Askari *et al.*, "Presenting a mathematical model for estimating the deep percolation due to irrigation," *Int. J. Hydraul. Eng.*, vol. 4, pp. 17-21, Jan. 2015.

[60] M. N. Yazdi *et al.*, "Water quality characterization of storm and irrigation runoff from a container nursery," *Science of the Total Environment*, vol. 667, pp. 166-178, Jun. 2019.

[61] M. Reichstein *et al.*, "Deep learning and process understanding for data-driven earth system science," *Nature*, vol. 566, pp. 195–204, Feb. 2019.

[62] Z. Wang *et al.*, "Deep learning for image super-resolution: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, pp. 3365-3387, Oct. 2021.

[63] D. W. Otter *et al.*, "A survey of the usages of deep learning for natural language processing," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, pp. 604-624, Feb. 2021.

[64] Y. Xu *et al.*, "Spatially explicit model for statistical downscaling of satellite passive microwave soil moisture," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, pp. 1182-1191, Feb. 2020.

[65] X. Zhipeng *et al.*, "Improved subsurface soil moisture prediction from surface soil moisture through the integration of the (de) coupling effect," *Journal of Hydrology*, vol. 608, pp. 1-12, May 2022.

[66] N. Lazarovitch *et al.*, "The root zone: Soil physics and beyond," *Vadose Zone Journal*, vol. 17, pp. 1-6, Feb. 2018.

[67] C. Coleen *et al.*, "Root zone soil moisture estimation with random forest," *Journal of hydrology*, vol. 593, pp. 1-14, Feb. 2021.

[68] J. R. Cordova *et al.*, "Physically based probabilistic models of infiltration, soil moisture, and actual evapotranspiration," *Water Resources Research*, vol. 17, pp. 93–106, Feb. 1981.

[69] P. R. Houser *et al.*, "Integration of soil moisture remote sensing and hydrologic modeling using data assimilation," *Water Resources Research*, vol. 34, pp. 3405–3420, Dec. 1998.

[70] M. Pezij *et al.*, "Applying transfer function-noise modelling to characterize soil moisture dynamics: A data-driven approach using remote sensing data," *Environmental Modelling Software*, vol. 131, pp. 1-15, Sep. 2020.

[71] J. M. Sabater *et al.*, "From near surface to root zone soil moisture using different assimilation techniques," *J. Hydrometeorol.*, vol. 8, pp. 94–206, Apr. 2007.

[72] X. Gao *et al.*, "Testing of observation operators designed to estimate profile soil moisture from surface measurements," *Hydrological Processes*, vol. 33, pp. 575–584, Nov. 2018.

[73] P. Jian and A. Loew *et al.*, "Recent advances in soil moisture estimation from remote sensing," *Water*, vol. 9, pp. 1-5, Jul. 2017.

[74] K. Zhang *et al.*, "Challenges and opportunities in precision irrigation decision-support systems for center pivots," *Environ. Res. Lett.*, vol. 16, pp. 1-26, Apr. 2021.

[75] G. P. Petropoulos *et al.*, "Surface soil moisture retrievals from remote sensing: Current status, products future trends," *Physics and Chemistry of the Earth*, vol. 83, pp. 36-56, Jan. 2015.

[76] L. S. Pereira *et al.*, "Soil water balance models for determining crop water and irrigation requirements and irrigation scheduling focusing on the FAO56 method

and the dual Kc approach, " *Agricultural Water Management*, vol. 241, pp. 1-22, Nov. 2020.

[77] E. Wang *et al.*, "Development of a closed-loop irrigation system for sugarcane farms using the internet of things, " *Computers and Electronics in Agriculture*, vol. 172, pp. 1-10, May 2020.

[78] S. Rubo and J. Zinkernagel, "Exploring hyperspectral reflectance indices for the estimation of water and nitrogen status of spinach," *Biosystems Engineering*, vol. 214, pp. 58-71, Feb. 2022.

[79] F. Viani, "Experimental validation of a wireless system for theirrigation management in smart farming applications, " *Microw.Opt. Technol. Lett.* vol. 58, pp. 2186–2189, Jun. 2016.

[80] K. Zhang *et al.*, "Estimating natural nitrous oxide emissions from the Qinghai–Tibetan Plateau using a process-based model: Historical spatiotemporal patterns and future trends," *Ecological Modelling*, vol. 466, pp. 1-12, Apr. 2022.

[81] L. Ma *et al.*, "Calibrating RZWQM2 model for maize responses to deficit irrigation," *Agric. Water Manag.*, vol. 103, pp. 140–149, Jan. 2012.

[82] R. Linker *et al.*, "Optimal model-based deficit irrigation scheduling using AquaCrop: A simulation study with cotton, potato and tomato," *Agric. Water Manag.*, vol. 163, pp. 236–243, Jan. 2016.

[83] C. O. Stockle *et al.*, "CropSyst, a cropping systems simulation mode," *Eur J Agron*, vol. 18, pp. 289-307, Jan. 2003.

[84] B. A. Keating *et al.*, "An overview of APSIM, a model designed for farming systems simulation," *Eur. J. Agron.*, vol. 18, pp. 267-288, Jan. 2003.

[85] M. Šťastná and Z. Žalud, "Sensitivity analysis of soil hydrologic parameters for two crop growth simulation models, " *Soil Till. Res.*, vol. 50, pp. 305-318, May 1999.

[86] B. Gabrielle *et al.*, “A priori parameterisation of the CERES soil-crop models and tests against several European data sets,” *Agronomie*, vol. 22, pp. 119-132, Oct. 2021.

[87] W. Ferrel, “An essay on the winds and currents of the ocean,” *Nashville J. Med. Surg.*, vol. 11, pp. 287–301, 1856.

