This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

Spatial Soil Moisture Prediction from In-situ Data Upscaled to Landsat Footprint: Assessing Area of Applicability of Machine Learning Models

Yi Yu, *Member, IEEE*, Brendan P. Malone, Luigi J. Renzullo, Chad A. Burton, Siyuan Tian, Ross D. Searle, Thomas F. A. Bishop, Jeffrey P. Walker, *Fellow, IEEE*

***Abstract*—The inherent spatial mismatch between satellite-derived and ground-observed near-surface soil moisture (SM) data necessitates cautious interpretation of point-to-pixel comparisons. Though data-driven upscaling of point-scale SM may enable statistically sound comparisons, the uncertainty across a spatial domain was less explored in previous studies. This gap underscores the need of addressing the spatial prediction uncertainties when extrapolating SM information to a broader spatial scale. Accordingly, this study presents a spatial prediction approach integrating machine learning (ML) and spatiotemporal fusion, which enables the characterisation of SM variability at the Landsat satellite footprint. Spatially clustered SM from 28 in-situ stations was extrapolated to a 100 km × 100 km area at 100 m resolution over a cross-validation period (2016-2019) and an independent test period (2020-2021). The area of applicability (AOA), which represents the spatial extent within which a prediction model is considered reliable, was determined for two ML models; Random Forests (RF) and eXtreme Gradient Boosting (XGB). The AOA of RF and XGB models encompassed 43.1% and 41.5% of the study area, respectively. The spatial SM predictions were further evaluated against multiple independent datasets, including field campaign data, in-situ SM from different networks, and satellite retrievals. Specifically, RF-predicted SM achieved a spatial R of 0.62-0.64 against field campaign data, temporal R of 0.84-0.91 against network-recorded data, and spatiotemporal R of 0.87 against SMAP L2 data during the cross-validation period. SM predictions within the AOA showed markedly lower uncertainties, which were further validated across an extended area (300 km × 300 km) with diverse physiographic conditions. Overall, this study demonstrated the use of AOA in delineating the statistically reliable spatial extent for ML-based SM predictions.**

***Index Terms*—Soil moisture; Spatial prediction; Upscaling; Landsat; Machine learning; Area of applicability.**

Yi Yu is with Fenner School of Environment & Society, The Australian National University, Canberra, ACT 2601, Australia, and also with CSIRO Agriculture and Food, Canberra, ACT 2601, Australia (e-mail: u6726739@anu.edu.au; yi.yu@csiro.au).

Brendan P. Malone is with CSIRO Agriculture and Food, Canberra, ACT 2601, Australia (e-mail: brendan.malone@csiro.au).

Luigi J. Renzullo and Siyuan Tian are with Bureau of Meteorology, Parkes, ACT 2600, Australia, and also with Fenner School of Environment & Society, The Australian National University, Canberra, ACT 2601, Australia (e-mail: luigi.renzullo@bom.gov.au; siyuan.tian@bom.gov.au).

Chad A. Burton is with Fenner School of Environment & Society, The Australian National University, Canberra, ACT 2601, Australia (e-mail: chad.burton@anu.edu.au).

Ross D. Searle is with CSIRO Agriculture and Food, St Lucia, QLD 4067, Australia (e-mail: ross.searle@csiro.au).

Thomas F. A. Bishop is with Sydney Institute of Agriculture, School of Life and Environmental Sciences, The University of Sydney, Eveleigh, NSW 2015, Australia (e-mail: thomas.bishop@sydney.edu.au).

Jeffrey P. Walker is with Department of Civil Engineering, Monash University, Clayton, VIC 3800, Australia (e-mail: jeff.walker@monash.edu).

I. INTRODUCTION

SOIL moisture (SM) is vital to global water, energy, and biogeochemical cycles [1], influencing the distribution of water and radiation at the land-atmosphere interface [2], [3]. Spatiotemporal SM datasets have been extensively developed and explored through remote sensing techniques [4], [5]. The remotely sensed SM data is typically acquired from microwave sensors, such as the Microwave Imaging Radiometer with Aperture Synthesis (MIRAS) on board the Soil Moisture and Ocean Salinity (SMOS) [6], the Advanced SCATterometer (ASCAT) on board the series of METeorological OPerational (METOP) satellites [7], and the L-band radiometer on board the Soil Moisture Active Passive (SMAP) [8]. Alternatively, the simulations can be obtained from land surface models (LSMs) embedded in data assimilation systems that incorporate the remotely sensed SM [9], often from retrospective reanalyses such as the European Centre for Medium-Range Weather Forecasts (ECMWF)'s fifth-generation climate ReAnalysis (ERA5) [10] and the National Aeronautics and Space Administration (NASA)'s Modern-Era Retrospective analysis for Research and Applications (MERRA) land data products [11]. These data have played a crucial role in studying SM dynamics within terrestrial environments.

An accurate evaluation of satellite-derived or model-estimated SM values is usually dependent on comparisons against in-situ SM measurements [12], [13]. At a farming or catchment scale, there is a necessity to enhance the comprehension and application of in-situ SM observations. This knowledge is often accumulated and refined over time by farm operators, while it can be supplemented through the deployment of SM sensors or networks. Some well-known SM ground networks, such as the TERrestrial ENvironmental Observatories (TERENO) [14], the UK COsmic-ray Soil Moisture Observing System (COSMOS-UK) [15], the Australian Cosmic-ray Soil Moisture Sensor Network (CosmOz) [16], the OzNet Hydrological Monitoring Network [17], play a pivotal role in evaluating SM from model simulations or satellite retrievals. However, the representativeness of ground measurements is constrained to a specific area and their distribution is typically sparse [18], [19]. The inherent mismatch in spatial support therefore necessitates a cautious interpretation of point-to-pixel comparisons. Researchers have consistently questioned how to effectively align the footprints between in-situ data and spatial grids. As a result, extensive investigations

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

into upscaling techniques have been undertaken [20].

Previous studies have implemented diverse approaches to upscale in-situ SM data to the spatial resolution of satellite footprints, primarily leveraging regression models while incorporating remotely sensed predictor variables. The MODIS-derived apparent thermal inertia (ATI) approach [20] is a representative semi-physical upscaling approach that ensures consistency between upscaled SM estimates and local hydro-meteorological characteristics. Concurrently, empirical methodologies, encompassing both geostatistical techniques and machine learning (ML) algorithms, have been extensively investigated, such as the Spatio-Temporal Regression Block Kriging (STRBK) [21], Least Absolute Shrinkage and Selection Operator (LASSO) [22] and Random Forests (RF) [23]. Despite their methodological diversity, most methods impose stringent requirements on the quantity and spatial arrangement of SM sensors. Furthermore, these studies exhibited limitations in validation methodology, lacking evaluations against independent datasets (e.g., field campaign) and assessment of spatial representativeness of models. Consequently, extant upscaling approaches typically demonstrated site-specificity with limited transferability across broader spatial domains [22]. A detailed comparison can be found in Table IV.

Especially, the discussion around uncertainty magnitude within the spatial prediction remains limited. The majority of studies have been confined to localised conditions, specifically upscaling in-situ SM at the precise geolocation of satellite pixels, and have primarily concentrated on cross-validation metrics, which are inadequate for comprehensively capturing the error distribution across a spatial domain. Considering that all upscaling strategies inherently carry nonnegligible errors [19], and are potentially influenced by spatial autocorrelation [24], it is crucial to assess the spatial distribution of error magnitudes and extrapolate the applicability of SM prediction at a relatively broader spatial scale. Given the scarcity of independent validation data for spatial prediction results in both space and time, emerging spatial analysis tools can play a pivotal role in quantifying such spatial uncertainty [25], [26].

The Area of Applicability (AOA) metric [25] describes the spatial region where a prediction model is expected to maintain cross-validation performance. This performance is learned from a spatially distributed training data that are usually collected from field sampling and/or ground stations. The AOA metric has been employed to determine the reliable spatial prediction extents for multiple geophysical variables, including soil properties [27], [28], soil organic carbon [29] and LST [30]. However, relatively little attention has been given to its use in SM prediction [31]. Results from field campaigns have consistently shown that variations in local topography, vegetation, and soil can systematically influence the relative bias of localised SM conditions when compared to broader-scale areal averages [19], [32]. The implementation of AOA can consequently help identify the reliable prediction area by accounting for the spatial variation of these predictor variables. Additionally, it is advisable to assess model performance against independent data sources, such as field campaigns or other networks [23]. These complementary assessments provide valuable context for interpreting the AOA metrics.

To enhance the extrapolation of spatial uncertainty, another area that requires improvement is the resolution of predictor variables used in spatial SM prediction. Previous studies have often relied on regression relationships between point-scale SM and predictor variables with coarse- to moderate-resolution. However, given the marked spatial variation in SM, the information provided by predictor variables at such resolutions (e.g., 1 km for MODIS) may not be directly applicable in building regression relationships with point-scale SM data. Matching the resolution of predictor variables to a level comparable to ground measurements is therefore expected to improve the quantification of spatiotemporal dynamics of SM. Moreover, as efforts have evolved to the production of fine-resolution (equivalent to or finer than 1 km) SM estimates (e.g., [33]–[36]), there is a growing need to quantify the error dynamics of SM at fine-resolution (e.g., 100 m). It holds promise for capturing the intricate variations in SM across fields, leading to more accurate assessments of spatial uncertainty.

Given the above, the objectives of this study are to: (1) spatially predict SM at 100 m resolution with daily frequency in an agricultural region using data-driven approaches; (2) conduct multi-fold spatial cross-validation to assess the performance of two alternate ML models during 2016-2019; (3) determine AOA of the employed models based on spatial cross-validation performance and feature importance during 2016-2019; and (4) evaluate the spatial SM prediction against independent data from field campaigns, additional monitoring networks, and satellite retrievals during both a cross-validation period (2016-2019) and independent test period (2020-2021), to complement the AOA metric.

II. STUDY AREA AND DATA

Table I summarises the data used in this study, including in-situ SM measurements, gridded predictor variables and independent SM datasets. The OzNet SM measurements were used as the response variable in the ML-based spatial prediction. The gridded predictor variables represent the multi-scale physical processes that align with recent findings on the importance of precipitation, vegetation, and temperature controls on SM variability across landscapes [37], including:

- MODerate Resolution Imaging Spectroradiometer (MODIS) products and Landsat-resolution products: which capture surface energy balance and vegetation water stress through albedo, NDVI, LST and ET;
- ANUClimate: which quantify atmospheric water flux components through air temperature (T_{air}), solar radiation (S_{rad}), vapor pressure deficit (VPD), and precipitation;
- Smoothed Digital Elevation Model (DEM-S) and the Soil and the Landscape Grid Australia (SLGA): which represent the underlying hydrological controls on water redistribution and retention.

The point-scale SM observations acquired by the Hydraprobe Data Acquisition System (HDAS) were used to evaluate the spatial pattern of the predicted SM. The independent in-situ SM observations from CosmOz and OzFlux networks and SM retrieval from the Soil Moisture Active

TABLE I
SUMMARY OF DATA USED IN THIS STUDY.

Category	Dataset	Variable(s)	Spatial resolution	Temporal frequency	Period	Reference(s)
Ground measurements	OzNet	SM	Point	Hourly	01 Jan 2016 – 31 Dec 2021	[17]
	CosmOz	SM	Sub-hectare	Hourly	01 Jan 2016 – 31 Dec 2021	[16]
	OzFlux	SM	Point	Hourly	01 Jan 2016 – 31 Dec 2021	[38]
Gridded predictor variables	MCD43A4	Albedo and NDVI	500 m	Daily	01 Jan 2016 – 31 Dec 2021	[39]
	MOD11A1	LST	1 km	Daily	01 Jan 2016 – 31 Dec 2021	[40]
	Landsat 8	Albedo, NDVI and LST	30 m / 100 m	16-day	01 Jan 2016 – 31 Dec 2021	[41]
	CMRSET	ET_a	30 m	Monthly	01 Jan 2016 – 31 Dec 2021	[42]
	ANUClimate	T_{air}, S_{rad}, VPD, and precipitation	1 km	Daily	01 Jan 2016 – 31 Dec 2021	[43]
	DEM-S	Elevation	30 m	Static	/	[44]
	SLGA	AWC, clay, sand and silt	90 m	Static	/	[45], [46]
Observed and retrieved SM	HDAS	SM	Point	Daily	30 Sep 2019 – 18 Oct 2019; 08 Mar 2021 – 26 Mar 2021	[47], [48]
	SMAP	SM	9 km	Daily	01 Jan 2016 – 31 Dec 2021	[8]

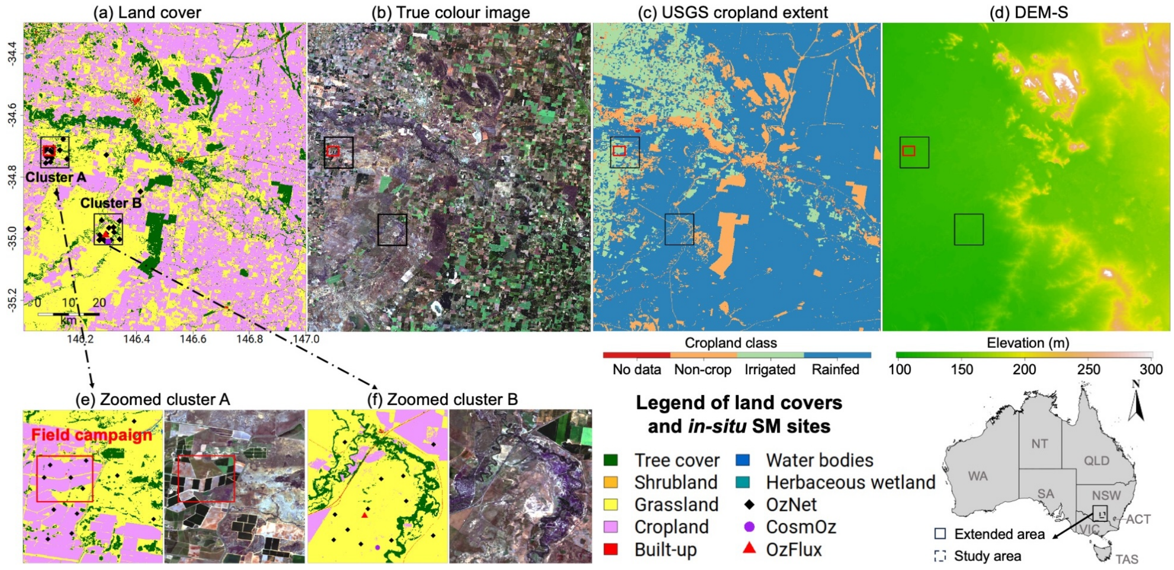

Fig. 1. (a) The land cover information of the Yanco agricultural region (highlighted in the dashed rectangle within state of NSW) in 2021 [49]; (b) a Landsat true-colour composite image acquired on 13 Oct 2020; (c) the USGS cropland extent [50]; (d) DEM-S [44]; and (e-f) zoomed land cover and true-colour images for cluster A and B, respectively. The black dots, purple circle and red triangle indicate in-situ sites from OzNet, CosmOz and OzFlux, respectively. The study area spans 146.00-147.00 °E and 34.30-35.30 °S (100 km × 100 km). The cluster A spans 146.06-146.16 °E and 34.67-34.77 °S (10 km × 10 km), with the field campaign in the Coleambally Irrigation Area (CIA; [51]) covering 146.07-146.11 °E and 34.70-34.73 °S (4 km × 3 km). Cluster B spans 146.25-146.35 °E and 34.92-35.02 °S (10 km × 10 km). The AOA assessment is also extended to a broader region (144.50-147.50 °E and 33.00-36.00 °S; 300 km × 300 km) surrounding the Yanco agricultural region. See Fig. A1 for more details about the extended area.

Passive (SMAP) mission were jointly used to evaluate the temporal dynamics of the spatial SM prediction.

