Remote Sensing of Environment 251 (2020) 112099

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Stochastic ensemble methods for multi-SAR-mission soil moisture retrieval

Liujun Zhu*, Jeffrey P. Walker, Xiaoji Shen

Department of Civil Engineering, Monash University, Clayton, Vic. 3800, Australia

ARTICLE INFO

Keywords:
Soil moisture
Ensemble learning
Synthetic aperture radar
Multi-frequency
Multi-temporal

ABSTRACT

The recent and projected investments across the world on radar satellite missions (e.g., Sentinel-1, SAOCOM, BIOMASS and NISAR) provide a great opportunity for operational radar soil moisture mapping with high spatial and temporal resolution. However, there is no retrieval algorithm that can make complementary use of the multi-frequency data from those missions, due to the large uncertainties in observations collected by the different sensors, different validity regions of the forward models, and the fact that inversion algorithms have been designed for specific data sources. In this study, the principle of ensemble learning was introduced to provide two general soil moisture retrieval frameworks accounting for these issues. Instead of trying to find an optimal global solution, multiple soil moisture retrievals (termed sub-retrievals) with moderate performance were first obtained using different channels and/or time instances randomly selected from the available data, with the retrieved ensemble of results being the final output. The ensemble retrievals, taking one existing snapshot method and two multi-temporal methods as the base retrieval algorithms, were evaluated using a synthetic data set with the effectiveness confirmed under various uncertainty sources. An evaluation using the Fifth Soil Moisture Active Passive Experiment (SMAPEx-5) data set showed that the ensemble retrieval outperformed the non-ensemble retrieval in most cases, with a decrease of 0.004 to 0.014 m^3/m^3 in Root Mean Square Error (RMSE) and an increase of 0.01 to 0.16 in correlation coefficient (R). Weakly biased and correlated sub-retrievals were confirmed to be the basic requirement of an effective ensemble retrieval, being consistent with use of ensemble learning in other applications.

1. Introduction

Information on soil moisture at high spatial and temporal resolution is in demand from a variety of sectors. Synthetic Aperture Radar (SAR) has shown promising results at field scale (Kornelsen and Coulibaly, 2013), with large scale mapping becoming possible because of the recent investment in radar satellites by many countries (e.g., Sentinel-1, SAOCOM and NovaSAR-S). Limited by the poor global coverage of individual SAR satellites (> 10 days), joint use of multiple satellites has been suggested and tested by Zhu et al. (2019a) to reach the 1–5 days temporal repeat requirement of most applications (Walker and Houser, 2004).

Multi-temporal and/or multi-configuration (incidence angle, polarization and frequency) data has a demonstrated potential in solving ill-posed soil moisture inversions (Baghdadi et al., 2006; Balenzano et al., 2011; Bindlish and Barros, 2000; Kim et al., 2014; Ouellette et al., 2017; Pierdicca et al., 2008; Zhu et al., 2019b; Zribi et al., 2005) with assumptions or prior knowledge of roughness, vegetation and/or soil moisture. However, joint soil moisture mapping from multiple SAR missions can be more challenging than from a single SAR, with potentially larger measurement and model errors. Current airborne and spaceborne radars commonly have an absolute calibration accuracy of better than 1 dB (Zhu et al., 2018), being sufficient for reliable soil moisture mapping. However, many studies have demonstrated the large calibration imbalance among different sensors operating at the same frequency (Baghdadi et al., 2014; Gorrab et al., 2015; Pettinato et al., 2013), and different beams or imaging modes of an individual sensor (Schmidt et al., 2018; Shimada et al., 2009).

Similarly, the available surface and vegetation scattering models have varying performances at different frequencies, polarizations, and incidence angles. For instance, the widely used Integral Equation Model (IEM) is prone to large errors at high incidence angles (Mancini et al., 1999) and short wavelengths (Choker et al., 2017). The Distorted Born Approximation (DBA) tends to have a large underestimation at C-band VV polarization over vertically-dominant vegetation layers (Cookmartin et al., 2000; Huang et al., 2017). Moreover, the single-look-complex (SLC) data collected by different sensors, beams, and imaging modes (e.g., ScanSAR and StripMap) have varying spacing. A different number of looks should therefore be applied in the multi-look to have a consistent retrieval grid, resulting in different residual speckle

* Corresponding author.
E-mail address: Liujun.zhu@monash.edu (L. Zhu).

https://doi.org/10.1016/j.rse.2020.112099
Received 31 March 2020; Received in revised form 3 August 2020; Accepted 11 September 2020

and thermal noise level. The varying geometric accuracy of different data sources (Shimada et al., 2009) is another uncertainty source, especially for a heterogeneous area.

To address these uncertainties, many studies have applied empirical calibration or correction factors to forward scattering models (Baghdadi et al., 2016; Lievens et al., 2011; Panciera et al., 2014), being an effective way to partly remove the biases among different data sources and the mismatch between measured and predicted radar data. The speckle and thermal noise of radar data have also been considered in the cost function of soil moisture inversion in many studies where the noise was commonly assumed to follow a zero-mean Gaussian distribution (Kim et al., 2012; Mattia et al., 2006; Notarnicola et al., 2008; Pierdicca et al., 2008; Pierdicca et al., 2010). Consequently, Zhu et al. (2019b) integrated the soil moisture temporal trend with the cost function to reduce the anomaly fluctuations in time series radar measurements resulting in an improved soil moisture retrieval. The use of those methods, however, is cumbersome for multi-SAR-retrieval, because a reliable estimation or assumption of the measurement and model errors is not readily available for data with various frequencies, polarizations, incidence angles and spacing.

In this study, the concept of ensemble machine learning was introduced to multi-SAR-mission soil moisture retrieval, with the expectation of reducing the effect of model and measurement errors. In ensemble learning, a set of alternative models are generated for the same task, with the outputs of those models (e.g. classifiers and predictors) being ensembled as the result. A straightforward example of why ensemble learning can lead to improved prediction/classification is provided in Fig. 1. A total of 21 independent binary classifiers were generated to vote for an output using the rule of "winner takes all". In other words, the ensemble of those classifiers would have an incorrect prediction when more than half of them (> 10) have incorrect predictions. For moderate accuracy classifiers (the bottom panel of Fig. 1) with an error rate of 0.3, the probability that more than 10 classifiers predict incorrectly is ~0.03 ($\Sigma_{i=11}{}^{21}C_{21}{}^{i}0.7^{21-i} \cdot 0.3^{i}$, where $C_{21}{}^{i}$ denotes choosing i from 21) and thus the ensemble of those classifiers results in an accuracy of 0.97. To have a successful ensemble, those binary classifiers should have independent predictions and should be better than a random guess (i.e. error rate = 0.5). However, the condition of independent predictions is commonly hard to meet. In real applications, a necessary and sufficient condition for the success of ensemble methods is that individual models be better than a random guess and have different errors for the same task (Hansen and Salamon, 1990). Many ensemble methods have been proposed for various applications, with the main difference being the way to generate the diversity of individual models. The diversity can be introduced using different subsets of the input data, e.g., the Bootstrap aggregating (Bagging, Breiman, 1996), and using random model parameters, or structures, e.g., the random forest (Breiman, 2001).

These ensemble methods have been widely used in remote sensing classification (Belgiu and Drăguţ, 2016; Healey et al., 2018), downscaling (Abbaszadeh et al., 2019; Hutengs and Vohland, 2016) and geophysical parameter regression (Huang et al., 2019; Stojanova et al., 2010; Zhang et al., 2018a). Different from those data-driven methods with a supervised training process, the ensemble principle, that multiple diverse and moderate solutions can produce a robust solution, was introduced here for multi-SAR-retrieval without a training process. The multi-SAR-mission data was randomly split into multiple subsets, followed by independent soil moisture retrieval (sub-retrievals) on each subset using existing retrieval methods, with the averaged results as output. This process is similar to the multi-model forecasts and decision fusion method which dates back to the 1960s (Bates and Granger, 1969) followed by many recent studies (Baez-Villanueva et al., 2020; Li et al., 2017; Quets et al., 2019), where the combination of multiple forecasts yields a lower root mean square error (RMSE) than any single forecast. Different from those methods, the principle of ensemble learning was used here to guide the construction of multiple sub-retrievals. Consequently, a synthetic radar data set with various uncertainty sources was generated in this study for making a comprehensive evaluation, followed by a real case study using the Fifth Soil Moisture Active and Passive Experiment (SMAPEx-5) data set.

2. Methodology

2.1. Ensemble frameworks

The ensemble retrieval framework proposed here was designed to suit most existing soil moisture retrieval methods, which can be grouped into snapshot methods and multi-temporal methods depending on their time instances. Accordingly, two separate frameworks are presented in Fig. 2. In the snapshot ensemble (Fig. 2a), N_e subsets were randomly selected from the available data. An existing snapshot method was then applied to retrieve soil moisture using each subset respectively, being sub-retrievals in Fig. 2a. The retrieved soil moisture values of N_e sub-retrievals were then ensemble-averaged as output, being the ensemble retrieval. To reflect the benefit from the ensemble, the retrieval using all available data was treated as the benchmark in this study.

Similar to the classic ensemble learning paradigm of Bagging (Breiman, 1996), the difference required in sub-retrievals is introduced by manipulating the input data feature (channel) space. Two parameters are required, being the ensemble size (N_e) and the number of channels (N_c) selected in each sub-retrieval, where a channel refers to an independent observation, e.g., C-band HH at an incidence angle of 30 deg. A snapshot ensemble retrieval with $N_e = 5$ and $N_c = 3$ indicates that 5 sub-retrievals were made. Each sub-retrieval used 3 channels randomly selected from the available channels (4 in Fig. 2 a) with replacement, while the corresponding snapshot benchmark retrieval used all available channels as shown in Fig. 2 a. The five retrieved soil moisture values were then averaged as a single value as the output.

