Science of the Total Environment 776 (2021) 145924

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Estimating catchment scale soil moisture at a high spatial resolution: Integrating remote sensing and machine learning

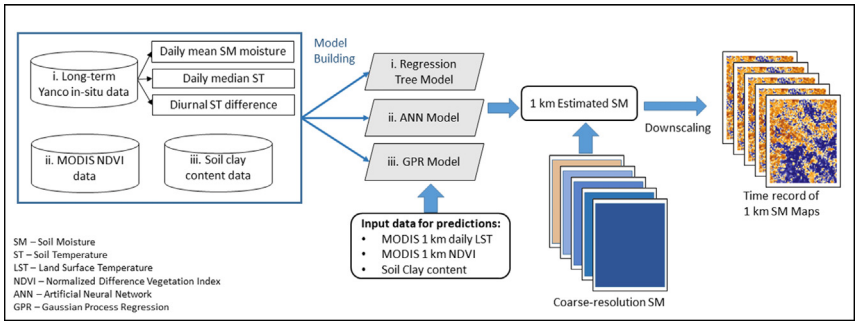
I.P. Senanayake [a,*], I.-Y. Yeo [a,*], J.P. Walker [b], G.R. Willgoose [a]

[a] School of Engineering, College of Engineering, Science and Environment, The University of Newcastle, Callaghan, NSW 2308, Australia
[b] Department of Civil Engineering, Monash University, Clayton, Victoria 3800, Australia

HIGHLIGHTS

- High spatial resolution soil moisture is important for a number of applications.
- A regression tree, an ANN and a GPR model were developed to downscale soil moisture.
- Downscaling models were developed based on the soil thermal inertia theory.
- Coarse spatial resolution soil moisture was downscaled to 1 km.
- The results, especially from regression tree and GPR models, are very encouraging.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 10 November 2020
Received in revised form 21 January 2021
Accepted 11 February 2021
Available online 18 February 2021

Editor: Manuel Esteban Lucas-Borja

Keywords:
Artificial neural network
Downscaling
Gaussian process regression
Regression tree model
Soil moisture

ABSTRACT

Soil moisture information is important for a wide range of applications including hydrologic modelling, climatic modelling and agriculture. L-band passive microwave satellite remote sensing is the most feasible option to estimate near-surface soil moisture (~0–5 cm soil depth) over large extents, but its coarse resolution (~10s of km) means that it is unable to capture the variability of soil moisture in detail. Therefore, different downscaling methods have been tested as a solution to meet the demand for high spatial resolution soil moisture. Downscaling algorithms based on the soil thermal inertia relationship between diurnal soil temperature difference (ΔT) and daily mean soil moisture content (μ_{SM}) have shown promising results over arid and semi-arid landscapes. However, the linearity of these algorithms is affected by factors such as vegetation, soil texture and meteorology in a complex manner. This study tested a (i) Regression Tree (RT), an Artificial Neural Network (ANN), and a Gaussian Process Regression (GPR) model based on the soil thermal inertia theory over a semi-arid agricultural landscape in Australia, given the ability of machine learning algorithms to capture complex, non-linear relationships between predictors and responses. Downscaled soil moisture from the RT, ANN and GPR models showed root mean square errors (RMSEs) of 0.03, 0.09 and 0.07 cm^3/cm^3 compared to airborne retrievals and unbiased RMSEs (ubRMSEs) of 0.07, 0.08 and 0.05 cm^3/cm^3 compared to in-situ observations, respectively. The study showed encouraging results to integrate machine learning techniques in estimating near-surface soil moisture at a high spatial resolution.

Abbreviations: AMSR-E, Advanced Microwave Scanning Radiometer on the Earth Observing System; ANN, artificial neural network; CABLE, Community Atmosphere Biosphere Land Exchange; CONUS, Continental United States; CSIRO, Commonwealth Scientific and Industrial Research Organisation; DisPATCh, DISaggregation based on Physical And Theoretical scale Change; EVI, enhanced vegetation index; GPR, Gaussian process regression; JULES, Joint UK Land Environment Simulator; LAI, leaf area index; LST, land surface temperature; MODIS, MODerate-resolution Imaging Spectroradiometer; MSMMN, Murrumbidgee Soil Moisture Monitoring Network; NDVI, normalized difference vegetation index; NLDAS, North American Land Data Assimilation System; NSSDC, National Soil Site Data Collection; NSVNIRD, National soil visible-near infrared database; PLMR, Polarimetric L-band Multi-beam Radiometer; RMSE, root mean square error; RT, regression tree; SMAP, Soil Moisture Active Passive; SMAPEx, Soil Moisture Active Passive Experiment; SMOS, Soil Moisture and Ocean Salinity; TI, thermal inertia; ubRMSE, unbiased RMSE.

* Corresponding authors.
E-mail addresses: Indishe.Senanayake@newcastle.edu.au (I.P. Senanayake), In-Young.Yeo@newcastle.edu.au (I.-Y. Yeo).

https://doi.org/10.1016/j.scitotenv.2021.145924
0048-9697/

1. Introduction

The demand for high spatial resolution soil moisture data is increasing rapidly for a broad range of applications at different scales including hydrology, climatology and agriculture (Engman, 1991; Kornelsen and Coulibaly, 2013; Schmugge, 1998). L-band passive microwave remote sensing has emerged as a feasible option to measue near-surface (~ 0–5 cm soil depth) soil moisture content across large extents with a good accuracy, but suffers from being low spatial resolution (Njoku and Entekhabi, 1996; Lakshmi, 2013; Mohanty et al., 2017; Schmugge, 1983). Downscaling satellite-based soil moisture retrievals is a viable option to address this problem (Lakshmi, 2013; Peng et al., 2017; Sabaghy et al., 2018). With the advancement in remote sensing of soil moisture and the launch of L-band soil moisture satellite missions such as Soil Moisture and Ocean Salinity (SMOS) (Kerr et al., 2010) and Soil Moisture Active Passive (SMAP) (Entekhabi et al., 2014), a number of downscaling models have been tested to improve the spatial resolution of the radiometric soil moisture retrievals (Sabaghy et al., 2020). Generally, the performance of these methods are site specific and not universally applicable, due to the spatial heterogeneity in climatic conditions and land surface properties such as topography, vegetation, and soil texture over different regions.

Due to the high spatial resolutions of optical/thermal datasets, the downscaling methods based on thermal/optical satellite datasets can deliver downscaled products at a high spatial resolution (Chauhan et al., 2003; Piles et al., 2011, 2014, 2016; Portal et al., 2018; Sánchez-Ruiz et al., 2014). Among them, methods based on the universal triangle between soil moisture, land surface temperature and vegetation index (Carlson et al., 1994) have been developed and tested by a number of researchers and exhibited good performance over arid and semi-arid landscapes (Peng et al., 2017). Merlin et al. (2012) developed the DISaggregation based on Physical And Theoretical scale Change (DisPATCh) model to downscale satellite soil moisture retrievals using 1 km spatial resolution MODerate-resolution Imaging Spectroradiometer (MODIS) products based on the 'universal triangle' concept. The DisPATCh model has shown an accuracy with the root mean square error (RMSE) temporally varying from 0.06 to 0.18 cm^3/cm^3 in Austral summer and winter, respectively, over the Murrumbidgee River catchment in Australia (Merlin et al., 2012; Sabaghy et al., 2018; Sabaghy et al., 2020).

Fang et al. (2013, 2018) and Fang and Lakshmi (2014) have related soil moisture to the diurnal soil temperature difference (ΔT) based on the changes in soil thermal inertia as a result of soil moisture content. They used soil temperature and soil moisture estimates from the North American Land Data Assimilation System (NLDAS) at 1/8° spatial resolution in those studies. They built regression algorithms between ΔT and diurnal temperature difference of soil (μ_{SM}) and modulated their model by using Normalized Difference Vegetation Index (NDVI) values. The downscaled soil moisture using this method showed RMSE varying from 0.02 to 0.06 cm^3/cm^3 in the Little Washita Watershed, Oklahoma, United States (Fang and Lakshmi, 2014). Senanayake et al., 2019a, 2021 developed a similar regression tree models based on soil thermal inertia relationship between ΔT and μ_{SM} over the Goulburn River catchment in the Upper Hunter Region of New South Wales, Australia, to downscale coarse resolution satellite soil moisture products. Here, they used point-scale in-situ observations from a soil moisture monitoring network and simulated soil moisture and soil temperature outputs from the Global Land Data Assimilation System (GLDAS) at 0.25° spatial resolution to downscale SMOS and SMAP satellite soil moisture products into 1 km resolution. These studies showed encouraging results with unbiased RMSE (ubRMSE) ranging from 0.05 to 0.10 cm^3/cm^3.

Until recent years, little progress has been made to develop downscaling models by employing machine learning models with multiple data sources. Machine learning algorithms are capable of recognizing complex and highly non-linear patterns between the input (predictors) and output (targets) variables from large datasets (Carbonell et al., 1983; El Bouchefry and de Souza, 2020; Jin et al., 2020; Shobha and Rangaswamy, 2018). However, the reliability of machine learning methods and choice of algorithms rely highly on the size and quality of the dataset (Du and Swamy, 2013; Okut, 2016; Rasmussen and Williams, 2006). There are a few studies where machine learning methods were used to estimate high spatial resolution soil moisture. Chai et al. (2011) employed the Broyden-Fletcher-Goldfarb-Shanno (BFGS) training algorithm based on the relationship between soil moisture and MODIS derived soil evaporative efficiency in the Goulburn River catchment, Australia on two days in November 2005 and achieved RMSE values varying from 0.018 to 0.035 cm^3/cm^3. Senanayake et al. (2019b) tested Levenberg-Marquardt algorithm to estimate soil moisture at a high spatial resolution over the same study area and achieved RMSEs varying from 0.058 to 0.088 cm^3/cm^3. Abbaszadeh et al. (2019) used the random forest model to downscale SMAP passive soil moisture products in the Continental United States (CONUS) area from April to December 2015 based on the properties of the top 5 cm soil profile. They observed ubRMSEs varying from 0.02 to 0.06 cm^3/cm^3 in the Little Washita watershed, and 0.02 to 0.07 cm^3/cm^3 against in-situ data at the Walnut Gulch Experimental Watershed, United States in 2015. Moreover, they showed that including parameters such as soil texture and topography can improve the accuracy of the downscaled soil moisture products. Alemohammad et al. (2018) used an artificial neural network (ANN) to build a relationship between coarse resolution soil moisture products and NDVI estimates and then, high resolution NDVI values were used to retrieve the soil moisture outputs at a high spatial resolution. Srivastava et al. (2013) employed three artificial intelligence techniques; (i) an ANN, (ii) a support vector machine (SVM), and (iii) a relevance vector machine (RVM) with a generalized linear model to enhance the spatial resolution of SMOS soil moisture products. Here, the ANN outperformed other two models and the downscaling method based on the season worked well compared to continuous time series. Im et al. (2016) have tested a random forest model, boosted regression trees, and a rule-based regression technique to downscale Advanced Microwave Scanning Radiometer on the Earth Observing System (AMSR-E) soil moisture products by using MODIS 1 km products including surface albedo, land surface temperature (LST), NDVI, enhanced vegetation index (EVI), leaf area index (LAI) and evapotranspiration. They achieved RMSEs of 0.049/0.057, 0.052/0.078, and 0.051/0.063 m^3/m^3, respectively, for those three algorithms over two regions in South Korea and Australia. Jin et al. (2020) tested a support vector area-to-area regression kriging (SVATARK) model by employing support vector regression and area-to-area kriging to downscale European Space Agency's (ESA) 25-km-resolution surface soil moisture product in the Naqu region on the Tibetan Plateau of China. The downscaled soil moisture of this study showed RMSEs ranging from 0.04 to 0.076 m^3/m^3 when compared against in-situ observations.

Although the downscaling models based on soil thermal inertia relationship between ΔT and μ_{SM} showed promising results over arid and semi-arid regions in estimating soil moisture at a high spatial resolution, the nature of this relationship is highly affected by multiple biogeophysical factors such as vegetation, soil texture and coldness/warmness of a day in a complex and non-linear manner (Lu et al., 2009; Senanayake et al., 2021; Van de Griend et al., 1985). Machine learning provides a robust approach to identify such complex, non-linear patterns compared to the traditional statistical methods (Lange and Sippel, 2020; Recknagel, 2001). In general, the assumptions on normally distributed, continuous data are not required in machine learning approaches, allowing the integration of data from multiple sources with unknown contributions. The study presented in this paper evaluates three downscaling models, (i) a regression tree, (ii) an ANN (with Bayesian regularization backpropagation algorithm) and (iii) a Gaussian Process Regression (GPR), based on the soil thermal inertia relationship between ΔT and μ_{SM} over the Yanco area of the Murrumbidgee River catchment, Australia. Here, Bayesian regularization backpropagation algorithm and GPR model were chosen by considering the size and noisiness of the datasets, whereas the regression tree model (Senanayake

et al., 2019a) was used as a baseline. The downscaled soil moisture products (1 km) were then compared against the airborne soil moisture retrievals from the Soil Moisture Active Passive Experiment-5 (SMAPEx-5) and in-situ soil moisture observations. The structure of this paper is as follows. The study location and its characteristics are explained in Section 2. Section 3 introduces datasets used in this study. Methodology used to build the three downscaling algorithms and validation scheme are described in Section 4 along with a brief review on the thermal inertia relationship between ΔT and μ_{SM} and associated physical processes affecting the soil moisture content. Section 5 provides a discussion on results, limitations and potential future work, followed by conclusion in Section 6.

2. Study area – Yanco area, Murrumbidgee River catchment

The study mainly focused on the Yanco region, a 60 × 60 km area located in the Murrumbidgee River catchment in Eastern Australia (Fig. 1). The Yanco region is located between latitudes 34.561° S and 35.170°0 S and longitudes 145.826°6 E and 146.439° E. In general, this is a flat land area with elevations ranging from 117 m to 150 m with negligible slope (Young et al., 2008; Ye et al., 2020). Soils in the Yanco area consist of mainly clay, red brown earth, transitional red brown earth, sand over clay, and deep sands. The average annual rainfall over the area is 418.5 mm, mostly falling during late autumn and winter (Smith et al., 2012; Yee et al., 2016). The mean daytime temperature ranges from 32.1° C in January (Austral summer) to 13.5° C in July (Austral winter). The Yanco area and its in-situ observations (see Section 3.1.) have been used in a number of soil moisture and remote sensing studies (e.g., Colliander et al., 2017; Merlin et al., 2008; Piles et al., 2011; Ye et al., 2019; Zhu et al., 2018a, 2018b).