A. Study area and in-situ measurements

The Yanco agricultural region was designated as the study area (Fig. 1). Yanco is situated within the Murrumbidgee Catchment, with a Mediterranean climate that is representative for the majority of agricultural regions in southeastern Australia (i.e., hot dry summer and cool moist winter). Fig. 1 shows (a) the land cover information in 2021 [49]; (b) a Landsat true-colour composite image acquired on 13 Oct 2020; (c) a map of cropland extent produced by the United States Geological Survey (USGS); and (d) the Smoothed Digital Elevation Model (DEM-S) [44]. The annual total precipitation in this region averages approximately 400 mm. The in-situ SM data were collected from the OzNet Hydrological Monitoring Network [17] between 01 Jan 2016 and 31 Dec 2021. OzNet site locations are indicated as black squares in Fig. 1 (a) and detailed in Table II. The OzNet sites were categorised into four groups to facilitate a four-fold spatial cross-validation (Table II). Moreover, two spatial clusters, where OzNet sites have a relatively dense distribution, were chosen, referred to as cluster A (Fig. 1 e) and cluster B (Fig. 1 f), respectively.

Two field campaigns were conducted in the Coleambally Irrigation Area (CIA) in 2019 [47] and 2021 [48], respectively. The spatial coverage of the field campaigns spans 146.07-

TABLE II
SUMMARY OF THE SITE INFORMATION FROM OZNET. THE 'FOLD' COLUMN REPRESENTS THE NUMBERED GROUP OF THE FOUR-FOLD SPATIAL CROSS-VALIDATION. THE 'CLUSTER' COLUMN REPRESENTS THE SPATIAL CLUSTER WHERE THE SITES ARE LOCATED. COSMOZ-YANCO AND OZFLUX-YANCO ARE LOCATED WITHIN THE CLUSTER B REGION (SEE FIG. 1) BUT THEY WERE NOT INCLUDED IN THE TRAINING.

Site name	Latitude (°S)	Longitude (°E)	Fold	Cluster
ya3	34.6772	146.1397	1	A
ya4a	34.7060	146.0794	2	A
ya4b	34.7031	146.1053	3	A
ya4c	34.7142	146.0943	4	A
ya4d	34.7142	146.0751	1	A
ya4e	34.7214	146.1030	4	A
ya5	34.7129	146.1277	3	A
ya7a	34.7352	146.0820	1	A
ya7b	34.7378	146.0987	4	A
ya7d	34.7544	146.0778	2	A
ya7e	34.7507	146.0949	1	A
ya9	34.7414	146.1536	2	A
yb1	34.9412	146.2765	4	B
yb3	34.9427	146.3401	2	B
yb5a	34.9653	146.3026	4	B
yb5b	34.9634	146.3184	3	B
yb5d	34.9848	146.2930	2	B
yb5e	34.9797	146.3205	3	B
yb7a	34.9885	146.2694	1	B
yb7c	34.9984	146.2785	2	B
yb7d	35.0050	146.2685	3	B
yb7e	35.0077	146.2880	4	B
yb9	35.0022	146.3398	3	B
y10	35.0053	146.3099	1	B
y5	34.7283	146.2932	3	N/A*
y7	34.8518	146.1153	4	N/A
y8	34.8470	146.4140	1	N/A
y9	34.9678	146.0163	2	N/A
CosmOz-Yanco	35.0100	146.3000	N/A	N/A
OzFlux-Yanco	34.9893	146.2907	N/A	N/A

* N/A represents they do not belong to any spatial cluster.

146.11 °E and 34.70-35.73 °S (Fig. 1 e; 4 km × 3 km), which was included within the cluster A. Additionally, in-situ SM data were collected from two other SM monitoring networks to facilitate independent evaluations, including (i) the Australian Cosmic-Ray Neutron Soil Moisture Monitoring Network (CosmOz) [16] and (ii) the Australian and New Zealand Flux Research and Monitoring (OzFlux) [38], shown as a purple dot and a red triangle in Fig. 1 (a) and (f).

The study domain is also extended to a mosaic of six Landsat tiles (path: 092-093, row: 083-085). The extended area spans 144.50-147.50 °E and 33.00-36.00 °S (300 km × 300 km; Fig. 1). See Fig. A1 for more details.

B. Gridded predictor variables

1) MODIS data: The MODIS products have been extensively used in moderate-resolution land surface monitoring [52]. The 500 m resolution daily Nadir Bidirectional reflectance distribution function Adjusted Reflectance (NBAR) product (MCD43A4; version 6.1) [39] and the 1 km resolution daily LST product (MOD11A1; version 6.1) [40] were collected over 01 Jan 2016 – 31 Dec 2021. The 500 m daily albedo and Normalised Difference Vegetation Index (NDVI) were derived from the surface reflectance bands of MCD43A4 dataset. The MODIS albedo was calculated using the Liang method [53]. These two collections were acquired from NASA's Earthdata Search platform (https://search.earthdata.nasa.gov/search; accessed 28 Apr 2025).

2) Landsat-resolution data: The Landsat 8 mission provides data from visible to thermal infrared bands, with a spatial resolution ranging from 30 to 100 m. The 30 m resolution 16-day Landsat 8 NBAR product (version 3.0.0) [41] was collected for the period spanning 01 Jan 2016 – 31 Dec 2021. With the application of the Bidirectional Reflectance Distribution Function (BRDF) it ensures consistency with the MCD43A4 product. The 30 m resolution 16-day albedo and NDVI were derived using its surface reflectance bands. The Landsat albedo was calculated using the Liang method [53]. The 100 m resolution 16-day Landsat LST was retrieved using a split-window algorithm [54]. A 30 m resolution monthly CSIRO MODIS Reflectance-based Scaling EvapoTranspiration (CMRSET; version 2.2) Landsat actual EvapoTranspiration (ET_a) product [42] was also collected between 01 Jan 2016 – 31 Dec 2021. The Landsat 8 data for Australia is publicly available from the Digital Earth Australia (DEA; https://dea.ga.gov.au/; accessed 28 Apr 2025).

3) ANUClimate: ANUClimate (version 2.0) is a climatic dataset spanning the Australian continent, comprising gridded daily and monthly climate variables at a spatial resolution of 0.01° [43]. This dataset was generated by interpolating Australia's national point climate data using trivariate thin plate smoothing spline functions [55], which account for spatial variations in longitude, latitude, and enhanced elevation. The average air temperature (T_air), incoming shortwave radiation (S_rad), vapour pressure deficit (VPD) and precipitation were collected during 01 Jan 2016 – 31 Dec 2021 to serve as predictor variables, all of which have a spatial resolution of 1 km and daily frequency.

4) Static data: The elevation and soil data were collected as static predictor variables. Specifically, the Australian Smoothed Digital Elevation Model (DEM-S; version 1.0) [44] was collected, which eliminates the impacts of ground vegetation and has undergone a smoothing process to mitigate the influence of noise, with a spatial resolution of 1 arcsecond (approximately 30 m). The available water capacity (AWC), clay, sand and silt data were acquired from the Soil and Landscape Grid Australia (SLGA; version 2.0) collection [45], [46]. Their spatial resolution is 3 arcsecond (approximately 90 m) and the depth of soil data used herein is 0-5 cm, consistent with the near-surface layer of SM.

C. Observed and retrieved SM

1) Hydraprobe Data Acquisition System (HDAS): HDAS is an integrated spatial data acquisition tool functioning within a mobile GIS environment [56]. HDAS was utilised to collect SM data during the P-band Radiometer Inferred Soil Moisture (PRISM) field campaigns in CIA in 2019 and 2021 [47], [48]. Throughout the campaigns, routine HDAS measurements were conducted across regular grids spaced at 50 m intervals to capture the spatial distribution of near-surface SM (approximately 5 cm). The acquired near-surface HDAS SM data spans 30 Sep 2019 – 18 Oct 2019 and 08 Mar 2021 – 26 Mar 2021. The

HDAS data is available at https://prism.monash.edu/index.html (accessed 28 Apr 2025).

2) SMAP: The SMAP mission was launched by NASA in 2015 and measures L-band microwave brightness temperature at 1.4 GHz [8]. The L2 product (version 6) was collected between 01 Jan 2016 and 31 Dec 2021, which provides radiometer SM derived from level 1 observations and ancillary information. It has an original spatial resolution of 36 km and a temporal revisit frequency of 2-3 days. The data is gridded at the 9 km Equal Area Scalable Earth-2 (EASE2) grids using the Backus-Gilbert interpolation method. The data is publicly available from the National Snow and Ice Data Centre (https://nsidc.org/data/smap; accessed 28 Apr 2025).

III. METHODOLOGY

Fig. 2 presents the experimental design in four key steps. Firstly, a nearest neighbour method was utilised to resample all predictor variables to 0.001 ° resolution grids under the World Geodetic System 1984 (WGS84) datum. Concurrently, a spatiotemporal fusion was implemented between MODIS and Landsat data to downscale MODIS albedo, NDVI, and LST to a 100 m resolution with daily frequency (Section III-A). Secondly, the ML models were trained using point-scale SM from 28 OzNet sites between 2016 and 2019 as the response variable, with the collected predictor variables serving as input features. The performance of two ML models, specifically the Random Forests (RF) and eXtreme Gradient Boosting (XGB), were compared through a four-fold spatial cross-validation (CV) and a cross-cluster (CC) validation to assess the transferability of the spatial prediction approach while considering regional nuances (Section III-B). Thirdly, the AOA metrics of RF and XGB models were determined based on their respective performance and feature importance in CV and CC (Section III-C). Fourthly, the spatial SM prediction within AOA was evaluated against multiple independent datasets, including data from field campaigns, CosmOz and OzFlux networks, and SMAP SM retrievals (Section III-D).

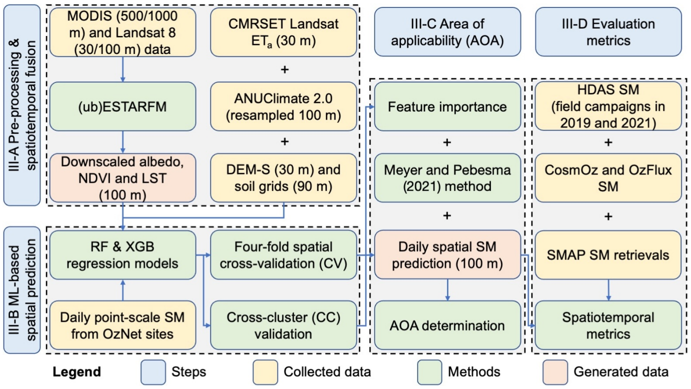

Fig. 2. The experimental design. The 'Steps' polygons correspond to Section III-A - III-D.

A. Spatiotemporal fusion

The implementation of spatiotemporal fusion is to address the resolution mismatch between key geophysical indicators and surface characteristics in SM prediction. While MODIS products (>500 m) provide frequent data essential for capturing daily surface dynamics, their spatial resolution is insufficient to characterise field-scale heterogeneity in SM. Through fusion with Landsat data (30-100 m), it achieves daily estimates of these predictor variables at resolutions commensurate with underlying soils (90 m) and topographic features (30 m).

The Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM) [57] has been extensively applied to generate fine-resolution surface reflectance [51], [57] and demonstrated an exceptional performance in fusing surface reflectance. Furthermore, an unbiased variant (ubESTARFM) [58] has demonstrated more effectiveness in fusing LST data. Hence, ESTARFM was implemented to fuse MODIS and Landsat surface reflectance data (i.e., albedo and NDVI), while ubESTARFM was implemented to fuse MODIS and Landsat LST data, to generate daily 100 m estimates of these predictor variables. The formula of ESTARFM is given as:

$$\begin{aligned} F(x_{w/2}, y_{w/2}, t_p) &= F(x_{w/2}, y_{w/2}, t_0) \\ &+ \sum_{i=1}^{N} W_i \times V_i \times (C(x_i, y_i, t_p) - C(x_i, y_i, t_0)) \end{aligned} \quad (1)$$

where $F(x_{w/2}, y_{w/2}, t_0)$ is the central pixel $(x_{w/2}, y_{w/2})$ of the fine-resolution image within a search window at time t_0; t_p is the prediction time; N is the number of similar pixels within the search window; W_i is the weight of the ith similar pixel; V_i is a regression coefficient between the chosen fine- and coarse-resolution pixels; $C(x_i, y_i)$ is the pixel at the location (x_i, y_i) of a coarse-resolution image. For ubESTARFM, a local bias correction step is implemented to the central and similar fine-resolution pixels within the search window. For more details, please see [57], [58].

B. ML-based spatial prediction

ML models and algorithms have been well established and applied in geoscience and remote sensing [59]. Two widely recognised ML models were employed, namely the Random Forests (RF) [60] and the eXtreme Gradient Boosting (XGB) [61], to perform a comparative study. These ensemble algorithms have demonstrated robust performance in capturing non-linear relationships in environmental data, particularly in SM estimation applications [62]. RF's bootstrap aggregation mechanism provides protection against overfitting, while XGB's gradient boosting framework offers complementary strengths in handling complex feature interactions. Both ML approaches offer built-in feature importance assessment capabilities, enabling systematic quantification of predictor variable contributions - a critical requirement for implementing AOA. They can also be interpreted using the SHapley Additive exPlanations (SHAP), which provide insights into the average marginal contribution of a feature value across all possible coalitions [63], [64]. The methodological selection balances algorithmic sophistication, interpretability, computational efficiency, and prediction accuracy.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

1) Four-fold spatial cross-validation (CV): A 'leave-seven-sites-out' CV strategy was performed to assess the fitting performance of the ML models between 2016 and 2019. The 28 OzNet sites were randomly divided into four folds. In each fold, seven sites were reserved for validation purposes, while the remaining 21 sites were utilised for training the models. Specifically, each fold contains one site outside both clusters and six sites with close neighbours (Fig. 1 and Table II). The specifics of each fold are outlined in the 'fold' column of Table II. This approach was chosen to ensure that the trained model captures the spatiotemporal distribution of SM data across a relatively large spatial domain.

2) Cross-cluster (CC) validation: The validation strategy was expanded beyond random selection of sites by conducting a CC validation. This approach was crucial due to the spatial clustering of OzNet sites (Fig. 1), necessitating an evaluation of model performance using a 'leave-cluster-out' approach commonly employed in spatial ML applications [65]. Details regarding each cluster (i.e., A and B) can be found in the 'cluster' column of Table II. The cluster A and B, each of which contains 12 OzNet sties, were employed reciprocally for a CC validation. Specifically, sites within cluster A were utilised for training while sites within cluster B were utilised for validation, and vice versa.

C. Area of applicability (AOA)

The AOA metric [25] serves as a tool to evaluate the spatial extent within which a ML model can be reliably applied. In spatial prediction studies, ML models are typically trained using spatially distributed data obtained from ground stations. However, the model needs to make prediction in new geographic areas, which may possess different feature combinations and relationships to those present in the model training data. AOA delineates the area where the model can apply relationships learned from training data, ensuring that the estimated cross-validation performance remains valid. It begins with a standardisation procedure that ensures equivalent scaling across all predictor variables through the implementation of z-score normalisation:

$$X_{i,j}^{s} = \frac{X_{i,j} - \bar{X}_j}{\sigma_j} \tag{2}$$

where $X_{i,j}^{s}$ denotes the standardised value of the jth predictor variable for the ith observation (i.e., $X_{i,j}$); $\bar{X}_j$ and σ_j represent the arithmetic mean and standard deviation of the jth predictor variable, respectively. Furthermore, ML models reveal differential predictive power across variables [66]. To incorporate this heterogeneity, an importance-weighted standardisation is implemented:

$$X_{i,j}^{sw} = w_j X_{i,j}^{s} \tag{3}$$

where w_j is the weight derived from the variable importance for the jth predictor variable. The Euclidean distance between a given prediction location k and a training location i in the predictor variable space can be expressed as:

$$d(k,i) = \sqrt{\sum_{j=1}^{p} (X_{k,j}^{sw} - X_{i,j}^{sw})^2} \tag{4}$$

where d denotes the Euclidean distance; p is the number of predictor variables. The minimal distance is then given as:

$$d_k = arg_i min(d(k,i)) \tag{5}$$

The dissimilarity index (DI) for a given prediction location (i.e., DI_k) is defined as:

$$\mathrm{DI}_k = d_k / \bar{d} \tag{6}$$

where $\bar{d}$ is the arithmetic mean value of all pairwise distances between training data. The DI for training instances should be computed following a cross-validation, where distances are calculated to the nearest training point outside the instance's validation fold. The value of DI ranges from 0 to ∞, where $\mathrm{DI}_k = 0$ represents the prediction location k has no dissimilarity (i.e., identical characteristics) with a training location (Fig. 3).