[88] P. N. Edwards, “History of climate modeling,” *Wiley Interdisciplinary Reviews: Climate Change*, vol. 2, pp. 128-139, Jan. 2011.

[89] S. Wolfert *et al.*, “Big data in smart farming: A review,” *Agricultural Systems*, vol. 153, pp. 69-80, May 2017.

[90] A. Kamilaris *et al.*, “Deep learning in agriculture: A survey,” *Computers and Electronics in Agriculture*, vol. 147, pp. 70-90, Apr. 2018.

[91] N. Ganatra, “Deep learning methods and applications for precision agriculture, ” *Machine Learning for Predictive Analysis*, vol. 141, pp. 1-12, Oct. 2020.

[92] P. K. Kashyap *et al.*, “Towards precision agriculture: IoT-enabled intelligent irrigation systems using deep learning neural network, ” *IEEE Sensors Journal*, vol. 21, pp. 17479-17491, Aug. 2021.

[93] C. Wang *et al.*, “A review of deep learning used in the hyperspectral image analysis for agriculture,” *Artif. Intell. Rev.*, vol. 54, pp. 5205–5253, May 2021.

[94] A. K. Choudhary *et al.*, “Drought analysis using standardized precipitation index and supplemental irrigation planning in Banka, Bihar,” *Hydrological Modeling*, pp. 125-138, Jan. 2022.

[95] P. Zhang *et al.*, “Short-term rainfall forecasting using multi-layer perceptron,” *IEEE Transactions on Big Data*, vol. 6, pp. 93-106, Mar. 2020.

[96] S. Aswin *et al.*, “Deep learning models for the prediction of rainfall,” in *Proc. IEEE 2018 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, Apr. 2018, pp. 657-661.

[97] K. McGuffie and A. Henderson-Sellers, "Forty years of numerical climate modelling," *International Journal of Climatology: A Journal of the Royal Meteorological Society*, vol. 21, pp. 1067-1109, Jul. 2001.

[98] G. Flato *et al.*, "Evaluation of climate models," *Climate change 2013: The physical science basis*, Cambridge, United Kingdom, C.U.P, ch. 9, sec. 9.1, pp. 746-754.

[99] E. Moreno-Chamarro *et al.*, "Impact of increased resolution on long-standing biases in HighResMIP-PRIMAVERA climate models," *Geoscientific Model Development*, vol. 15, pp. 269-289, Jan. 2022.

[100] E. B. Wetherley *et al.*, "Remote sensing and energy balance modeling of urban climate variability across a semi-arid megacity," *Urban Climate*, vol. 35, pp. 1-21, Jan. 2021.

[101] W. C. Wang and P. H. Stone, "Effect of ice-albedo feedback on global sensitivity in a one-dimensional radiative-convective climate model," *J. Atmos. Sci.*, vol. 37, pp. 545-52, Mar. 1980.

[102] P. McIntosh *et al.*, "Improving seasonal forecasts for SWWA, " *Final report on the managing climate variability project*, GRDC, pp. 1-59, Mar. 2011.

[103] R. Ranjan *et al.*, "Deep learning for understanding faces: machines may be just as good, or better, than humans," *IEEE Signal Process. Mag.*, vol. 35, pp. 66-83, Jan. 2018.

[104] J. Abbot and J. Marohasy, "Application of artificial neural networks to rainfall forecasting in Queensland, Australia," *Advances in Atmospheric Sciences*, vol. 29, pp. 717-730, Jan. 2012.

[105] K. Kawaguchi and L. Kaelbling, "Elimination of all bad local minima in deep learning," in *Proc. International Conference on Artificial Intelligence and Statistics*, Online, Italy, Jun. 2020, pp. 853-863.

[106] S. Lawrence and C. L. Giles, "Overfitting and neural networks: conjugate gradient and backpropagation," in *Proc. IEEE-INNS-ENNS International Joint Conference on Neural Networks*, Como, Italy, Jul. 2000, pp. 114-119.

[107] M. Cogswell *et al.*, "Reducing overfitting in deep networks by decorrelating representations," in *Proc. International Conference on Learning Representations*, San Diego, CA, Nov. 2015, pp. 1-12.

[108] H. Song *et al.*, "Learning from noisy labels with deep neural networks: A survey," *IEEE Transactions on Neural Networks and Learning Systems*, (early access 2022).

[109] M. P. McGuire *et al.*, "Prediction of tornado days in the United States with deep convolutional neural networks," *Computers Geosciences* , vol. 159, pp.1-10 , Feb. 2022.

[110] S. Rasheed *et al.*, "High-resolution remote sensing image classification through deep neural network," in *Proc. IEEE 2021 International Conference on Digital Futures and Transformative Technologies*, Islamabad, Pakistan, May 2021, pp. 1-6.

[111] N. Kriegeskorte and T. Golan, "Neural network models and deep learning," *Current Biology*, vol. 29, pp. 231-236, Apr. 2019.

[112] L. Ali *et al.*, "Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures," *Sensors*, vol. 21, pp. 1-22, Mar. 2022.

[113] K. Srinivasan *et al.*, "Performance comparison of deep CNN models for detecting driver's distraction," *Computers, Materials Continua*, vol. 68, pp. 4109-4124, Jun. 2021.

[114] S. P. Van *et al.*, "Deep learning convolutional neural network in rainfall–runoff modelling," *Journal of Hydroinformatics*, vol. 22, pp. 541-561, Apr. 2020.

[115] B. Athiwaratkun and J. W. Stokes, "Malware classification with LSTM and GRU language models and a character-level CNN," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, New Orleans, LA, Jun. 2017, pp. 2482-2486.