Similarly, N_e time series subsets of the available time series radar data are randomly selected for soil moisture retrieval in the multi-temporal ensemble (Fig. 2b). Despite the random selection of radar configurations (frequency, incidence angle and polarization) for each time instance, a random selection of time instances from the full time series can also be made, referred to here as temporal sampling. For example, no data was selected for day 2 in the second subset of Fig. 2b. An additional parameter of temporal sampling size (N_t) is thus required to determine the number of time instances involved in each sub-retrieval. Given a time series of 8 days, a multi-temporal ensemble retrieval with $N_e = 5$, $N_c = 1$ and $N_t = 7$ means that 5 sub-retrievals were made on the randomly selected 7 of the 8 dates with replacement (temporal sampling). For each selected date, one of the available channels (4 and 5 for day 1 and 2 in Fig. 2 b respectively) was selected with replacement. The corresponding benchmark retrieval used all channels as shown in Fig. 2 b. Since $N_t = 7$, each sub-retrieval only retrieved the soil moisture on the 7 selected days. However, the ensembled results still covered the entire time series as in the example depicted in Fig. 3.

Each sub-retrieval inherits part of the measurement and model uncertainty of the single retrieval (benchmark) in Fig. 2 because of the random sampling according to sensor configuration and/or time. Similar to the ensemble of binary classifiers presented in Fig. 1, the effect of uncertainty is expected to be reduced at the final step of averaging multiple sub-retrievals. The random selection in radar configuration can reduce the effect of data and model imbalance among different frequencies, incidence angles and polarizations, while the random selection in time can partly remove the effect of outliers over time. As the user-defined parameters (N_e, N_c and N_t) could have a significant effect on the ensemble performance, the impact of these parameters was analyzed in detail herein. To facilitate others to build on this work, the code of the proposed ensemble framework is provided at https://github.com/rszlj/MultiSAR-Soil-Moisture-Retrieval.

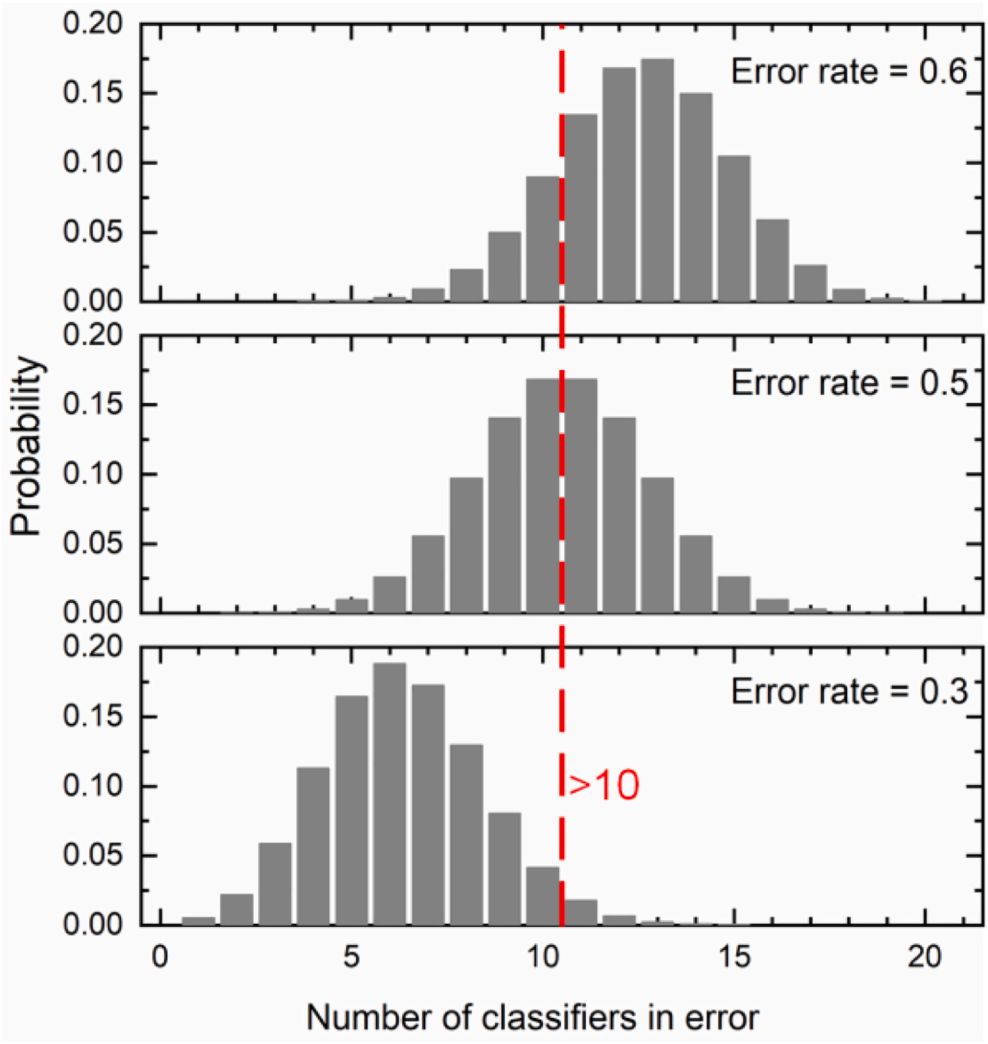

Fig. 1. The probability that 21 independent binary classifiers will make an error, assuming the error rate of each classifier is 0.6 (top), 0.5 (middle) and 0.3 (bottom). The ensemble of the 21 classifiers will make the incorrect prediction when more than 10 of the classifiers predict incorrectly using the rule of "winner takes all". Thus, the summation of bins on the right side of the red line is the error rate of the ensemble model, being 0.82, 0.5 and 0.03 for the top, middle and bottom, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

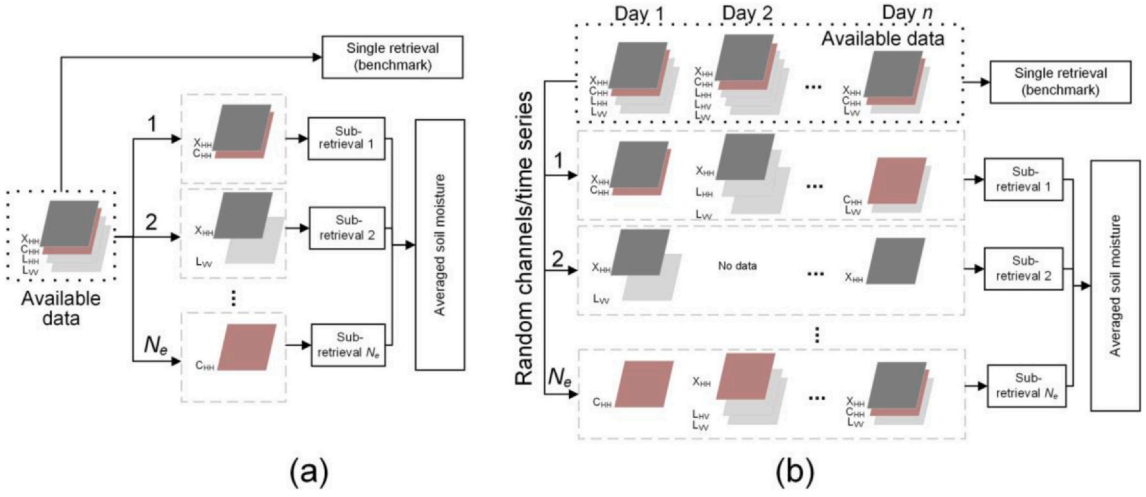

Fig. 2. The ensemble retrieval framework for (a) snapshot methods and (b) multi-temporal methods. The two frameworks share a similar process that N_e sub-retrievals are made through selecting different input channels with replacement and having the N_e retrieval results averaged as the output.

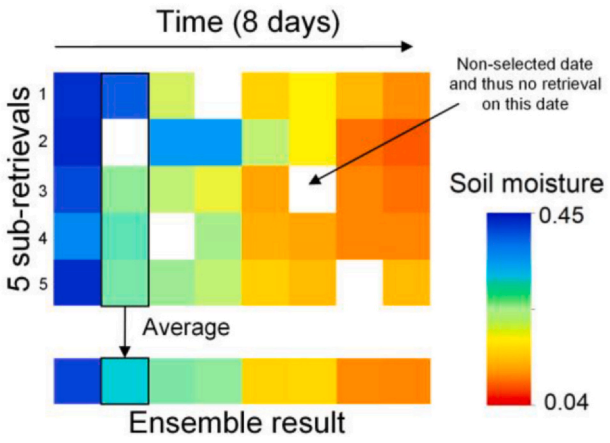

Fig. 3. Conceptual figure showing the ensemble of 5 sub-retrievals with an $N_t = 7$ given a time series of 8 days. Since $N_t = 7$, each sub-retrieval is made on the randomly selected 7 of 8 dates, with the unselected date being no data.