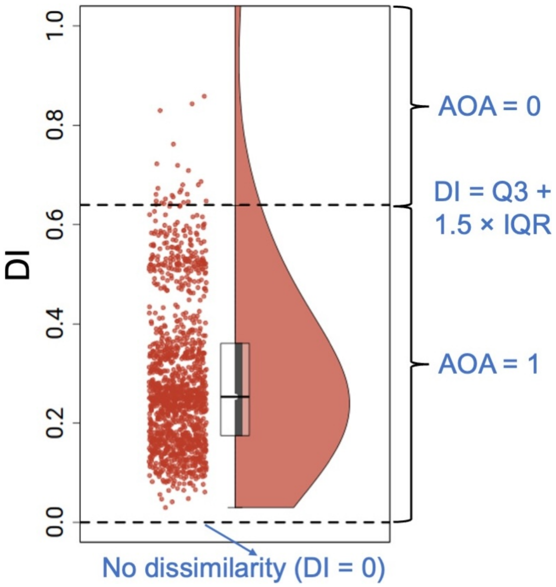

Fig. 3. An illustration of the binary determination of AOA using RF model based on a four-fold spatial CV.

The outlier-removed maximum DI is used as a threshold to determine a binary AOA, which is given as:

$$\mathrm{AOA}_k = \begin{cases} 1, & \text{if } \mathrm{DI}_k \leq \mathrm{DI}_{\mathrm{Q3}} + 1.5 \times \mathrm{DI}_{\mathrm{IQR}} \\ 0, & \text{if } \mathrm{DI}_k > \mathrm{DI}_{\mathrm{Q3}} + 1.5 \times \mathrm{DI}_{\mathrm{IQR}} \end{cases} \tag{7}$$

where Q3 and IQR denote the third quartile and interquartile range of a boxplot, respectively. When $\mathrm{AOA}_k = 1$, it represents the characteristics of predictor variables at location k are considered similar to those of training data (Fig. 3), thus the prediction model is considered applicable at location k.

The exhibition of DI estimations and AOA determinations was advocated in conjunction with the prediction of ML models [25]. These metrics can help encapsulate spatial constraints

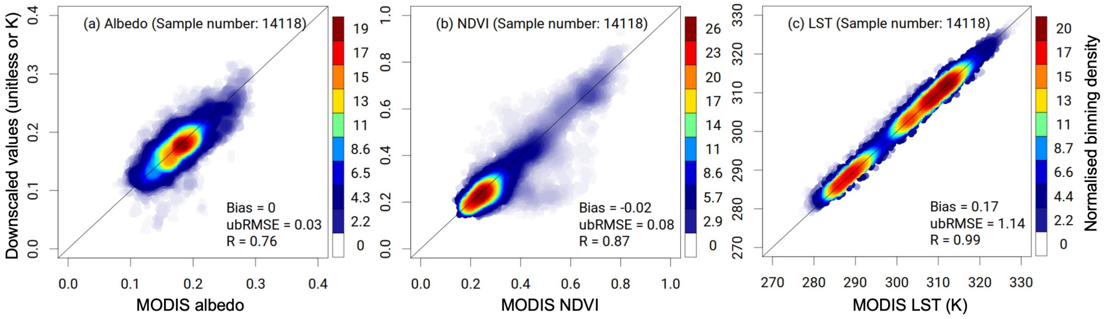

Fig. 4. Binned density scatterplots of (a) albedo, (b) NDVI, and (c) LST between 01 Jan 2016 and 31 Dec 2019 (excluding training dates of spatiotemporal fusion).

and zones of reliable prediction for ML models, thereby complementing the traditional validation metrics and prevalent estimates of uncertainty. For more details, please see [25].

D. Evaluation metrics

Bias, unbiased root mean squared error (ubRMSE) and Pearson correlation coefficient (R) were used to evaluate the performance of spatiotemporal fusion and ML models, which are respectively given as:

$$\text{Bias} = \frac{\sum_{i=1}^{N}(y_i - x_i)}{N} \tag{8}$$

$$\text{ubRMSE} = \sqrt{\frac{\sum_{i=1}^{N}((y_i - \bar{y}_i) - (x_i - \bar{x}_i))^2}{N}} \tag{9}$$

$$R = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{10}$$

where y_i denotes the time series of estimated data; x_i denotes the time series of reference data; $\bar{y}_i$ and $\bar{x}_i$ are the mean values of the estimated and reference data, respectively; N is the number of observations of the time series.

IV. RESULTS

A. Evaluation of downscaled predictor variables

Fig. 4 shows the downscaled performance of (a) albedo, (b) NDVI, and (c) LST versus MODIS data during 01 Jan 2016 – 31 Dec 2019 using binned density scatterplots. The validation employed a point-based comparison strategy, where both MODIS and downscaled data were extracted at precisely matched OzNet monitoring locations (n = 14,118) through coordinate-based spatial matching. Downscaled albedo had a bias of 0, ubRMSE of 0.03, and R of 0.76, with values clustering around 0.2. NDVI showed a slight negative bias of -0.02, ubRMSE of 0.08, and R of 0.87, with values mostly between 0.2 and 0.3. LST had a bias of 0.17 K, ubRMSE of 1.14 K, and R of 0.99, with values concentrated around 290 K and 305-315 K. Overall, the downscaled predictor variables demonstrated consistency with MODIS data, as evidenced by the established metrics and concentration patterns on the 1:1 line as observed in the scatterplots.

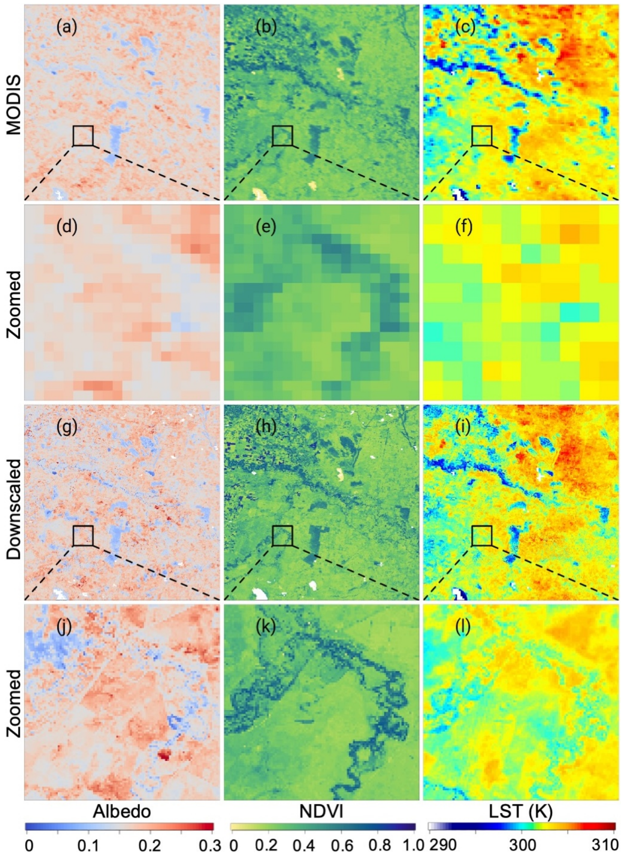

Fig. 5. Spatial comparison between MODIS and downscaled predictor variables on 02 Apr 2017. (a-c) are MODIS albedo, NDVI and LST, respectively; (d-f) are zoomed windows of MODIS predictor variables spanning 146.25 to 146.35 °E and 34.92 to 35.02 °S (i.e., cluster B); (g-i) are the downscaled albedo, NDVI and LST, respectively; and (j-l) are zoomed windows of downscaled predictor variables covering the same zoomed area with (d-f).

Fig. 5 provides an illustrative example from 02 Apr 2017 showcasing the spatial comparison between MODIS data and downscaled predictor variables. Fig. 5 (g-i) present downscaled patterns of predictor variables across the entire study area, demonstrating consistency with MODIS data but exhibiting sharpened features in comparison (Fig. 5 a-c). Upon closer examination of a zoomed area (Fig. 5 j-l), the downscaled predictor variables revealed enhanced details that are more visually discernible compared to the MODIS data (Fig. 5 d-f). Notably, features such as the vegetated areas (observed in Fig.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

5 k) are more distinct in the downscaled predictor variables. In general, the spatiotemporal fusion enabled a better capture of spatial details of predictor variables at the field scale, offering improved visualisation and representation of the landscape.

B. *Evaluation of model performances*

Fig. 6 presents boxplots of (a) bias, (b) ubRMSE and (c) R for a four-fold spatial CV of RF and XGB models, respectively, between 01 Jan 2016 and 31 Dec 2019. The median bias was approximately 0.00 m^3/m^3 for RF and -0.01 m^3/m^3 for XGB, with bias values ranging from -0.05 to 0.05 m^3/m^3 for both models. The median ubRMSE was around 0.05 m^3/m^3 for both models, with values between 0.03 and 0.12 m^3/m^3. The median R was approximately 0.79 for RF and 0.78 for XGB, with RF showing slightly better performance. These metrics indicate that both models performed comparably in estimating SM contents across the OzNet sites.

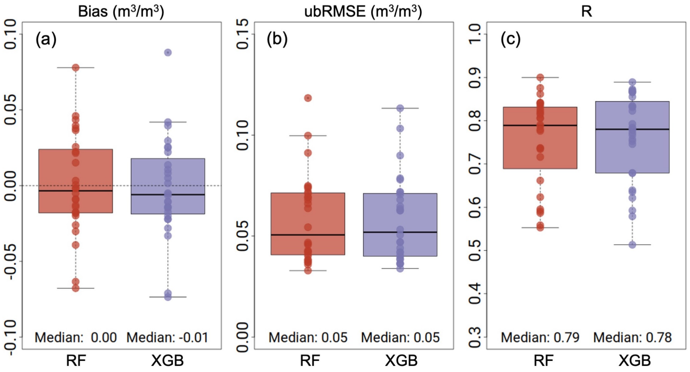

Fig. 6. Boxplots of (a) bias, (b) ubRMSE and (c) R for a four-fold spatial CV of RF and XGB models, respectively, between 01 Jan 2016 and 31 Dec 2019. The details of sites included in each fold are given in Table II. The scatters represent the metrics of individual sites.

Fig. 7 presents boxplots of bias, ubRMSE and R for the CC validation of (a-c) RF and (e-f) XGB model, respectively, between 01 Jan 2016 and 31 Dec 2019. In cluster A, RF and XGB had similar performance, with median bias, ubRMSE, and R values of about -0.02 m^3/m^3, 0.08 m^3/m^3, and 0.70, respectively (Fig. 7; cluster A columns). In cluster B, RF slightly outperformed XGB with median bias, ubRMSE, and R values of 0.05 m^3/m^3, 0.05 m^3/m^3, and 0.71, respectively, compared to XGB's 0.06 m^3/m^3, 0.06 m^3/m^3, and 0.68 (Fig. 7; cluster B columns). When trained on data from both clusters A and B, both models captured SM patterns outside these clusters (Fig. 7; remaining 4 sites columns), with RF (XGB) showing bias, ubRMSE, and R values of 0.02 (0.01) m^3/m^3, 0.05 (0.05) m^3/m^3, and 0.78 (0.80), respectively. In summary, these boxplots showcase the spatial transferability of the ML models, demonstrating the capability of both models to maintain consistent performance across different spatial clusters within a distance of roughly 20 km (see Fig. 1).

Fig. 8 shows the violin plots of SHAP values of six most important predictor variables of (a-c) RF model and (d-f) XGB model between 01 Jan 2016 and 31 Dec 2019 for all sites, cluster A and cluster B, respectively. For all sites, both models identified albedo, VPD, NDVI, ET, LST, and DEM as the top predictor variables, though their order differed (Fig. 8 a, d). In cluster B for RF (Fig. 8 c) and cluster A for XGB (Fig. 8 e), these predictor variables remained most important. However, for RF in cluster A (Fig. 8 b) and XGB in cluster B (Fig. 8 f), S_{rad} emerged as a new predictor, ranking 5th and 6th, respectively. The colour-coded normalised feature values (yellow to purple) reveal each predictor variable's contribution: moderate to high VPD and LST generally contribute to the reduction of predicted SM values, while high ET and NDVI would contribute to the increase. Conversely, the values of albedo, DEM and S_{rad} exhibited

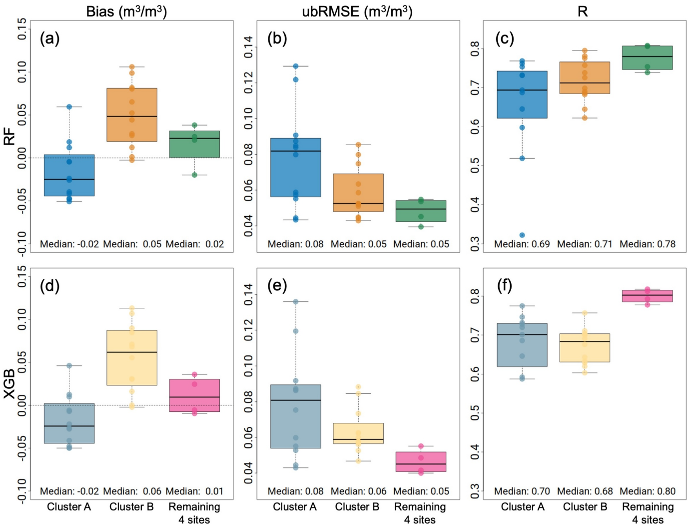

Fig. 7. Boxplots of bias, ubRMSE and R for the CC validation of (a-c) RF and (e-f) XGB model, respectively, between 01 Jan 2016 and 31 Dec 2019. Metrics within cluster A were derived from models trained using data within cluster B, and vice versa. Metrics for the remaining 4 sites (see Fig. 1 and Table II) were derived from models trained using data from both clusters A and B. The details of sites included in each cluster are given in Table II. The scatters represent the metrics of individual sites.

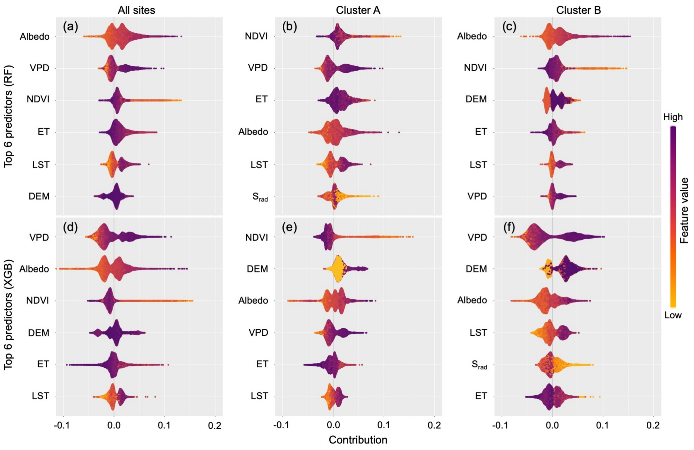

Fig. 8. Violin plots of SHAP values of six most important predictor variables of (a-c) RF and (d-f) XGB between 01 Jan 2016 and 31 Dec 2019 for all sites, cluster A and cluster B, respectively. The predictor variables are displayed in the order of importance.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

more variability and lacked specific trends in predicting SM values. Overall, while predictor rankings varied by region, the most important predictor variables generally remained consistent. These variations highlight the models' reliance on regional land cover and landscape characteristics, influencing the importance and impact of specific predictor variables.

Fig. 9 presents a Spearman correlation matrix illustrating the pairwise relationships for all predictor variables used in this study. Notable observations include strong negative correlations between clay and silt/sand content, and strong positive correlations among LST, T_{avg}, VPD, S_{rad}, reflecting their interconnected nature. NDVI demonstrated moderate negative correlations with LST, albedo, T_{avg}, VPD, and S_{rad}. Conversely, ET demonstrated moderate positive correlations with LST, NDVI, T_{avg}, VPD, and S_{rad}. This matrix identified potential collinearities among predictors, thereby complementing the feature importance as given by SHAP values.