3. Data

3.1. The Murrumbidgee soil moisture monitoring network

The Murrumbidgee Soil Moisture Monitoring Network (MSMMN) consists of thirty-eight monitoring stations installed over the Murrumbidgee River catchment (Fig. 1). MSMMN was established primarily to measure the soil moisture content at various soil depths across the root zone (0–90 cm). Soil temperature is also measured across the soil depths with precipitation and other auxiliary data. The eighteen monitoring stations established in 2001 are called the first generation sites. Eight of them were grouped with automatic weather stations of the Australia Bureau of Meteorology. The other ten stations were grouped into two 150 km^2 clusters, Kyeamba and Adelong Creek catchments (Smith et al., 2012). The second generation monitoring stations were established two years later also to measure the soil moisture across the root zone along with soil temperature at a single depth and precipitation. These monitoring stations were located in two focus areas; (i) extending the Kyeamba Creek catchment monitoring stations towards the confluence of the Murrumbidgee River, and (ii) in a grid over the Yanco area (Fig. 1), to assist with assessing remotely sensed soil moisture retrievals (Smith et al., 2012). These sites were upgraded in 2006 to measure the soil moisture of the top 5 cm and soil temperature at the top 2.5 cm soil profiles. The sites were instrumented with Campbell scientific water content reflectometers installed vertically at 0–30, 30–60, and 60–90 cm soil depths, surface soil moisture probes at 0–5 or 0–8 cm soil depths, and soil temperature probes at 2.5 or 4 cm and 15 cm soil depths (Smith et al., 2012). Additional surface soil moisture monitoring stations were set up in this area around 2010 (Panciera et al., 2014; Smith et al., 2012).

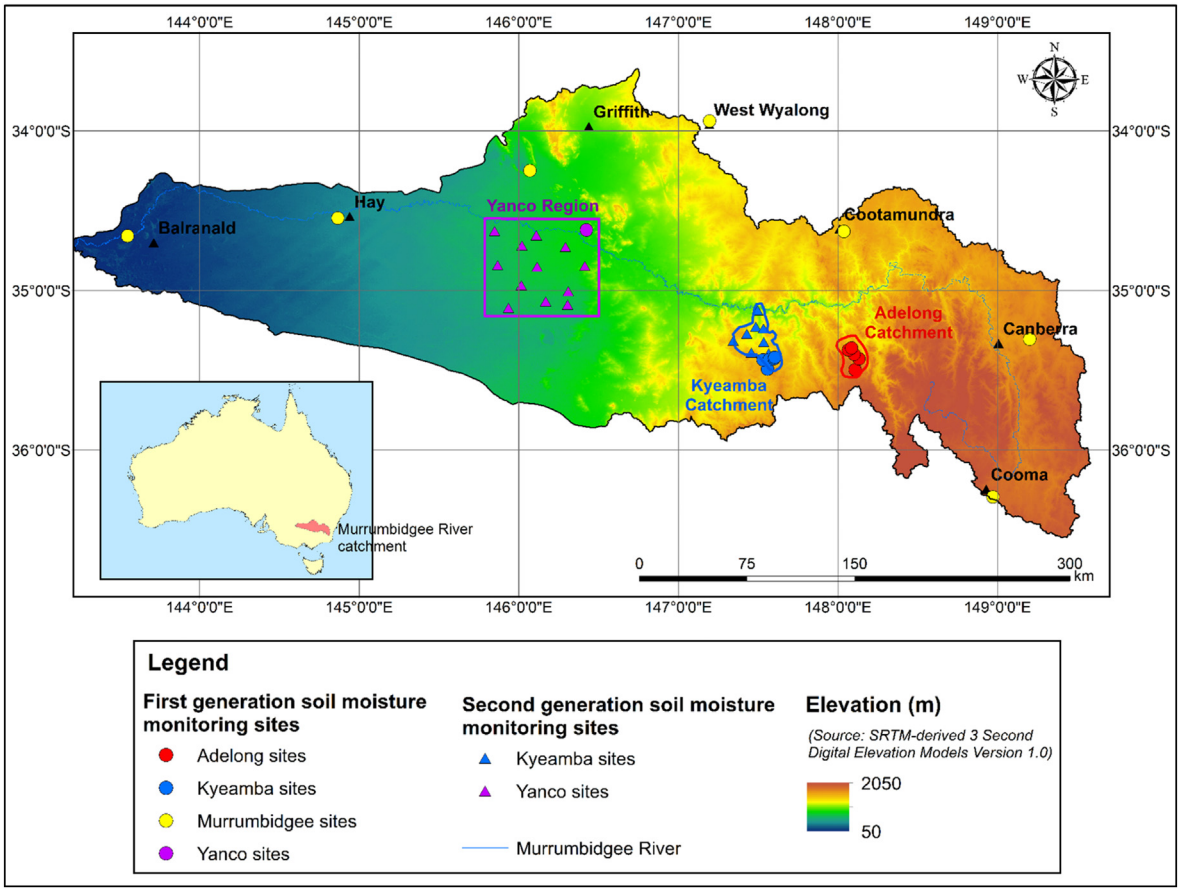


Fig. 1. Location and the digital elevation model (DEM) of the Murrumbidgee River Catchment, Australia and distribution of the Murrumbidgee soil moisture monitoring network (MSMMN) monitoring stations.

The MSNNN provides high quality, site-representative datasets to build downscaling algorithms and to validate the downscaled satellite soil moisture products since the monitoring stations were installed at the Catchment Average Soil Moisture Monitoring (CASMM) sites as described by Grayson and Western (1998). The CASMM sites were carefully located on the mid-slope areas in order to represent a wider extent around them (Grayson and Western, 1998). The MSMMN datasets have been used to validate both coarse resolution satellite soil moisture retrievals (Colliander et al., 2017; Smith et al., 2012), and downscaled soil moisture products (e.g., Malbéteau et al., 2016; Molero et al., 2016; Piles et al., 2012; Sabaghy et al., 2020; Yee et al., 2016) from multiple soil moisture satellite missions including SMAP, SMOS, and AMSR-E.

Near-surface in-situ soil moisture and soil temperature observations obtained from the Yanco monitoring stations (Fig. 2) were used in this work to build the downscaling models and to validate the downscaled soil moisture products. A summary of the MSMMN site characteristics is given in Table 1. Note that only the details of near-surface soil moisture measurements are given in the table, since it is the focus soil depth of this study (i.e., the approximate depth observed by the L-band satellite soil moisture retrievals). A detailed description on MSMMN is given in Smith et al. (2012) and Young et al. (2008). The MSMMN dataset can be accessed through the OzNet web site (www.oznet.org.au).

3.2. SMAPEx-5 airborne observations

The Soil Moisture Active Passive Experiments (SMAPEx) consist of a series airborne field campaigns carried out over the 60 km × 60 km Yanco study area located in the Murrumbidgee River catchment (www.

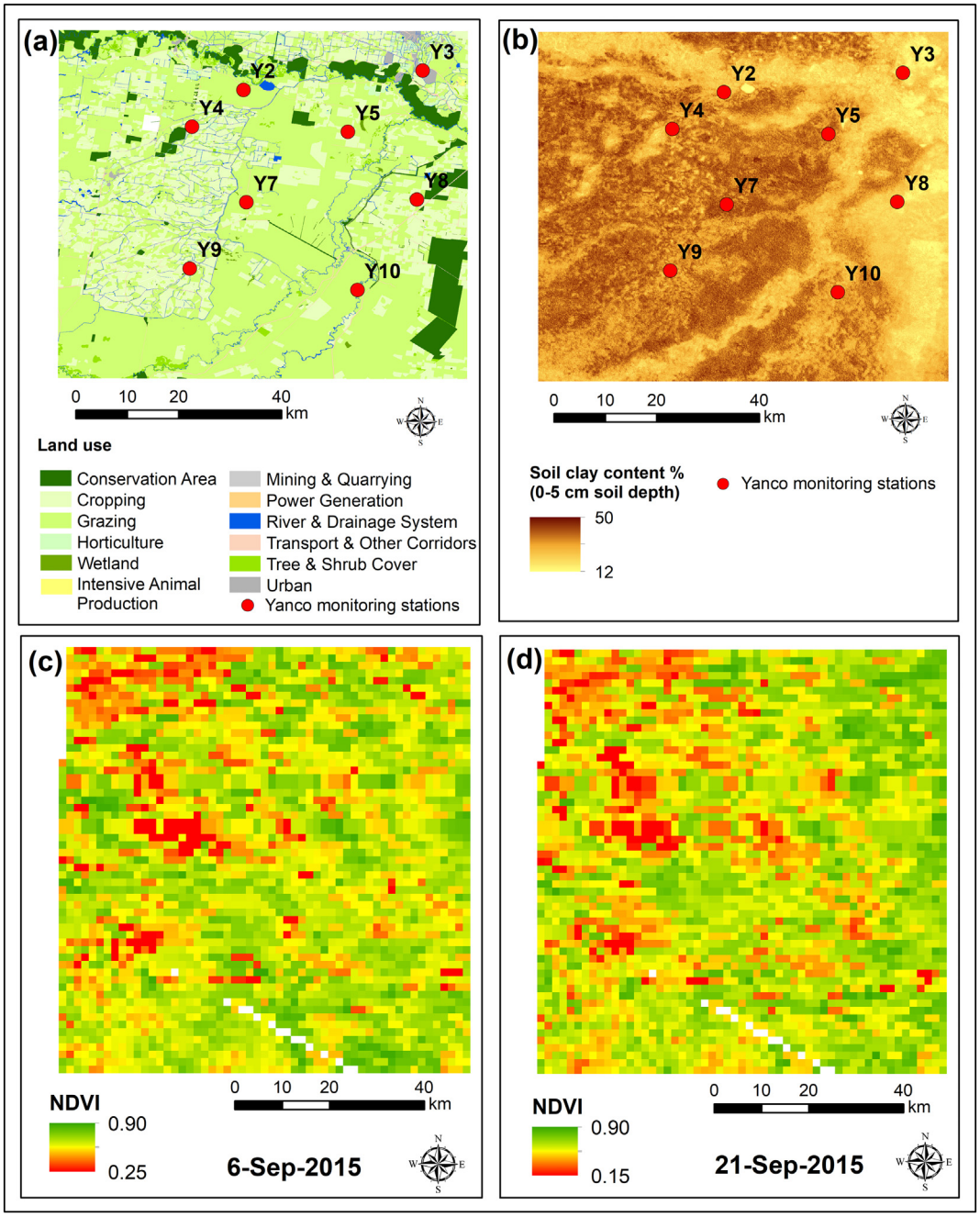


Fig. 2. (a) Land use/land cover over the Yanco study area (Source: The Department of Environment and Climate Change, NSW). (b) Soil clay content of the Yanco study area (Source: National Soil and Landscape Grid, Australia), with the locations of the soil moisture monitoring stations. (c) Normalized Difference Vegetation Index (NDVI) values over the Yanco study area on 6th and (d) 21st September 2015, as retrieved by MODIS 16-day NDVI composites (MYD13A2).

Table 1
Characteristics of the study site: soil moisture monitoring stations in Yanco area of the Murrumbidgee River catchment (Modified from: Young et al., 2008).

Site	Name	Latitude (WGS84[a])	Longitude (WGS84)	Near surface measurements			Data availability	Old site (O) or New site (N)	Soil clay content at top 5 cm (%) (SLGA[d])	Elevation (m AHD[e])	Annual precipitation (mm)	Annual PET[f] (mm)	Vegetation	Land use
				SM[b] depth (cm)	ST[c] depth (cm)	Data frequency (min)								
Y1	Uri Park	−34.62888	145.84895	2.5	5	20	2003–2018	N	26	120	416	1188	Improved pasture	Grazing
Y2	Banandra	−34.65478	146.11028	2.5	5	20	2003–2016	N	30	130	426	1202	Improved pasture	Grazing
Y3	Yanco Research Station	−34.621	146.424	7	4	30	2001–2011	O	25	144	442	1209	Grass	Grassland
Y4	Eulo	−34.71943	146.02003	2.5	5	20	2003–2018	N	33	130	417	1198	Improved pasture	Irrigated crop/ grazing
Y5	Dry Lake	−34.72835	146.29317	2.5	5	20	2003–2018	N	35	136	435	1210	Improved pasture	Grazing
Y6	S. Coleambally	−34.84262	145.86692	2.5	5	20	2003–2018	N	36	121	411	1200	Oats	Irrigated crop
Y7	Yamma Rd.	−34.85183	146.11530	2.5	5	20	2003–2018	N	42	128	424	1207	Improved pasture	Grazing
Y8	Wynella	−34.84697	146.41398	2.5	5	20	2003–2018	N	27	149	451	1216	Improved pasture	Grazing
Y9	Yammacoona	−34.96777	146.01632	2.5	5	20	2003–2018	N	33	122	416	1213	Oats	Illrrigated crop
Y10	Cheverelis	−35.00535	146.30988	2.5	5	20	2003–2018	N	34.65	119	437	1220	Improved pasture	Grazing
Y11	Bundure	−35.10975	145.93553	2.5	5	20	2003–2018	N	35.95	113	406	1215	Improved pasture	Grazing
Y12	Spring Bank	−35.06960	146.16893	2.5	5	20	2003–2018	N	34.56	120	428	1221	Oats/improved pasture	Crop/grazing
Y13	Widgiewa	−35.09025	146.30648	2.5	5	20	2003–2018	N	34.59	121	439	1222	Improved pasture	Grazing

[a] WGS: World Geodetic System 1984.
[b] SM: Soil Moisture.
[c] ST: Soil Temperature.
[d] SLGA: Soil and Landscape Grid of Australia.
[e] AHD: Australian Height Datum.
[f] PET: Potential Evaporation.

smapex.monash.edu.au) (Panciera et al., 2014; Ye et al., 2016; Ye et al., 2020). These experiments were conducted over the time frame of 2010–2015 in different seasons. SMAPEx-5 was conducted in September 2015, in parallel to the launch of SMAP satellite mission. The objectives of SMAPEx-5 included evaluation, validation and comparison of satellite soil moisture products, and development soil moisture retrieval algorithms (Ye et al., 2016). Accordingly, the SMAPEx-5 dataset has been used for a number of soil moisture studies (Sabaghy et al., 2020; Ye et al., 2016, 2019; Zhu et al., 2016, 2018a, 2018b). The same SMAP frequencies were used in the SMAPEx-5 airborne campaign, while the observations were timed approximately at the local SMAP overpass times. A Polarimetric L-band Multi-beam Radiometer (PLMR) providing dual polarized (vertical and horizontal) L-band (1.41 GHz) brightness temperature (T_b) measurements with six beams at 7, 21.5 and 38.5 degree incident angles at each side of the flight line was used in the SMAPEx-5 airborne campaign. These observations were taken at ~3000 m above ground level between 03:00 and 09:00 h, i.e., centred on SMAP descending overpass using eight replicated flights (Ye et al., 2016, 2019). Throughout the SMAPEx-5, an accuracy of over 1.4 K was observed at both vertical and horizontal polarizations (Ye et al., 2019). The SMAPEx-5 PLMR brightness temperatures showed a good correlation coefficient of 0.97 against the SMAP 36 km radiometer L1C data (Ye et al., 2019). This shows the ability of SMAPEx-5 soil moisture retrievals to provide good simulation for coarse spatial resolution satellite soil moisture products. The land use/land cover of the SMAPEx-5 footprint consisted of mainly cropping and grazing. Wheat was cultivated in cropping areas and dense grasses were observed in grazing areas. There were series of rainfall events at the beginning of SMAPEx-5 without any further rainfall throughout the campaign, providing ideal drying down conditions (Ye et al., 2019).