[116] G. F. Lin and M. C. Wu, "A hybrid neural network model for typhoon-rainfall forecasting," *J. of Hydrology*, vol. 375, pp. 450-458, Sep. 2009.

[117] P. Zhang *et al.*, "A deep belief network based precipitation forecast approach using multiple environmental factors," *Intelligent Data Analysis*, vol. 22, pp. 843-866, Jun. 2018.

[118] F. Mekanik *et al.*, "Multiple regression and artificial neural network for long-term rainfall forecasting using large scale climate modes," *Journal of Hydrology*, vol. 503, pp. 11-21, Oct. 2013.

[119] X. Shi *et al.*, "Deep learning for precipitation nowcasting: A benchmark and a new model," *Adv. in Neural Info. Processing Systems*, vol. 30, pp. 5622-5632, Dec. 2017.

[120] S. Ravuri *et al.*, "Skilful precipitation nowcasting using deep generative models of radar," *Nature*, vol. 597, pp. 672-677, Sep. 2021.

[121] A. Koesdwiady *et al.*, "Improving traffic flow prediction with weather information in connected cars: a deep learning approach," *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 9508-9517, Dec. 2016.

[122] W. Wu and M. Peng, "A data mining approach combining *K*-means clustering with bagging neural network for short-term wind power forecasting," *IEEE Internet of Things J.*, vol. 4, pp. 979-986, Aug. 2017.

[123] S. Mark *et al.*, "Estimating root zone soil moisture at distant sites using MODIS NDVI and EVI in a semi-arid region of southwestern USA," *Ecological Informatics*, vol. 5, pp. 400-409, May. 2010.

[124] J. Tian *et al.*, "Improving surface soil moisture retrievals through a novel assimilation algorithm to estimate both model and observation errors," *Remote Sensing of Environment*, vol. 269, pp. 1-13, Feb. 2022.

[125] L. Xu *et al.*, "In-situ and triple-collocation based evaluations of eight global root zone soil moisture products," *Remote Sensing of Environment*, vol. 254, pp. 1-16, Mar. 2021.

[126] J. Peng *et al.*, "A review of spatial downscaling of satellite remotely sensed soil moisture," *Rev. Geophys.*, vol. 55, pp. 341–366, Jun. 2017.

[127] T. W. Collow *et al.*, "Evaluation of SMOS retrievals of soil moisture over the central United States with currently available in situ observations," *J. Geophys. Res., Atmos.*, vol. 117, pp. 1-15, May 2012.

[128] S. S. Chai *et al.*, "Use of soil moisture variability in artificial neural network retrieval of soil moisture," *Remote Sens.*, vol. 2, pp. 166–190, Jan. 2010.

[129] G. Baroni *et al.*, "The role of vegetation and soil properties on the spatio-temporal variability of the surface soil moisture in a maize-cropped field," *Journal of Hydrology*, vol. 489, pp. 148-159, May 2013.

[130] Y. Qi *et al.*, "Seismic microwave brightness temperature anomaly detection using multitemporal passive microwave satellite images: Ideas and limits," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 6792-6806, Jun. 2021.

[131] T. Kawanishi *et al.*, "The Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E), NASDA's contribution to the EOS for global energy and water cycle studies," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 184-194, Apr. 2003.

[132] Z. Bartalis *et al.*, "Initial soil moisture retrievals from the METOP-A Advanced Scatterometer," *Geophysical Research Letters*, vol. 34, pp. 1-5, Oct. 2007.

[133] A. Okuyama and K. Imaoka, "Intercalibration of advanced microwave scanning radiometer-2 (AMSR2) brightness temperature," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, pp. 4568-4577, Mar. 2015.

[134] Y. H. Kerr *et al.*, "The SMOS mission: New tool for monitoring key elements of the global water cycle," *Proceedings of the IEEE*, vol. 98, pp. 666-687, May 2010.

[135] R. Zhang *et al.*, "A comprehensive validation of the SMAP enhanced level-3 soil moisture product using ground measurements over varied climates and landscapes," *Remote sensing of environment*, vol. 223, pp. 82-94, Mar. 2019.

[136] J. Zhang *et al.*, "Evaluation of the ASCAT surface soil moisture product for agricultural drought monitoring in USA," in *Proc. IEEE International Geoscience and Remote Sensing Symposium*, Milan, Italy, Jul. 2015, pp. 669-672.

[137] T. J. Jackson *et al.*, "Validation of advanced microwave scanning radiometer soil moisture products," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, pp. 4256-4272, Dec. 2010.

[138] R. Bindlish *et al.*, "AMSR2 soil moisture product validation," in *Proc. IEEE International Geoscience and Remote Sensing Symposium*, Fort Worth, TX, Jul. 2017, pp. 5637-5640.

[139] R. Bindlish *et al.*, "Global soil moisture from the Aquarius/SAC-D satellite: description and initial assessment," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, pp. 923-927, May 2015.

[140] H. Zhang *et al.*, "NDVI dynamic changes and their relationship with meteorological factors and soil moisture," *Environmental Earth Sciences*, vol. 77, pp. 1-11, Aug. 2018.

[141] M. Piles *et al.*, "Remote sensing of vegetation dynamics in agro-ecosystems using smap vegetation optical depth and optical vegetation indices," in *Proc. IEEE International Geoscience and Remote Sensing Symposium*, TX, USA, Jul. 2017, pp. 4346-4349.

[142] H. Kim *et al.*, "Approximate estimation of soil moisture from NDVI and land surface temperature over Andong region, Korea, " *Korean Journal of Remote Sensing*, vol. 30, pp. 375-381, Jun. 2014.

[143] W. Ha *et al.*, "A review of downscaling methods for remote sensing-based irrigation management: Part I," *Irrigation Science*, vol. 31, pp. 831-850, Jul. 2013.