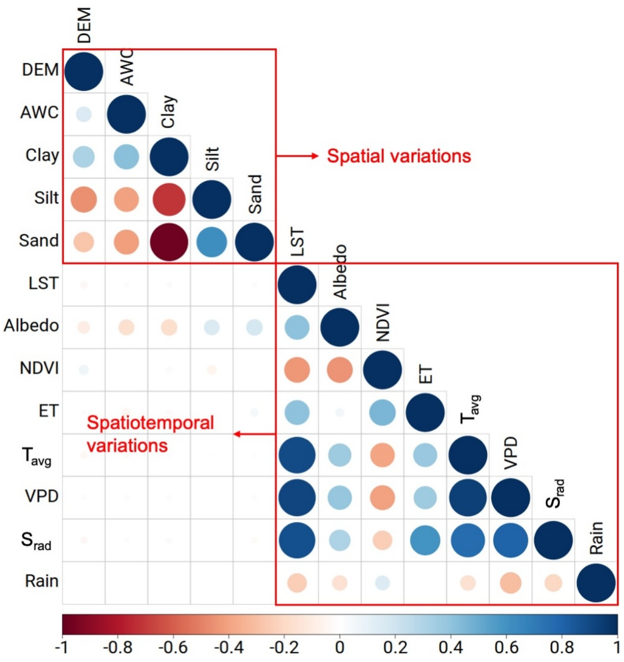

Fig. 9. Spearman correlation matrix of the pairwise relationships for the full spectrum of 13 predictor variables.

C. AOA of spatial SM predictions

Fig. 10 shows the spatial distribution of the median DI (with values exceeding 5 masked out) and AOA based on CV and CC for (a-d) RF and (e-h) XGB, respectively, between 01 Jan 2016 and 31 Dec 2019. Both models showed low DI in the western region, regardless of CV or CC validation (Fig. 10 a, c, e, g). However, DI increased drastically in the eastern area due to no training sites and distinct landscape patterns (see Fig. 1). The DI thresholds for AOA determination were 0.64 (CV) and 0.87 (CC) for RF (Fig. 10 a, c), almost double those for XGB, which were 0.34 (CV) and 0.44 (CC) (Fig. 10 e, g). AOA for RF (XGB) based on CV was 43.1% (41.5%), while for CC it was 25.4% (26.9%). Predictor variables within the determined AOA were deemed comparable to those used in training. It is expected that spatial SM prediction within AOA can demonstrate comparable performance metrics to that of the CV (see Fig. 6) and CC (see Fig. 7). The CV approach (using 21 sites each time) also demonstrated capability to effectively expand AOA compared to CC (using 12 sites each time).

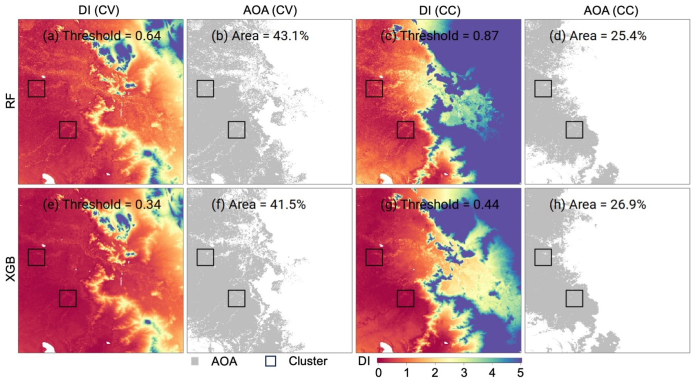

Fig. 10. Spatial distribution of median DI (with values exceeding 5 masked out) and AOA based on performance of CV and CC for (a-d) RF and (e-h) XGB, respectively, between 01 Jan 2016 and 31 Dec 2019. The derived DI and AOA had an identical temporal frequency with the spatial SM prediction (i.e., daily). DI was composited using its median values, and AOA was composited using its majority values (i.e., >50%). Threshold values indicate the maximum permissible DI for reliable predictions, and area percentages represent the proportion of the study region within the AOA. Black rectangles indicate the cluster locations. The mask value of 5 was empirically determined to optimise visualisation clarity while maintaining representation of all relevant DI gradients that influence AOA delineation.

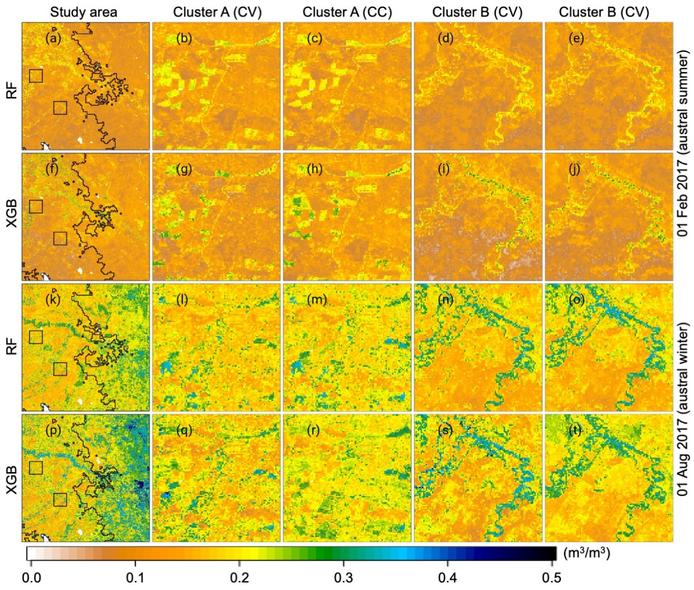

Fig. 11. Spatial examples of predicted SM using RF and XGB models on 01 Feb 2017 (austral summer) and 01 Aug 2017 (austral winter) for the study area (first column), zoomed areas based on CV (second and fourth columns), and zoomed areas based on CC (third and fifth columns), respectively. AOA was displayed as the western area of the black boundaries, encompassing both spatial clusters.

Fig. 11 provides spatial examples of predicted SM using RF and XGB models on 01 Feb 2017 (austral summer) and 01 Aug 2017 (austral winter) for the entire study area, and zoomed areas based on CV or CC, respectively. AOA was

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

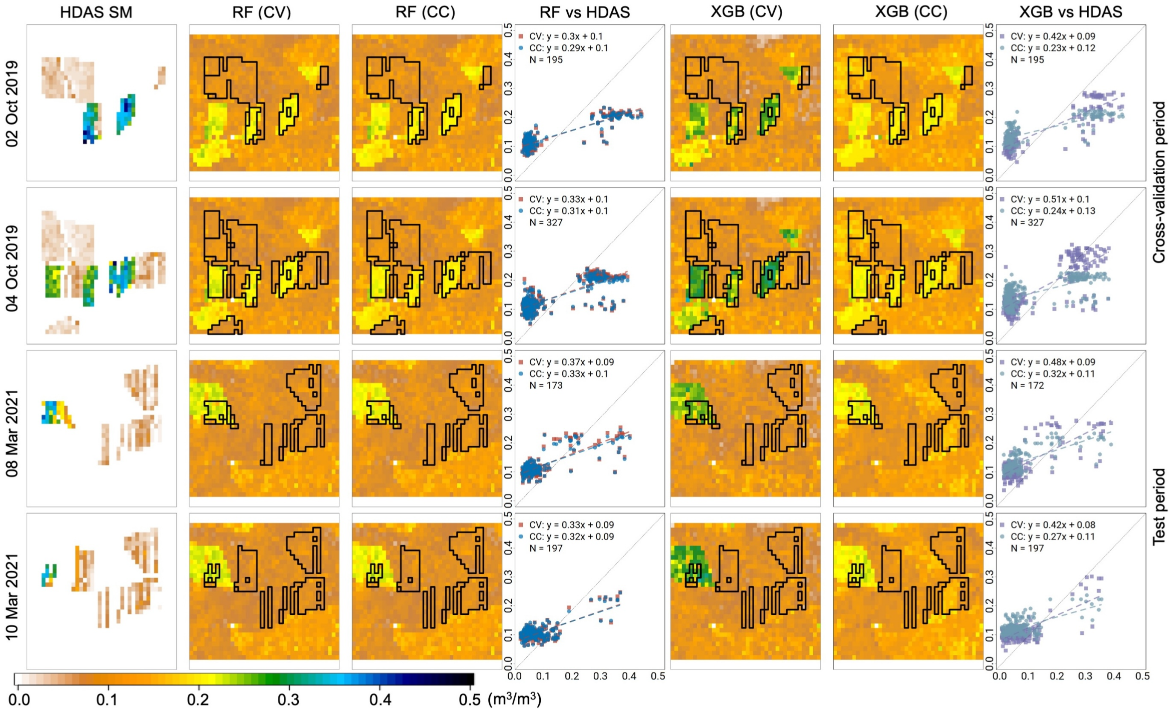

Fig. 12. Spatial examples of HDAS SM and predicted SM using RF and XGB on 02 Oct 2019 and 04 Oct 2019 (cross-validation period) and 08 Mar 2021 and 10 Mar 2021 (test period), respectively. The first column shows the interpolated HDAS SM at 100 m resolution grids. The black polygons within second, third, fifth and sixth columns from left are derived from the spatial extent of HDAS SM. The fourth and seventh columns from left are scatterplots of predicted SM derived from CV and CC, using RF and XGB, against HDAS SM, respectively.

displayed as the western area within the black boundaries (Fig. 11 a, f, k, p), covering both spatial clusters. On 01 Feb 2017 (austral summer), predicted SM values were relatively dry (0.1 to 0.2 m^3/m^3), with both models yielding similar results (Fig. 11 a, f), though XGB showed slightly higher SM values. Zoomed areas derived from CV and CC also appeared visually comparable (e.g., Fig. 11 b, c, g, h). On 01 Aug 2017 (austral winter), predicted SM values were higher (0.2 to 0.5 m^3/m^3) and more variable, especially in the eastern regions outside AOA (Fig. 11 k, p). These disparities can be attributed to both models learning only from data within the AOA, rendering conditions outside the AOA as ‘unknown space’ and thus more uncertain. The results within zoomed areas, derived from both CV and CC, also exhibited greater variability (e.g., Fig. 11 l, m, q, r). In summary, this detailed visualisation highlights the capability of AOA in assessing the reliability of predicted SM at the field scale. The observed variations in SM across landscapes underscore the importance of considering regional nuances in training strategies for accurate and context-specific predictions.

D. Evaluations against independent data

The AOA metric indicated that the RF and XGB models were anticipated to reliably predict the SM dynamics within approximately 43.1% and 41.5% of the study region, respectively, as determined by CV metrics. To validate AOA, both models were assessed against independent data from field campaigns, CosmOz, OzFlux, and SMAP. Fig. 12 shows the spatial examples of HDAS SM collected during field campaigns and predicted SM using the RF and XGB models on 02 Oct 2019 and 04 Oct 2019 (within the cross-validation period) and 08 Mar 2021 and 10 Mar 2021 (within the test period), respectively. The 4 km × 3 km field campaign area was within the AOA (see Fig. 1 and Fig. 10). Both models captured the spatial patterns of wet regions. For RF, predictions from CV and CC were nearly identical (Fig. 12, second to fourth columns), but both underestimated SM in wet regions. For XGB, CV predictions had higher values than CC (Fig. 12, fifth to seventh columns) and better matched HDAS reference grids (Fig. 12, first column). In summary, both the RF and XGB models demonstrated capabilities to capture the correct spatial patterns of SM within the HDAS field campaign area, with the XGB model based on CV providing the closest prediction to HDAS SM values.

Table III details bias, ubRMSE, and R metrics for spatial SM prediction against HDAS SM on specific field campaign dates in September to October 2019 (within the cross-validation period) and March 2021 (within the test period). Both models maintained stable mean bias values of 0.04-0.05 m^3/m^3 in both periods. The ubRMSE of both models even reduced within the test period, with mean values of 0.06 m^3/m^3 compared to 0.08-0.09 m^3/m^3 within the cross-validation period. The most marked difference was in R performance. The CV-based RF consistently showed better or comparable R values (0.60-0.62) than CV-based XGB (0.53-0.58). The CC-based RF also performed slightly better (0.64-0.65) than CC-based XGB (0.62-0.64). In summary, both the RF and XGB models demonstrated comparable performances in bias and ubRMSE within both the cross-validation and test periods. However, RF showed more robust R performance

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

TABLE III

METRICS OF BIAS, UBRMSE AND R OF SPATIAL SM PREDICTION AGAINST HDAS SM ON SPECIFIC FIELD CAMPAIGN DATES IN SEPTEMBER TO OCTOBER IN 2019 (WITHIN CROSS-VALIDATION PERIOD) AND MARCH IN 2021 (WITHIN TEST PERIOD). THE METRICS ARE BASED ON PIXELWISE COMPARISONS.

Period	Date	Bias (m^3/m^3)				ubRMSE (m^3/m^3)				R				Sample numbers
		CV		CC		CV		CC		CV		CC		
		RF	XGB	RF	XGB	RF	XGB	RF	XGB	RF	XGB	RF	XGB	
Cross-validation	30 Sep 2019	0.02	0.04	0.03	0.03	0.10	0.11	0.11	0.11	0.81	0.81	0.86	0.85	147
	02 Oct 2019	0.02	0.03	0.01	0.02	0.09	0.11	0.10	0.10	0.89	0.81	0.92	0.90	195
	04 Oct 2019	0.05	0.05	0.03	0.03	0.07	0.10	0.08	0.09	0.83	0.69	0.81	0.80	327
	07 Oct 2019	0.06	0.05	0.04	0.04	0.09	0.09	0.09	0.09	0.51	0.41	0.51	0.51	320
	11 Oct 2019	0.07	0.06	0.05	0.05	0.09	0.09	0.08	0.08	0.37	0.30	0.42	0.42	314
	14 Oct 2019	0.04	0.03	0.03	0.03	0.08	0.07	0.07	0.07	0.34	0.30	0.46	0.45	157
	18 Oct 2019	0.06	0.05	0.06	0.06	0.05	0.06	0.05	0.05	0.63	0.54	0.63	0.65	242
	All	0.05	0.05	0.04	0.04	0.08	0.09	0.08	0.08	0.62	0.53	0.64	0.64	1702
Test period	08 Mar 2021	0.05	0.06	0.04	0.04	0.07	0.07	0.07	0.07	0.72	0.72	0.77	0.76	173
	10 Mar 2021	0.05	0.06	0.04	0.04	0.05	0.05	0.05	0.05	0.73	0.61	0.74	0.73	197
	17 Mar 2021	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.48	0.48	0.55	0.47	162
	19 Mar 2021	0.03	0.02	0.02	0.02	0.06	0.06	0.06	0.06	0.41	0.48	0.49	0.46	138
	All	0.05	0.05	0.04	0.04	0.06	0.06	0.06	0.06	0.60	0.58	0.65	0.62	670

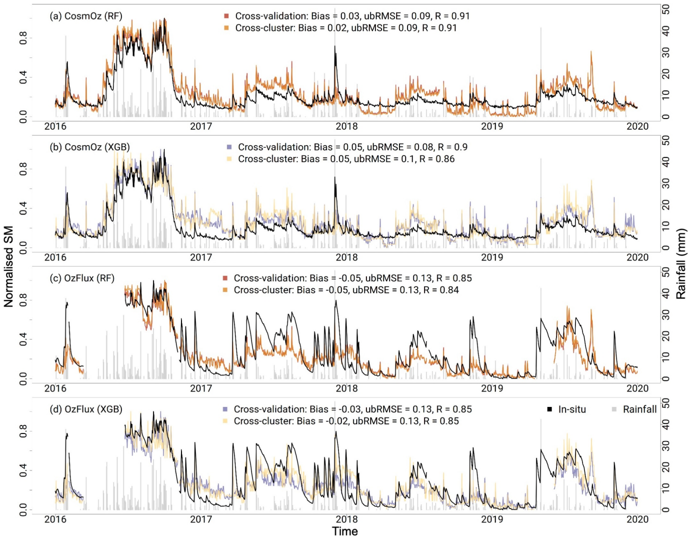

Fig. 13. Time series of predicted SM (both CV and CC), CosmOz and OzFlux SM, and rainfall during the cross-validation period (2016-2019). The unit of SM herein is wetness (unitless), which is different with that of OzNet and HDAS SM (m^3/m^3). This is because CosmOz only provides normalised SM data ranging between 0 and 1.

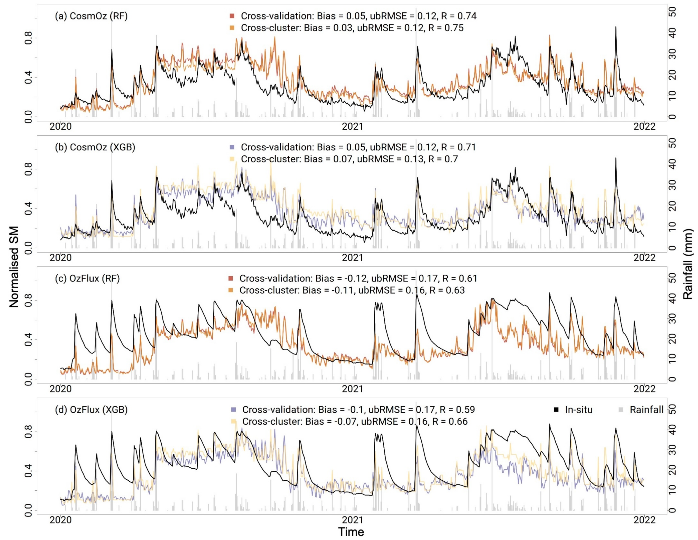

Fig. 14. Same with Fig. 13 but for the test period (2020-2021).

with moderate values (0.60-0.65), while XGB exhibited more variability (0.53-0.64).