2.2. Soil moisture retrieval methods

Several soil moisture retrieval methods were compared with their ensemble version presented in Fig. 2 for a comprehensive evaluation of the proposed concept. In the snapshot multi-frequency method, soil moisture of a bare soil surface was determined by minimizing the cost function:

$$f = \sum_{j=1}^{m} \sqrt{\frac{1}{m}(\sigma_j^0 - \sigma_{\mathrm{model},j}^0(s,l,mv))^2} \quad (1)$$

where $\sigma_{\mathrm{model},\,j}{}^0(s,l,mv)$ and σ_j^0 are the backscattering coefficients from forward models and the observation in dB, respectively, with the subscript j being the order of m input channels. Parameters s, l and mv denote root mean square roughness height [cm], roughness correlation length [cm] and soil moisture [m^3/m^3] respectively. A genetic algorithm was then applied to search the optimal solution of Eq. 1, being similar to previous studies using multi-frequency data (e.g. Bindlish and Barros, 2000, 2001; Zhang et al., 2018b).

For the multi-temporal retrieval, the cost function was extended to include time series data with assumed time-invariant roughness and/or vegetation such that:

$$f = \sum_{i=1}^{n} \sqrt{\frac{1}{m_i} \sum_{j=1}^{m_i} (\sigma_{ij}^0 - \sigma_{\mathrm{model},ij}^0(s,l,mv_i))^2} \quad (2)$$

where subscript i refers to the ith date of a time series of n days. Notably, m_i can vary in time as shown in Fig. 2b because the available satellite acquisitions vary in time. Soil moisture retrieval is only undertaken on dates with data. The same genetic algorithm was used to derive the time series soil moisture (mv_1, mv_2, …, mv_n) and the time-invariant roughness parameters s and l. A similar cost function has been used by Kim et al. (2012, 2014) and confirmed to outperform the snapshot methods. Moreover, a dry-down trend of soil moisture can be integrated into Eq. 2 for improved results (Zhu et al., 2019a, 2019b). In this study, the snapshot method (Eq. 1), the multi-temporal method (Eq. 2) and the multi-temporal method with a dry-down constraint were used, abbreviated SSM, MT and MTD hereafter.

2.3. Evaluation metrics

The main contribution of this study is applying existing soil moisture retrieval methods within the ensemble framework to provide improved soil moisture mapping without the knowledge of data and model uncertainties. Accordingly, the evaluation is focused on the potential improvement from using the ensemble concept rather than the performance of a specific forward model or soil moisture retrieval algorithm. The widely used bias, correlation coefficient (R) and root mean square error (RMSE) were selected as the indicators of retrieval improvement. These metrics were calculated for the benchmark and ensemble retrieval described in Fig. 2.

To further explore the impacts of the robustness and diversity of sub-retrieval on the ensemble retrieval, the mean squared error (MSE) of the ensemble retrieval was decomposed according to (Ueda and Nakano, 1996):

$$\mathrm{MSE} = \overline{\mathrm{bias}}^2 + \frac{1}{N_e}\overline{\mathrm{var}} + \left(1 - \frac{1}{N_e}\right)\overline{\mathrm{covar}} \quad (3)$$

where $\overline{\mathrm{bias}}$ and $\overline{\mathrm{var}}$ are the average bias and variance of individual sub-retrievals respectively. The $\overline{\mathrm{covar}}$ is the averaged covariance of sub-retrievals given by:

$$\overline{\mathrm{covar}} = \frac{1}{N_e(N_e-1)} \sum_{i=1}^{N_e} \sum_{j=1,i\neq j}^{N_e} \mathrm{cov}_{ij} \quad (4)$$

where cov_{ij} is the covariance of the ith and jth sub-retrieval. According to Eq. 3, for a large number of sub-retrievals N_e, MSE and RMSE are mainly dependent on $\overline{\mathrm{bias}}$ and $\overline{\mathrm{covar}}$. Generating multiple sub-retrievals with small $\overline{\mathrm{bias}}$ and $\overline{\mathrm{covar}}$ is thus key to a successful ensemble retrieval. Since $\overline{\mathrm{bias}}$ equals to the bias of ensemble retrieval, $\overline{\mathrm{covar}}$ was calculated to show the correlation (diversity) of sub-retrievals.

3. Data

3.1. SMAPEx-5 data set

The SMAPEx-5 was carried out in the Yanco area, NSW. During the three-week campaign (7th – 27th September 2015), L-band airborne radar data and intensive ground sampling including soil moisture, roughness and vegetation were collected for in-orbit calibration and validation of the NASA Soil Moisture Active and Passive (SMAP) mission (Ye et al., 2020). This data set has been used to develop and validate the multi-temporal retrieval with a dry down constraint (MTD) in Zhu et al. (2019 a, b, c). In this study, the same data set was used to evaluate the ensemble version of MTD. Accordingly, only a brief summary of the SMAPEx-5 data set used in this analysis is provided here, with the details in Zhu et al. (2019a).

Extensive ground soil moisture sampling (0–5 cm) was made across eight flying dates. These measurements were made on a uniform grid of 250 m spacing over six 3 × 3 km focus areas. The soil moisture sampling dates and the time series soil moisture variation from 36 OzNet soil moisture monitoring stations (http://www.oznet.org.au/) are depicted in Fig. 4. The average soil moisture of OzNet was observed to have a dry-down trend from 0.3 to 0.15 m^3/m^3. Roughness was measured interleaved with the soil moisture sampling dates in selected paddocks using a sequence of pin profiler measurements over 3 m transects. The measured s and l for isotropic surfaces ranged from 0.5 to 3 cm and 5 to 35 cm respectively, with large s values (up to 9 cm) observed in paddocks with periodic row features. Vegetation water content (VWC) was measured destructively on selected paddocks. The VWC of the two main vegetation types (wheat and grass) ranged from 1.17 to 3.72 kg/m^2 and 0.53 to 1.62 kg/m^2 respectively.

A time series of twenty SAR acquisitions at L-, C- and X-band were collected for fifteen different dates of the three-week campaign (Fig. 4). The eight L-band (1.26 GHz) acquisitions were collected concurrently with the ground soil moisture sampling, using the airborne Polarimetric L-band Imaging SAR (PLIS, Zhu et al., 2018). The seven C-band RADARSAT-2 (5.4 GHz) were acquired on DoY 252, 254, 255, 257, 262, 264 and 269 respectively, including three quad-polarization and four dual-polarization products. The X-band data set consisted of five COSMO-SkyMed STRIPMAP HIMAGE, collected on DoY 251, 253, 261, 263 and 269 respectively. Different multi-looks were applied to get a similar

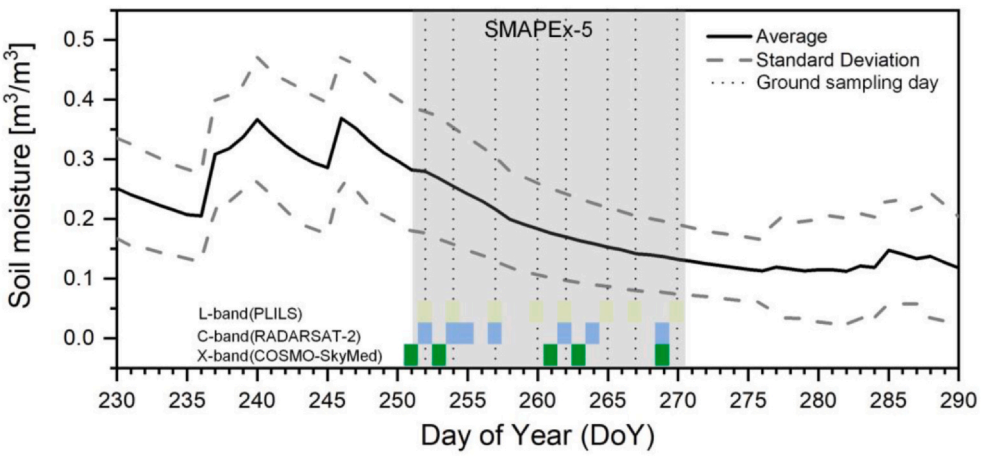

Fig. 4. Time series of the top 5 cm soil moisture from all OzNet sites over the Yanco area in 2015 and the acquisition dates of L-band PLIS, C-band RADARSAT-2 and X-band COSMO-SkyMed during the SMAPEx-5 period. The average and standard deviation of soil moisture of each sampling date was used to generate the time series soil moisture in the synthetic data set.

grid size, being then resampled to a uniform grid of ~25 m. Please refer to Zhu et al. (2019a) for more details.

3.2. Synthetic data set

A synthetic radar data set (Table 1) over bare soil surfaces with various potential uncertainties was generated using the Improved IEM (IIEM, Fung et al., 2002) for a comprehensive evaluation. Although the IIEM is not perfect, especially in simulating the cross-polarization backscatter, this has not degraded the synthetic study because the IIEM was also used in the inversion. Three noise sources were considered, representing various uncertainties encountered in the soil moisture retrieval:

i) Speckle and thermal noise. Radar data has inherent speckle noise because of the coherent nature of a radar system. While the speckle noise can be partly removed by multi-look, the remaining noise can still have a significant effect on soil moisture retrieval over a satisfactory spatial resolution of < 100 m. Following the simulation of Sentinel-1 and SMAP radar data by Balenzano et al. (2011) and Kim et al. (2012), the speckle noise was described as Gaussian with a zero mean and a standard deviation of 0.7 dB in this study. This assumption is equivalent to the residual speckle noise after a multi-look operation of ~39 looks. For Sentinel-1 Interferometric Wide Swath data (5 × 20 m), the equivalent spacing of the multi-looked data is ~60 m.

ii) Imbalance among different polarizations. Both radar measurements and forward models can have an imbalance among different polarizations. The measurements are commonly well calibrated with a negligible bias, e.g. 0.3 dB in intensity for Sentinel-1 (Schmidt et al., 2018), 0.05 dB for PALSAR (Shimada et al., 2009), and 0.18–0.34 dB for several airborne radar systems (Zhu et al., 2018). However, the biases of current forward models among different polarizations can be on the order of 1 dB (Choker et al., 2017). Accordingly, a maximum imbalance of 1 dB was used by applying a systematic bias of +0.5 dB, −0.5 dB and 0 dB to the HH, HV and VV simulations, respectively.

iii) Imbalance among different frequencies and incidence angles. This results from the data calibration and failure of models to capture the frequency and angular effects. Since a detailed knowledge of this source is not available, the biases of the IEM measured in a laboratory experiment (Mancini et al., 1999) were used here. A systematic bias of −5 dB, −1 dB, −1.5 dB and − 2 dB was applied to C-band 23°, C-band 35°, L-band 23° and L-band 35° simulations, respectively.