[144] Y. Yong *et al.*, "Ensemble downscaling of soil moisture data using BMA and ATPRK, " *Korean Journal of Remote Sensing*, vol.36, pp.587-607, Aug. 2020.

[145] T. Yu *et al.*, "Deep precipitation downscaling," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, Jan. 2021.

[146] A. B. Abbes *et al.*, "Soil moisture estimation from SMAP observations using long short-term memory (LSTM)," in *Proc. IEEE International Geoscience and Remote Sensing Symposium*, Yokohama, Japan, Nov. 2019, pp. 1590-1593.

[147] N. Das *et al.*, "The SMAP and copernicus sentinel-1 A/B microwave active-passive high resolution surface soil moisture product, " *Remote Sens. Environ.*, vol. 233, pp. 1-17, Nov. 2019.

[148] D. Liu and A. K. Mishra, "Performance of AMSRE soil moisture data assimilation in CLM4.5 model for monitoring hydrologic fluxes at global scale," *Journal of Hydrology*, vol. 547, pp. 67-79, Apr. 2017.

[149] Y. Song *et al.*, "Soil moisture memory and its effect on the surface water and heat fluxes on seasonal and inter-annual time scales," *J. Geophys. Res. Atmos.*, vol. 124, pp. 10730–10741, Oct. 2019.

[150] Q. Li *et al.*, "GANs-LSTM model for soil temperature estimation from meteorological: a new approach," *IEEE Access*, vol. 8, pp. 59427-59443, Mar. 2020.

[151] M. Piles *et al.*, "Downscaling SMOS-derived soil moisture using MODIS visible/infrared data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, pp. 3156-3166, Sep. 2011.

[152] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on Information Theory*, vol. 14, pp. 55-63, Jan. 1968.

[153] G. Gabriel and V. Virginia, "Downscaling SMAP soil moisture images with wavelets transform and PCA method," in *Proc. IEEE International Symposium of Geoscience and Remote Sensing*, Valdivia, Chile, Jun. 2017, pp. 1-4.

[154] L. Fang *et al.*, "Extinction profiles fusion for hyperspectral images classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, pp. 1803–1815, Mar. 2018.

[155] W. Xu *et al.*, "Downscaling SMAP soil moisture products with convolutional neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 4051-4062, Mar. 2021.

[156] P. K. Srivastava *et al.*, "Machine learning techniques for downscaling SMOS satellite soil moisture using MODIS land surface temperature for hydrological application," *Water Resour. Manage.*, vol. 27, pp. 3127–3144, Jun. 2013.

[157] J. Im *et al.*, "Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches," *Environ. Earth Sci.*, vol. 75, pp. 1-19, Jul. 2016.

[158] F. Wen *et al.*, "A value-consistent method for downscaling SMAP passive soil moisture with MODIS products using self-adaptive window," *IEEE Trans. Geosci. Remote Sens*., vol. 58, pp. 913-924, Feb. 2020.

[159] A. J. Rigden *et al.*, "Combined influence of soil moisture and atmospheric evaporative demand is important for accurately predicting us maize yields," *Nature Food*, vol. 1, pp. 127–133, Feb. 2020.

[160] W. Dorigo *et al.*, "Global automated quality control of in situ soil moisture data from the international soil moisture network," *Vadose Zone J.*, vol. 12, pp. 1-21, Jan. 2013.

[161] H. Vereecken *et al.*, "On the value of soil moisture measurements in vadose zone hydrology: A review Water," *Resources Research*, vol. 44, pp. 1-21, Apr. 2008.

[162] P. Dobriyal *et al.*, "A review of the methods available for estimating soil moisture and its implications for water resource management," *Journal of Hydrology*, vol. 458, pp. 110–117, Aug. 2012.

[163] A. Colliander *et al.*, "Validation of SMAP surface soil moisture products with core validation sites," *Remote Sens. Environ.*, vol. 191, pp. 215-231, Mar. 2017.

[164] P. O'Neill *et al.*, "SMAP L3 radiometer global daily 36 km EASE-grid soil moisture," *Natl. Snow Ice Data Cent*, version. 5, pp. 1-82, 2018.

[165] D. Entekhabi *et al.*, "The Soil Moisture Active Passive (SMAP) Mission," *Proceedings of the IEEE*, vol. 98, pp. 704-716, May 2010.

[166] S. Yueh *et al.*, "NASA Soil Moisture Active Passive mission status and science performance"in *Proc. IEEE International Geoscience and Remote Sensing Symposium*, Beijing, China, Jul. 2016, pp. 116-119.

[167] T. J. Jackson *et al.*, "Validation of advanced microwave scanning radiometer soil moisture products," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, pp. 4256-4272, Jun. 2010.

[168] W.T. Crow *et al.*, "Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products," *Rev. Geophys.*, vol. 50, pp. 1-20, Apr. 2012.

[169] L. Fang *et al.*, "An inter-comparison of soil moisture data products from satellite remote sensing and a land surface model," *International journal of applied earth observation and geoinformation*, vol. 48, pp. 37-50, Jun. 2016.

[170] F. Chen *et al.*, "Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture," *Adv. Water Resour*, vol. 34, pp. 526–536, Jan. 2011.

[171] J. Munoz-Sabater *et al.*, "From near-surface to root-zone soil moisture using different assimilation techniques," *J. Hydrometeorol*, vol. 8, pp. 194–206, Apr. 2007.

[172] F. T. Ulaby *et al.*, "Radar mapping of surface soil moisture," *Journal of Hydrology*, vol. 184, pp. 57–84, Oct. 1996.

[173] F. Chen *et al.*, "Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture," *Adv. Water Resour.*, vol. 34, pp. 526-536, Apr. 2011.