Fig. 13 and Fig. 14 present the time series of predicted SM (both CV and CC), CosmOz and OzFlux SM, and rainfall throughout the cross-validation period (2016-2019) and test period (2020-2021), respectively. In the cross-validation period, the temporal variations of RF based on CV and CC were visually identical, exhibiting high agreement with CosmOz SM (R = 0.91; Fig. 13 a) and OzFlux SM (R = 0.84-0.85; Fig. 13 c). As a comparison, the temporal variations of CV-based XGB differed from those based on CC. The R performance of XGB ranged from 0.86 to 0.90 against CosmOz SM (Fig. 13 b) and was 0.85 against OzFlux SM (Fig. 13 d), respectively, indicating its sensitivity to training samples.

In the test period, the temporal variations of SM predicted by RF based on CV and CC remained visually similar, exhibiting moderate agreement with CosmOz SM (R = 0.74-0.75; Fig. 14 a) and OzFlux SM (R = 0.61 - 0.63; Fig. 14 c). Comparatively, the temporal variations of SM predicted by XGB based on CV still diverged from those based on CC. The R performance of SM predicted by XGB was 0.70-0.71 against CosmOz SM (Fig. 14 b) and was 0.59-0.66 against OzFlux SM (Fig. 14 d), respectively. In summary, these temporal metrics indicated that RF exhibited slightly better agreement with independent SM time series and displayed more robustness irrespective of variations in training sample numbers.

Fig. 15 presents the spatial distribution of bias, ubRMSE and R of SM predicted by the RF and XGB models against SMAP L2 data. The evaluation covers the entire study area, cluster A and cluster B, spanning both the cross-validation period (2016-2019; Fig. 15 a-r) and test period (2020-2021; Fig. 15 s-aj), respectively. Spatial bias and ubRMSE for both models were comparable in both periods, ranging from -0.02

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

to 0.00 m^3/m^3 and 0.06 to 0.07 m^3/m^3, respectively. Notably, during the cross-validation period, the spatial SM prediction showed a negative bias (approximately -0.03 m^3/m^3) and lower ubRMSE (<0.06 m^3/m^3) within AOA compared to positive bias (0 to 0.03 m^3/m^3) and higher ubRMSE (>0.06 m^3/m^3) outside AOA (Fig. 15 a, b, d, e). Regarding the spatial R performance, SM predicted by RF demonstrated slightly better agreement with SMAP data, with a median value of 0.87 (Fig. 15 c), compared to XGB prediction with a median value of 0.86 (Fig. 15 f) within the cross-validation period. However, in the test period, both models achieved relatively moderate R agreement (0.75) with SMAP SM. Unlike bias and ubRMSE, the R performance within AOA did not show marked differences compared to areas outside the AOA. In summary, these findings highlight the capability of AOA in delineating a spatial extent conducive to reliable prediction, particularly evident in bias and ubRMSE metrics.

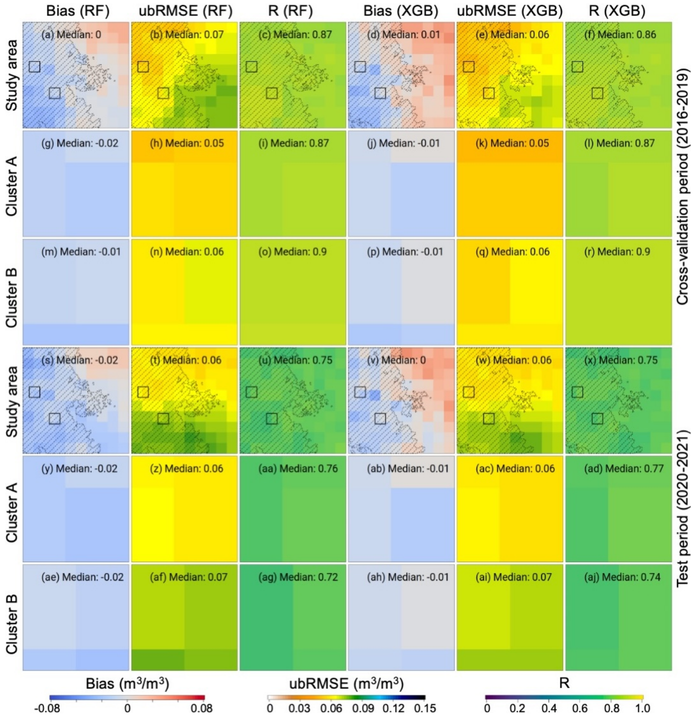

Fig. 15. Spatial distribution of bias, ubRMSE and R of SM predicted by the RF and XGB models against SMAP L2 for the entire study area, cluster A and cluster B within cross-validation period (2016-2019; a-r) and test period (2020-2021; s-aj), respectively. The spatial SM prediction was aggregated to 0.1 ° (∼10 km) resolution grids. AOA was displayed as the western area of the black boundaries, encompassing both spatial clusters.

V. DISCUSSION

Upscaling emerges as an attractive strategy to reconcile the spatial mismatch between point-scale data and satellite-derived SM grids [21]–[23], [67]. However, previous studies were often restricted to localised conditions and have overlooked the quantification of spatial uncertainty in SM estimates. In this study, a data-driven approach was introduced to spatially predict SM at a 100 m resolution in an Australian agricultural region. The AOA of two ML models were determined, both of which covered roughly 40% of the study region, and evaluated against multiple independent datasets. A few discussion points are summarised as follows.

A. *A further assessment of AOA*

The spatial pattern of AOA, as shown in Fig. 10, exhibited noticeable sensitivity to land cover and topographic features. To fully explore the spatial characteristics of AOA in SM prediction, a further assessment using an extended region in 2019 was conducted. The extended AOA analysis, encompassing a broader spatial extent beyond the primary study area, demonstrates alignment with natural landscape boundaries (Fig. 16), particularly evident in the correspondence between AOA delineation and cropland extent shown in Fig. A1 (a, b). This alignment is further reinforced by the distinct gradient in DI values that traces the transition from agricultural to non-agricultural lands. The spatial validation against SMAP L2 retrievals substantiates the AOA's effectiveness (Fig. 17), particularly evident in capturing the spatial area with relatively lower bias (Fig. 17 a), compared to positive/negative values in peripheral zones. This spatial coherence between AOA delineation and error structure demonstrates the metric's capacity to identify domains where the model's characteristics remain consistent with the training conditions, thereby providing a quantitative basis for constraining prediction reliability across heterogeneous landscapes. The DI distributions for both CV and CC validation approaches are constrained by these landscape features, with threshold values systematically higher in regions of homogeneous land cover and gentle topography (see Fig. 16 and Fig. A1). This systematic correspondence between AOA boundaries and model performance metrics validates the capability of AOA to identify reliable prediction zones based on landscape characteristics, while simultaneously highlighting the challenges of extrapolating SM predictions across distinct landscape units.

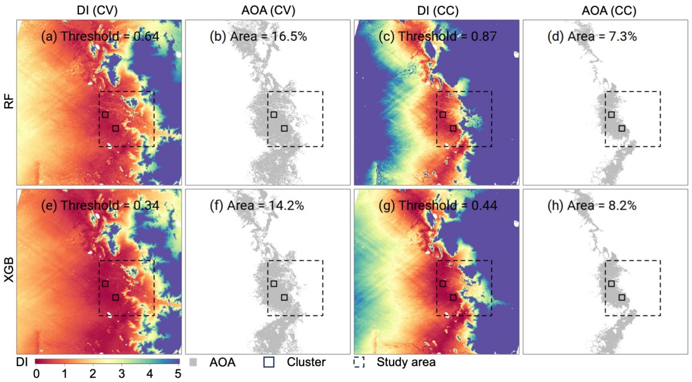

Fig. 16. Spatial distribution of median DI (with values exceeding 5 masked out) and AOA based on performance of CV and CC for (a-d) RF and (e-h) XGB for the extended region, respectively, between 01 Jan 2019 and 31 Dec 2019. The explanations of AOA and DI are the same as in Fig. 10 caption. Black and dashed rectangles indicate the cluster locations and the primary study area, respectively. The detailed spatial information of the extended area can be found in Fig. A1.

A quantitative analysis was also conducted to characterise surface properties within and outside the AOA boundaries,

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

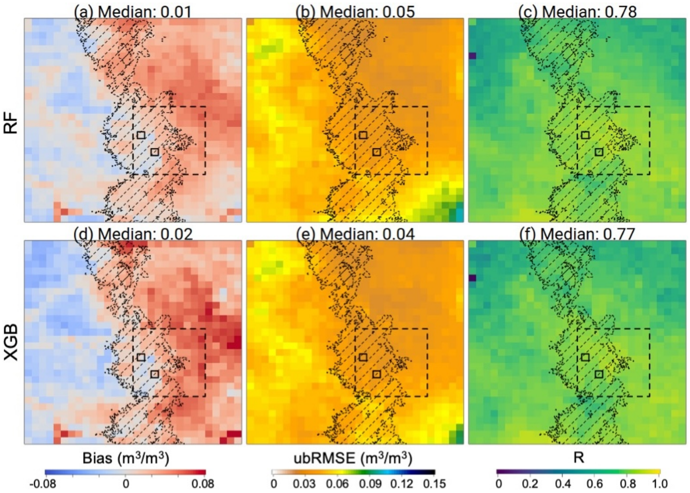

Fig. 17. Spatial distribution of bias, ubRMSE and R metrics of SM predicted by the RF and XGB models against SMAP L2 for the extended region in 2019, respectively. The spatial SM prediction was aggregated to 0.1 ° (∼10 km) resolution grids. AOA was displayed as the black polygons filled with dashed lines, encompassing the primary study area (dashed rectangle) and both spatial clusters (solid rectangles). The detailed spatial information of the extended area can be found in Fig. A1.

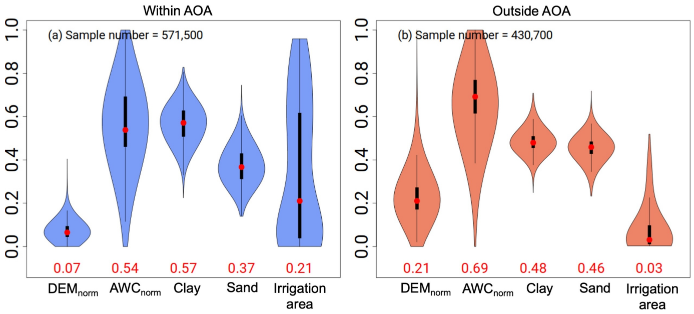

Fig. 18. Violin plots of normalised DEM (DEM_{norm}), normalised AWC (AWC_{norm}), clay, sand and area equipped for irrigation [68] for the regions (a) within AOA and (b) outside AOA. The red dots within the violins represent the median, with the decimal numbers at the bottom indicating the numeric values. The analysis is restricted to the primary study area (Fig. 1; 100 km × 100 km) that exhibits comparable pixel distributions within and outside AOA boundaries.

enabling the assessment of the model's applicability across distinct physiographic domains. Fig. 18 presents the distribution of normalised DEM (DEM_{norm}), normalised AWC (AWC_{norm}), clay, sand and area equipped for irrigation (Food and Agriculture Organisation) [68] for the regions (a) within AOA and (b) outside AOA. Regions with higher elevation and steeper slopes (Fig. 18 b) may experience different precipitation patterns and runoff behaviours compared to low-lying, flat regions (Fig. 18 a), thereby impacting the spatial distribution of SM. The flat region within AOA had more intensive irrigation activities (Fig. 18 a), with the normalised irrigation area (0.21) being seven times of that outside AOA (0.03). These irrigation activities are likely to homogenise SM patterns within AOA through reduced sub-pixel heterogeneity, making them more consistent and predictable. While explicit incorporation of irrigation factors remains methodologically challenging due to the complex interplay between management practices and natural hydrological processes, the AOA metric effectively captures these anthropogenic influences through their manifestation in predictor variables and resulting SM patterns. Additionally, the AOA pattern provided clear insights into the regional heterogeneity within the study area (as illustrated in Fig. 1). Unlike the regions within AOA, the eastern region, marked by more rugged and varied topography, was demonstrated to have higher uncertainty (as in comparison against SMAP L2; Fig. 15). The heterogeneity in the eastern region introduces variability in microclimates and soil properties, posing challenges for the ML models to generate reliable SM prediction. The regional differences highlighted by the AOA suggest the need for localised calibration of spatial prediction models to account for spatial heterogeneity. Furthermore, the land cover heterogeneity can also be an important factor to be considered in future efforts [13], as AOA delineates regions where environmental conditions are considered similar with those represented in the training dataset. The degradation of ML model performance across heterogeneous regions can be attributed to insufficient representation of specific multivariate environmental conditions, consistent with recent findings [13] that site representativeness improves with increased land cover homogeneity. Future studies should incorporate explicit land cover similarity metrics as auxiliary variables to enhance model interpretation across diverse physiographic domains.

The ML modelling processes leveraged the SHAP method [63], [64], which offered an avenue for gaining insights into the underlying principles governing SM interactions with various predictor variables from a data-driven perspective. The implementation of SHAP integrated with the underlying model structures: RF's ensemble averaging mechanism across independently constructed trees versus XGB's sequential gradient boosting approach optimising residual errors. These intrinsic algorithmic architectures, through which the two models construct feature-response relationships, are reflected in the observed differential ranking patterns of predictor variables between RF and XGB models. Especially, these differences manifested in the violin plots' distribution morphologies, with RF exhibiting more symmetrical contribution patterns (8 a-c) compared to XGB's pronounced multi-modality and wider contribution ranges (8 d-f). Despite the algorithmic variations in SHAP value distributions, both models identified consistent dominant predictor variables (i.e., albedo, DEM, ET, LST, NDVI and VPD) in the spatial CV (8 a, d), though their ranking orders exhibited nuances within different clusters characterised by varying environmental conditions.

Regarding the absence of soil texture variables (i.e., clay, sand, silt) among the most important predictor variables, several factors may contribute to this phenomenon. Firstly, there is a covarying relationship between AWC and soil texture data, leading to redundant information that may reduce the importance of soil texture predictor variables in ML models. Secondly, since this study focused on the top 0-5 cm of soil, the homogeneity of soil mapping for AWC and texture within this shallow depth range across the study

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

areas, particularly within the AOA envelope, could diminish the distinct contribution of soil texture variables. Thirdly, given the spatiotemporal variabilities inherent in SM dynamics, static soil texture data may not adequately capture the dynamic nature of SM variations over time, leading to their lower importance compared to the dynamically sensed predictor variables. If the objective was to predict long-term mean SM, static soil texture data might prove more contribution, as they reflect the overall capacity of a soil type to hold moisture. However, when modelling seasonal and interannual variability, the signal from static soil texture data is likely overwhelmed by the covarying predictors that capture dynamic changes. Finally, considering that the training data clusters cover small areas, it is plausible that the soil properties are relatively homogeneous within these clusters, further reducing the distinct contribution of soil texture variables in the models.

B. Strengths and prospects

Traditional SM upscaling approaches encompass a spectrum of spatial interpolation methodologies, from simple geometric interpolators such as arithmetic averaging [19] and Thiessen polygons [69], to distance-weighted techniques including inverse distance weighting [70] and various kriging implementations [71]. While these methods effectively utilize point-scale SM measurements, they operate independently of auxiliary remote sensing data that could potentially enhance spatial predictions. More recently, physically based approaches incorporating apparent thermal inertia (ATI), derived from the relationship between LST dynamics and soil thermal properties [20], have emerged as complementary techniques. However, current upscaling methodologies face several fundamental challenges. Geostatistical methods such as kriging, while highly effective near observation points, require densely distributed training data to establish reliable spatial dependence structures and assume stationarity in their underlying covariance functions. When applied beyond the range defined by the variogram model or with limited data availability, kriging essentially becomes an extrapolator, producing uniform predictions that may not accurately reflect true spatial variability. Additionally, these methods demonstrate limited capacity in capturing non-linear relationships across heterogeneous landscapes where multiple environmental factors influence SM distributions. Table IV summarises various representative SM upscaling studies from diverse ground networks.