Three synthetic data sets, A, B and C, were generated with different uncertainty sources (Table 1). For each subset, 200 simulations of eight-day time series soil moisture on the eight SMAPEx-5 ground sampling days were randomly generated from a normal distribution N(a_i, b_i). The a_i and b_i refers to the average value and standard deviation of OzNet observations on the ith date (Fig. 3). The soil surface root mean square height s of each simulation was randomly generated from a uniform distribution U(0.5, 3) ranging from 0.5 to 3 cm, while the correlation length l was generated from U(5, 35). Notably, only one s and l were generated for each simulation (a time series of 8 days), denoting that roughness was constant over time. An exponential correlation function was used in all simulations for simplicity.

Time series speckle-free backscattering coefficients (HH, HV, and VV) were then calculated using the IEM and the simulated soil moisture and roughness. Two frequencies (1.26 GHz and 5.4 GHz) and two incidence angles (23° and 35°) were considered, being consistent with the uncertainty source iii above. These radar configurations were considered to cover the most commonly used radar data for soil moisture retrieval provided by current L- and C-band satellites. The uncertainty sources were finally added to the speckle-free backscattering coefficients (Table 1). Taking the C-band HH polarization at 23° as an example, apart from the speckle noise, a systematic bias of −4.5 dB (+0.5 dB for HH and − 5 dB for C-band 23°) was used in data set C. Accordingly, in each data set there are 200 time series of backscattering

Table 1
Summary of the inputs and radar configurations for generating the synthetic radar data set. The autocorrelation function required in the IIEM was fixed as an exponential function. An eight-day time series of soil moisture was randomly generated according to the statistics in Fig. 3, being a simulation of the eight ground sampling dates of SMAPEx-5. U(a, b) refers to a uniform distribution ranging from a to b, while N(a, b) is a normal distribution with the average and standard deviation being a and b respectively. The s and l are the soil surface root mean square roughness height and roughness correlation length, respectively.

Data set	s [cm]	l [cm]	Frequency [GHz]	Incidence angle [°]	Uncertainty source [dB]	Number of simulations
A	U(0.5, 3)	U(5, 35)	1.26 and 5.4	23 and 35	Speckle N(0, 0.7)	200
B	U(0.5, 3)	U(5, 35)	1.26 and 5.4	23 and 35	Uncertainty source of set A + polarization imbalance (0.5 for HH, −0.5 for HV)	200
C	U(0.5, 3)	U(5, 35)	1.26 and 5.4	23 and 35	Uncertainty source of set B + frequency and incidence angle imbalance (−5, −1, −1.5 and − 2 dB at C-band 23°, C-band 35°, L-band 23° and L-band 35°)	200

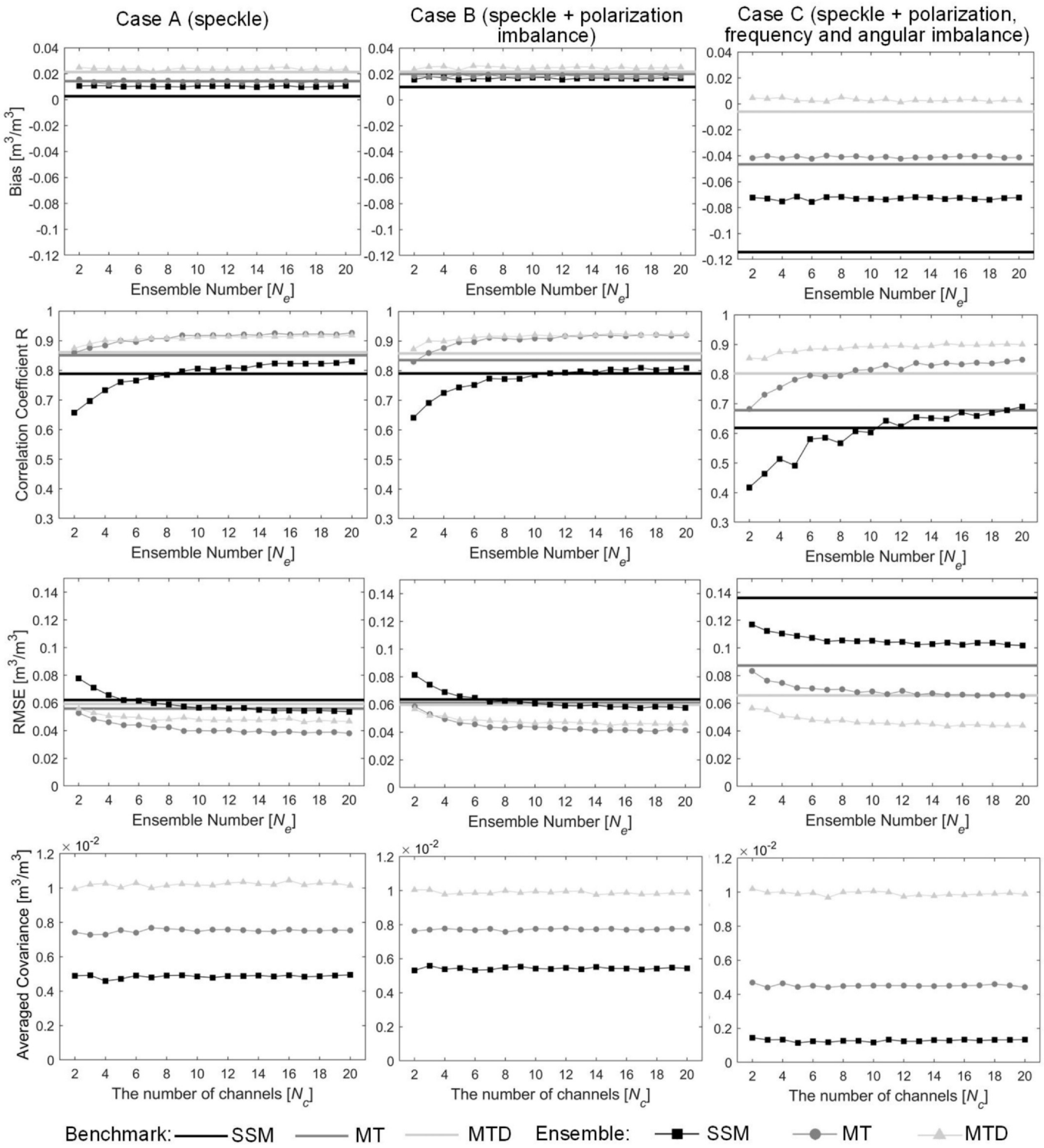

Fig. 5. The effect of ensemble number (N_e) on retrieval accuracy of data set A (left), B (middle) and C (right). The number of channels (N_c) is 3 for all methods. The temporal sampling size (N_t) is 8 for both MT and MTD. The SSM, MT, and MTD refer to the snapshot method, multi-temporal method, and multi-temporal method with a dry-down constraint. Notably, the line representing the R of the SSM benchmark on set A (middle left) is overlapped by that of MTD.

coefficients having 8 dates, with each date having 12 channels (3 polarization × 2 frequency × 2 incidence angle).

4. Results

4.1. Evaluation using the synthetic data sets

A comprehensive evaluation of the ensemble methods was first carried out using the synthetic data set. Soil moisture was retrieved using three methods individually (i.e. SSM, MT and MTD as defined