[174] J. Muñoz-Sabater *et al.*, "From near-surface to root-zone soil moisture using different assimilation techniques," *J. Hydrometeorol.*, vol. 8, pp. 194-206, Apr. 2007.

[175] W. Wagner *et al.*, "A method for estimating soil moisture from ers scatterometer and soil data," *Remote Sensing of Environment*, vol. 70, pp. 191-207, Nov. 1999.

[176] C. Albergel *et al.*, "From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations," *Hydrol. Earth Syst. Sci.*, vol. 12, pp. 1-15,Dec. 2008.

[177] R. Zhuang *et al.*, "Quantifying long-term land surface and root zone soil moisture over tibetan plateau, " *Remote Sensing*, vol. 12, pp. 1-20, Feb. 2020.

[178] G. Bonaccorso, *Mach. Learn. Alg.*, Birmingham, U.K.: Packt, 2017.

[179] J. Shiri *et al.*, "Using soil easily measured parameters for estimating soil water capacity: Soft computing approaches," *Computers and Electronics in Agriculture*, vol. 141, pp. 327-339, Sep. 2017.

[180] M. Ali *et al.*, "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," *Journal of hydrology*, vol. 576, pp. 164-184, Sep. 2019.

[181] F. Kratzert *et al.*, "Nearing Toward improved predictions in ungauged basins: Exploiting the power of machine learning," *Water Resources Research*, vol. 55, pp. 11344-11354, Nov. 2019.

[182] R. C. Spear *et al.*, "An example of augmenting regional sensitivity analysis using machine learning software," *Water Resources Research*, vol. 56, pp. 1-16, Apr. 2020.

[183] A.T. Teweldebrhan *et al.*, "Coupled machine learning and the limits of acceptability approach applied in parameter identification for a distributed hydrological model," *Hydrology and Earth System Sciences*, vol. 24, pp. 4641-4658, Apr. 2020.

[184] D. Shrestha *et al.*, "A novel approach to parameter uncertainty analysis of hydrological models using neural networks," *Hydrology and Earth System Sciences*, vol. 13 , pp. 1-14, Jul. 2009.

[185] N. Kayastha *et al.*, "Prediction of hydrological models uncertainty by a committee of machine learning-models," in *Proc. 11th International Conference on Hydroinformatics*, New York, USA, pp. 1-5, Jan. 2014.

[186] K. C. Kornelsen *et al.*, "Root-zone soil moisture estimation using data-driven methods," *Water Resources Research*, vol. 50, pp. 2946-2962, Mar. 2014.

[187] S. Ahmad *et al.*, "Estimating soil moisture using remote sensing data: A machine learning approach," *Advances in Water Resources*, vol. 33, pp. 69-80, Jan. 2010.

[188] P. K. Srivastava *et al.*, "Machine learning techniques for downscaling smos satellite soil moisture using modis land surface temperature for hydrological application," *Water Resources Management*, vol. 27, pp. 3127-3144, Apr. 2013.

[189] D. Zhang *et al.*, "Upscaling of surface soil moisture using a deep learning model with viirs rdr, " *ISPRS International Journal of Geo-Information*, vol. 6, pp. 130-150, Apr. 2017.

[190] R. Prasad *et al.*, "Soil moisture forecasting by a hybrid machine learning technique: Elm integrated with ensemble empirical mode decomposition," *Geoderma.*, vol. 330, pp. 136-161, Nov. 2018.

[191] O. Matei *et al.*, "A data mining system for real time soil moisture prediction," *Procedia Engineering*, vol. 181, pp. 837-844, Feb. 2017.

[192] E. J. Coopersmith *et al.*, "Using machine learning to produce near surface soil moisture estimates from deeper in situ records at us climate reference network (uscrn) locations: Analysis and applications to amsr-e satellite validation," *Advances in Water Resources*, vol. 98, pp. 122-131, Dec. 2016.

[193] F. Karandish and J. Šimnek, "A comparison of numerical and machine-learning modeling of soil water content with limited input data," *Journal of Hydrology*, vol. 543, pp. 892-909, Dec. 2016.

[194] N. Tang *et al.*, "Solar power generation forecasting with a LASSO-based approach," *IEEE Internet Things J.*, vol. 5, pp. 1090-1099, Apr. 2018.

[195] R. S. Sangari and M. Balamurugan, "A survey on rainfall prediction using data mining," *Int. J. Comput. Sci. & Mobile Application*, vol. 2, pp. 84-88, Feb. 2014.

[196] G. Xu and V. Chandrasekar, "Operational feasibility of neural-network-based radar rainfall estimation," *IEEE Geosci. & Remote Sens. Lett.*, vol. 2, pp. 13-17, Jan. 2005.

[197] Y. Tu *et al.*, "Speech enhancement based on teacher–student deep learning using improved speech presence probability for noise-robust speech recognition," *IEEE/ACM Trans. Audio, Speech, & Language Processing*, vol. 27, pp. 2080-2091, Dec. 2019.

[198] F. Beritelli *et al.*, "Rainfall estimation based on the intensity of the received signal in a LTE/4G mobile terminal by using a probabilistic neural network," *IEEE Access*, vol. 6, pp. 30865-30873, Jun. 2018.

[199] K. C. Luk *et al.*, "An application of artificial neural networks for rainfall forecasting," *Math. Comput. Model.*, vol. 33, pp. 683-693, Apr. 2001.

[200] The Bureau of Meteorology, *Meteorological Verification Data Technical Reference*, Melbourne, Australia, 2015. Accessed on: Feb. 28, 2020. [Online]. Available: https://data.gov.au/search?q=Rainfall%20and%20temperature%20forecast%20and%20observations.