3.3. MODIS LST and NDVI products

MODIS/Aqua LST and Emissivity (LST/E) Daily L3 Global 1 km Grid V006 dataset at 1 km spatial resolution (MYD11A1) (Wan et al., 2015) was employed to extract daily daytime and night time LST values across the Yanco area in September 2015 (i.e., covering the period of SMAPEx-5). NDVI values from 2004 to 2017 over the Yanco area were acquired from the MODIS/Aqua Vegetation Indices 16-Day L3 Global 1 km SIN Grid V005 (MYD13A2) products at 1 km spatial resolution (Didan, 2015) for both model development and predictions.

3.4. Soil and landscape grid National Soil Attributes Maps-Australia

The clay content of the top 5 cm soil profile over the study area was obtained from the National Soil Attributes Maps of the Soil and Landscape Grid of Australia (Grundy et al., 2015). This dataset is available at the data access portal of the Commonwealth Scientific and Industrial Research Organisation (CSIRO) (https://data.csiro.au). This database was released for Australia in 2014 in connection to the GlobalSoilMap initiative. The dataset consists of quantitative soil properties for the entirety of Australia on a 90 m grid. It was developed by using the site data and spectroscopic observations in Australia. The site data used to build this soil database had been collected during 1931–2013 by the CSIRO National Soil Archive and National Soil Database (NatSoil) and state and territory government agencies as a part of National Soil Site Data Collection (NSSDC). The spectroscopic observations were collected with the National Soil Visible-Near Infrared Database (NSVNIRD) to determine the properties of soils, with the aid of soil sampling data collected during the National Geochemical Survey of Australia (Rossel et al., 2015). The clay content layers of the soil grid database were developed by using the data from 15,192 NSSDC and 1113 NSVNIRD sites over Australia.

4. Methodology

This section consists of two major parts: (i) developing the downscaling models and estimating soil moisture at a high spatial

resolution, and (ii) validating downscaled soil moisture products. A summary of the methodology is given in Fig. 3. In-situ soil moisture and temperature datasets were thoroughly examined using visual inspection and other statistical exploratory methods. Then, those datasets along with NDVI and soil clay content data were used to build three downscaling models, i.e., a regression tree model, Bayesian regularization algorithm based ANN model, and an exponential GPR model.

4.1. Developing the downscaling models based on the thermal inertia relationship

All three downscaling models developed in this study are based on the soil thermal inertia relationship of ΔT and μ_{SM}. This section presents the theoretical background of the soil thermal inertia theory. Thermal inertia is defined as the resistance of the temperature of an object to the fluctuations of its surrounding temperature (Sellers, 1965). Accordingly, the temperature of an object with a high thermal inertia varies more slowly than an object with a low thermal inertia. Therefore, the ΔT of soil shows an inverse relationship with thermal inertia (Engman, 1991), such that:

$$\Delta T = f\left(\frac{1}{TI}\right) \tag{1}$$

$$\Delta T = T_{PM} - T_{AM} \tag{2}$$

where TI is thermal inertia, ΔT is the diurnal temperature difference, T_{PM} is the afternoon and T_{AM} is the early morning soil temperatures. Thermal inertia can be expressed as (Wang et al., 2010):

$$TI = \sqrt{\rho c k} \tag{3}$$

where ρ is the bulk density (in kgm^{-3}), c is the specific heat capacity (in $Jkg^{-1}\,K^{-1}$) and k is the thermal conductivity (in $Wm^{-1}\,K^{-1}$) of the object. Since the specific heat capacity of water is significantly higher than of the dry soil, there is a clear contrast between the thermal inertia between wet and dry soils. Accordingly, wet soils exhibit a low ΔT compared to dry soils due to the presence of water (Verstraeten et al., 2006). This leads to an inverse relationship between the ΔT and μ_{SM}. This ΔT and μ_{SM} is affected by the factors such as vegetation and soil texture in a complex, non-linear manner (Lu et al., 2009; Senanayake et al., 2021; Van de Griend et al., 1985).

The ΔT - $\mu_{SM,}$ regression tree models (Fang et al., 2013; Fang and Lakshmi, 2014; Senanayake et al., 2019a) have a limited ability of dealing with these complex and non-linear relationships due to knowledge driven manual classification. This work tested the ability of machine learning algorithms to capture these complex relationships between ΔT, μ_{SM} and modulating factors to develop downscaling models to estimate soil moisture at a high spatial resolution. The ΔT and μ_{SM} values calculated from the Yanco in-situ datasets of the month of September from 2004 to 2017 excluding 2015 (in total, 1802 data records) were employed in training the models along with daily median temperature data (T_{med}),

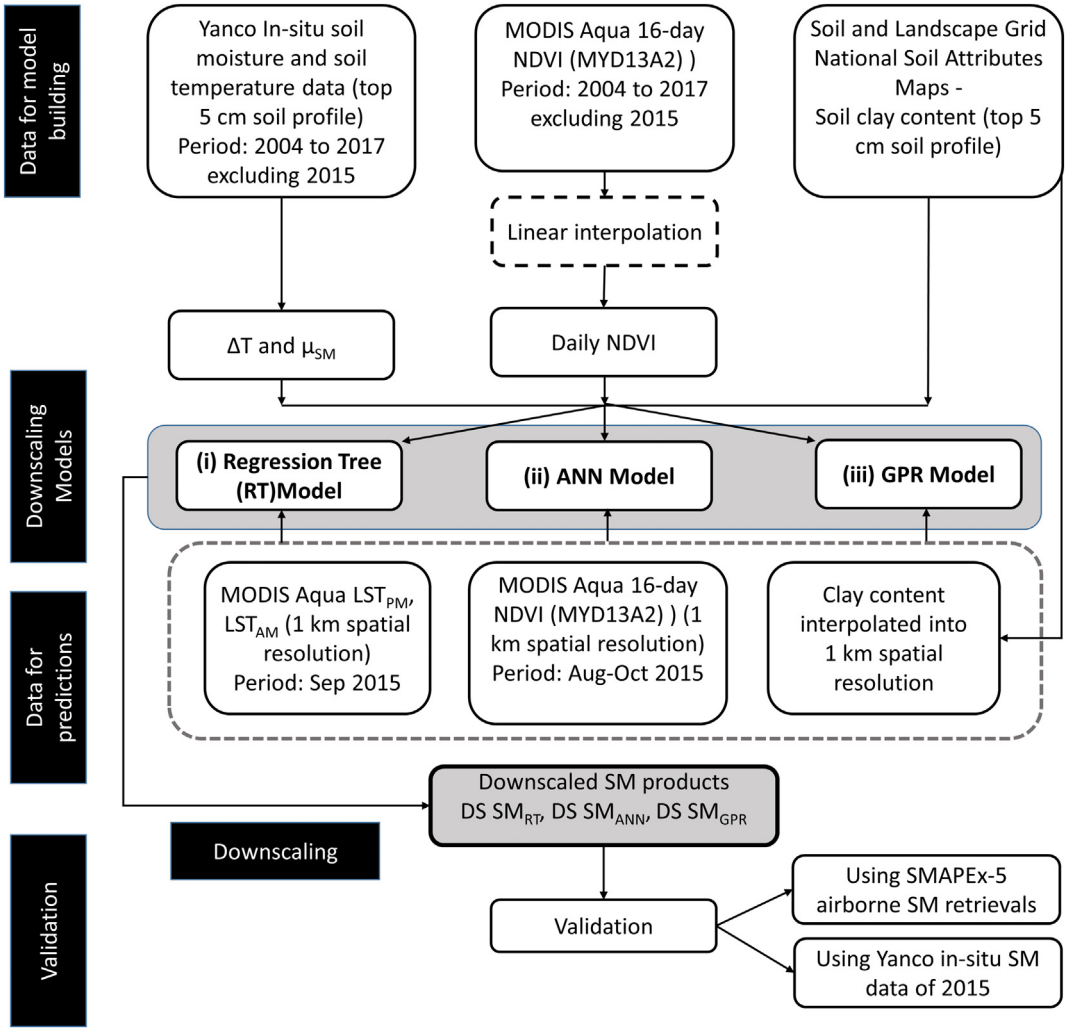


Fig. 3. Flow chart showing a summary of the methodology used in this study.

MODIS derived NDVI data, and soil clay content values obtained from the National Soil and Landscape Grid. Here, September was chosen due to the availability of SMAPEx-5 airborne observations (data from September 2015 was used for validation). Table 2 provides a short summary of the three models. NDVI and clay content values were extracted at point-scale for each Yanco monitoring station.

4.1.1. Method 1: regression tree model

The ΔT-μ_{SM} regression tree model was developed by employing Yanco in-situ observations. The ΔT - μ_{SM} relationship was modulated by vegetation density, soil clay content and daily median temperature. Vegetation density affects the evapotranspiration, and therefore regulates both surface temperature and soil moisture content (Jackson et al., 1982; Rodriguez-Iturbe et al., 1999; Soliman et al., 2013). De Alcântara Silva et al. (2016) classified NDVI values into three classes based on the vegetation density as, (i) NDVI<0.4 (areas covered by grass or no vegetation), (ii) 0.4 < NDVI<0.6 (areas with abundant and vigorous vegetation), and (iii) NDVI>0.6 (areas with dense and vigorous vegetation). Clay content affects the soil thermal inertia based on the impact of soil texture to the thermal conductivity (Wang et al., 2010). Including clay content to the downscaling model has a good potential of increasing its accuracy (Abbaszadeh et al., 2019). Bonan (2015) classifies soils with clay content>35% as heavy clays. This classification was used to set the break value for clay content in this model. T_{med} was used in this model considering the effect of general warmness or coolness of the day to the ΔT. Accordingly, the regression tree was classified into three classes based on the NDVI value (i.e., NDVI≤0.4, 0.4 < NDVI<0.6 and NDVI≥0.6), two each on the soil clay content (clay≤35% and clay>35%) and daily median soil temperature ($T_{med} \leq 15$ °C and $T_{med} > 15$ °C). The average T_{med} value of the dataset, 15 °C, was chosen as the break value for T_{med}. In summary, the regression tree was classified into 12 classes as shown in Fig. 4.

4.1.2. Method 2: artificial neural network (ANN)

ANNs are capable of adaptively learning complex functional forms and capturing highly non-linear relationships between the inputs (predictors) and targets (responses) (Haykin, 1994). An ANN is generated by interconnecting artificial neurons in the input, hidden and output layers. The input layer composed of independent variables, which are connected to the hidden layer. The hidden layer consists of activation functions. The activation functions calculate weights for the variables to determine the effects of predictors on the target variables. The prediction process is ended at the output layer, where the results are presented with an estimation error (Alaloul and Qureshi, 2020). The training and testing modes are the two phases of an ANN. During the training mode, the ANN is trained to recognize patterns in the given input datasets to achieve the desired output. The testing phase is where the pattern recognized in the training phase is employed to produce the associated output based on the inputs (Alaloul and Qureshi, 2020; Haykin, 1994).

For complex ANNs, the training might indicate good accuracy due to overfitting the data. Early stopping is a regularization method commonly utilized in ANNs to avoid overfitting when training a learner using an iterative method. Here, the original training dataset was split into two, as training set and validation set. In determining when overfitting had begun, the error on the validation set was used as a proxy for the generalization error, which indicates how accurately the algorithm can predict outputs to an unseen input (Prechelt, 1998). In this study, the Neural Net Fitting Toolbox in Matlab® R2017b was used to train and simulate the ANN using the Bayesian regularization backpropagation algorithm. Bayesian regularization backpropagation uses an objective function consisting residual sum of squares and the sum of squared weights in order to reduce estimation errors and to obtain a good generalized model (Burden and Winkler, 2008; MacKay, 1992). In backpropagation, when an input pattern is fed into the ANN during training phase, the ANN tries to learn and compares the predicted output values with the expected output values. The biases between the predicted and expected output values were then backpropagated through the network. Then, the weights in the hidden and output layer nodes were adjusted based on a gradient descent algorithm. This creates a network which maps the input and output values through the hidden neurons (Han et al., 2012). Since Bayesian Regularization does not need a validation dataset, it can use the entire dataset for training the model. This provides better generalization performance compared to the methods using early stopping as the regularization method for a small datasets (Du and Swamy, 2013; Okut, 2016).

The Bayesian regularization backpropagation algorithm was used in building the ANN model in this study considering the smaller size ($n = 1802$) of the dataset, and its automated regularization ability. ΔT, T_{med}, clay content and NDVI values were used as the inputs and μSM values as the targets when simulating the model. Trial and error method was used to find the best correlation by changing the number of hidden neurons using a neural network with two hidden layers between predictors and targets to map the non-linear relationships.

4.1.3. Model 3: exponential Gaussian process regression (GPR) model

GPR models are nonparametric kernel-based probabilistic models which are capable of recognizing complex relationships between input and output data (Rasmussen, 2004). Non-parametric models assume that the data distribution cannot be defined by a finite set of parameters. Instead, they define the data distribution by assuming an infinite dimensionless function. Therefore, at first, a prior probability distribution is defined over the functions in Gaussian Processes. This can be converted into a posterior over functions. Then, the covariance function assumes the points with similar predictor values (x_i) to have closer response values (y_i) (Rasmussen and Williams, 2006). Accordingly, it defines the covariance between the two latent variables (variables inferred from other observed variables), $f(x_i)$ and $f(x_j)$, where x_i and x_j are vectors of $d \times 1$ dimension. This defines how the response of a point, x_i, gets affected by the responses at other points, x_j, $i \neq j$, $i = 1, 2, 3\ldots, n$

In Gaussian processes, different kernel functions can be used to define the covariance function, $k(x_i, x_j)$. The covariance function can be given as $k(x_i, x_{j\,|\,\theta})$, since it can be parameterized in terms of the kernel parameters in vector θ. The kernel parameters are often based on the signal standard deviation, σ_f, and the characteristic length scale, σ_l. The σ_l briefly defines the required distance between input values x_i for the response values to become uncorrelated. Both σ_f and σ_l are greater than 0. This can be imposed by the unconstrained parametrization vector θ, so that, $\theta_1 = \log \sigma_l$, and $\theta_2 = \log \sigma_f$ (Rasmussen and Williams, 2006). The basic idea is that, if x_i and x_j are determined by the kernel to be similar, a similar output of the functions at those points can be expected (Murphy, 2012).

Table 2
Summary of the three downscaling models used in this study.