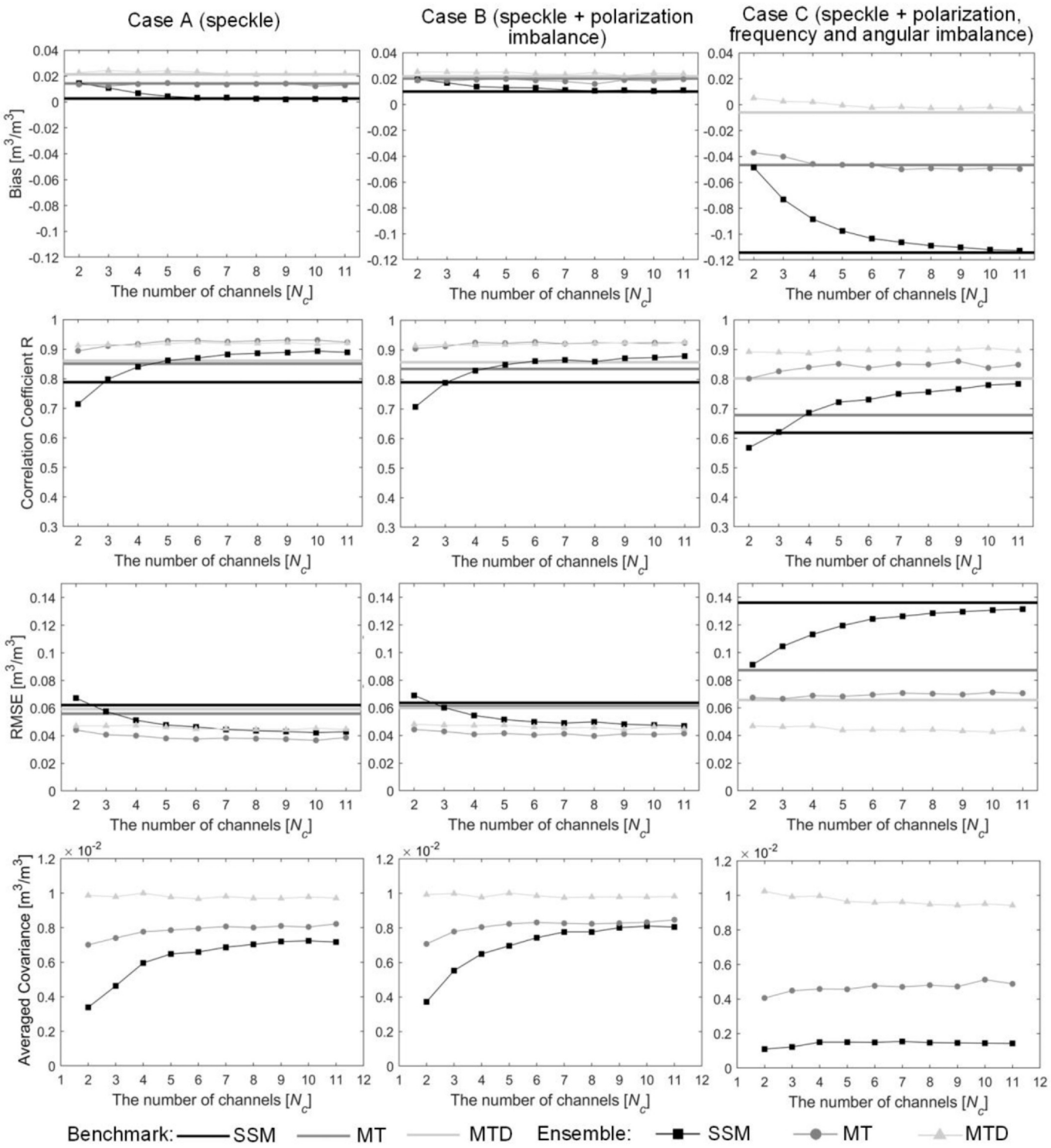

Fig. 6. The effect of number of channels N_c on retrieval accuracy. The ensemble number (N_e) is 10 for all methods. The temporal sampling size (N_t) is 8 for both MT and MTD.

previously) and their ensemble versions, and then compared with the truth (simulated soil moisture). The ensemble number (N_e) and the number of channels (N_c) are required in the snapshot ensembles, with an additional parameter of temporal sampling size (N_t) for multi-temporal ensembles. The effect and the selection of those parameters were investigated as follows.

4.1.1. The effect of ensemble number (N_e)

The effect of N_e was analyzed first with a fixed N_c of 3 for all methods and a fixed N_t of 8 for multi-temporal methods. The results of the three data sets are shown in Fig. 5. In the benchmark retrievals, the two multi-temporal benchmarks achieved a slightly smaller RMSE on data sets A and B than the SSM benchmark, with the difference being less than 0.006 m^3/m^3. A relatively large difference of 0.07 in R was observed between two multi-temporal benchmarks and the SSM benchmark. This can result from the reduced sensitivity of multi-temporal methods to random noise, given data sets A and B were mainly affected by random speckle noise. Small biases (< 0.02 m^3/m^3) were

observed for all three benchmark retrievals on data sets A and B, in line with the zero bias of data set A and the small systematic bias (+0.5 dB for HH and − 0.5 dB for HV) of data set B. In contrast, large negative retrieval biases were observed for the SSM benchmark (-0.114 m^3/m^3) and MD benchmark (-0.048 m^3/m^3) on data set C because of the large systematic biases among different frequency and incidence angles (Table 1). The MTD benchmark performed much better than the MT benchmark in all three metrics on set C, confirming that additional constraints (e.g., the dry down constraint) are still required for complex multi-temporal retrieval scenarios.

The RMSE of the SSM ensemble was reduced as N_e increased on the three data sets, being nearly constant for $N_e > 10$. A relatively large N_e of up to 15 was able to improve the R on all data sets. In contrast, the bias of the SSM ensemble and the averaged covariance of the SSM sub-retrievals changed little for different N_e. The different responses of bias and averaged covariance to N_e can be explained by Eq. 3. The ensemble bias and the averaged covariance.

only relate to the expectation of bias and covariance for individual sub-retrieval's, being independent on N_e. The RMSE on the other hand reduced gradually as the effect of sub-retrievals' total variance ($\overline{\text{var}}/N_e + (N_e - 1)\overline{\text{covar}}/N_e$) reduced gradually for an increased N_e. A similar relationship between N_e and performance was observed for the MT and MTD ensemble but having reduced benefit from using a large N_e compared to the SSM ensemble.

The SSM ensemble outperformed the benchmark for $N_e > 6$ on data sets A and B in terms of RMSE, while it achieved a lower RMSE than the benchmark on data set C in all cases of N_e. An increasing N_e of 9, 10 and 11 was required for the SSM ensemble to outperform the benchmark R on set A, B and C, respectively. The MT ensemble achieved lower RMSE than the benchmark on all three data sets. However, the required N_e for the MT ensemble to achieve a higher R than the benchmark increased from 2 on data set A to 9 on data set C. The MTD outperformed the benchmark on all data sets in terms of RMSE and R.

The different behaviors of these three methods regarding N_e can be related to data noise and the sensitivity of sub-retrievals to the noise. The SSM sub-retrievals are most sensitive to noise, followed by the MT and MTD sub-retrievals (Zhu et al., 2019b). This means that the result of each SSM sub-retrieval can be more unstable than that of multi-temporal methods. Consequently, the SSM sub-retrievals had the largest averaged variance ($\overline{\text{var}}$) and the smallest averaged covariance ($\overline{\text{covar}}$, the fourth row of Fig. 5), while the MTD sub-retrievals had the smallest averaged variance and the largest averaged covariance. According to Eq. 3, a large N_e can reduce the contribution of averaged variance ($\overline{\text{var}}/N_e$) but increase the effect of averaged covariance () on MSE and thus RMSE and R. As a result, the SSM ensemble can benefit most using a large N_e, followed by MT and MTD. The data noise level played a similar role of the algorithm's sensitivity to noise. For example, the SSM and MT sub-retrievals had lower averaged covariance on data set C than data set A because of the larger uncertainty of data set C. As a result, the SSM and MT ensemble required a larger N_e to have higher R than the benchmark on set C. Notably, for data set C that was mainly affected by the systematic bias, the RMSE of ensemble retrieval was dominated by the averaged bias of sub-retrievals independent of N_e. Although the optimal N_e related to the noise, an N_e of 10 is suggested considering the limited changes using a larger N_e and the fact that the required computation cost of an ensemble retrieval is N_e times that of the benchmark and each sub-retrieval.

4.1.2. The effect of channel number (N_c)

Fig. 6 shows the effect of N_c on ensemble retrieval accuracy. The performance of SSM ensemble was improved as N_c increased on data set A and B, being superior to the benchmark for $N_c > 2$. The major improvement (0.15 in R, 0.02 m^3/m^3 in RMSE and 0.01 in bias) was made by increasing N_c from 2 to 6, with little and even negative contribution of using a larger N_c. For the MT ensemble, it achieved better results than the benchmark in all cases of N_c on data set A and B. However, negligible improvement (< 0.03 in R, < 0.006 m^3/m^3 in RMSE and 0 in bias) was observed as N_c increased from 2 to 11. Similarly, the MTD ensemble was found to be independent on N_c and outperform the benchmark in all cases on data set A and B.

On data set C, the R of the SSM ensemble and MT ensemble increased as N_c increased, being consistent with the results on data sets A and B. However, a contrary trend was observed for RMSE, which increased from 0.091 to 0.131 m^3/m^3 for the SSM ensemble and from 0.067 to.

0.071 m^3/m^3 for the MT ensemble. This can be explained by the increased bias of sub-retrievals using a larger N_c. Using more input channels with large negative biases led to larger underestimation in SSM and MT sub-retrievals (first row of Fig. 6). In contrast, the MTD sub-retrievals were less sensitive to the varying biases among different channels and thus the RMSE of the MTD ensemble changed only a little for different N_c on data set C.

The effect of N_c on ensemble retrieval is more complex than that of N_e because it relates to all three terms of Eq. 3. For data sets A and B that were mainly affected by random noise, the SSM sub-retrievals were more reliable using more channels (a large N_c). Specially, the averaged bias of SSM sub-retrievals had the largest decrease when increased N_c from 2 to 4 because three unknows (mv, l and s) need to be determined in individual SSM sub-retrievals. Conversely, a larger N_c means that individual sub-retrievals are more likely to share the same channels and thus results in a higher covariance (the last row of Fig. 5), reducing the benefit of the ensemble. For the two multi-temporal methods, the number of input channels ($N_c \times 8$ days here) can be much larger than that of unknowns (2 roughness +8 soil moisture values) as surface roughness was assumed to be time-invariant. Consequently, the MT and MTD sub-retrievals already had good performance for a small N_c (e.g. 2) and so a large N_c can only slightly change the averaged variance and slightly reduce the covariance.

For data set C, more input channels led to large biases in individual sub-retrievals (the first row of Fig. 5). The RMSE of the SSM ensemble and MT ensemble was dominated by the bias of sub-retrievals. However, the R still increased with more channels, being independent of the systematic bias. It is clear that the "optimal" N_c varies with different noise level, noise type and the soil moisture retrieval methods. A small N_c of 1 or 2 is suggested for the two multi-temporal methods regardless of the noise source and type. A relatively large N_c of 4–6 is suggested to have a competitive R and for the potential systematic bias to be removed in post-processing.