[201] J. Yang *et al.*, "Feature selection for MLP neural network: The use of random permutation of probabilistic outputs," *IEEE Trans. on Neural Networks*, vol. 20, pp. 1911-1922, Dec. 2009.

[202] J. Yang and C. Ong, "Feature selection using probabilistic prediction of support vector regression," *IEEE Trans. on Neural Networks*, vol. 22, pp. 954-962, Jun. 2011.

[203] A. Graves *et al.*, "Hybrid speech recognition with deep bidirectional LSTM," *IEEE* Workshop on automatic speech recognition and understanding, Olomouc, pp. 273-278, Dec. 2013.

[204] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, pp. 281–305, Feb. 2012.

[205] Y. A. Liou and A. W. England, “A land surface process/radiobrightness model with coupled heat and moisture transport in soil,” *IEEE Trans. Geosci. Remote Sens.*, vol. 36, pp. 273-286, Jan. 1998.

[206] T. J. Jackson and T. J. Schmugge, “Passive microwave remote sensing system for soil moisture: Some supporting research,” *IEEE Trans. Geosci. Remote Sens.*, vol. 27, pp. 225–235, Mar. 1989.

[207] S. Oprea *et al.*, “A review on deep learning techniques for video prediction,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, pp. 2806-2826, Dec. 2020.

[208] K. Zhang *et al.*, “Prediction of 3-D ocean temperature by multilayer convolutional LSTM,” *IEEE Geosci. Remote Sens. Lett.*, vol. 17, pp. 1303-1307, Aug. 2020.

[209] H. Zang *et al.*, “Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatio-temporal correlations,” *Renewable Energy*, vol. 160, pp. 26-41, Nov. 2020.

[210] E. Lima *et al.*, “Learning and transferring convolutional neural network knowledge to ocean front recognition,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, pp. 354-358, Mar. 2017.

[211] T. Akilan *et al.*, “A 3D CNN-LSTM-based image-to-image foreground segmentation,” *IEEE Trans. Intell. Transp. Syst.*, vol. 21, pp. 959-971, Mar. 2020.

[212] G. Wang *et al.*, “Light field multi-view video coding with two-directional parallel inter-view prediction,” *IEEE Trans. Image Process.*, vol. 25, pp. 5104-5117, Nov. 2016.

[213] J. Zhang *et al.*, “Evaluation of the ASCAT surface soil moisture product for agricultural drought monitoring in USA,” in *Proc. IEEE International Geoscience and Remote Sensing Symposium*, Milan, Italy, Jul. 2015, pp. 669-672.

[214] T. J. Jackson *et al.*, “Validation of advanced microwave scanning radiometer soil moisture products,” *IEEE Trans. Geosci. Remote Sens.*, vol. 48, pp. 4256-4272, Dec. 2010.

[215] R. Bindlish *et al.*, "AMSR2 soil moisture product validation," in *Proc. IEEE International Geoscience and Remote Sensing Symposium* (IGARSS), Fort Worth, TX, Jul. 2017, pp. 5637-5640.

[216] R. Bindlish *et al.*, "Global soil moisture from the Aquarius/SAC-D satellite: Description and initial assessment," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, pp. 923-927, May 2015.

[217] M. Piles *et al.*, "Downscaling SMOS-derived soil moisture using MODIS visible/infrared data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, pp. 3156-3166, Sep. 2011.

[218] Q. Li *et al.*, "GANs-LSTM model for soil temperature estimation from meteorological: A new approach," *IEEE Access*, vol. 8, pp. 59427-59443, Mar. 2020.

[219] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on Information Theory*, vol. 14, pp. 55-63, Jan. 1968.

[220] L. Fang *et al.*, "Extinction profiles fusion for hyperspectral images classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, pp. 1803–1815, Mar. 2018.

[221] W. Xu *et al.*, "Downscaling SMAP soil moisture products with convolutional neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 4051-4062, Mar. 2021.

[222] P. K. Srivastava *et al.*, "Machine learning techniques for downscaling SMOS satellite soil moisture using MODIS land surface temperature for hydrological application," *Water Resour. Manage.*, vol. 27, pp. 3127–3144, Jun. 2013.

[223] J. Im *et al.*, "Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches," *Environ. Earth Sci.*, vol. 75, p. 1120, Jul. 2016.

[224] K. Jordahl, "Geopandas: Python tools for geographic data," *Web: https://github.com/geopandas/geopandas*, 2014.

[225] D. J. Marchette and C. E. Priebe, "The adaptive kernel neural network," *Mathematical and Computer Modelling*, vol. 14, pp. 328-333, Jan. 1990.

[226] A. Colliander *et al.*, "SMAP/In Situ core validation site land surface parameters match-up data version 1," Boulder, Colorado USA, NASA national snow and ice data center distributed active archive center, 2017.

[227] M. Moghaddam *et al.*, "Soil moisture profiles and temperature data from SoilSCAPE sites, USA," ORNL DAAC, Oak Ridge, Tennessee, USA, 2016.

[228] H. McNairn *et al.*, "SMAPVEX16 Manitoba station soil moisture data version 1, "Boulder, Colorado USA, NASA national snow and ice data center distributed active archive center, 2016.

[229] M. Cosh *et al.*, "SMAPVEX19-21 Massachusetts Temporary Soil Moisture Network version 1," Boulder, Colorado USA, NASA national snow and ice data center distributed active archive center, 2021.

[230] D. Sun and R. T. Pinker, "Case study of soil moisture effect on land surface temperature retrieval," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, pp. 127-130, Apr. 2004.

[231] W. Hively *et al.*, "Mapping crop residue and tillage intensity using worldview-3 satellite shortwave infrared residue indices," *Remote Sens.*, vol. 10, pp. 1-22, Oct. 2018.