Conversely, the proposed data-driven approach herein, distinguished by its AOA delineation, advances beyond these constraints. The AOA metric employs standardised predictor variables, importance-weighted Euclidean distances, and outlier-removal thresholds to establish prediction reliability bounds. This approach captures complex, non-linear patterns in predictor variables while providing quantitative metrics for prediction reliability - capabilities not readily available in traditional geostatistical frameworks. The effectiveness of AOA is evidenced by the alignment between AOA boundaries and validation metrics against SMAP retrievals (remarkably lower bias and ubRMSE within AOA), particularly in regions where traditional interpolation methods would be constrained by their inherent assumptions about spatial dependence structures.

The incorporation of diverse predictor variables enhanced its potential for application at a broader scale. Burton et al. [73] demonstrated the effectiveness of ML-based approaches in spatially predicting point-scale eddy covariance flux data across a continental scale. While acknowledging the inherent differences between SM and carbon flux data (e.g., greater spatiotemporal variabilities in SM), this study suggested a viable prediction approach of SM at larger geographic extent or higher temporal frequency. However, when dealing with large spatial extents, the representativeness and applicability of the relationship between in-situ SM measurements and predictor variables for successful prediction remain uncertain [13]. Ensuring the robustness of these relationships across diverse landscapes and climatic conditions is critical for the success of spatial prediction efforts. Careful consideration of the selection and evaluation of predictor variables is crucial for improving the generalisability and reliability of SM prediction at broader scales [74]. Apart from data sourced from polar-orbiting platforms (i.e., MODIS herein), it is also worth considering data obtained from geostationary platforms, such as Himawari-8/9 [75] and Fengyun-4A [76]. These geostationary platforms present opportunities to monitor more detailed sub-daily vegetation and phenology dynamics [77], LST variations [78], and gross primary productivity [79]. Integrating data from such platforms can help enhance the diurnal predictive capability of ML models. Nevertheless, the choice of predictor variables should undergo iterations of assessment to ensure the reliability and accuracy of the prediction results [74]. Accurately quantifying the contribution of essential predictor variables not only improves the operational efficiency of this approach but also aids in reducing computational resources, making it more practical for widespread use in larger-scale SM assessments.

The proposed approach integrated spatiotemporal fusion with ML models to spatially predict in-situ SM measurements, achieving a spatial resolution of 100 m. This resolution, closely approximating the point-scale, is particularly well-suited for agricultural applications, allowing for the effective capture of landscape heterogeneity. However, the adoption of spatiotemporal fusion techniques could introduce additional uncertainties. These uncertainties arise from the reliance on regression coefficients derived from coarse-resolution image pairs, leading to concerns about their applicability at finer resolutions. While internal bias correction steps for LST data have been employed to mitigate such errors [58], quantifying the temporal variations of surface reflectance data (albedo and NDVI), particularly in heterogeneous agricultural regions, remains challenging. The inherent complexities of these regions can introduce variability that may not be fully captured or corrected by existing methods, emphasising the need for careful consideration and potential refinement of fusion techniques in such environments.

Several critical aspects warrant attention and potential refinement in future studies. Firstly, the reliability and representativeness of in-situ SM measurements could influence the accuracy of the spatial prediction results [13]. The quality of in-situ data is contingent on sensor calibration, monitoring techniques, and adherence to network standards. Meanwhile,

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

TABLE IV

AN OVERVIEW OF REPRESENTATIVE SM UPSCALING STUDIES. STUDIES ARE ORDERED CHRONOLOGICALLY BY YEAR OF PUBLICATION.

Ground network	Method and predictor variables	Objectives	Key findings	Potential limitations	Reference
Soil Moisture / Temperature Monitoring Network (SMTMN), Tibet Autonomous Region, China	Bayesian linear regression with MODIS-derived apparent thermal inertia (ATI).	Upscale in-situ SM to an area-averaged scale that is comparable to satellite footprints.	The upscaling approach reduced RMSE from 0.023 to 0.013 m^3/m^3, with the temporal pattern of upscaled SM agreeing with local hydro-meteorological knowledge.	The results lacked evaluations against independent data (e.g., from satellite missions or field campaigns).	[20]
Heihe Water Allied Telemetry Experimental Research experiment (HiWATER), Gansu Province, China	Spatiotemporal regression block kriging (STRBK) model with information of vegetation and LST.	Upscale in-situ SM observations to validate the Polarimetric L-band Multi-beam Radiometer (PLMR) SM data.	STRBK can yield upscaled SM with satisfactory accuracy for validation purposes. PLMR SM showed an RMSE of 0.03 m^3/m^3 against upscaled SM and was considered applicable for watershed-scale SM monitoring.	The collected in-situ SM only spanned a 28-day period and the evaluations of PLMR SM were only conducted on two days.	[21]
Soil moisture Sensing Controller And oPtimal Estimator (SoilSCAPE) network, California, United States	RF model with various landscape parameters.	Upscale in-situ SM to validate SMAP products at various scales, with additional validation efforts against data from the SMAPVEX12 campaign and an additional network.	The proposed approach demonstrated better capability to predict mean SM than other upscaling methods in comparison against field measurements.	The distribution of in-situ sties was relatively clustered and the AOA of RF model was not assessed.	[23]
Soil Moisture / Temperature Monitoring Network (SMTMN), Tibet Autonomous Region, China	Bayesian data fusion method with MODIS ATI data.	Upscale in-situ SM to microwave remote sensing footprint.	The upscaled SM derived using data from randomly selected three stations was comparable with the mean values of SM from 20 stations, with the mean RMSE being 0.02 m^3/m^3.	The results lacked evaluations against independent data (e.g., from satellite missions or field campaigns).	[67]
A network with 20 soil monitoring stations, Twente region, Netherlands	Dutch National Hydrological model with scaling parameters derived from observed and simulated SM.	Develop upscaling methods to translate the spatial mean of point-scale SM to SMAP domain, hence to assess the performance of SMAP data retrieved using different algorithms.	The afternoon-acquired Single Channel Algorithm at Vertical Polarization (SCA-V) SMAP SM retrievals had best agreement with native spatial mean of in-situ SM, being an ubRMSE of 0.059 m^3/m^3, while higher bias was found when compared against upscaled SM.	The results indicated none of the SMAP products met the target accuracy using the Twente measurements, which may be attributed to frozen conditions and antecedent precipitation.	[72]
Marena, Oklahoma, In Situ Sensor Testbed (MOISST), Oklahoma, United States	Least Absolute Shrinkage and Selection Operator (LASSO) regression model with predictor variables.	Develop a broadly applicable upscaling approach that has transferrable capability between sites.	The proposed model only required point-scale SM and sand data, and achieved a reasonable accuracy, being an RMSE between 0.006 and 0.031 m^3/m^3 among six sites.	The results were cross-validated while lacked evaluations against independent data (e.g., from satellite missions or field campaigns).	[22]
OzNet Hydrological Monitoring Network, Yanco, Australia	ML models and spatiotemporal fusion, with multi-source predictor variables.	Spatially predict SM at 100 m resolution and assess its AOA, with evaluation against independent data from field campaigns, external networks and SMAP data.	Spatial cross-validation demonstrated an ubRMSE (R) of 0.05 m^3/m^3 (0.78). Independent evaluations demonstrated mean R values of 0.62-0.64 against field data, 0.84-0.91 against additional SM networks, and 0.87 against satellite retrievals, respectively.	The in-situ sites were spatially clustered, potentially limiting their representativeness of the general conditions of study area. The results were especially impacted by topographic features.	This study

the spatial prediction approach also relies on the spatial representativeness of the selected clusters for cross-validation. Inadequate capturing of variability in land cover, soil types, or topography within these clusters may hinder the generalisation of results to other regions. External factors, such as climate extremes, can also impact data continuity, as exemplified by the discontinuity caused by the 2019 wildfires in Australia affecting the OzNet network. Secondly, the transferability of this approach requires careful consideration. Achieving success in spatial prediction within one region or cluster does not guarantee similar performance in areas with distinct environmental conditions. The efficacy of this approach may vary when applied to regions characterised by different land cover types, climatic regimes, or soil characteristics. Future research should focus on evaluating the adaptability of these approaches in diverse environments, including vegetated or arid regions. Thirdly and finally, the challenge of insufficient validation sites poses a potential limitation. In regions with a sparse distribution of in-situ measurement sites or an uneven distribution among clusters, the robustness of cross-validation results may be compromised. This issue is particularly pronounced in the southern hemisphere like Australia, where non-agricultural regions lack adequate representation, further complicating the application of this approach. Efforts to address these limitations and refine the methodology should be a priority in future research.

VI. CONCLUSIONS

Previous SM upscaling studies were often restricted to localised conditions and overlooked the quantification of spatial uncertainty in SM estimates. This study introduced a spatial prediction approach that integrates ML and spatiotemporal fusion, while implementing the AOA metric to assess the prediction capabilities of two ML models. Point-scale SM from 28 OzNet sites was extrapolated to a 100 km × 100 km area at 100 m resolution with daily frequency, spanning both a cross-validation period (2016-2019) and a test period (2020-2021). The AOA metrics of two ML models, RF and XGB, were determined based on their spatial CV performance and feature importance.

The AOA of RF and XGB covered 43.1% and 41.5% of the study area, respectively, with an expected ubRMSE of 0.05 m^3/m^3 and an R of 0.78. The spatial SM prediction delineated by AOA was then evaluated against multiple independent datasets (HDAS, CosmOz, OzFlux, and SMAP L2). The results indicated that RF-predicted SM demonstrated more robustness and better agreement with independent data. Specifically, during the cross-validation period, RF-predicted SM achieved mean R values of 0.62-0.64 against HDAS SM, 0.84-0.91 against independent SM networks, and 0.87 against SMAP L2 data, respectively. Notably, spatial SM prediction within AOA showed lower uncertainties against SMAP L2 compared to areas outside AOA, which was further assessed across an extended spatial domain (Fig. A1).

In summary, this study offered a reliable approach for extrapolating point-scale SM data to match the resolution of Landsat satellite footprint, influenced by regional nuances in predictor variables. Future research directions may explore the variability of AOA by examining diverse predictor variables, and assess the transferability of the proposed approach across various environmental conditions and land cover types.

APPENDIX A
ILLUSTRATION OF THE EXTENDED REGION

The extended area comprises a mosaic of six Landsat tiles (path: 092-093, row: 083-085) to evaluate westward extensibility of AOA. Fig. A1 presents (a) land cover classification information; (b) USGS cropland extent dataset differentiating between irrigated and rainfed agricultural systems; and (c) elevation gradients ranging from 0 to 400 m. The domain encompasses a physiographic transition from the agricultural core in the east (146.0-147.0 °E) to predominantly non-agricultural regions in the west (145.0-146.0 °E), providing a test case for assessing AOA performance across distinct landscape boundaries.

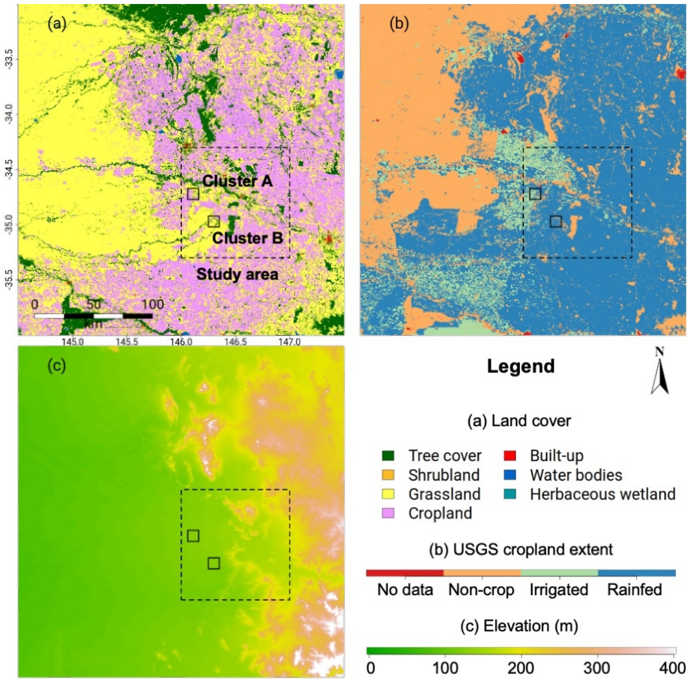

Fig. A1. (a) The land cover information of the extended area [49]; (b) the USGS cropland extent [50]; and (c) DEM-S [44]. The spatial extent of the extended area spans 144.50-147.50 °E and 33.00-36.00 °S (300 km × 300 km).

ACKNOWLEDGEMENTS

This research was performed as part of a PhD by the first author (Y. Yu) under an academic collaboration between the Australian National University (ANU) and the Commonwealth Scientific and Industrial Research Organisation (CSIRO). This research was undertaken while supported by the ANU University Research Scholarship and an ANU-CSIRO Digital Agriculture PhD Supplementary Scholarship. This research was partially supported by the Grains Research and Development Corporation (GRDC) project *SoilWaterNow: Soil water* *nowcasting for the grains industry* (UOS2002-001RTX). The PRISM field campaigns were supported by the Australian Research Council through the Discovery Project *Towards P-Band Soil Moisture Sensing from Space* (DP170102373). This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS). This research was supported by resources and expertise provided by CSIRO IMT Scientific Computing. The authors acknowledge the OzNet (https://www.oznet.org.au), CosmOz (https://cosmoz.csiro.au), and OzFlux (https://www.ozflux.org.au) networks for provision and processing of data. The authors thank the TGRS editorial team and the two anonymous reviewers for their constructively critical comments which provided the catalyst to improve an earlier version of this research.

CODE AND DATA AVAILABILITY

The analysis scripts for this research are publicly available at https://github.com/yuyi13/OzNet_AOA, with a permalink registered at Zenodo (https://doi.org/10.5281/zenodo.15290777). The experimental data are publicly available from the CSIRO Data Access Portal (https://doi.org/102.100.100/705658).

REFERENCES

[1] Sonia I Seneviratne, Thierry Corti, Edouard L Davin, Martin Hirschi, Eric B Jaeger, Irene Lehner, Boris Orlowsky, and Adriaan J Teuling, “Investigating soil moisture–climate interactions in a changing climate: A review,” *Earth-Science Reviews*, vol. 99, no. 3, pp. 125–161, 2010, https://doi.org/10.1016/j.earscirev.2010.02.004.

[2] Corrado Corradini, Renato Morbidelli, and Florisa Melone, “On the interaction between infiltration and Hortonian runoff,” *Journal of Hydrology*, vol. 204, no. 1, pp. 52–67, 1998, https://doi.org/10.1016/s0022-1694(97)00100-5.

[3] Shmuel Assouline, “Infiltration into soils: Conceptual approaches and solutions,” *Water Resources Research*, vol. 49, no. 4, pp. 1755–1772, 2013, https://doi.org/10.1002/wrcr.20155.

[4] Gianpaolo Balsamo, Anna Agusti-Panareda, Clement Albergel, Gabriele Arduini, Anton Beljaars, Jean Bidlot, Eleanor Blyth, Nicolas Bousserez, Souhail Boussetta, Andy Brown, Roberto Buizza, Carlo Buontempo, Frédéric Chevallier, Margarita Choulga, Hannah Cloke, Meghan F Cronin, Mohamed Dahoui, Patricia De Rosnay, Paul A Dirmeyer, Matthias Drusch, Emanuel Dutra, Michael B Ek, Pierre Gentine, Helene Hewitt, Sarah P E Keeley, Yann Kerr, Sujay Kumar, Cristina Lupu, Jean-François Mahfouf, Joe McNorton, Susanne Mecklenburg, Kristian Mogensen, Joaquín Muñoz-Sabater, Rene Orth, Florence Rabier, Rolf Reichle, Ben Ruston, Florian Pappenberger, Irina Sandu, Sonia I Seneviratne, Steffen Tietsche, Isabel F Trigo, Remko Uijlenhoet, Nils Wedi, R Iestyn Woolway, and Xubin Zeng, “Satellite and In Situ Observations for Advancing Global Earth Surface Modelling: A Review,” *Remote Sensing*, vol. 10, no. 12, pp. 2038, 2018, https://doi.org/10.3390/rs10122038.