4.1.3. The effect of temporal sampling size (N_t)

The effect of N_t on the MTD ensemble was investigated in Fig. 7, with the N_e and N_c being 10 and 3, respectively. Notably, N_t is subject to the length of available time series data that meet the assumption of time-invariant roughness. Since the length of simulated time series was 8 days with constant roughness, N_t was varied from 2 to 7 to ensure a multi-temporal retrieval with random sampling in time. The $N_t = 8$ was also included to show the retrieval without temporal sampling.

The RMSE showed a "smiling" curve on data set A and B, with the smallest RMSE (~0.036 m^3/m^3) achieved at N_t = 4–6. This is partly related to the relatively larger bias observed at $N_t = 1$ (~0.015 m^3/m^3) and 8 (~0.028 m^3/m^3). The R of the MTD ensemble increased first from ~0.875 at $N_t = 1$ to ~0.935 at $N_t = 5$ and then decreased to ~0.918 at $N_t = 8$ on data set A and B. Individual MTD sub-retrievals using a longer time series are more robust, as the temporal evolution of longer moisture series can better constrain the soil moisture retrieval (Zhu et al., 2019b). Accordingly, the MTD ensemble based on larger N_t can have better results in R. An increasing N_t, on the other hand, means increased similarity of data used in sub-retrievals. Unfortunately, MTD only retrieved soil moisture on the selected dates which are different in various sub-retrievals. The averaged covariance of sub-retrievals thus cannot be calculated. Although only a limited improvement in RMSE (~

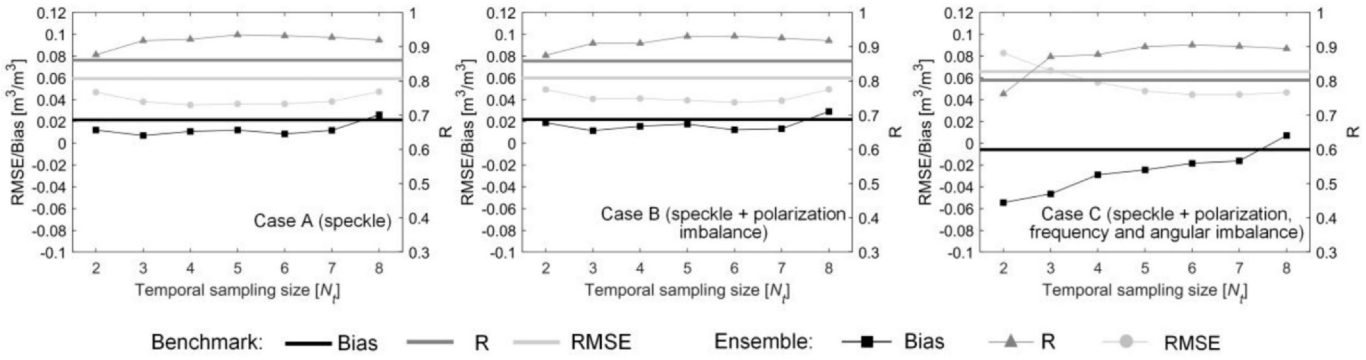

Fig. 7. The effect of temporal sampling size (N_t) on the ensemble performance for the multi-temporal algorithm with a dry-down constraint (MTD).

0.012 m^3/m^3) and R (0.016) was observed for using an $N_t < 8$, the effectiveness of using temporal sampling in the multi-temporal retrieval was confirmed on data sets A and B. A different relationship between N_t and retrieval accuracy was found on data set C because of the large and varying bias among different channels. Longer time series are therefore required to obtain more reliable sub-retrieval in view of RMSE and bias. However, the highest R (~0.90) was achieved using $N_t = 6$ and 7, confirming the effectiveness of temporal sampling (using $N_t < 8$ here) on data set C.

4.2. Evaluation using the SMAPEx-5 dataset

The MTD ensemble was also evaluated using the SMAPEx-5 dataset, with the retrieval results being compared with that of the MTD benchmark. Notably, this subsection is focused on comparison of the MTD benchmark and MTD ensemble, with the detailed performance of MTD benchmark that is already presented in Zhu et al. (2019b). Similar to the synthetic study, the MTD benchmark used different combinations of all available data, as shown in Table 2. Following the synthetic studies, the parameters of N_c and N_e were 1 and 10 respectively. N_t was set to 80% of the total time instances, whose values are listed in Table 2. For example, $N_t = 6$ for the retrieval using L-band only (8 acquisitions collected on different dates).

Table 2 lists the performance of MTD and its ensemble version using various radar data. For each algorithm, a total of 5 retrievals were made at the paddock scale (~0.1–0.5 km) using different data configurations: L-band only as well as two- and three-frequency combinations, simulating the single and multi-radar retrievals. In general, the ensemble version of MTD outperformed the MTD in all cases, except the retrieval over wheat using C- and X-band data. The improvement in RMSE and R ranged from 0.004 to 0.014 m^3/m^3 and 0.01 to 0.16 respectively, confirming the effectiveness of the ensemble strategy.

Importantly, the MTD ensemble seemed to be more suitable for soil moisture retrieval from data collected by multiple satellites than the MTD benchmark. Specifically, the L-band-only retrieval had the best results in the MTD benchmark, with the retrieved results being slightly deteriorated by combining the C-band and/or X-band data. This was ascribed to the failure of modeling uncertainties in multi-frequency data (Zhu et al., 2019a). Conversely, the MTD ensemble seemed to be less sensitive to such uncertainties. Dual-frequency retrieval (L + C and L + X) had similar results to that using L-band only. Retrieval using L-, C- and X-band data achieved the best results over bare and wheat. The positive effect of the additional surface information contained in the multi-frequency data overwhelmed the negative effect of the accompanied uncertainty.

Fig. 8 shows the retrieved soil moisture at the paddock scale of both algorithms using L-, C- and X-band. The main improvement of using the ensemble retrieval here was to have fewer 'outliers' outside the ± 0.06 m^3/m^3 target. For instance, the MTD had many more points in the red circles than its ensemble version. The reason for this difference is that multiple sub-retrievals in the MTD ensemble are less likely to have consistent large overestimations or underestimations using different input data, and thus result in a more reliable retrieval.

5. Discussion

5.1. Conditions for effective ensemble retrievals

The principle behind the ensemble machine learning was introduced to provide a simple but effective way of utilizing the increasingly available radar data in soil moisture retrieval. Similar to the conditions of successful ensemble learning in various applications (e.g., Lessmann et al., 2019; Minku et al., 2009; Zhang et al., 2016a; Zhu et al., 2016), ensemble retrievals require robust and diverse sub-retrievals as demonstrated in the synthetic study. While it is straightforward to post-check the robustness and diversity of sub-retrievals using

Table 2
Comparison of MTD and its ensemble version using the SMAPEx-5 data set. The bold values are the best results in each landcover type, while those with underline are cases where MTD performed better than its ensemble version. The N_t/N_d denotes N_t dates were randomly selected from the N_d dates with data.

	Data	Bare soil			Grass			Wheat			N_t/N_d
		RMSE	R	Bias	RMSE	R	Bias	RMSE	R	Bias	
Benchmark	L	0.061	0.74	−0.012	0.047	0.91	−0.014	0.058	0.84	−0.004	−/8
(MTD)	L + C	0.061	0.73	−0.004	0.050	0.89	−0.022	0.059	0.82	−0.003	−/11
(Zhu et al., 2019a)	L + X	0.062	0.73	−0.007	0.055	0.87	0.006	0.058	0.83	0.006	−/13
	C + X	0.067	0.61	−0.009	0.056	0.87	−0.026	0.071	0.80	−0.014	−/12
	L + C + X	0.062	0.75	−0.007	0.054	0.87	−0.019	0.058	0.83	0.006	−/15
Ensemble MTD	L	0.057	0.77	0.010	0.040	0.92	−0.005	0.054	0.84	0.001	6/8
	L + C	0.057	0.79	0.016	**0.038**	**0.92**	**0.003**	0.055	0.83	0.001	9/11
	L + X	0.057	0.75	0.009	0.042	0.90	−0.005	0.052	0.85	0.006	10/13
	C + X	0.054	0.77	−0.014	0.046	0.93	−0.035	0.072	0.84	−0.057	10/12
	L + C + X	**0.053**	**0.80**	**0.001**	0.040	0.92	−0.013	**0.052**	**0.86**	**−0.005**	12/15