Model	Inputs (predictors)	Targets (responses)	Model details
Regression tree	ΔT, T_{med}, NDVI, Soil clay content	μ_{SM}	Two T_{med} classes: $T_{med} \leq 15$ °C and $T_{med} > 15$ °C Three NDVI classes: NDVI≤0.4, 0.4 < NDVI<0.6 and NDVI≥0.6 Two soil clay content classes: clay≤35% and clay>35%
ANN	ΔT, T_{med}, NDVI, Soil clay content	μ_{SM}	5 hidden neurons with 15% of random data records for testing Bayesian regularization backpropogation algorithm
GPR	ΔT, T_{med}, NDVI, Soil clay content	μ_{SM}	GPR with an exponential kernel

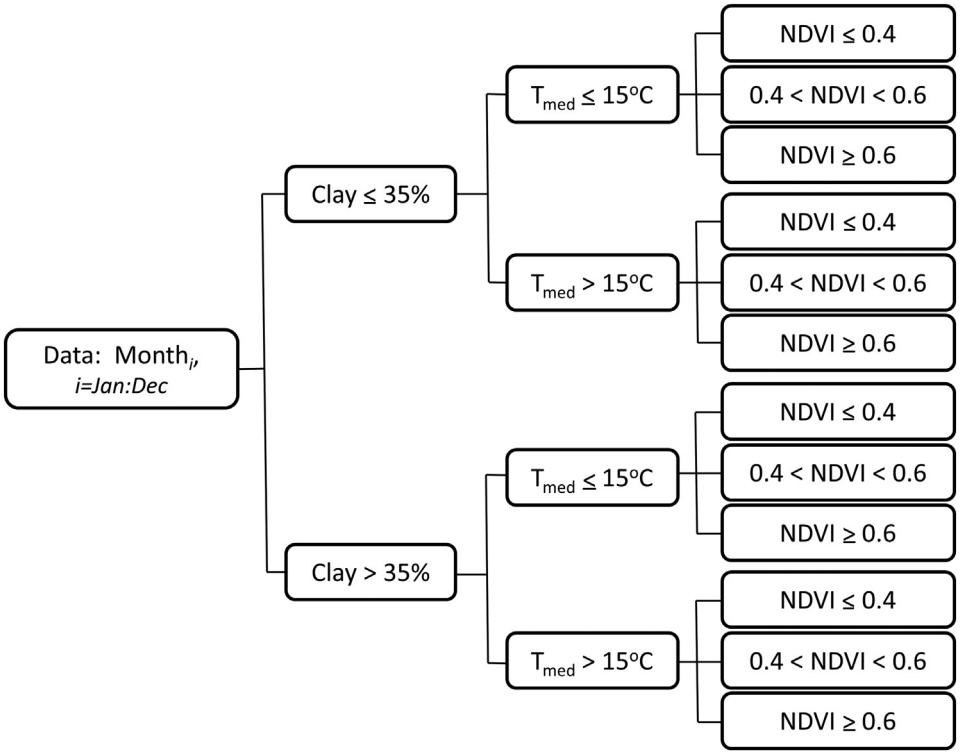


Fig. 4. The classification used in the regression tree model.

GPR models have been used to capture the relationships between complex datasets in various sectors (Hultquist et al., 2014; Mehdipour et al., 2014; Nguyen-Tuong and Peters, 2008; Taki et al., 2018; Wu et al., 2012), which can be useful to test the ability of GPR models in downscaling soil moisture. However, GPR models have not been used in previous studies to downscale coarse-spatial resolution soil moisture products. Another reason to choose GPR in this study is its good performance with small datasets (Murphy, 2012; Rasmussen and Williams, 2006; Wang and Hu, 2015; Zhang et al., 2018). An exponential kernel in Matlab R2017b Regression Learner App was chosen for building the downscaling model. The exponential kernel can be defined as (Rasmussen and Williams, 2006)

$$k(x_i, x_j|\theta) = \sigma_f^2 \ exp\left(-\frac{r}{\sigma_l}\right) \tag{4}$$

where r is the Euclidean distance between x_i and x_j. This can be defined as

$$r = \sqrt{(x_i - x_j)^T (x_i - x_j)} \tag{5}$$

The same datasets which were used in the previous models, i.e., ΔT, NDVI, T_{med} and soil clay content values as inputs and μ_{SM} values as targets, were employed here to develop the model.

4.2. Downscaling coarse resolution soil moisture product

The models were run over the Yanco area in September 2015. Here, the ΔT, NDVI, T_{med} and clay content values over the Yanco area were input to the models (Sections 4.1.1 to 4.1.3) at 1 km spatial resolution to estimate soil moisture at a spatial resolution of 1 km (SM_{est}). In downscaling, ΔT values were derived from MODIS Aqua LST values (spatial resolution of 1 km). MODIS Aqua overpass times over the Yanco area (01:30 h and 13:30 h) were assumed as reasonable approximators of the lowest and highest soil temperature values to show the diurnal soil temperature difference (Eq. (2)). It is assumed that any bias in LST data are nullified when calculating ΔT. Daily NDVI values were obtained by interpolating MODIS 16-day NDVI products temporally using bilinear interpolation. The mean of MODIS T_{AM} and T_{PM} values were considered as close approximates for the T_{med} values. Clay content obtained from the 90 m resolution National Soil and Landscape Grid (Grundy et al., 2015) was averaged over MODIS 1 km grid (i.e., 1 km) and used as the clay content inputs.

To replicate SMAP data for developing downscaling algorithms and validation is one of the main objective of SMAPEx airborne experiments. SMAPEx-5 retrievals provided very good approximates of SMAP data for developing and testing SMAP related algorithms (Ye et al., 2019). The SMAPEx-5 airborne soil moisture retrievals aggregated over the study area, was used as a proxy for a coarse-spatial resolution, L-band satellite soil moisture pixel in this study. This coarse-spatial resolution satellite soil moisture product was downscaled using the SM_{est} values obtained from the three models. The downscaled soil moisture ($SM_{ds,p}$) at a 1 km pixel p was calculated as:

$$SM_{ds,p} = SM_{est,p} + \left[SM_{SMAPEX,ag} - \frac{1}{n}\sum_{i=1}^{n} SM_{est,i}\right] \tag{6}$$

where $SM_{est,p}$ and $SM_{est,i}$ are soil moisture values (1 km) estimated by the models at pixels p and i (i = 1:n), $SM_{SMAPEX,ag}$ is the coarse spatial resolution satellite soil moisture value simulated by aggregating the SMAPEX-5 soil moisture over the study area, and n is the number of 1 km pixels inside the coarse resolution pixel (Fang et al., 2013). Here, the difference of values between the coarse resolution soil moisture product and estimated soil moisture averaged over the coarse resolution footprint was used for correcting the bias in SM_{est} with respect to the satellite soil moisture products. In other words, the coarse spatial resolution satellite soil moisture product was used to capture the overall temporal pattern of soil moisture, while the downscaling model was used to delineate sub-pixel spatial patterns within the coarse resolution satellite footprint. SM_{est} derived from each of the three models were used separately to downscale the simulated coarse resolution soil moisture

product (Eq. 6) over the study area for the SMAPEx-5 airborne campaign dates (i.e., 8th, 10th, 13th, 16th, 18th, 21st, 23rd and 26th Sep 2015).

4.3. Validation of the downscaled soil moisture products

The downscaled soil moisture products were validated by using two methods. First, they were compared against the SMAPEx-5 airborne soil moisture retrievals at 1 km spatial resolution. This allowed one to one comparison between the downscaled soil moisture against airborne soil moisture retrievals on the SMAPEx-5 days. Therefore, RMSE was used for the comparison between the downscaled soil moisture products and SMAPEx-5 airborne retrievals (Colliander et al., 2018; Entekhabi et al., 2010). Second, the downscaled soil moisture products were compared against the Yanco in-situ soil moisture observations. Seven soil moisture monitoring stations located over the study area (Y2, Y4, Y5, Y7, Y8, Y9, Y10; see Fig. 2) were used in this comparison based on the data

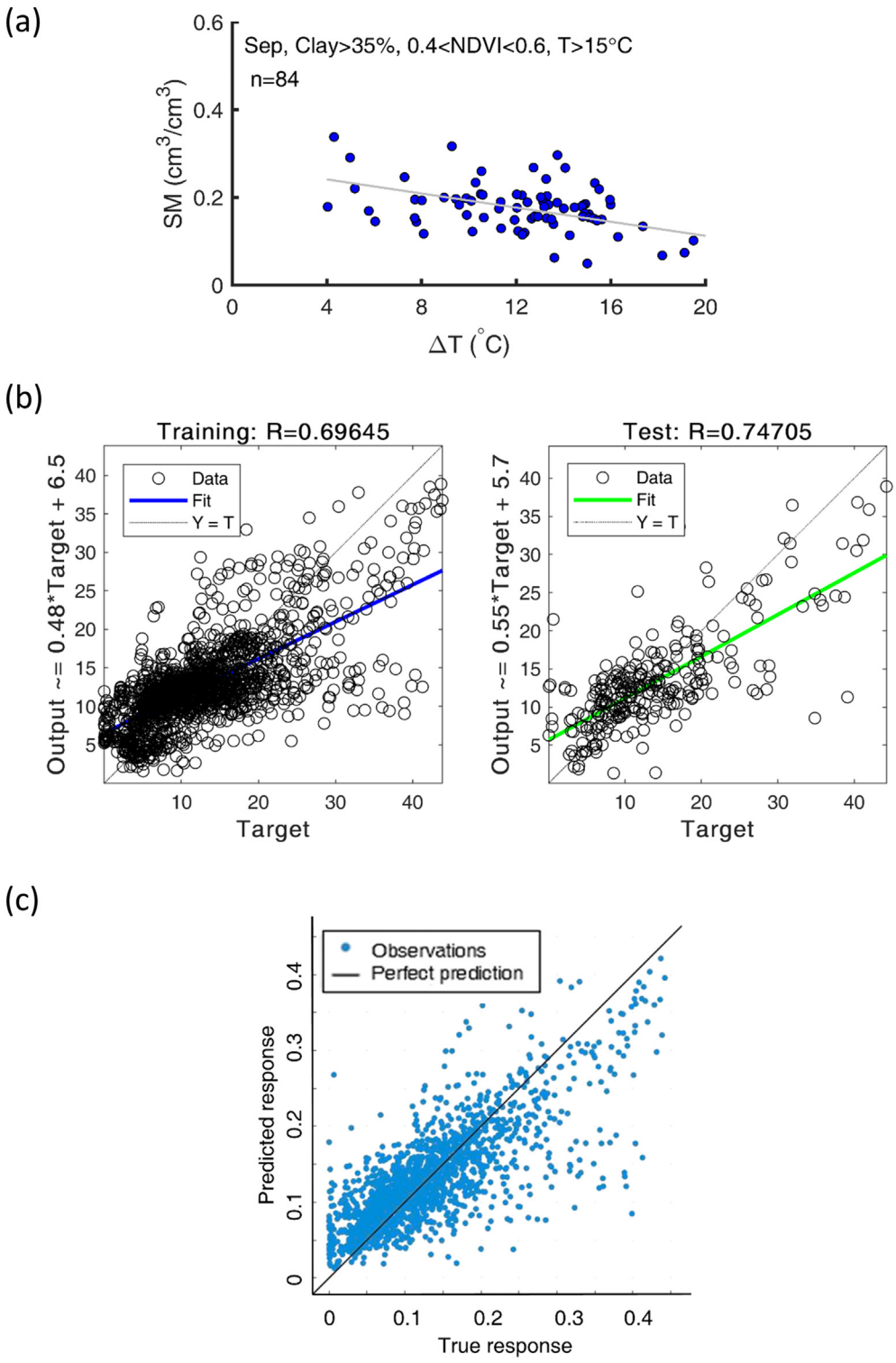


Fig. 5. (a) Regression between ΔT and μ_{SM} for the category Clay>35%, 0.4 < NDVI<0.6, T_{med} > 15 °C. (b) Results from the model validation of the ANN. (c) The plot showing soil moisture predicted by the GPR model against the actual in-situ soil moisture observations during model development.

availability during the SMAPEx-5 period. Note that for each of these seven instances, there was only one station per 1 km pixel of the downscaled soil moisture. As previously explained, it was assumed that these sites can provide good soil moisture approximations over the 1 km pixels based on their positioning at CASMM sites (Smith et al., 2012). Yee et al. (2016) have also identified that these sites can provide areal average soil moisture measurements to calibrate and validate satellite soil moisture retrievals and hydrological models. Here, a relative metric, ubRMSE, was used to statistically evaluate the agreement between the downscaled and in-situ soil moisture by taking the scale mismatch between point scale in-situ observations and 1 km downscaled products into account (Colliander et al., 2018; Entekhabi et al., 2010).

5. Results and discussion

5.1. Evaluating the variability of the Yanco in-situ soil moisture dataset

The monthly variability of soil moisture over the period of 2004 to 2017 (from 2002 to 2012 for Y3) is given in Appendix 1. Appendix 2 shows the mean near-surface soil moisture content at Yanco stations. It is clear that all the Yanco stations, generally, show dry conditions with a mean soil moisture value below 0.2 cm^3/cm^3 (Appendix 2). The stations have demonstrated slightly wet conditions in the Austral winter (i.e., June to August) compared to the extremely low soil moisture content in the Austral summer (i.e., December to February) with the catchment average soil moisture contents of 0.18 to 0.09 cm^3/cm^3 in Austral winter and summer, respectively (see Appendix 1). Note that, the data from Y3 was not used in building the downscaling models, due to the mismatch in measured depths for soil moisture and temperature compared to other sites (see Table 1 for further details).

5.2. Results from model development and validation

The regression plotted between ΔT and μ_{SM} for the category with clay content>35%, 0.4 < NDVI<0.6 and T_{med} > 15 °C is shown in Fig. 5 a as an example for the regressions of the first method. Results from the validation of the ANN model is shown in Fig. 5 b. The model showed a good correlation for both training and testing data for 5 hidden neurons as shown in the figure with 15% of random data records used for testing. Fig. 5 c shows the results of the GPR model validation. The model showed an RMSE of 0.052 cm^3/cm^3 and a coefficient of determination (R^2) of 0.6 (Fig. 5) between the predicted response from the model and observed soil moisture data.