[232] M. A. Edwards *et al.*, "Evaluating endemic ecosystem with the aid of optical remote sensing and geographical information systems (GIS) techniques," in *Proc. 6th FIG Regional Conference*, San José, Costa Rica, pp. 1-10, Nov. 2007.

[233] P. Ceccato *et al.*, "Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1: Theoretical approach," *Remote Sensing of Environment*, vol. 82, pp. 188-197, Oct. 2002.

[234] M. H. Cosh *et al.*, "Estimating vegetation water content during the soil moisture active passive validation experiment in 2016," *J. Applied Remote Sensing*, vol. 13, pp. 1-12, Feb. 2019.

[235] H. Sahlaoui *et al.*, "Predicting and interpreting student performance using ensemble models and shapley additive explanations," *IEEE Access*, vol. 9, pp. 152688-152703, Nov. 2021.

[236] R. Ragab, "Towards a continuous operational system to estimate the root-zone soil moisture from intermittent remotely sensed surface moisture," *Journal of Hydrology*, vol. 173, pp. 1-25, Dec. 1995.

[237] S. Azadeh *et al.*, "Developing pedotransfer functions using Sentinel-2 satellite spectral indices and machine learning for estimating the surface soil moisture," *Journal of Hydrology*, vol. 606, pp. 1-17, Mar. 2022.

[238] F. T. Ulaby *et al.*, "Radar mapping of surface soil moisture," *Journal of Hydrology*, vol. 184, pp.57–84, Oct. 1996.

[239] Z. Mehrez and M. Dechambre, "A new empirical model to retrieve soil moisture and roughness from C-band radar data," *Remote Sensing of Environment*, vol. 84, pp. 42-52, Jan. 2003.

[240] E. Dara and I. Rodriguez-Iturbe., "Analytical framework for the characterization of the space-time variability of soil moisture," *Advances in water resources*, vol. 17, pp. 35-45, Jan. 1994.

[241] S. Morteza *et al.*, "An analytical model for estimation of land surface net water flux from near-surface soil moisture observations," *Journal of hydrology*, vol. 570, pp. 26-37, Mar. 2019.

[242] A. Ritter *et al.*, "Using inverse methods for estimating soil hydraulic properties from field data as an alternative to direct methods," *Agricultural Water Management*, vol. 59, pp. 77–96, Mar. 2003.

[243] D. J. Wilson *et al.*, "A terrain and data-based method for generating the spatial distribution of soil moisture," *Advances in Water Resources*, vol. 28, pp. 43-54, Jan. 2005.

[244] A. Mustafa, "Modelling the root zone soil moisture using artificial neural networks, a case study," *Environmental Earth Sciences*, vol. 75, pp. 1-12, Jul. 2016.

[245] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feed-forward neural networks," in *Proc. of the thirteenth international conference on artificial intelligence and statistics*, Sardinia Italy, May 2010, pp. 249–256.

[246] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85–117, Jan. 2015.

[247] L. Xunjian and Y. Cui, "Spatially downscaling a global evapotranspiration product for end user using a deep neural network: A case study with the GLEAM product," *Remote Sensing*, vol. 14, pp. 1-13, Jan. 2022.

[248] Z. Jing *et al.*, "Convolutional neural networks for water content classification and prediction with ground penetrating radar," *IEEE Access*, vol. 7, pp. 185385-185392, Dec. 2019.

[249] B. Rohit *et al.*, "A novel approach of transfer learning for satellite image classification," *Proceedings of Data Analytics and Management*, vol. 90, pp. 379-392, Jan. 2022.

[250] O. Adeyemi *et al.*, "Dynamic neural network modelling of soil moisture content for predictive irrigation scheduling," *Sensors*, vol. 18, pp. 1-22, Oct. 2018.

[251] B. Lim *et al.*, "Temporal fusion transformers for interpretable multi-horizon time series forecasting," *International Journal of Forecasting*, vol. 37, pp. 1748–1764, Dec. 2021.

[252] The Victorian government city of Melbourne open data portal, *Soil Sensor Readings*, Melbourne, Australia, 2021. Accessed on: Jan. 28, 2022. [Online]. Available: https://data.melbourne.vic.gov.au/Environment/Soil-Sensor-Readings/mv4n-8k4v.

[253] E. Gasca *et al.*, "Eliminating redundancy and irrelevance using a new MLP-based feature selection method," *Pattern Recognition*, vol. 39, pp. 313-315, Feb. 2006.

[254] R. Rajendran *et al.*, "Hybrid optimization algorithm based feature selection for mammogram images and detecting the breast mass using multilayer perceptron classifier," *Computational Intelligence*, vol. 38, pp. 1559-1593, Aug. 2022.

[255] The Victorian government city of Melbourne open data portal, *Soil Sensor Locations*, Melbourne, Australia, 2021. Accessed on: Jan. 28, 2022. [Online]. Available: https://data.melbourne.vic.gov.au/Environment/Soil-Sensor-Locations/dxvp-runr.

[256] The Victorian government city of Melbourne open data portal, *Soil types by area (Urban Forest)*, Melbourne, Australia, 2021. Accessed on: Jan. 28, 2022. [Online]. Available: https://data.melbourne.vic.gov.au/Environment/Soil-types-by-area-Urban-Forest-/t3zn-qgfx.

[257] F. O. Nachtergaele *et al.*, "New developments in soil classification: world reference base for soil resources," *Geoderma*, vol. 96, pp. 345-357, Jul. 2000.

[258] J. D Olden and D. A. Jackson, "Illuminating the "black box": A randomization approach for understanding variable contributions in artificial neural networks," *Ecological Modelling*, vol. 154, pp. 135-150, Aug. 2002.