[5] Jian Peng, Clement Albergel, Anna Balenzano, Luca Brocca, Oliver Cartus, Michael H Cosh, Wade T Crow, Katarzyna Dabrowska-Zielinska, Simon Dadson, Malcolm W J Davidson, Patricia de Rosnay, Wouter Dorigo, Alexander Gruber, Stefan Hagemann, Martin Hirschi, Yann H Kerr, Francesco Lovergine, Miguel D Mahecha, Philip Marzahn, Francesco Mattia, Jan Pawel Musial, Swantje Preuschmann, Rolf H Reichle, Giuseppe Satalino, Martyn Silgram, Peter M van Bodegom, Niko E C Verhoest, Wolfgang Wagner, Jeffrey P Walker, Urs Wegmüller, and Alexander Loew, “A roadmap for high-resolution satellite soil moisture applications – confronting product characteristics with user requirements,” *Remote Sensing of Environment*, vol. 252, pp. 112162, 2021, https://doi.org/10.1016/j.rse.2020.112162.

[6] Y H Kerr, P Waldteufel, J P Wigneron, J Martinuzzi, J Font, and M Berger, “Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 8, pp. 1729–1735, 2001, https://doi.org/10.1109/36.942551.

[7] Zoltan Bartalis, Wolfgang Wagner, Vahid Naeimi, Stefan Hasenauer, Klaus Scipal, Hans Bonekamp, Julia Figa, and Craig Anderson, “Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT),” *Geophysical Research Letters*, vol. 34, no. 20, 2007, https://doi.org/10.1029/2007gl031088.

[8] D Entekhabi, E G Njoku, P E O’ Neill, K H Kellogg, W T Crow, W N Edelstein, J K Entin, S D Goodman, T J Jackson, J Johnson, J Kimball, J R Piepmeier, R D Koster, N Martin, K C McDonald, M Moghaddam, S Moran, R Reichle, J C Shi, M W Spencer, S W Thurman, L Tsang, and J Van Zyl, “The Soil Moisture Active Passive (SMAP) Mission,” *Proceedings of the IEEE*, vol. 98, no. 5, pp. 704–716, 2010, https://doi.org/10.1109/jproc.2010.2043918.

[9] Jianzhi Dong, Wade T Crow, Kenneth J Tobin, Michael H Cosh, David D Bosch, Patrick J Starks, Mark Seyfried, and Chandra Holifield Collins, “Comparison of microwave remote sensing and land surface modeling for surface soil moisture climatology estimation,” *Remote Sensing of Environment*, vol. 242, pp. 111756, 2020, https://doi.org/10.1016/j.rse.2020.111756.

[10] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Saleh Abdalla, Xavier Abellan, Gianpaolo Balsamo, Peter Bechtold, Gionata Biavati, Jean Bidlot, Massimo Bonavita, Giovanna De Chiara, Per Dahlgren, Dick Dee, Michail Diamantakis, Rossana Dragani, Johannes Flemming, Richard Forbes, Manuel Fuentes, Alan Geer, Leo Haimberger, Sean Healy, Robin J Hogan, Elías Hólm, Marta Janisková, Sarah Keeley, Patrick Laloyaux, Philippe Lopez, Cristina Lupu, Gabor Radnoti, Patricia de Rosnay, Iryna Rozum, Freja Vamborg, Sebastien Villaume, and Jean-Noël Thépaut, “The ERA5 global reanalysis,” *Quarterly Journal of the Royal Meteorological Society*, vol. 146, no. 730, pp. 1999–2049, 2020, https://doi.org/10.1002/qj.3803.

[11] Rolf H Reichle, Clara S Draper, Q Liu, Manuela Girotto, Sarith P P Mahanama, Randal D Koster, and Gabrielle J M De Lannoy, “Assessment of MERRA-2 Land Surface Hydrology Estimates,” *Journal of Climate*, vol. 30, no. 8, pp. 2937–2960, 2017, https://doi.org/10.1175/jcli-d-16-0720.1.

[12] Jiangyuan Zeng, Zhen Li, Quan Chen, Haiyun Bi, Jianxiu Qiu, and Pengfei Zou, “Evaluation of remotely sensed and reanalysis soil moisture products over the Tibetan Plateau using in-situ observations,” *Remote Sensing of Environment*, vol. 163, pp. 91–110, 2015, https://doi.org/10.1016/j.rse.2015.03.008.

[13] Chenchen Peng, Jiangyuan Zeng, Kun-Shan Chen, Hongliang Ma, Husi Letu, Xiang Zhang, Pengfei Shi, and Haiyun Bi, “Spatial Representativeness of Soil Moisture Stations and Its Influential Factors at a Global Scale,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 63, pp. 1–15, 2025, https://doi.org/10.1109/TGRS.2024.3523484.

[14] Steffen Zacharias, Heye Bogena, Luis Samaniego, Matthias Mauder, Roland Fuß, Thomas Pütz, Mark Frenzel, Mike Schwank, Cornelia Baessler, Klaus Butterbach-Bahl, Oliver Bens, Erik Borg, Achim Brauer, Peter Dietrich, Irena Hajnsek, Gerhard Helle, Ralf Kiese, Harald Kunstmann, Stefan Klotz, Jean Charles Munch, Hans Papen, Eckart Priesack, Hans Peter Schmid, Rainer Steinbrecher, Ulrike Rosenbaum, Georg Teutsch, and Harry Vereecken, “A Network of Terrestrial Environmental Observatories in Germany,” *Vadose Zone Journal*, vol. 10, no. 3, pp. 955–973, 2011, https://doi.org/10.2136/vzj2010.0139.

[15] J G Evans, H C Ward, J R Blake, E J Hewitt, R Morrison, M Fry, L A Ball, L C Doughty, J W Libre, O E Hitt, D Rylett, R J Ellis, A C Warwick, M Brooks, M A Parkes, G M H Wright, A C Singer, D B Boorman, and A Jenkins, “Soil water content in southern England derived from a cosmic-ray soil moisture observing system – COSMOS-UK,” *Hydrological Processes*, vol. 30, no. 26, pp. 4987–4999, 2016, https://doi.org/10.1002/hyp.10929.

[16] Aaron Hawdon, David McJannet, and Jim Wallace, “Calibration and correction procedures for cosmic-ray neutron soil moisture probes located across Australia,” *Water Resources Research*, vol. 50, no. 6, pp. 5029–5043, 2014, https://doi.org/10.1002/2013wr015138.

[17] A B Smith, J P Walker, A W Western, R I Young, K M Ellett, R C Pipunic, R B Grayson, L Siriwardena, F H S Chiew, and H Richter, “The Murrumbidgee soil moisture monitoring network data set,” *Water Resources Research*, vol. 48, no. 7, 2012, https://doi.org/10.1029/2012wr011976.

[18] L Brocca, R Morbidelli, F Melone, and T Moramarco, “Soil moisture spatial variability in experimental areas of central Italy,” *Journal of Hydrology*, vol. 333, no. 2, pp. 356–373, 2007, https://doi.org/10.1016/j.jhydrol.2006.09.004.

[19] Wade T Crow, Aaron A Berg, Michael H Cosh, Alexander Loew, Binayak P Mohanty, Rocco Panciera, Patricia de Rosnay, Dongryeol Ryu, and Jeffrey P Walker, “Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products,” *Reviews of Geophysics*, vol. 50, no. 2, 2012, https://doi.org/10.1029/2011rg000372.

[20] Jun Qin, Kun Yang, Ning Lu, Yingying Chen, Long Zhao, and Menglei Han, “Spatial upscaling of in-situ soil moisture measurements based on MODIS-derived apparent thermal inertia,” *Remote Sensing of Environment*, vol. 138, pp. 1–9, 2013, https://doi.org/10.1016/j.rse.2013.07.003.

[21] Jianghao Wang, Yong Ge, Gerard B M Heuvelink, and Chenghu Zhou, “Upscaling In Situ Soil Moisture Observations to Pixel Averages with Spatio-Temporal Geostatistics,” *Remote Sensing*, vol. 7, no. 9, pp. 11372–11388, 2015, https://doi.org/10.3390/rs70911372.

[22] William G Brown, Michael H Cosh, Jingnuo Dong, and Tyson E Ochsner, “Upscaling soil moisture from point scale to field scale: Toward a general model,” *Vadose Zone Journal*, vol. 22, no. 2, pp. e20244, 2023, https://doi.org/10.1002/vzj2.20244.

[23] D Clewley, J B Whitcomb, R Akbar, A R Silva, A Berg, J R Adams, T Caldwell, D Entekhabi, and M Moghaddam, “A Method for Upscaling In Situ Soil Moisture Measurements to Satellite Footprint Scale Using Random Forests,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 6, pp. 2663–2673, 2017, https://doi.org/10.1109/jstars.2017.2690220.

[24] Yaping Xu, Cuiling Liu, Lei Wang, and Lei Zou, “Exploring the Spatial Autocorrelation in Soil Moisture Networks: Analysis of the Bias from Upscaling the Texas Soil Observation Network (TxSON),” *Water*, vol. 15, no. 1, pp. 87, 2023, https://doi.org/10.3390/w15010087.

[25] Hanna Meyer and Edzer Pebesma, “Predicting into unknown space? Estimating the area of applicability of spatial prediction models,” *Methods in Ecology and Evolution*, vol. 12, no. 9, pp. 1620–1633, 2021, https://doi.org/10.1111/2041-210x.13650.

[26] Hanna Meyer and Edzer Pebesma, “Machine learning-based global maps of ecological variables and the challenge of assessing them,” *Nature Communications*, vol. 13, no. 1, pp. 2208, 2022, https://doi.org/10.1038/s41467-022-29838-9.

[27] Tomislav Hengl, Matthew A E Miller, Josip Križan, Keith D Shepherd, Andrew Sila, Milan Kilibarda, Ognjen Antonijević, Luka Glušica, Achim Dobermann, Stephan M Haefele, Steve P McGrath, Gifty E Acquah, Jamie Collinson, Leandro Parente, Mohammadreza Sheykhmousa, Kazuki Saito, Jean-Martial Johnson, Jordan Chamberlin, Francis B T Silatsa, Martin Yemefack, John Wendt, Robert A MacMillan, Ichsani Wheeler, and Jonathan Crouch, “African soil properties and nutrients mapped at 30 m spatial resolution using two-scale ensemble machine learning,” *Scientific Reports*, vol. 11, no. 1, pp. 6130, 2021, https://doi.org/10.1038/s41598-021-85639-y.

[28] Daniel Žížala, Robert Minařík, Jan Skála, Hana Beitlerová, Anna Juřicová, Jessica Reyes Rojas, Vít Penížek, and Tereza Zádorová, “High-resolution agriculture soil property maps from digital soil mapping methods, Czech Republic,” *CATENA*, vol. 212, pp. 106024, 2022, https://doi.org/10.1016/j.catena.2022.106024.

[29] Tom Broeg, Michael Blaschek, Steffen Seitz, Ruhollah Taghizadeh-Mehrjardi, Simone Zepp, and Thomas Scholten, “Transferability of Covariates to Predict Soil Organic Carbon in Cropland Soils,” *Remote Sensing*, vol. 15, no. 4, 2023, https://doi.org/10.3390/rs15040876.

[30] Lilian-Maite Lezama Valdes, Marwan Katurji, and Hanna Meyer, “A Machine Learning Based Downscaling Approach to Produce High Spatio-Temporal Resolution Land Surface Temperature of the Antarctic Dry Valleys from MODIS Data,” *Remote Sensing*, vol. 13, no. 22, 2021, https://doi.org/10.3390/rs13224673.

[31] Theodora Lendzioch, Jakub Langhammer, Lukáš Vlček, and Robert Minařík, “Mapping the Groundwater Level and Soil Moisture of a Montane Peat Bog Using UAV Monitoring and Machine Learning,” *Remote Sensing*, vol. 13, no. 5, 2021, https://doi.org/10.3390/rs13050907.

[32] Jian Peng, Alexander Loew, Olivier Merlin, and Niko E C Verhoest, “A review of spatial downscaling of satellite remotely sensed soil moisture,” *Reviews of Geophysics*, vol. 55, no. 2, pp. 341–366, 2017, https://doi.org/10.1002/2016rg000543.

[33] Ahmed Samir Abowarda, Liangliang Bai, Caijin Zhang, Di Long, Xueying Li, Qi Huang, and Zhangli Sun, “Generating surface soil moisture at 30 m spatial resolution using both data fusion and machine learning toward better water resources management at the field scale,” *Remote Sensing of Environment*, vol. 255, pp. 112301, 2021, https://doi.org/10.1016/j.rse.2021.112301.

[34] M. Stenson, R. Searle, B. P. Malone, A. Sommer, L. J. Renzullo, and H. Di, “Australia wide daily volumetric soil moisture estimates. Version 1.0 [Dataset],” 2021, https://doi.org/https://doi.org/10.25901/b020-nm39.

[35] Yi Yu, Brendan P. Malone, and Luigi J. Renzullo, “Empirical Upscaling of Point-Scale Soil Moisture Measurements for Spatial Evaluation of Model Simulations and Satellite Retrievals,” in *2024 IEEE International Geoscience and Remote Sensing Symposium*, Athens, Greece, July 2024, pp. 11496–11501, https://doi.org/10.1109/IGARSS53475.2024.10642763.

[36] Elaheh Ghafari, Jeffrey P Walker, Liujun Zhu, Andreas Colliander, and Alireza Faridhosseini, “Spatial downscaling of SMAP radiometer soil moisture using radar data: Application of machine learning to the SMAPEx and SMAPVEX campaigns,” *Science of Remote Sensing*, vol. 9, pp. 100122, 2024, https://doi.org/10.1016/j.srs.2024.100122.

[37] Chenchen Peng, Jiangyuan Zeng, Kun-Shan Chen, Zhen Li, Hongliang Ma, Xiang Zhang, Pengfei Shi, Tiantian Wang, Lu Yi, and Haiyun Bi, “Global spatiotemporal trend of satellite-based soil moisture and its influencing factors in the early 21st century,” *Remote Sensing of Environment*, vol. 291, pp. 113569, 2023, https://doi.org/10.1016/j.rse.2023.113569.

[38] Jason Beringer, Lindsay B Hutley, Ian McHugh, Stefan K Arndt, David Campbell, Helen A Cleugh, James Cleverly, Víctor Resco de Dios, Derek Eamus, and Bradley Evans, “An introduction to the Australian and New Zealand flux tower network - OzFlux,” *Biogeosciences*, vol. 13, no. 21, pp. 5895–5916, 2016, https://doi.org/10.5194/bg-13-5895-2016.

[39] Crystal Schaaf and Zhuosen Wang, “MCD43A4 MODIS/Terra+Aqua Nadir BRDF-Adjusted Reflectance Daily L3 Global - 500m,” 2015, https://doi.org/10.5067/MODIS/MCD43A4.006.

[40] Zhengming Wan, “New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product,” *Remote Sensing of Environment*, vol. 140, pp. 36–45, 2014, https://doi.org/10.1016/j.rse.2013.08.027.

[41] F Li, D L B Jupp, S Reddy, L Lymburner, N Mueller, P Tan, and A Islam, “An Evaluation of the Use of Atmospheric and BRDF Correction to Standardize Landsat Data,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 3, no. 3, pp. 257–270, 2010, https://doi.org/10.1109/jstars.2010.2042281.

[42] Juan P Guerschman, Tim R McVicar, Jamie Vleeshower, Thomas G Van Niel, Jorge L Peña-Arancibia, and Yun Chen, “Estimating actual evapotranspiration at field-to-continent scales by calibrating the CMRSET algorithm with MODIS, VIIRS, Landsat and Sentinel-2 data,” *Journal of Hydrology*, vol. 605, pp. 127318, 2022, https://doi.org/10.1016/j.jhydrol.2021.127318.

[43] Michael Hutchinson, Tingbao Xu, Jennifer Kesteven, Ian Marang, and Bradley Evans, “ANUClimate 2.0,” 2021, https://doi.org/https://doi.org/10.25914/60A10AA56DD1B.

[44] Geoscience Australia, “Digital Elevation Model (DEM) of Australia derived from LiDAR 5 Metre Grid,” 2015, Available at: https://developers.google.com/earth-engine/datasets/catalog/AU_GA_DEM_1SEC_v10_DEM-S.