5.3. Comparison of downscaled soil moisture products with SMAPEx-5 airborne soil moisture retrievals

Fig. 6 shows the comparison between (a) SMAPEx-5 retrievals and the downscaled soil moisture products from (b) regression tree, (c) ANN and (d) GPR models on six SMAPEx-5 airborne campaign dates (i.e., 13th, 16th, 18th, 21st, 23rd and 26th September 2015). Data from 8th and 10th September 2015 were excluded from the figure due to the data gaps caused by the cloud contamination and any negative soil moisture estimates were filtered. In general, the downscaled soil moisture from all three models were able to capture the NW-SE oriented dry patch across the study area which can be seen in the SMAPEx-5 retrievals as a NW-SE oriented stripe (Fig. 6). The dry patch on the N-E corner of the study area can also be seen in both the SMAPEx-5 and the downscaled soil moisture maps. As per visual inspection, downscaled products from the regression tree model showed the spatial patterns closest to the SMAPEx-5 soil moisture maps among the three model outputs. Some wet pixels can be observed

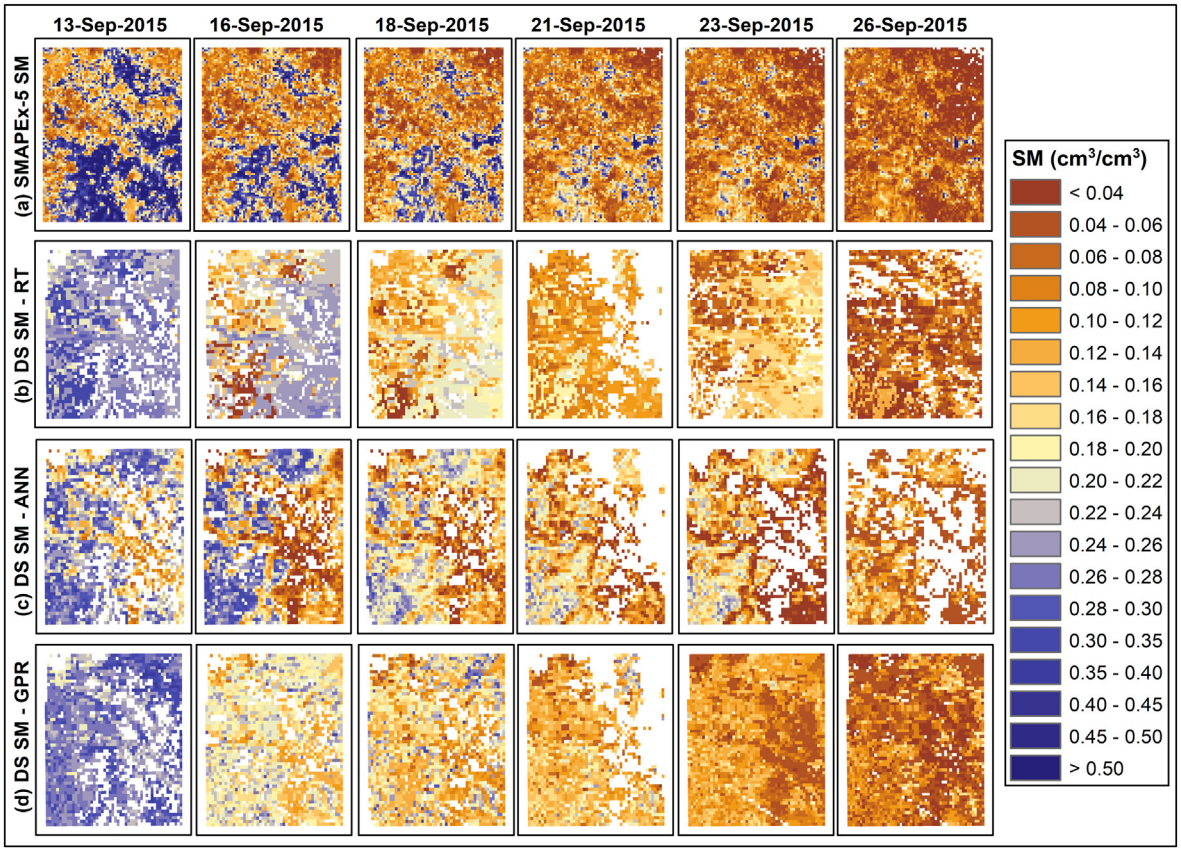


Fig. 6. Comparison between (a) SMAPEx-5 soil moisture observations and downscaled soil moisture from (b) the regression tree model, (c) artificial neural network (ANN) and, (d) Gaussian Process Regression (GPR) model.

over the south-eastern region of the Yanco area at the beginning of SMAPEx-5 campaign in SMAPEx-5 airborne retrievals and downscaled soil moisture from the regression tree model. Downscaled soil moisture from ANN and GPR models showed slightly drier patches over this region.

Fig. 7 shows the variability of the SMAPEx-5 soil moisture retrievals and downscaled soil moisture from the regression tree model, ANN, and the GPR model in boxplots. Note that the cloud affected pixels in the downscaled soil moisture products and their respective SMAPEx-5 pixels were omitted from the boxplots and the error calculation. The boxplots between SMAPEx-5 and the downscaled soil moisture product from the regression tree model (Fig. 7a and b) showed a good agreement in terms of both mean and variability. It is evident that the drying out condition from 8th September to 26th September 2015 over the Yanco region has been successfully captured by all three downscaled products (Fig. 7). Both the standing water during the beginning of SMAPEx-5 and the low mean soil moisture value over the area at the latter half of the SMAPEx-5 experiment (~0.2 cm^3/cm^3) might have affected the accuracy of the comparison. The low mean soil moisture content makes it difficult to interpret spatial patterns within a small value range.

The RMSE values between the SMAPEx-5 retrievals and downscaled soil moisture products for each SMAPEx-5 day are shown in Table 3. Downscaled soil moisture from the regression tree model, ANN and the GPR model show average RMSE values of 0.03, 0.09 and 0.07 cm^3/cm^3, respectively, with respect to the SMAPEx-5 retrievals.

5.4. Comparison of downscaled data with in-situ soil moisture observations

Fig. 8 shows the comparison between in-situ observations and downscaled soil moisture products generated by the three models during the SMAPEx-5 days at four Yanco stations, Y2, Y4, Y7 and Y8. A good temporal agreement can be seen between downscaled products from all three models against the in-situ observations. The drying trend after the precipitation event at the beginning of SMAPEx-5 has been captured successfully by all the downscaled products.

The comparison between in-situ soil moisture observations and the downscaled products from the three models (i.e., regression tree, ANN and GPR) showed ubRMSEs of 0.07, 0.08 and 0.05 cm^3/cm^3, respectively (Fig. 9).

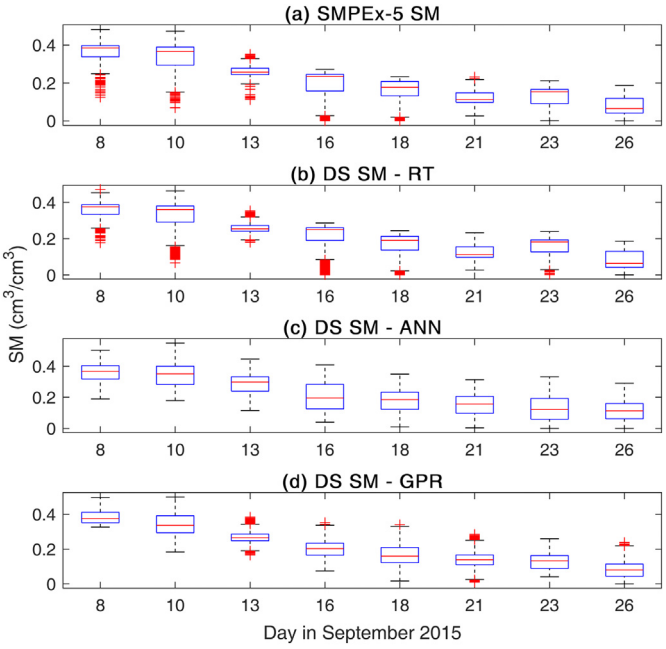


Fig. 7. Daily spatial soil moisture variability across the study area captured by (a) SMAPEx-5 airborne retrievals, b) downscaled soil moisture products from regression tree model, (c) artificial neural network (ANN) model, and (d) Gaussian Process Regression (GPR) model.

Table 3
Comparison between the SMAPEx-5 soil moisture and the downscaled soil moisture from the Regression Tree (RT) model, Artificial Neural Network (ANN) and the Gaussian Process Regression (GPR) models.

	RMSE (cm^3/cm^3)							
	Sep8	Sep 10	Sep 13	Sep 16	Sep 18	Sep 21	Sep 23	Sep 26
RT	0.04	0.04	0.02	0.04	0.03	0.03	0.04	0.02
ANN	0.10	0.12	0.06	0.14	0.10	0.07	0.10	0.07
GPR	0.07	0.09	0.05	0.08	0.07	0.06	0.06	0.06

5.5. Discussion

The downscaled soil moisture products from all three models showed a good agreement with the Yanco in-situ observations and SMAPEx-5 airborne soil moisture retrievals. Soil moisture estimated by the regression tree and the GPR model have showed better results compared to the downscaled soil moisture from the ANN model. Overfitting can be a possible reason for the errors in the downscaled soil moisture estimated with the ANN model (Okut, 2016).

There can be a few sources of errors affecting the comparison between downscaled soil moisture and the SMAPEx-5 soil moisture retrievals, such as the effects from standing water and agricultural activities (Ye et al., 2020). Here, it is noteworthy to mention that a large fraction of the study area is covered by agricultural lands with frequent irrigation applications. Nearly one third of the Yanco region belongs to the Coleambally Irrigation Area. This agricultural area consists of over 500 farms with a dense irrigation channel network (Panciera et al., 2014). Fig. 10 shows the distribution of irrigated agriculture over the Yanco area (NSW DPIE, 2017). Though this may not be the exact conditions specific to 2015, the map shows a general picture of the irrigation lands in the area. The amount of water release in this area varies with the cropping cycles and affect soil moisture measurements consequently (Bretreger et al., 2020; Panciera et al., 2014). Generally, the in-situ monitoring stations are established on locations which are not subjected to flooding or irrigation. Therefore, the

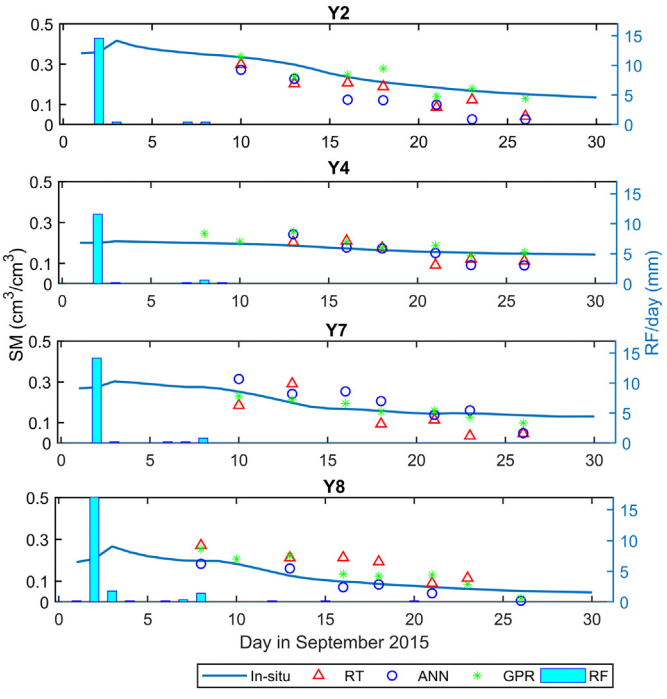


Fig. 8. Comparison between in-situ soil moisture data and downscaled soil moisture products from the regression tree (RT), artificial neural network (ANN) and exponential Gaussian process regression (GPR) models at four Yanco monitoring stations after rainfall (RF) in September 2015.

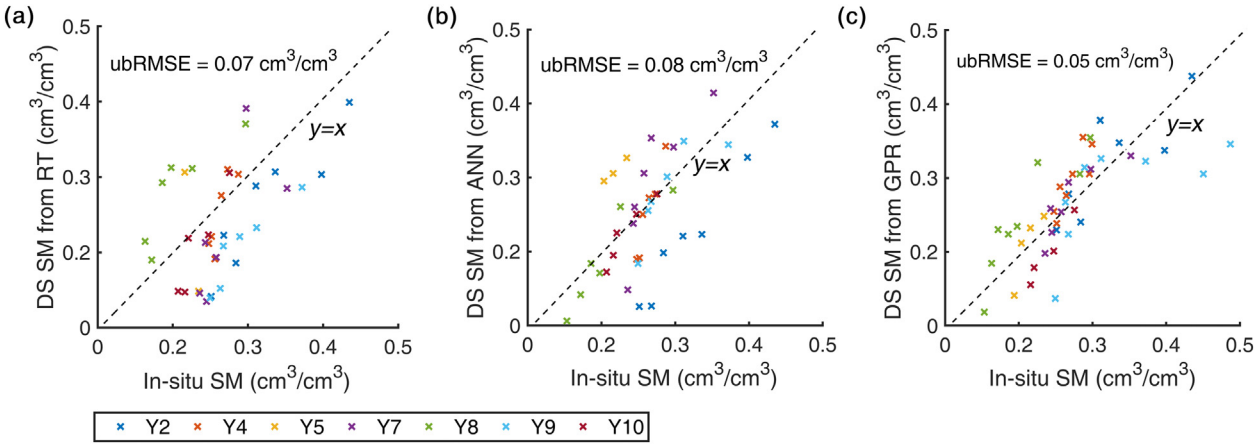


Fig. 9. Comparison between the Yanco in-situ observations against downscaled soil moisture products from (a) the regression tree model, (b) artificial neural network, and (c) Gaussian process regression model on SMAPEx-5 dates in September 2015.

point-scale in-situ observations used for model development were unable to capture all different water management cases and land use practices affecting soil moisture variability over the study extent. This causes an attenuation of model efficiency over some of the land use/land cover classes which are not represented in the model due to lack of data.

On the other hand, the standing water from the precipitation events at the beginning of SMAPEx-5, agricultural activities and a temporary pond located in the south-eastern part of the area can induce uncertainties to the L-band soil moisture observations due to the effect of water fraction on the T_b measurements (Ye et al.,

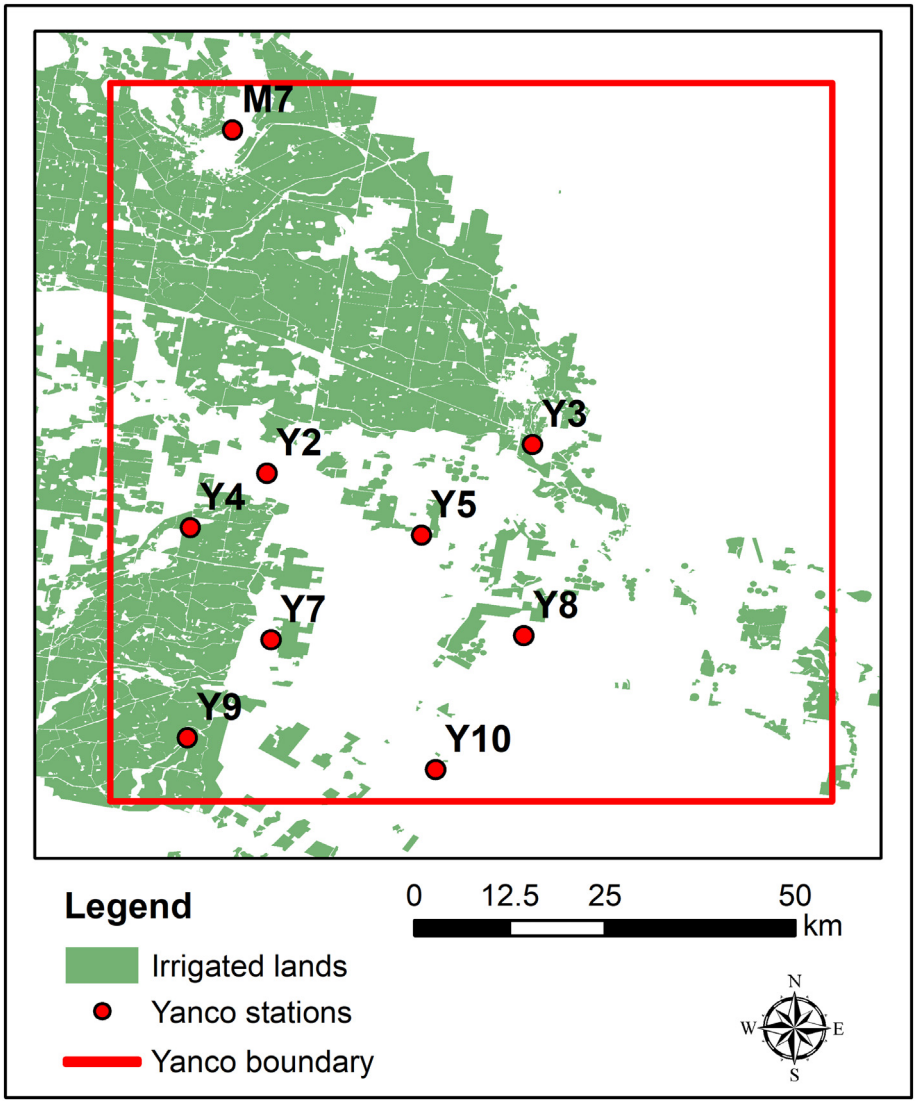


Fig. 10. Distribution of irrigated agriculture over the Yanco area.