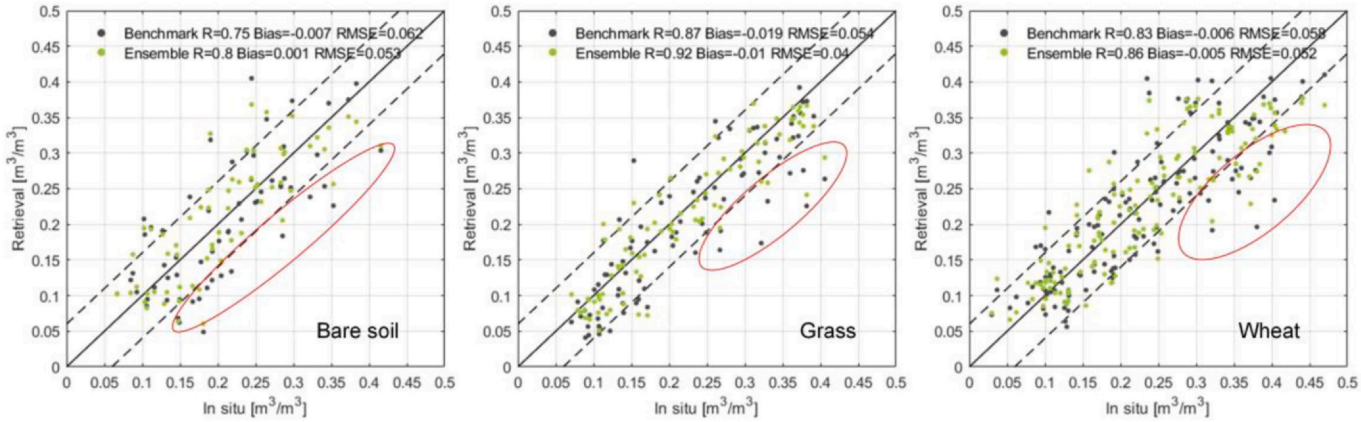

Fig. 8. In situ versus retrieved soil moisture at the paddock scale from MTD (benchmark) and MTD ensemble using time series L-, C- and X-band data. The solid line is the 1:1 relationship while the dashed lines denote the ± 0.06 m^3/m^3 target accuracy. The red circles are examples where MTD ensemble has fewer outliers than MTD. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

bias, variance, and covariance, it is hard to select channels for the optimal ensemble before soil moisture retrieval because of the varying complexity of retrieval scenarios. Generally, more channels are suggested for sub-retrievals when i) the input data and the forward model have large uncertainties; and ii) the base algorithm used is sensitive to noise. Fig. 6 suggests a minimum of 2 channels in each sub-retrieval for effective snapshot ensembles on data set A and B. This is not easy to be met even with the forthcoming "big radar" data, because many more channels are required to generate multiple (e.g., 10) different sub-retrievals. Assumptions, such as a time-invariant surface condition in a short period (Zhang et al., 2016b) are thus required to have sufficient independent observations for the target date. In contrast, multi-temporal algorithms are less sensitive to data uncertainties and can have a longer time series in soil moisture retrieval (i.e., more channels) for areas with limited roughness and vegetation changes, thus enabling numerous robust sub-retrievals.

In this study, the difference or diversity of sub-retrievals was generated using random combinations of input channels and/or time instances, being similar to the widely used ensemble learning paradigm of Bagging (Breiman, 1996). The difference of sub-retrievals can also result from 1) using different retrieval algorithms, e.g. randomly taking the MTD and MT as the base algorithm; and 2) using different forward models, e.g. the random use of the Oh model and IEM over bare soil. Those options are expected to reduce the uncertainty of individual models, surface parameterizations in model calibration or inversion algorithms. For example, the IEM predictions have been shown to have a bias of 0.3–4.2 dB at X-band HH pol (Choker et al., 2017) as a result of a roughness parameterization failure and the over-simplified model assumptions, resulting in an overestimated sub-retrieval. This can be partly corrected by another sub-retrieval using the Oh model (Oh et al., 2002) which has a bias of −1 dB bias for the same channel (Choker et al., 2017). The joint use of those options is also possible, with their counterparts in ensemble learning well documented (Zhang and Ma, 2012). Moreover, strategies used in ensemble forecasting (Li et al., 2017) can also lend support to the ensemble soil moisture retrieval.

5.2. Relationship between the ensemble and benchmark single retrieval

The proposed ensemble framework is compatible with most existing soil moisture retrieval algorithms. However, the two random selection strategies of input data (Fig. 2) cannot be used directly. Especially, time series data at the same polarization, incidence angle and frequency should be used in order to have temporal difference/ratio in several multi-temporal methods (e.g., Balenzano et al., 2011; Ouellette et al., 2017) and thus different sub-retrievals can only be generated by selecting N_t time instances from the full time series.

Since an ensemble retrieval consists of multiple sub-retrievals using existing soil moisture retrieval algorithms, it inherits the same assumptions of the retrieval algorithms used. For the snapshot multi-frequency retrieval (Eq. 1), all the input data should be collected on the same date with negligible changes in surface conditions. All sub-retrievals had the same task to derive the average or effective soil moisture of the target date. However, the different penetration depths of different radar configurations (frequency, incidence angle, polarization) would make the sensed soil moisture vary among different sub-retrievals. This challenge is from the multi-frequency retrieval approach itself and is also true for the benchmark retrieval using all multi-frequency data. A commonly used assumption is that the soil moisture is uniform for the top 5 cm and that all available data at L-, C- and X-band have the same penetration depth. For multi-temporal retrieval, the ensemble version also shares the same assumptions made for traditional multi-temporal retrievals, including time-invariant roughness and vegetation, and uniform soil moisture in top 5 cm if multi-configuration data was used. Accordingly, one should be careful to use different retrieval algorithms in the same ensemble retrieval.

An important difference between the ensemble retrieval and the single retrieval is the technical choice of solving the complex retrieval from irregular multi-frequency data collected by multiple satellites. As aforementioned, soil moisture retrieval from multi-frequency data is still challenging, commonly requiring a careful combination of the multi-frequency data (Ulaby et al., 2014), and needing reliable forward models to cover the available radar configurations (Zhu et al., 2019a). Although those challenges also need to be addressed in the ensemble retrieval, it is not necessary to have an "optimal" solution to the whole retrieval process because each sub-retrieval only requires a moderate accuracy. Accordingly, ensemble retrieval can significantly reduce the complexity of algorithm development, at the expense of slightly increased computational cost in running the multiple sub-retrievals. Fortunately, it is straightforward to parallelise the ensemble retrieval because all sub-retrievals are independent without requiring any communication. Moreover, the requirement of sub-retrieval with moderate accuracy also means a reduced sensitivity to data uncertainties introduced across the whole preprocessing chain.

6. Conclusion

Stochastic ensemble methods based on existing soil moisture retrieval algorithms were presented to retrieve soil moisture from radar data collected by multiple satellites, being a flexible option for soil moisture mapping with the coming era of big SAR data. The proposed ensemble methods were comprehensively evaluated using a synthetic data set for bare soil and the SMAPEx-5 data set with three landcover

types. Results confirmed that ensemble retrieval outperformed the benchmark retrievals in most cases. The improvement in RMSE on the real data set was up to 0.015 m^3/m^3, with its potential on other data sets confirmed in the synthetic study. A global ensemble size of 10 was found to be sufficient to provide satisfactory results for all retrieval methods. The principle that multiple moderate soil moisture retrieval can be merged into an improved retrieval is expected to provide a simple but effective alternative for multi-frequency soil moisture retrieval.

Declaration of Competing Interest

None.

Acknowledgments

The SMAPEx-5 field campaign was supported by an Australian Research Council Discovery Project (DP140100572).