[45] Brendan Malone and Ross Searle, “Updating the Australian digital soil texture mapping (Part 2): spatial modelling of merged field and lab measurements,” *Soil Research*, vol. 59, no. 5, pp. 435–451, 2021, https://doi.org/10.1071/sr20284.

[46] R. D. Searle, P. D. S. N. Somarathna, and B. P. Malone, “Soil and Landscape Grid National Soil Attribute Maps - Available Volumetric Water Capacity (Percent) (3 arc second resolution) Version 2,” 2022, https://doi.org/https://doi.org/10.25919/4jwj-na34.

[47] X Wu, N Ye, J Walker, I.-Y. Yeo, T Jackson, Y Kerr, E Kim, and A McGrath, “The P-band Radiometer Inferred Soil Moisture Experiment 2019 WORKPLAN,” 2019, Available at: https://prism.monash.edu/PRISM-19.html.

[48] X Wu, N Ye, J Walker, I.-Y. Yeo, T Jackson, Y Kerr, E Kim, and A McGrath, “The P-band Radiometer Inferred Soil Moisture Experiment 2021 WORKPLAN,” 2021, Available at: https://prism.monash.edu/PRISM-21.html.

[49] Daniele Zanaga, Ruben Van De Kerchove, Dirk Daems, Wanda De Keersmaecker, Carsten Brockmann, Grit Kirches, Jan Wevers, Oliver Cartus, Maurizio Santoro, Steffen Fritz, Myroslava Lesiv, Martin Herold, Nandin-Erdene Tsendbazar, Panpan Xu, Fabrizio Ramoino, and Olivier Arino, “ESA WorldCover 10 m 2021 v200,” Oct. 2022, https://doi.org/10.5281/ZENODO.7254221.

[50] Pardhasaradhi Teluguntla, Prasad S Thenkabail, Adam Oliphant, Jun Xiong, Murali Krishna Gumma, Russell G Congalton, Kamini Yadav,

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

and Alfredo Huete, "A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 144, pp. 325–340, 2018, https://doi.org/10.1016/j.isprsjprs.2018.07.017.

[51] Irina V Emelyanova, Tim R McVicar, Thomas G Van Niel, Ling Tao Li, and Albert I J M van Dijk, "Assessing the accuracy of blending Landsat–MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: A framework for algorithm selection," *Remote Sensing of Environment*, vol. 133, pp. 193–209, 2013, https://doi.org/10.1016/j.rse.2013.02.007.

[52] C O Justice, E Vermote, J R G Townshend, R Defries, D P Roy, D K Hall, V V Salomonson, J L Privette, G Riggs, A Strahler, W Lucht, R B Myneni, Y Knyazikhin, S W Running, R R Nemani, Wan Zhengming, A R Huete, W van Leeuwen, R E Wolfe, L Giglio, J Muller, P Lewis, and M J Barnsley, "The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 4, pp. 1228–1249, 1998, https://doi.org/10.1109/36.701075.

[53] Shunlin Liang, "Narrowband to broadband conversions of land surface albedo I: Algorithms," *Remote Sensing of Environment*, vol. 76, no. 2, pp. 213–238, 2001, https://doi.org/10.1016/s0034-4257(00)00205-4.

[54] Juan C Jiménez-Muñoz, José A Sobrino, Dražen Skoković, Cristian Mattar, and Jordi Cristóbal, "Land surface temperature retrieval methods from Landsat-8 thermal infrared sensor data," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 10, pp. 1840–1843, 2014, https://doi.org/10.1109/lgrs.2014.2312032.

[55] M F Hutchinson, "Interpolating mean rainfall using thin plate smoothing splines," *International Journal of Geographical Information Systems*, vol. 9, no. 4, pp. 385–403, 1995, https://doi.org/10.1080/02693799508902045.

[56] Olivier Merlin, Jeffrey P Walker, Rocco Panciera, R Young, Jetse Kalma, and E Kim, "Soil Moisture Measurement in Heterogeneous Terrain," in *MODSIM 2007 International Congress on Modelling and Simulation*, Christchurch, New Zealand, 2007, Available at: https://mssanz.org.au/MODSIM07/papers/46_s60/SoilMoistures60_Merlin_.pdf.

[57] Xiaolin Zhu, Jin Chen, Feng Gao, Xuehong Chen, and Jeffrey G Masek, "An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions," *Remote Sensing of Environment*, vol. 114, no. 11, pp. 2610–2623, 2010, https://doi.org/10.1016/j.rse.2010.05.032.

[58] Yi Yu, Luigi J Renzullo, Tim R McVicar, Brendan P Malone, and Siyuan Tian, "Generating daily 100 m resolution land surface temperature estimates continentally using an unbiased spatiotemporal fusion approach," *Remote Sensing of Environment*, vol. 297, pp. 113784, 2023, https://doi.org/10.1016/j.rse.2023.113784.

[59] David J Lary, Amir H Alavi, Amir H Gandomi, and Annette L Walker, "Machine learning in geosciences and remote sensing," *Geoscience Frontiers*, vol. 7, no. 1, pp. 3–10, 2016, https://doi.org/10.1016/j.gsf.2015.07.003.

[60] Leo Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, https://doi.org/10.1023/a:1010933404324.

[61] Jerome H Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001, Available at: http://www.jstor.org/stable/2699986.

[62] Hongliang Ma, Jiangyuan Zeng, Xiang Zhang, Jian Peng, Xiaojun Li, Peng Fu, Michael H Cosh, Husi Letu, Shaohua Wang, Nengcheng Chen, and Jean-Pierre Wigneron, "Surface soil moisture from combined active and passive microwave observations: Integrating ASCAT and SMAP observations based on machine learning approaches," *Remote Sensing of Environment*, vol. 308, pp. 114197, 2024, https://doi.org/10.1016/j.rse.2024.114197.

[63] Scott M Lundberg and Su-In Lee, "A unified approach to interpreting model predictions," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017, pp. 4768–4777, Curran Associates Inc., https://doi.org/10.48550/arxiv.1705.07874.

[64] Scott M Lundberg, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee, "From local explanations to global understanding with explainable AI for trees," *Nature Machine Intelligence*, vol. 2, no. 1, pp. 56–67, 2020, https://doi.org/10.1038/s42256-019-0138-9.

[65] Hanna Meyer, Christoph Reudenbach, Stephan Wöllauer, and Thomas Nauss, "Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction," *Ecological Modelling*, vol. 411, pp. 108815, 2019, https://doi.org/10.1016/j.ecolmodel.2019.108815.

[66] Max Kuhn, "Building Predictive Models in R Using the caret Package," *Journal of Statistical Software*, vol. 28, no. 5, pp. 1 – 26, Nov. 2008, https://doi.org/10.18637/jss.v028.i05.

[67] Shengguo Gao, Zhongli Zhu, Haiteng Weng, and Jinshui Zhang, "Upscaling of sparse in situ soil moisture observations by integrating auxiliary information from remote sensing," *International Journal of Remote Sensing*, vol. 38, no. 17, pp. 4782–4803, 2017, https://doi.org/10.1080/01431161.2017.1320444.

[68] Stefan Siebert, Verena Henrich, Karen Frenken, and Jacob Burke, "Update of the digital global map of irrigation areas to version 5," 2013, https://doi.org/https://doi.org/10.13140/2.1.2660.6728.

[69] Alfred H Thiessen, "Precipitation averages for large areas," *Monthly Weather Review*, vol. 39, no. 7, pp. 1082–1089, 1911, https://doi.org/10.1175/1520-0493(1911)39⟨1082b:pafla⟩2.0.co;2.

[70] Patrick M Bartier and C.Peter Keller, "Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW)," *Computers & Geosciences*, vol. 22, no. 7, pp. 795–799, 1996, https://doi.org/10.1016/0098-3004(96)00021-0.

[71] R Webster and M A Oliver, *Geostatistics for Environmental Scientists*, Wiley, 2007, Available at: https://books.google.com.au/books?id=WBwSyvIvNY8C.

[72] R van der Velde, A Colliander, M Pezij, H J F Benninga, R Bindlish, S K Chan, T J Jackson, D M D Hendriks, D C M Augustijn, and Z Su, "Validation of SMAP L2 passive-only soil moisture products using upscaled in situ measurements collected in Twente, the Netherlands," *Hydrology and Earth System Sciences*, vol. 25, no. 1, pp. 473–495, 2021, https://doi.org/10.5194/hess-25-473-2021.

[73] C A Burton, L J Renzullo, S W Rifai, and A I J M Van Dijk, "Empirical upscaling of OzFlux eddy covariance for high-resolution monitoring of terrestrial carbon uptake in Australia," *Biogeosciences*, vol. 20, no. 19, pp. 4109–4134, 2023, https://doi.org/10.5194/bg-20-4109-2023.

[74] Yi Yu, Luigi J Renzullo, and Siyuan Tian, "Continental scale downscaling of AWRA-L analysed soil moisture using random forest regression," in *MODSIM2021, 24th International Congress on Modelling and Simulation*, Sydney, Australia, 2021, pp. 498–504, https://doi.org/10.36334/modsim.2021.j10.yu.

[75] Kotaro Bessho, Kenji Date, Masahiro Hayashi, Akio Ikeda, Takahito Imai, Hidekazu Inoue, Yukihiro Kumagai, Takuya Miyakawa, Hidehiko Murata, and Tomoo Ohno, "An introduction to Himawari-8/9 - Japan's new-generation geostationary meteorological satellites," *Journal of the Meteorological Society of Japan. Ser. II*, vol. 94, no. 2, pp. 151–183, 2016, https://doi.org/10.2151/jmsj.2016-009.

[76] Jun Yang, Zhiqing Zhang, Caiying Wei, Feng Lu, and Qiang Guo, "Introducing the new generation of Chinese geostationary weather satellites, Fengyun-4," *Bulletin of the American Meteorological Society*, vol. 98, no. 8, pp. 1637–1658, 2017, https://doi.org/10.1175/bams-d-16-0065.1.

[77] Tomoaki Miura, Shin Nagai, Mika Takeuchi, Kazuhito Ichii, and Hiroki Yoshioka, "Improved characterisation of vegetation and land surface seasonal dynamics in central Japan with Himawari-8 hypertemporal data," *Scientific Reports*, vol. 9, no. 1, pp. 15692, 2019, https://doi.org/10.1038/s41598-019-52076-x.

[78] Yi Yu, Luigi J Renzullo, Tim R McVicar, Thomas G Van Niel, Dejun Cai, Siyuan Tian, and Yichuan Ma, "Solar zenith angle-based calibration of Himawari-8 land surface temperature for correcting diurnal retrieval error characteristics," *Remote Sensing of Environment*, vol. 308, pp. 114176, 2024, https://doi.org/10.1016/j.rse.2024.114176.

[79] A M Khan, P C Stoy, J Joiner, D Baldocchi, J Verfaillie, M Chen, and J A Otkin, "The Diurnal Dynamics of Gross Primary Productivity Using Observations From the Advanced Baseline Imager on the Geostationary Operational Environmental Satellite-R Series at an Oak Savanna Ecosystem," *Journal of Geophysical Research: Biogeosciences*, vol. 127, no. 3, pp. e2021JG006701, 2022, https://doi.org/10.1029/2021jg006701.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3565818

Yi Yu (Member, IEEE) received a B.Mgmt. in Land Resources Management from Southwest University, China, in 2018, an M.Sc. in Environment (Advanced) from the Australian National University (ANU), Australia, in 2020, and a Ph.D. in Hydrology and Remote Sensing from ANU in 2024. His Ph.D. research investigated the role of remotely sensed land surface temperature in improving spatiotemporal monitoring capability of agricultural drought. He is currently a postdoctoral researcher at the University of Sydney. His research interests focus on developing hybrid approaches that integrate physical and data-driven methods to better understand land-atmosphere interactions across various spatiotemporal scales, particularly in the context of climate extremes such as drought. He received the Commonwealth Scientific and Industrial Research Organisation (CSIRO) Agriculture and Food Director's Award - Next Generation Science in 2023. He was among the Student Paper Competition Finalists at the 2024 IEEE International Geoscience and Remote Sensing Symposium (IGARSS).

Brendan P. Malone received his B.Sc. in Agriculture in 2008 with 1st Class Honours from the University of Sydney, Australia, and his Ph.D. in Digital Soil Mapping from the same University in 2012. His Ph.D. thesis introduced new methods for rescuing legacy soil data to create thematic, highly granular and comprehensive soil mapping products relevant across multiple spatial scales.

Luigi J. Renzullo received the B.Sc. degree in Mathematics and Computing, the PostGrad.Dip. in Applied Physics, and the Ph.D. degree in Remote Sensing Science from Curtin University of Technology, Perth, Australia, in 1996, 1997, and 2004, respectively. He has over 25 years' experience in satellite remote sensing with research spanning applications of visible, thermal and microwave sensing systems to environmental modelling problems. He was a senior research scientist with the Commonwealth Scientific and Industrial Research Organisation (CSIRO) for 15 years and with the Australian National University (ANU) for 5 years. Currently Dr Renzullo is a researcher at the Australian Bureau of Meteorology and Honorary Senior Lecturer at ANU. Dr Renzullo is a member of the Australian Mathematical Society and Australian Institute of Physics.

Chad A. Burton received a B.Sc. (Honours) in Geosciences from Monash University, and an M.Sc. in Environmental Change and Management from the University of Oxford. He is currently a Ph.D. candidate at the Australian National University's Fenner School of Environment and Society, where he is researching data-driven methods to enhance the understanding of Australia's terrestrial carbon cycle and ecosystem dynamics. Chad also serves as an Earth Observation Scientist for Geoscience Australia where he develops Earth Observation-based geospatial products for monitoring environmental change.

Siyuan Tian received the Master's degree in Engineering (Geomatics) from the University of Melbourne, Carlton, Australia, and the Ph.D. degree in Earth Science from the Australian National University, Canberra, Australia, in 2013 and 2019, respectively. Since her Ph.D., her research has been focused on using the Earth Observations together with land surface models to understand the interaction between the vegetation and hydrological cycle. She is currently a research scientist at the Bureau of Meteorology, Australia. She received the Modelling and Simulation Society of Australia and New Zealand Inc. (MSSANZ) Early Career Research Excellence Award in 2023.

Ross D. Searle received a B.Sc. in Agricultural Science from the University of Queensland (1990) majoring in Soil Science. He currently works for the Commonwealth Scientific and Industrial Research Organisation as a Senior Experimental Scientist. Ross Searle is a member of the Soil Science Australia.

Thomas F. A. Bishop received a B.Sc. in Agriculture (Honours) in 1997 and a Ph.D. in Precision Agriculture in 2002, both from the University of Sydney (USYD), Sydney, NSW, Australia. His main teaching is related to applied statistics, environmental science, and GIS. He is the Academic Director of the Sydney Informatics Hub which is a Core Research Facility at USYD. He is a Professor with the School of Life and Environmental Science at USYD. Prior to starting work at USYD in 2007, he held postdoctoral positions with the University of Florida, FL, USA, Rothamsted Research, Harpenden, U.K., and the University of New South Wales, Sydney, NSW, Australia. His research interests include modelling and predicting the variation of environmental properties in space and time with an emphasis on applying this to the domains of soil, agriculture and hydrology. Prof. Bishop is an Associate Editor for the European Journal of Soil Science and Soil Research and on the Editorial Board of Geoderma and Pedosphere.

Jeffrey P. Walker (Fellow, IEEE) received his B.E. (Civil) and B.Surveying degrees in 1995 with Hons 1 and University Medal from the University of Newcastle, Australia, and his Ph.D. in Water Resources Engineering from the same University in 1999. His Ph.D. thesis was among the early pioneering research on estimation of root-zone soil moisture from assimilation of remotely sensed surface soil moisture observations. He then joined NASA Goddard Space Flight Centre to implement his soil moisture work globally. In 2001 he moved to the Department of Civil and Environmental Engineering at the University of Melbourne as Lecturer, where he continued his soil moisture work, including development of the only Australian airborne capability for simulating new satellite missions for soil moisture. In 2010 he was appointed as Professor in the Department of Civil Engineering and Environmental Engineering at Monash University where he has continued this research. He is contributing to soil moisture satellite missions at NASA, ESA and JAXA, as a Science Team member for the Soil Moisture Active Passive (SMAP) mission and Cal/val Team member for the Soil Moisture and Ocean Salinity (SMOS) and Global Change Observation Mission – Water (GCOM-W) respectively. He is an IEEE Fellow and Australian Research Council Laureate Fellow.