2015; Ye et al., 2020). Ye et al. (2020) explained that some of the wheat paddocks in the north-western part of the Yanco area were subjected to flood irrigation at the end of SMAPEx-5 and this water fraction has caused overestimation in the airborne soil moisture retrievals. This has caused a slight increase in relative soil moisture content over some of the irrigated lands (Fig. 10) at the latter part of SMAPEx-5 (Fig. 6). Furthermore, the highly variable vegetation water content driven by different vegetation types and growth rates (Ye et al., 2020) was another factor which had affected the T_b measurements.

Downscaling algorithms built with the in-situ data collected over one area might not be applicable over a different spatial domain due to the disparities in the biogeophysical properties. Therefore, it would be worthwhile to test these machine learning algorithms using simulated land surface model outputs despite their different spatial resolutions. Since such datasets are available globally or regionally, they can be adopted over catchments where in-situ datasets are unavailable or combined with in-situ data. Senanayake et al. (2021) showed that the downscaling algorithms built by using both a point scale in-situ data and gridded land surface model (LSM, at ~25 km resolution) time series outputs can provide robust methods, despite their significant differences in spatial scales. The Global Land Data Assimilation System (GLDAS) (Rodell et al., 2004), the Joint UK Land Environment Simulator (JULES) (Best et al., 2011), Community Atmosphere Biosphere Land Exchange (CABLE) (Kowalczyk et al., 2006), Australian Water Resource Assessment-Landscape (AWRA-L) (Frost et al., 2015) and Australian Water Availability Project-Waterdyn (AWAP-Waterdyn) (Raupach et al., 2009) can be given as example for potential LSM outputs for such an approach. Introducing factors such as topography, surface albedo and wind speed to the model also has a good potential in improving its accuracy.

The effect from the cloud cover to the thermal/optical data is a major problem encountered in thermal/optical data based downscaling methods. Employing thermal data from geostationary satellites such as the Multi-Functional Transport Satellite (MTSAT)-1R (Himawari-6) can be a potential solution for this problem (Oyoshi et al., 2014; Yamamoto and Ishikawa, 2018). With their high temporal frequency geostationary satellites have a potential of providing cloud free approximates of T_{AM} and T_{PM} values. However, this cannot completely ensure cloud free imagery over each and every day. Methods of filling spatial and temporal data gaps such as merging MODIS Aqua and Terra LSTs (Crosson et al., 2012) can also be employed to address the effect of cloud contamination of thermal data.

Since a long-term time record of coarse resolution satellite soil moisture retrievals from SMOS and SMAP satellites are available from 2009, the models evaluated here can be used to develop a long-term time record of high spatial resolution (1 km) near-surface soil moisture for various applications. Here, it is noteworthy to mention about the ongoing work on developing a consistent multi-satellite soil moisture product with a 1-day temporal resolution by combining SMOS and SMAP soil moisture retrievals (Bindlish et al., 2017). This dataset has a good potential of being used effectively to produce a historic time-record of high spatial resolution soil moisture at a daily time-scale. Such a historical time record of daily high spatial resolution soil moisture can play a vital role in regional scale hydrologic, climatic and agricultural applications over arid and semi-arid landscapes.

6. Conclusion

This study compared three models (i.e., a regression tree, an ANN with Bayesian regularization algorithm, and an exponential GPR) based on the soil thermal inertia relationship between ΔT and μ_{SM} to estimate soil moisture at a high spatial resolution by downscaling the coarse-spatial resolution soil moisture product. The ΔT and μ_{SM} values extracted from the MSMNN dataset were used to build the models along with NDVI, soil clay content and T_{med} data. Machine learning algorithms were employed in this study by considering the complex, non-linear relationships between these factors and the ability of machine learning algorithms to capture such complex, non-linear relationships. The models were applied over the Yanco region of the Murrumbidgee River catchment to estimate soil moisture at 1 km spatial resolution. The downscaled soil moisture products from the three models were then compared against the SMAPEx-5 airborne soil moisture retrievals and Yanco in-situ soil moisture observations. Downscaled soil moisture from the regression tree model and the GPR showed better results compared to the ANN. The models provide encouraging results and insights for developing a long-term time record of high spatial resolution soil moisture over semi-arid agricultural landscapes by downscaling coarse spatial resolution satellite soil moisture products.

Nomenclature

ΔT	Diurnal temperature difference of soil (°C)
θ	Unconstrained parametrization vector
μ_{SM}	Daily mean soil moisture content (cm^3/cm^3)
ρ	Bulk density (kgm^{-3})
σ	Standard deviation
c	Specific heat capacity ($Jkg^{-1}\ K^{-1}$)
$f(x_i), f(x_j)$	Latent variables where x_i and x_j are vectors of $d \times 1$ dimension
k	Thermal conductivity ($Wm^{-1}\ K^{-1}$)
MYD11A1	MODIS/Aqua LST and Emissivity (LST/E) Daily L3 Global 1 km Grid V006 dataset
MYD13A2	MODIS/Aqua Vegetation Indices 16-Day L3 Global 1 km SIN Grid V005 dataset
n	Size of the dataset or number of pixels
r	Euclidean distance between two points
SM_{est}	Estimated soil moisture (cm^3/cm^3)
$SM_{est,p}$, $SM_{est,i}$	Estimated soil moisture at pixels p and i (cm^3/cm^3)
$SM_{SMAPEX,ag}$	Coarse spatial resolution satellite soil moisture value simulated by aggregating the SMAPEX-5 soil moisture over the study area (cm^3/cm^3)
T_{AM}	Early morning soil temperature (°C)
T_b	Brightness temperature (K)
T_{med}	Daily median temperature (°C)
T_{PM}	Afternoon soil temperature (°C)
x_i	Predictor values
y_i	Response values

Source: NSW Department of Planning, Industry and Environment, 2017.

CRediT authorship contribution statement

I.P. Senanayake: Conceptualization, Methodology, Software, Formal analysis, Validation, Investigation, Writing – original draft, Writing – review & editing. **I.-Y. Yeo:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Funding acquisition. **J.P. Walker:** Data curation, Resources, Writing – review & editing, Supervision. **G.R. Willgoose:** Conceptualization, Writing – review & editing, Supervision.

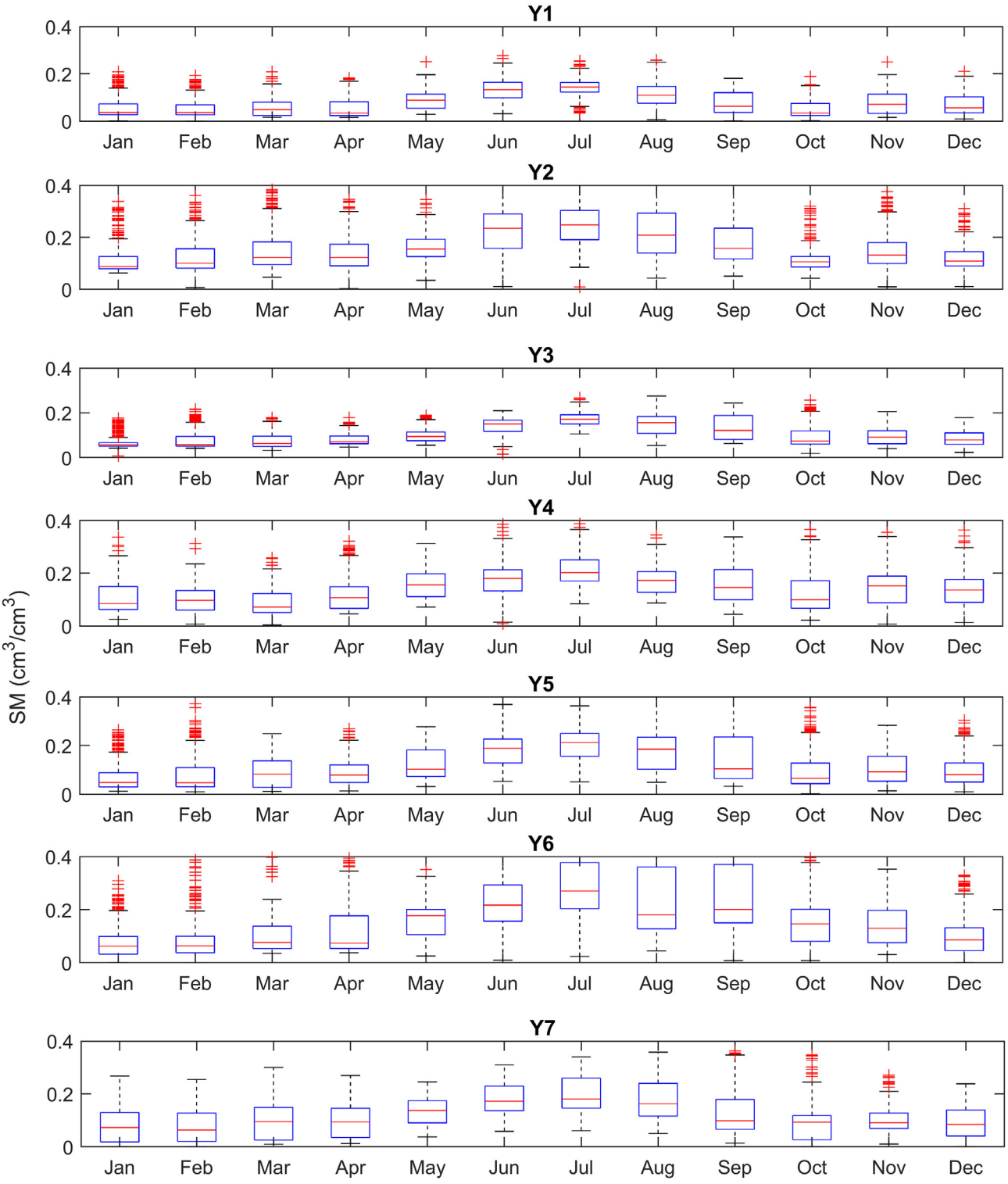
Declaration of competing interest

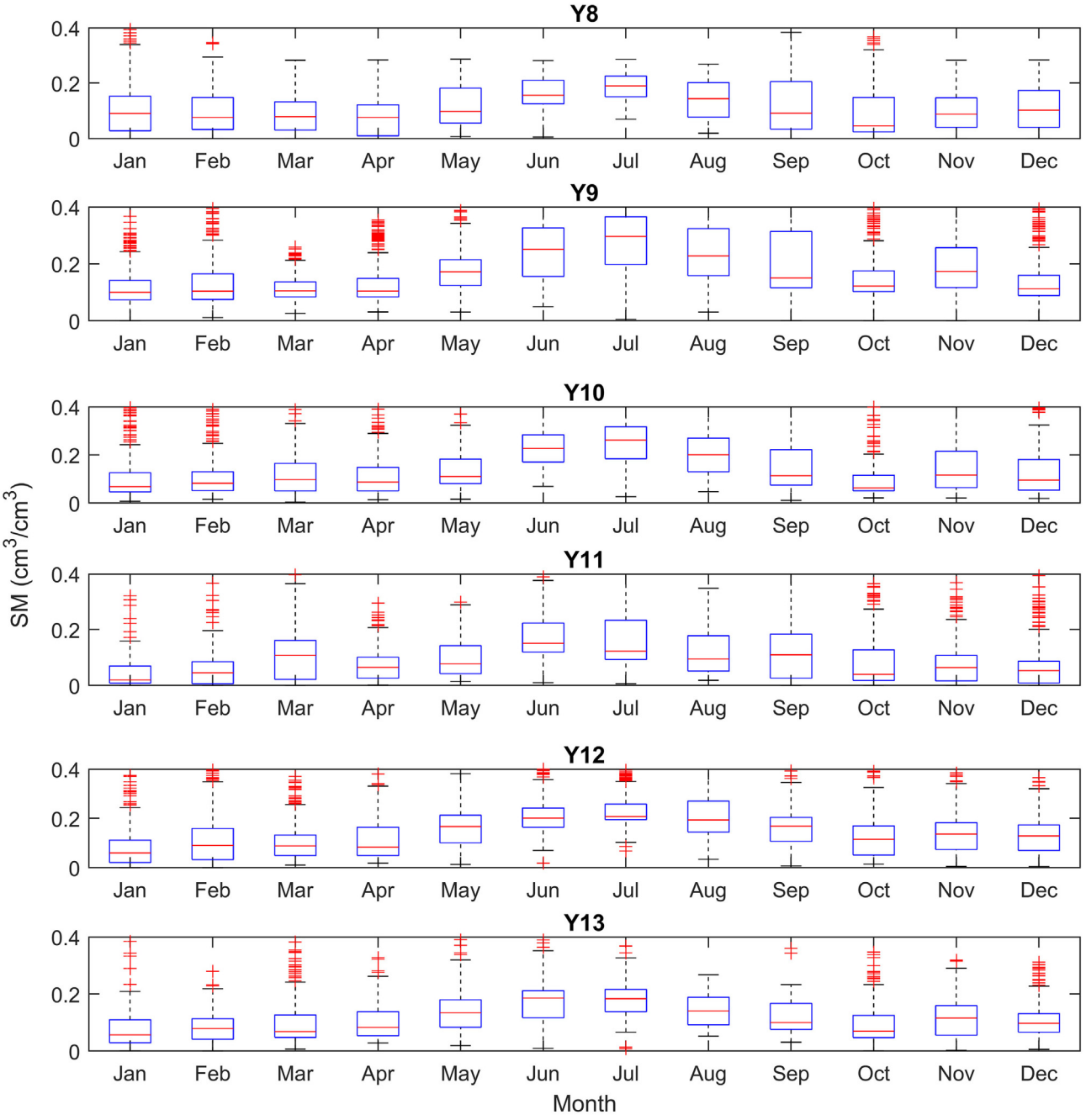
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by the University of Newcastle Postgraduate Research Scholarship (UNIPRS). We extend our gratitude to Dr. Nan Ye, Research Fellow at the Department of Civil Engineering, Monash University, Australia for his assistance related to the SMAPEx datasets.