[259] N. Aloysius *et al.*, "Incorporating relative position information in transformer-based sign language recognition and translation," *IEEE Access*, vol. 9, pp. 145929-145942, Oct. 2021.

[260] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Magazine*, vol. 13, pp. 47-60, Nov. 1996.

[261] J. B. Basara and K. C. Crawford, "Linear relationships between root-zone soil moisture and atmospheric processes in the planetary boundary layer," *Journal of Geophysical Research: Atmospheres*, vol. 107, pp. 1-18, Aug. 2002.

[262] T. B. Sampathkumar, "Soil moisture distribution and root characters as influenced by deficit irrigation through drip system in cotton–maize cropping sequence," *Agricultural water management*, vol. 103, pp. 43-53, Jan. 2012.

[263] E. M. O'Loughlin and E. K. S. Nambiar, " Plantations, farm forestry, and water: A discussion paper," in *Proc. of National Workshop*, Melbourne, Australia, Jun. 2000, pp. 1–25.

[264] J. Brenner *et al.*, "Chapter 3, Natural resources conservation service," *National Engineering Handbook: Irrigation Guide*, United States Department of Agriculture, vol. 652, pp. 56-87, Sep. 1997.

[265] J. W. Laurence *et al.*, "Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging," *Monthly Weather Rev.*, vol. 4, pp. 1364–1385, Apr. 2007.

[266] C. F. Chen *et al.*, "Crossvit: cross-attention multi-scale vision transformer for image classification," in *Proc. IEEE/CVF International Conference on Computer Vision*, Montreal, QC, Canada, Oct. 2021, pp. 347-356.

[267] E. Vories *et al.*, "Comparison of precision and conventional irrigation management of cotton and impact of soil texture," *Precision Agriculture*, vol. 22, pp. 414-431, Apr. 2021.

[268] E. D. Vories and S. R. Evett, "Irrigation challenges in the sub-humid US Mid-South," *International Journal of Water*, vol. 8, pp. 259–274, Aug. 2014.

[269] R. Sui *et al.*, "Effect of nitrogen application rate on yield and quality in irrigated and rainfed cotton," *Journal of Cotton Science*, vol. 21, pp. 113–121, Jan. 2017.

[270] E. D. Vories *et al.*, "Estimating the impact of delaying irrigation for Midsouth cotton on clay soil," *Transactions of the ASABE*, vol. 50, pp. 929–937, Feb. 2007.

[271] G.A. Lehrsch *et al.*, "Irrigation-Induced," in *Reference module in Earth Systems and Environmental Sciences*, Elsevier Ltd., Oxford, UK, 2014.

[272] G. Megersa and J. Abdulahi *et al.*, "Irrigation system in Israel: A review," *International Journal of water resources and environmental engineering*, vol. 7, pp. 29-37, Mar. 2015.

[273] C. B. Hedley *et al.*, "Water: Advanced irrigation technologies," in *Encyclopedia of Agriculture and Food Systems*, Elsevier Ltd., Oxford, UK, 2014.

[274] D. L. Bjorneberg *et al.*, "Irrigation methods," in *Encyclopedia of Soils in the Environment*, Elsevier Ltd., Oxford, UK, 2005.

[275] D. B. Lobell *et al.*, “Crop yield gaps: Their importance, magnitudes, and causes,” *Annual Review of Environment and Resources*, vol. 34, pp. 179–204, Nov. 2009.

[276] H. Navarro-Hellín *et al.*, “A decision support system for managing irrigation in agriculture,” *Comput. Electron. Agric.*, vol. 124, pp. 121–131, Mar. 2016.

[277] S. W. Tsang *et al.*, “Applying artificial intelligence modeling to optimize green roof irrigation,” *Energy Build.*, vol. 127, pp. 360–369, Jun. 2016.

[278] A. Hassan *et al.*, “Groundwater estimation from major physical hydrology components using artificial neural networks and deep learning,” *Water*, vol. 12, pp. 1-18, Dec. 2019.

[279] M. Rahimzad *et al.*, “Performance comparison of an LSTM-based deep learning model versus conventional machine learning algorithms for streamflow forecasting,” *Water Resources Management*, vol. 35, pp. 4167-4187, Sep. 2021.

[280] L. Ni *et al.*, “Streamflow and rainfall forecasting by two long short-term memory-based models,” *Journal of Hydrology*, vol. 583, pp. 1-10, Apr. 2020.

[281] M. Chhetri *et al.*, “Deep BLSTM-GRU model for monthly rainfall prediction: A case study of Simtokha, Bhutan,” *Remote sensing*, vol. 12, pp. 1-13, Sep. 2020.

[282] P. M. Atkinson *et al.*, “Downscaling in remote sensing,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 22, pp. 106-114, Jun. 2013.

[283] M. Sdraka *et al.*, “Deep learning for downscaling remote sensing images: Fusion and super-resolution,” *IEEE Geoscience and Remote Sensing Magazine*, (Early access 2022).

[284] F. Shahid *et al.*, “Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM,” *Chaos, Solitons and Fractals*, vol. 140, pp. 1-9, Nov. 2020.

[285] G. Jha *et al.*, "Irrigation decision support systems (IDSS) for California's water–nutrient–energy nexus," *Chaos, Solitons and Fractals*, vol. 12, pp. 1-31, Aug. 2022.

[286] S. Yang *et al.*, "LSTM and GRU neural network performance comparison study: Taking Yelp review dataset as an example," in *Proc. IEEE International Workshop on Electronic Communication and Artificial Intelligence*, Shanghai, China, Jun. 2020, pp. 98-101.

[287] H. Yin *et al.*, "Rainfall-runoff modeling using LSTM-based multi-state-vector sequence-to-sequence model," *Journal of Hydrology*, vol. 598, pp. 1-14, Jul. 2021.