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, 324–344.
Baez-Villanueva, O.M., Zambrano-Bigiarini, M., Beck, H.E., McNamara, I., Ribbe, L., Nauditt, A., Birkel, C., Verbist, K., Giraldo-Osorio, J.D., Thinh, N.X., 2020. RF-MEP: a novel random Forest method for merging gridded precipitation products and ground-based measurements. Remote Sens. Environ. 239, 111606.
Baghdadi, N., Holah, N., Zribi, M., 2006. Soil moisture estimation using multi-incidence and multi-polarization ASAR data. Int. J. Remote Sens. 27, 1907–1920.
Baghdadi, N., El Hajj, M., Dubois-Fernandez, P., Zribi, M., Belaud, G., Cheviron, B., 2014. Signal level comparison between TerraSAR-X and COSMO-SkyMed SAR sensors. IEEE Geosci. Remote Sens. Lett. 12, 448–452.
Baghdadi, N., Choker, M., Zribi, M., Hajj, M., Paloscia, S., Verhoest, N., Lievens, H., Baup, F., Mattia, F., 2016. A new empirical model for radar scattering from bare soil surfaces. Remote Sens. 8, 920.
Balenzano, A., Mattia, F., Satalino, G., Davidson, M.W., 2011. Dense temporal series of C- and L-band SAR data for soil moisture retrieval over agricultural crops. Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing 4, 439–450.
Bates, J.M., Granger, C.W., 1969. The combination of forecasts. J. Oper. Res. Soc. 20, 451–468.
Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31.
Bindlish, R., Barros, A.P., 2000. Multifrequency soil moisture inversion from SAR measurements with the use of IEM. Remote Sens. Environ. 71, 67–88.
Bindlish, R., Barros, A.P., 2001. Parameterization of vegetation backscatter in radar-based, soil moisture estimation. Remote Sens. Environ. 76, 130–137.
Breiman, L., 1996. Bagging predictors. Mach. Learn. 24, 123–140.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Choker, M., Baghdadi, N., Zribi, M., El Hajj, M., Paloscia, S., Verhoest, N.E., Lievens, H., Mattia, F., 2017. Evaluation of the Oh, Dubois and IEM backscatter models using a large dataset of SAR data and experimental soil measurements. Water 9, 38.
Cookmartin, G., Saich, P., Quegan, S., Cordey, R., Burgess-Allen, P., Sowter, A., 2000. Modeling microwave interactions with crops and comparison with ERS-2 SAR observations. IEEE Trans. Geosci. Remote Sens. 38, 658–670.
Fung, A., Liu, W., Chen, K., Tsay, M., 2002. An improved IEM model for bistatic scattering from rough surfaces. Journal of Electromagnetic Waves and Applications 16, 689–702.
Gorrab, A., Zribi, M., Baghdadi, N., Mougenot, B., Chabaane, Z.L., 2015. Potential of X-band TerraSAR-X and COSMO-SkyMed SAR data for the assessment of physical soil parameters. Remote Sens. 7, 747–766.
Hansen, L.K., Salamon, P., 1990. Neural network ensembles. IEEE Transactions on Pattern Analysis & Machine Intelligence 993–1001.
Healey, S.P., Cohen, W.B., Yang, Z., Brewer, C.K., Brooks, E.B., Gorelick, N., Hernandez, A.J., Huang, C., Hughes, M.J., Kennedy, R.E., 2018. Mapping forest change using stacked generalization: an ensemble approach. Remote Sens. Environ. 204, 717–728.
Huang, H., Tsang, L., Njoku, E.G., Colliander, A., Liao, T.-H., Ding, K.-H., 2017. Propagation and scattering by a layer of randomly distributed dielectric cylinders using Monte Carlo simulations of 3D Maxwell equations with applications in microwave interactions with vegetation. IEEE Access 5, 11985–12003.
Huang, H., Liu, C., Wang, X., Zhou, X., Gong, P., 2019. Integration of multi-resource remotely sensed data and allometric models for forest aboveground biomass estimation in China. Remote Sens. Environ. 221, 225–234.
Hutengs, C., Vohland, M., 2016. Downscaling land surface temperatures at regional scales with random forest regression. Remote Sens. Environ. 178, 127–141.
Kim, S.-B., Tsang, L., Johnson, J.T., Huang, S., Van Zyl, J.J., Njoku, E.G., 2012. Soil moisture retrieval using time-series radar observations over bare surfaces. IEEE Trans. Geosci. Remote Sens. 50, 1853–1863.
Kim, S.-B., Moghaddam, M., Tsang, L., Burgin, M., Xu, X., Njoku, E.G., 2014. Models of L-band radar backscattering coefficients over global terrain for soil moisture retrieval. IEEE Trans. Geosci. Remote Sens. 52, 1381–1396.
Kornelsen, K.C., Coulibaly, P., 2013. Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications. J. Hydrol. 476, 460–489.
Lessmann, S., Haupt, J., Coussement, K., De Bock, K.W., 2019. Targeting customers for profit: an ensemble learning framework to support marketing decision-making. Inf. Sci.
Li, W., Duan, Q., Miao, C., Ye, A., Gong, W., Di, Z., 2017. A review on statistical post-processing methods for hydrometeorological ensemble forecasting. Wiley Interdiscip. Rev. Water 4, e1246.
Lievens, H., Verhoest, N., De Keyser, E., Vernieuwe, H., Matgen, P., Álvarez-Mozos, J., De Baets, B., 2011. Effective roughness modelling as a tool for soil moisture retrieval from C-and L-band SAR. Hydrol. Earth Syst. Sci. Discuss. 7, 4995–5031.
Mancini, M., Hoeben, R., Troch, P.A., 1999. Multifrequency radar observations of bare surface soil moisture content: a laboratory experiment. Water Resour. Res. 35, 1827–1838.
Mattia, F., Satalino, G., Dente, L., Pasquariello, G., 2006. Using a priori information to improve soil moisture retrieval from ENVISAT ASAR AP data in semiarid regions. IEEE Trans. Geosci. Remote Sens. 44, 900–912.
Minku, L.L., White, A.P., Yao, X., 2009. The impact of diversity on online ensemble learning in the presence of concept drift. IEEE Trans. Knowl. Data Eng. 22, 730–742.
Notarnicola, C., Angiulli, M., Posa, F., 2008. Soil moisture retrieval from remotely sensed data: neural network approach versus Bayesian method. IEEE Trans. Geosci. Remote Sens. 46, 547–557.
Oh, Y., Sarabandi, K., Ulaby, F.T., 2002. Semi-empirical model of the ensemble-averaged differential Mueller matrix for microwave backscattering from bare soil surfaces. IEEE Trans. Geosci. Remote Sens. 40, 1348–1355.
Ouellette, J.D., Johnson, J.T., Balenzano, A., Mattia, F., Satalino, G., Kim, S.-B., Dunbar, R.S., Colliander, A., Cosh, M.H., Caldwell, T.G., 2017. A time-series approach to estimating soil moisture from vegetated surfaces using L-band radar backscatter. IEEE Trans. Geosci. Remote Sens. 55, 3186–3193.
Panciera, R., Tanase, M.A., Lowell, K., Walker, J.P., 2014. Evaluation of IEM, Dubois, and Oh radar backscatter models using airborne L-band SAR. IEEE Trans. Geosci. Remote Sens. 52, 4966–4979.
Pettinato, S., Santi, E., Paloscia, S., Pampaloni, P., Fontanelli, G., 2013. The intercomparison of X-band SAR images from COSMO-SkyMed and TerraSAR-X satellites: case studies. Remote Sens. 5, 2928–2942.
Pierdicca, N., Castracane, P., Pulvirenti, L., 2008. Inversion of electromagnetic models for bare soil parameter estimation from multifrequency polarimetric SAR data. Sensors 8, 8181–8200.
Pierdicca, N., Pulvirenti, L., Bignami, C., 2010. Soil moisture estimation over vegetated terrains using multitemporal remote sensing data. Remote Sens. Environ. 114, 440–448.
Quets, J., De Lannoy, G.J., Al Yaari, A., Chan, S., Cosh, M.H., Gruber, A., Reichle, R.H., Van der Schalie, R., Wigneron, J.-P., 2019. Uncertainty in soil moisture retrievals: an ensemble approach using SMOS L-band microwave data. Remote Sens. Environ. 229, 133–147.
Schmidt, K., Ramon, N.T., Schwerdt, M., 2018. Radiometric accuracy and stability of sentinel-1A determined using point targets. Int. J. Microw. Wirel. Technol. 10, 538–546.
Shimada, M., Isoguchi, O., Tadono, T., Isono, K., 2009. PALSAR radiometric and geometric calibration. IEEE Trans. Geosci. Remote Sens. 47, 3915–3932.
Stojanova, D., Panov, P., Gjorgjioski, V., Kobler, A., Džeroski, S., 2010. Estimating vegetation height and canopy cover from remotely sensed data with machine learning. Ecological Informatics 5, 256–266.
Ueda, N., Nakano, R., 1996. Generalization error of ensemble estimators. In: Proceedings of International Conference on Neural Networks (ICNN'96). IEEE, pp. 90–95.
Ulaby, F.T., Long, D.G., Blackwell, W.J., Elachi, C., Fung, A.K., Ruf, C., Sarabandi, K., Zebker, H.A., Van Zyl, J., 2014. Microwave Radar and Radiometric Remote Sensing. The University of Michigan Press, Ann Arbor.
Walker, J.P., Houser, P.R., 2004. Requirements of a global near-surface soil moisture satellite mission: accuracy, repeat time, and spatial resolution. Adv. Water Resour. 27, 785–801.
Ye, N., Walker, J.P., Wu, X., Jeu, R.D., Gao, Y., Jackson, T.J., Jonard, F., Kim, E., Merlin, O., Pauwels, V., Renzullo, L.J., Rüdiger, C., Sabaghy, S., Hebel, V., Yueh, S.H., Zhu, L., 2020. The Soil Moisture Active Passive Experiments: Validation of the SMAP Products in Australia. (Ieee Transactions on Geoscience and Remote Sensing).
Zhang, C., Ma, Y., 2012. Ensemble Machine Learning: Methods and Applications. Springer.
Zhang, W., Zou, H., Luo, L., Liu, Q., Wu, W., Xiao, W., 2016a. Predicting potential side effects of drugs by recommender methods and ensemble learning. Neurocomputing 173, 979–987.
Zhang, X., Chen, B., Fan, H., Huang, J., Zhao, H., 2016b. The potential use of multi-band SAR data for soil moisture retrieval over bare agricultural areas: Hebei, China. Remote Sens. 8, 7.
Zhang, C., Denka, S., Cooper, H., Mishra, D.R., 2018a. Quantification of sawgrass marsh aboveground biomass in the coastal Everglades using object-based ensemble analysis and Landsat data. Remote Sens. Environ. 204, 366–379.
Zhang, X., Chen, B., Zhao, H., Li, T., Chen, Q., 2018b. Physical-based soil moisture retrieval method over bare agricultural areas by means of multi-sensor SAR data. Int. J. Remote Sens. 39, 3870–3890.
Zhu, L., Xiao, P., Feng, X., Zhang, X., Huang, Y., Li, C., 2016. A co-training, mutual learning approach towards mapping snow cover from multi-temporal high-spatial resolution satellite imagery. ISPRS J. Photogramm. Remote Sens. 122, 179–191.
Zhu, L., Walker, J.P., Ye, N., Rüdiger, C., Hacker, J., Panciera, R., Tanase, M.A., Wu, X., Gray, D., Stacy, N., Goh, A., Yardley, H., Mead, J., 2018. 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, 4513–4525.
Zhu, L., Walker, J.P., Tsang, L., Huang, H., Ye, N., Rüdiger, C., 2019a. A multi-frequency framework for soil moisture retrieval from time series radar data. Remote Sens. Environ. 235, 111433.
Zhu, L., Walker, J.P., Tsang, L., Huang, H., Ye, N., Rüdiger, C., 2019b. Soil moisture retrieval from time series multi-angular radar data using a dry down constraint. Remote Sens. Environ. 231, 111237.
Zhu, L., Walker, J.P., Ye, N., Rüdiger, C., 2019c. Roughness and vegetation change detection: a pre-processing for soil moisture retrieval from multi-temporal SAR imagery. Remote Sens. Environ. 225, 93–106.
Zribi, M., Baghdadi, N., Holah, N., Fafin, O., 2005. New methodology for soil surface moisture estimation and its application to ENVISAT-ASAR multi-incidence data inversion. Remote Sens. Environ. 96, 485–496.