Appendix 1. The distribution of daily soil moisture over each month at Yanco monitoring stations, based on Yanco in-situ data from 2004 to 2017 (from 2002 to 2012 for Y3)





Appendix 2. Mean and standard deviation (STD) of near surface soil moisture in Yanco Sites over the period of 2004 to 2017 (from 2002 to 2012 for Y3)

	Mean SM (cm^3/cm^3)	STD SM (cm^3/cm^3)
Y1	0.081	0.054
Y2	0.168	0.088
Y3	0.107	0.051
Y4	0.145	0.071
Y5	0.121	0.085
Y6	0.168	0.123
Y7	0.122	0.080
Y8	0.114	0.083
Y9	0.176	0.132
Y10	0.149	0.112
Y11	0.108	0.106

References

Abbaszadeh, P., Moradkhani, H., Zhan, X., 2019. Downscaling SMAP radiometer soil moisture over the CONUS using an ensemble learning method. Water Resour. Res. 55 (1), 324–344. https://doi.org/10.1029/2018WR023354.

Alaloul, W.S., Qureshi, A.H., 2020. Data processing using artificial neural networks. In Dynamic Data Assimilation–Beating the Uncertainties. IntechOpen https://doi.org/10.5772/intechopen.91935.

Alemohammad, S.H., Kolassa, J., Prigent, C., Aires, F., Gentine, P., 2018. Global downscaling of remotely sensed soil moisture using neural networks. Hydrol. Earth Syst. Sci. 22 (10), 5341–5356. https://doi.org/10.5194/hess-22-5341-2018.

Best, M.J., Pryor, M., Clark, D.B., Rooney, G.G., Essery, R.L.H., Ménard, C.B., Edwards, J.M., Hendry, M.A., Porson, A., Gedney, N., Mercado, L.M., Sitch, S., Blyth, E., Boucher, O., Cox, P.M., Grimmond, C.S.B., Harding, R.J., 2011. The joint UK land environment simulator (JULES), model description–part 1: energy and water fluxes. Geosci. Model Dev. 4 (1), 677–699. https://doi.org/10.5194/gmd-4-677-2011.

Bindlish, R., Jackson, T. J., Chan, S., Colliander, A., & Kerr, Y. (2017). Integration of SMAP and SMOS L-band observations. Paper presented at the 2017 IEEE international geoscience and remote sensing symposium (IGARSS), Fort Worth, TX, 2017, pp. 2546–2549. doi: https://doi.org/10.1109/IGARSS.2017.8127514.

Bonan, G., 2015. Ecological Climatology: Concepts and Applications. Cambridge University Press, NY, USA.

Bretreger, D., Yeo, I.Y., Hancock, G., Willgoose, G., 2020. Monitoring irrigation using Landsat observations and climate data over regional scales in the Murray-Darling Basin. J. Hydrol. 590, 125356. https://doi.org/10.1016/j.jhydrol.2020.125356.

Burden, F., Winkler, D., 2008. Bayesian regularization of neural networks. In: Livingstone, D.J. (Ed.), Artificial Neural Networks. Humana Press, NJ, United States, pp. 23–42 https://doi.org/10.1007/978-1-60327-101-1.

Carbonell, J.G., Michalski, R.S., Mitchell, T.M., 1983. An overview of machine learning. In: Michalski, R.S., Carbonell, J.G., Mitchell, T.M. (Eds.), Machine learning, an artificial intelligence approach. volume I. Morgan Kaufmann, United States, pp. 3–23. https://doi.org/10.1016/B978-0-08-051054-5.50005-4.

Carlson, T.N., Gillies, R.R., Perry, E.M., 1994. A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover. Remote Sens. Rev. 9 (1–2), 161–173. https://doi.org/10.1080/02757259409532220.

Chai, S., Walker, J.P., Veenendaal, B., West, G., 2011. An artificial neural network model for downscaling of passive microwave soil moisture. Paper Presented at the 6th IASME/WSEAS International Conference on Water Resources, Hydraulics & Hydrology (WHH'11), Cambridge, UK.

Chauhan, N., Miller, S., Ardanuy, P., 2003. Spaceborne soil moisture estimation at high resolution: a microwave-optical/IR synergistic approach. Int. J. Remote Sens. 24 (22), 4599–4622. https://doi.org/10.1080/0143116031000156837.

Colliander, A., Jackson, T.J., Bindlish, R., Chan, S., Das, N., Kim, S.B., Cosh, M.H., Dunbar, R.S., Dang, L., Pashaian, L., Asanuma, J., Aida, K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K., Goodrich, D., Al Jassar, H., Lopez-Baeza, E., Martínez-Fernández, J., González-Zamora, A., Livingston, S., McNairn, H., Pacheco, A., Moghaddam, M., Montzka, C., Notarnicola, C., Niedrist, G., Pellarin, T., Prueger, J., Pulliainen, J., Rautiainen, K., Ramos, J., Seyfried, M., Starks, P., Su, Z., Zeng, Y., van der Velde, R., Thibeault, M., Dorigo, W., Vreugdenhil, M., Walker, J.P., Wu, X., Monerris, A., O'Neill, P.E., Entekhabi, D., Njoku, E.G., Yueh, S., 2017. Validation of SMAP surface soil moisture products with core validation sites. Remote Sens. Environ. 191, 215–231. https://doi.org/10.1016/j.rse.2017.01.021.

Colliander, A., Jackson, T.J., Chan, S.K., O'Neill, P., Bindlish, R., Cosh, M.H., Caldwell, T., Walker, J.P., Berg, A., McNairn, H., Thibeault, M., Martínez-Fernández, J., Jensen, K.H., Asanuma, J., Seyfried, M.S., Bosch, D.D., Starks, P.J., Holifield Collins, C., Prueger, J.H., Su, Z., Lopez-Baezar, E., Yueh, S.H., 2018. An assessment of the differences between spatial resolution and grid size for the SMAP enhanced soil moisture product over homogeneous sites. Remote Sens. Environ. 207, 65–70. https://doi.org/10.1016/j.rse.2018.02.006.

Crosson, W.L., Al-Hamdan, M.Z., Hemmings, S.N., Wade, G.M., 2012. A daily merged MODIS aqua–Terra land surface temperature data set for the conterminous United States. Remote Sens. Environ. 119, 315–324. https://doi.org/10.1016/j.rse.2011.12.019.

De Alcântara Silva, V.M., Gama, C.M., dos Santos, É.G., Nunes, S.H.P., dos Santos, C.A.C., 2016. Characterization NDVI space-time and surface and analysis phytosociologic albedo for São João do Cariri. Journal of Hyperspectral Remote Sensing 6 (6), 305–315. https://doi.org/10.5935/2237-2202.20160030.

Didan, K., 2015. MYD13A2 MODIS/aqua vegetation indices 16-day L3 Global 1km SIN Grid V006. NASA EOSDIS Land Processes DAAC https://doi.org/10.5067/modis/myd13a2.006.

Du, K.-L., Swamy, M.N., 2013. Neural Networks and Statistical Learning. Springer-Verlag, Londonhttps://doi.org/10.1007/978-1-4471-5571-3.

El Bouchefry, K., de Souza, R.S., 2020. Learning in big data: Introduction to machine learning. In: Škoda, P., Adam, F. (Eds.), Knowledge Discovery in Big Data from Astronomy and Earth Observation (Pp. 225–249). Elsevier https://doi.org/10.1016/B978-0-12-819154-5.00023-0.

Engman, E.T., 1991. Applications of microwave remote sensing of soil moisture for water resources and agriculture. Remote Sens. Environ. 35 (2–3), 213–226. https://doi.org/10.1016/0034-4257(91)90013-V.

Entekhabi, D., Reichle, R.H., Koster, R.D., Crow, W.T., 2010. Performance metrics for soil moisture retrievals and application requirements. J. Hydrometeorol. 11 (3), 832–840. https://doi.org/10.1175/2010JHM1223.1.

Entekhabi, D., Yueh, S., O'Neill, P.E., Kellogg, K.H., Allen, A., Bindlish, R., Brown, M., Chan, S., Colliander, A., Crow, W.T., Das, N., De Lannoy, G., Dunbar, R.S., Edelstein, W.N., Entin, J.K., Escobar, V., Goodman, S.D., Jackson, T.J., Jai, B., Johnson, J., Kim, E., Kim, S., Kimball, J., Koster, R.D., Leon, A., McDonald, K.C., Moghaddam, M., Mohammed, P., Moran, S., Njoku, E.G., Piepmeier, J.R., Reichle, R., Rogez, F., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Van Zyl, J., Weiss, B., West, R., 2014. SMAP Handbook–Soil Moisture Active Passive: Mapping Soil Moisture and Freeze/Thaw from Space, JPL Publication JPL. Jet Propulsion Laboratory, Pasadena, California, pp. 400–1567.

Fang, B., Lakshmi, V., 2014. Soil moisture at watershed scale: remote sensing techniques. J. Hydrol. 516, 258–272. https://doi.org/10.1016/j.jhydrol.2013.12.008.

Fang, B., Lakshmi, V., Bindlish, R., Jackson, T.J., Cosh, M., Basara, J., 2013. Passive microwave soil moisture downscaling using vegetation index and skin surface temperature. Vadose Zone J. 12 (3). https://doi.org/10.2136/vzj2013.05.0089.

Fang, B., Lakshmi, V., Bindlish, R., Jackson, T., 2018. Downscaling of SMAP soil moisture using land surface temperature and vegetation data. Vadose Zone J. 17 (1). https://doi.org/10.2136/vzj2017.11.0198.

Frost, A.J., Ramchurn, A., Hafeez, M., 2015. Evaluation of AWRA-L for national drought and soil moisture monitoring. Paper presented at the 36th Hydrology and Water Resources Symposium (HWRS 2015). Hobart, TAS, pp. 1496–1504.

Grayson, R.B., Western, A.W., 1998. Towards areal estimation of soil water content from point measurements: time and space stability of mean response. J. Hydrol. 207 (1–2), 68–82. https://doi.org/10.1016/S0022-1694(98)00096-1.

Grundy, M., Rossel, R.V., Searle, R., Wilson, P., Chen, C., Gregory, L., 2015. Soil and landscape grid of Australia. Soil Research 53 (8), 835–844. https://doi.org/10.1071/SR15191.

Han, J., Kamber, M., Pei, J., 2012. Classification: advanced methods. Data mining concepts and techniques. Morgan Kaufmann, United States, pp. 393–443 https://doi.org/10.1016/B978-0-12-381479-1.00009-5.

Haykin, S., 1994. Neural Networks: A Comprehensive Foundation: Prentice Hall PTR, NJ (United States).

Hultquist, C., Chen, G., Zhao, K., 2014. A comparison of Gaussian process regression, random forests and support vector regression for burn severity assessment in diseased forests. Remote Sensing Letters 5 (8), 723–732. https://doi.org/10.1080/2150704X.2014.963733.

Im, J., Park, S., Rhee, J., Baik, J., Choi, M., 2016. Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches. Environ. Earth Sci. 75 (15), 1120. https://doi.org/10.1007/s12665-016-5917-6.

Jackson, T.J., Schmugge, T.J., Wang, J.R., 1982. Passive microwave sensing of soil moisture under vegetation canopies. Water Resour. Res. 18 (4), 1137–1142. https://doi.org/10.1029/WR018i004p01137.

Jin, Y., Ge, Y., Liu, Y., Chen, Y., Zhang, H., Heuvelink, G.B., 2020. A machine learning-based geostatistical downscaling method for coarse-resolution soil moisture products. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14, 1025–1037. https://doi.org/10.1109/JSTARS.2020.3035386.

Kerr, Y.H., Waldteufel, P., Wigneron, J., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M., Font, J., Reul, N., Gruhier, C., 2010. The SMOS mission: new tool for monitoring key elements of the global water cycle. Proc. IEEE 98 (5), 666–687. https://doi.org/10.1109/JPROC.2010.2043032.

Kornelsen, K.C., Coulibaly, P., 2013. Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications. J. Hydrol. 476, 460–489. https://doi.org/10.1016/j.jhydrol.2012.10.044.

Kowalczyk, E.A., Wang, Y.P., Law, R.M., Davies, H.L., McGregor, J.L., Abramowitz, G., 2006. The CSIRO atmosphere biosphere land exchange (CABLE) model for use in climate models and as an offline model. CSIRO Marine and Atmospheric Research Paper. 13, p. 42.

Lakshmi, V., 2013. Remote sensing of soil moisture. ISRN Soil Science 2013. https://doi.org/10.1155/2013/424178.

Lange, H., Sippel, S., 2020. Machine learning applications in hydrology. In: Levia, D.F., Carlyle-Moses, D.E., Lida, S., Michalzik, B., Nanko, K., Tischer, A. (Eds.), Forest-Water Interactions. Springer, Cham, pp. 233–257 https://doi.org/10.1007/978-3-030-26086-6_10.

Lu, S., Ju, Z., Ren, T., Horton, R., 2009. A general approach to estimate soil water content from thermal inertia. Agric. For. Meteorol. 149 (10), 1693–1698. https://doi.org/10.1016/j.agrformet.2009.05.011.

MacKay, D.J., 1992. Bayesian interpolation. Neural Comput. 4 (3), 415–447. https://doi.org/10.1162/neco.1992.4.3.415.

Malbéteau, Y., Merlin, O., Molero, B., Rüdiger, C., Bacon, S., 2016. DisPATCh as a tool to evaluate coarse-scale remotely sensed soil moisture using localized in situ measurements: application to SMOS and AMSR-E data in southeastern Australia. Int. J. Appl. Earth Obs. Geoinf. 45, 221–234. https://doi.org/10.1016/j.jag.2015.10.002.

Mehdipour, P., Navidi, I., Parsaeian, M., Mohammadi, Y., Moradi, L.M., Rezaei, D.E., Nourijelyani, K., Farzadfar, F., 2014. Application of Gaussian process regression (GPR) in estimating under-five mortality levels and trends in Iran 1990-2013, study protocol. Archives of Iranian Medicine 17 (3), 189–192.

Merlin, O., Walker, J.P., Kalma, J.D., Kim, E.J., Hacker, J., Panciera, R., Young, R., Summerell, G., Hornbuckle, J., Hafeez, M., 2008. The NAFE'06 data set: towards soil moisture retrieval at intermediate resolution. Adv. Water Resour. 31 (11), 1444–1455. https://doi.org/10.1016/j.advwatres.2008.01.018.

Merlin, O., Rüdiger, C., Al Bitar, A., Richaume, P., Walker, J.P., Kerr, Y.H., 2012. Disaggregation of SMOS soil moisture in southeastern Australia. IEEE Trans. Geosci. Remote Sens. 50 (5), 1556–1571. https://doi.org/10.1109/TGRS.2011.2175000.

Mohanty, B.P., Cosh, M.H., Lakshmi, V., Montzka, C., 2017. Soil moisture remote sensing: State of the science. Vadose Zone J. 16 (1), 1–9. https://doi.org/10.2136/vzj2016.10.0105.

Molero, B., Merlin, O., Malbéteau, Y., Al Bitar, A., Cabot, F., Stefan, V., Kerr, Y., Bacon, S., Cosh, M.H., Bindlish, R., 2016. SMOS disaggregated soil moisture product at 1 km resolution: processor overview and first validation results. Remote Sens. Environ. 180, 361–376. https://doi.org/10.1016/j.rse.2016.02.045.

Murphy, K.P., 2012. Machine Learning: a Probabilistic Perspective. MIT Press, Cambridge, Massachusetts.

Nguyen-Tuong, D., Peters, J., 2008. Local Gaussian process regression for real-time model-based robot control. Paper presented at the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems https://doi.org/10.1109/IROS.2008.4650850.

Njoku, E.G., Entekhabi, D., 1996. Passive microwave remote sensing of soil moisture. J. Hydrol. 184 (1–2), 101–129. https://doi.org/10.1016/0022-1694(95)02970-2.

NSW DPIE, 2017. NSW Land Use 2013. Department of Planning, Industry and Environment, NSW, Australia https://datasets.seed.nsw.gov.au/dataset/nsw-landuse-2013.
Okut, H., 2016. Bayesian regularized neural networks for small n big p data. In: Rosa, J.L.G. (Ed.), Artificial Neural Networks-Models and Applications. InTech, Rijeka, Croatia, pp. 27–48 https://doi.org/10.5772/61493.
Oyoshi, K., Akatsuka, S., Takeuchi, W., Sobue, S., 2014. Hourly LST monitoring with Japanese geostationary satellite MTSAT-1R over the Asia-Pacific region. Asian Journal of Geoinformatics 14 (3), 1–13.
Panciera, R., Walker, J.P., Jackson, T.J., Gray, D.A., Tanase, M.A., Ryu, D., Alessandra, M., Yardley, H., Rüdiger, C., Wu, X., 2014. The soil moisture active passive experiments (SMAPEx): toward soil moisture retrieval from the SMAP mission. IEEE Trans. Geosci. Remote Sens. 52 (1), 490–507. https://doi.org/10.1109/TGRS.2013.2241774.
Peng, J., Loew, A., Merlin, O., Verhoest, N.E., 2017. A review of spatial downscaling of satellite remotely sensed soil moisture. Rev. Geophys. 55 (2), 341–366. https://doi.org/10.1002/2016RG000543.
Piles, M., Camps, A., Vall-Llossera, M., Corbella, I., Panciera, R., Rüdiger, C., Kerr, Y.H., Walker, J., 2011. Downscaling SMOS-derived soil moisture using MODIS visible/infrared data. IEEE Trans. Geosci. Remote Sens. 49 (9), 3156–3166. https://doi.org/10.1109/TGRS.2011.2120615.
Piles, M., Vall-Llossera, M., Laguna, L., Camps, A., 2012. A downscaling approach to combine SMOS multi-angular and full-polarimetric observations with MODIS VIS/IR data into high resolution soil moisture maps. Paper Presented at the 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich. 2012, pp. 1247–1250. https://doi.org/10.1109/IGARSS.2012.6351316.
Piles, M., Sánchez, N., Vall-llossera, M., Camps, A., Martínez-Fernández, J., Martínez, J., González-Gambau, V., 2014. A downscaling approach for SMOS land observations: evaluation of high-resolution soil moisture maps over the Iberian Peninsula. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7 (9), 3845–3857. https://doi.org/10.1109/JSTARS.2014.2325398.
Piles, M., Petropoulos, G.P., Sánchez, N., González-Zamora, Á., Ireland, G., 2016. Towards improved spatio-temporal resolution soil moisture retrievals from the synergy of SMOS and MSG SEVIRI spaceborne observations. Remote Sens. Environ. 180, 403–417. https://doi.org/10.1016/j.rse.2016.02.048.
Portal, G., Vall-llossera, M., Piles, M., Camps, A., Chaparro, D., Pablos, M., Rossato, L., 2018. A spatially consistent downscaling approach for SMOS using an adaptive moving window. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11 (6), 1883–1894. https://doi.org/10.1109/JSTARS.2018.2832447.
Prechelt, L., 1998. Early stopping - but when? In: Orr, G.B., Müller, K.R. (Eds.), Neural Networks: Tricks of the Trade. Lecture Notes in Computer Science. 1524. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-49430-8_3
Rasmussen, C.E., 2004. Gaussian processes in machine learning. In: Bousquet, O., von Luxburg, U., Rätsch, G. (Eds.), *Advanced Lectures on Machine Learning. ML* 2003. *Lecture Notes in Computer Science*. 3176. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-28650-9_4.
Rasmussen, C.E., Williams, C., 2006. Gaussian processes for machine learning. In: Dietterich, T. (Ed.), Adaptive Computation and Machine Learning. MIT Press, Cambridge, Massachusetts, pp. 40–43.
Raupach, M., Briggs, P., Haverd, V., King, E., Paget, M., Trudinger, C., 2009. Australian water availability project (AWAP): CSIRO marine and atmospheric research component: Final report for phase 3. Centre for Australian Weather and Climate Research, a Partnership between the Bureau of Meteorology and CSIRO, Melbourne, Australia.
Recknagel, F., 2001. Applications of machine learning to ecological modelling. Ecol. Model. 146 (1–3), 303–310. https://doi.org/10.1016/S0304-3800(01)00316-7.
Rodell, M., Houser, P., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., 2004. The global land data assimilation system. Bull. Am. Meteorol. Soc. 85, 381–394. https://doi.org/10.1175/BAMS-85-3-381.
Rodriguez-Iturbe, I., D'odorico, P., Porporato, A., Ridolfi, L., 1999. On the spatial and temporal links between vegetation, climate, and soil moisture. Water Resour. Res. 35 (12), 3709–3722. https://doi.org/10.1029/1999WR900255.
Rossel, R.A.V., Chen, C., Grundy, M.J., Searle, R., Clifford, D., Campbell, P.H., 2015. The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project. Soil Research 53 (8), 845–864. https://doi.org/10.1071/SR14366.
Sabaghy, S., Walker, J.P., Renzullo, L.J., Jackson, T.J., 2018. Spatially enhanced passive microwave derived soil moisture: capabilities and opportunities. Remote Sens. Environ. 209, 551–580. https://doi.org/10.1016/j.rse.2018.02.065.
Sabaghy, S., Walker, J.P., Renzullo, L.J., Akbar, R., Chan, S., Chaubell, J., Das, J., Dunbar, R.S., Entekhabi, D., Gevaert, A., 2020. Comprehensive analysis of alternative downscaled soil moisture products. Remote Sens. Environ. 239, 111586. https://doi.org/10.1016/j.rse.2019.111586.
Sánchez-Ruiz, S., Piles, M., Sánchez, N., Martínez-Fernández, J., Vall-llossera, M., Camps, A., 2014. Combining SMOS with visible and near/shortwave/thermal infrared satellite data for high resolution soil moisture estimates. J. Hydrol. 516, 273–283. https://doi.org/10.1016/j.jhydrol.2013.12.047.
Schmugge, T., 1998. Applications of passive microwave observations of surface soil moisture. J. Hydrol. 212, 188–197. https://doi.org/10.1016/S0022-1694(98)00209-1.
Schmugge, T.J., 1983. Remote sensing of soil moisture: recent advances. IEEE Trans. Geosci. Remote Sens. 3, 336–344. https://doi.org/10.1109/TGRS.1983.350563.
Sellers, W.D., 1965. Physical Climatology. University of Chicago Press, Chicago, Illinois, United States.
Senanayake, I., Yeo, I.-Y., Tangdamrongsub, N., Willgoose, G., Hancock, G., Wells, T., Fang, B., Lakshmi, V., Walker, J.P., 2019a. An in-situ data based model to downscale radiometric satellite soil moisture products in the upper hunter region of NSW, Australia. J. Hydrol. 572, 820–838. https://doi.org/10.1016/j.jhydrol.2019.03.014.
Senanayake, I.P., Yeo, I.-Y., Willgoose, G.R., Hancock, G.R., Bretreger, D., 2019b. Using an artificial neural network to enhance the spatial resolution of satellite soil moisture products based on soil thermal inertia. Paper presented at the 23rd International Congress on Modelling and Simulation 2019 (MODSIM2019). 2019, pp. 1049–1055 Canberra, ACT, Australia.
Senanayake, I.P., Yeo, I.Y., Willgoose, G.R., Hancock, G.R., 2021. Disaggregating satellite soil moisture products based on soil thermal inertia: a comparison of a downscaling model built at two spatial scales. J. Hydrol., 125894 https://doi.org/10.1016/j.jhydrol.2020.125894.
Shobha, G., Rangaswamy, S., 2018. Machine learning. In: Gudivada, V.N., Rao, C.R. (Eds.), Handbook of Statistics (Pp. 197–228). Elsevier https://doi.org/10.1016/bs.host.2018.07.004.
Smith, A.B., Walker, J.P., Western, A.W., Young, R.I., Ellett, K.M., Pipunic, R.C., Grayson, R.B., Siriwardena, L., Chiew, F.H.S., Richter, H., 2012. The Murrumbidgee soil moisture monitoring network data set. Water Resour. Res. 48 (7). https://doi.org/10.1029/2012WR011976.
Soliman, A., Heck, R., Brenning, A., Brown, R., Miller, S., 2013. Remote sensing of soil moisture in vineyards using airborne and ground-based thermal inertia data. Remote Sens. 5 (8), 3729–3748. https://doi.org/10.3390/rs5083729.
Srivastava, P.K., Han, D., Ramirez, M.R., Islam, T., 2013. Machine learning techniques for downscaling SMOS satellite soil moisture using MODIS land surface temperature for hydrological application. Water Resour. Manag. 27 (8), 3127–3144. https://doi.org/10.1007/s11269-013-0337-9.
Taki, M., Rohani, A., Soheili-Fard, F., Abdeshahi, A., 2018. Assessment of energy consumption and modeling of output energy for wheat production by neural network (MLP and RBF) and Gaussian process regression (GPR) models. J. Clean. Prod. 172, 3028–3041. https://doi.org/10.1016/j.jclepro.2017.11.107.
Van de Griend, A.A., Camillo, P.J., Gurney, R.J., 1985. Discrimination of soil physical parameters, thermal inertia, and soil moisture from diurnal surface temperature fluctuations. Water Resour. Res. 21 (7), 997–1009. https://doi.org/10.1029/WR021i007p00997.
Verstraeten, W.W., Veroustraete, F., van der Sande, C.J., Grootaers, I., Feyen, J., 2006. Soil moisture retrieval using thermal inertia, determined with visible and thermal spaceborne data, validated for European forests. Remote Sens. Environ. 101 (3), 299–314. https://doi.org/10.1016/j.rse.2005.12.016.
Wan, Z., Hook, S., Hulley, G., 2015. MOD11A2 MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1 km SIN Grid V006. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD11A2.006.
Wang, J., Hu, J., 2015. A robust combination approach for short-term wind speed forecasting and analysis–combination of the ARIMA (autoregressive integrated moving average), ELM (extreme learning machine), SVM (support vector machine) and LSSVM (Least Square SVM) forecasts using a GPR (Gaussian process regression) model. Energy 93, 41–56. https://doi.org/10.1016/j.energy.2015.08.045.
Wang, J., Bras, R., Sivandran, G., Knox, R., 2010. A simple method for the estimation of thermal inertia. Geophys. Res. Lett. 37 (5). https://doi.org/10.1029/2009GL041851.
Wu, Q., Law, R., Xu, X., 2012. A sparse Gaussian process regression model for tourism demand forecasting in Hong Kong. Expert Syst. Appl. 39 (5), 4769–4774. https://doi.org/10.1016/j.eswa.2011.09.159.
Yamamoto, Y., Ishikawa, H., 2018. Thermal land surface emissivity for retrieving land surface temperature from Himawari-8. Journal of the Meteorological Society of Japan. Ser. II 96, 43–58. https://doi.org/10.2151/jmsj.2018-004.
Ye, N., Walker, J., Guerschman, J., Ryu, D., Gurney, R.J., 2015. Standing water effect on soil moisture retrieval from L-band passive microwave observations. Remote Sens. Environ. 169, 232–242. https://doi.org/10.1016/j.rse.2015.08.013.
Ye, N., Walker, J., Wu, X., Jackson, T., Renzullo, L., Merlin, O., Rüdiger, C., Entekhabi, D., DeJeu, R., Kim, E., 2016. Towards validation of SMAP: SMAPEX-4 &-5. Paper presented at the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, 2016, pp. 3469–3472 https://doi.org/10.1109/IGARSS.2016.7729897.
Ye, N., Walker, J., Bindlish, R., Chaubell, J., Das, N., Gevaert, A.I., Jackson, T.J., Rüdiger, C., 2019. Evaluation of SMAP downscaled brightness temperature using SMAPEx-4/5 airborne observations. Remote Sens. Environ. 221, 363–372. https://doi.org/10.1016/j.rse.2018.11.033.
Ye, N., Walker, J.P., Wu, X., de Jeu, R., Gao, Y., Jackson, T.J., Jonard, F., Kim, E., Merlin, O., Pauwels, V.R.N., Renzullo, L.J., Rudiger, C., Sabaghy, S., von Hebel, C., Yueh, S.H., Zhu, L., 2020. (2020) the soil moisture active passive experiments: validation of the SMAP products in Australia. IEEE Trans. Geosci. Remote Sens. https://doi.org/10.1109/TGRS.2020.3007371.
Yee, M.S., Walker, J.P., Monerris, A., Rüdiger, C., Jackson, T.J., 2016. On the identification of representative in situ soil moisture monitoring stations for the validation of SMAP soil moisture products in Australia. J. Hydrol. 537, 367–381. https://doi.org/10.1016/j.jhydrol.2016.03.060.
Young, R., Walker, J., Yeoh, N., Smith, A., Ellett, K., Merlin, O., Western, A., 2008. Soil Moisture and Meteorological Observations from the Murrumbidgee Catchment. Department of Civil and Environmental Engineering, The University of Melbourne.
Zhang, N., Xiong, J., Zhong, J., Leatham, K., 2018. Gaussian process regression method for classification for high-dimensional data with limited samples. Paper Presented at the 2018 Eighth International Conference on Information Science and Technology (ICIST), Cordoba, 2018, pp. 358–363 https://doi.org/10.1109/ICIST.2018.8426077.
Zhu, L., Walker, J.P., Ye, N., Rudiger, C., 2016. The effect of radar configuration on effective correlation length. Paper Presented at the 2016 International Conference on Electromagnetics in Advanced Applications (ICEAA), Cairns, QLD, 2016, pp. 820–823 https://doi.org/10.1109/ICEAA.2016.7731525.
Zhu, L., Walker, J.P., Tsang, L., Huang, H., Ye, N., Rüdiger, C., 2018a. Soil moisture retrieval over agricultural fields from time series multi-angular L-band radar data. Paper presented at the IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, 2018, pp. 6139–6142 https://doi.org/10.1109/IGARSS.2018.8518119.
Zhu, L., Walker, J.P., Ye, N., Rüdiger, C., Hacker, J.M., Panciera, R., Tanase, M.A., Wu, X., Gray, D.A., Stacy, N., 2018b. The Polarimetric L-band imaging synthetic aperture radar (PLIS): description, calibration, and cross-validation. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11 (11), 4513–4525. https://doi.org/10.1109/JSTARS.2018.2873218.