A Multivariable Approach for Estimating Soil Moisture from Microwave Radiation Imager (MWRI)

Sibo ZHANG1*, Fuzhong WENG2, and Wei YAO3

1 Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology, Beijing 100094
2 State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, China Meteorological Administration, Beijing 100081

(Received December 2, 2019; in final form March 6, 2020)

ABSTRACT

Accurate measurements of soil moisture are beneficial to our understanding of hydrological processes in the earth system. A multivariable approach using the random forest (RF) machine learning technique is proposed to estimate the soil moisture from Microwave Radiation Imager (MWRI) onboard Fengyun-3C satellite. In this study, Soil Moisture Operational Products System (SMOPS) products disseminated from NOAA are used as a truth to train the algorithm with the input of MWRI brightness temperatures (TBs) at 10.65, 18.7, 23.8, 36.5, and 89.0 GHz, TB polarization ratios (PRs) at 10.65, 18.7, and 23.8 GHz, height in digital elevation model (DEM), and soil porosity. The retrieved soil moisture is also validated against the independent SMOPS data, and the correlation coefficient is about 0.8 and mean bias is 0.002 m³ m⁻³ over the period from 1 August 2017 to 31 May 2019. Our retrieval of soil moisture also has a higher correlation with ECMWF ERA5 soil moisture data than the MWRI operational products. In the western part of China, the spatial distribution of MWRI soil moisture is much improved, compared to the MWRI operational products.

Key words: soil moisture, Microwave Radiation Imager (MWRI), machine learning, microwave remote sensing

Citation: Zhang, S. B., F. Z. Weng, and W. Yao, 2020: A multivariable approach for estimating soil moisture from Microwave Radiation Imager (MWRI). J. Meteor. Res., 34(4), 732–747, doi: 10.1007/s13351-020-9203-x.

1. Introduction

Soil moisture controls the evapotranspiration (ET) process of land surfaces and plays an irreplaceable role in the interactions between water, energy, and carbon cycles over land (Albergel et al., 2012; Cao et al., 2019; Green et al., 2019; Rudd et al., 2019; Seager et al., 2019; Zhang et al., 2019). In the natural environment, soil moisture varies greatly with soil properties (i.e., porosity, texture, density, and structure), surface roughness, topography, land cover, land temperature, rainfall, and ET. Therefore, spatial–temporal heterogeneity is the major characteristic of soil moisture. Previous studies have shown the impact of soil moisture on atmospheric variables at various scales (Xu et al., 2012; Zhang et al., 2013; Parrens et al., 2014; Ruosteenoja et al., 2018; Pangaluru et al., 2019). Today, in addition to the point measurements of soil moisture from in situ probes, microwave remote sensing is widely used to determine large-scale surface soil moisture, typically in the low frequencies from 1 to 10 GHz (Wagner et al., 2007). In microwave remote sensing, both active and passive techniques have been deployed for making the soil moisture measurements. For example, Fengyun-3C (FY-3C) Microwave Radiation Imager (MWRI) measures the thermal emission from the surface, commonly expressed as brightness temperatures (TBs). These TBs are used to estimate the surface parameters such as soil temperature, soil moisture, and surface emissivity (Schmugge and Jackson, 1994).
Many studies have demonstrated that the L-band technology (frequency $f = 1–2$ GHz, wavelength $\lambda = 30–15$ cm) can be a better choice for soil moisture remote sensing since its radiation can better penetrate through the vegetation canopy than other shorter microwave wavelengths at X-/C-band and also the contribution from the atmosphere is negligible (Wigneron et al., 2003). Soil Moisture and Ocean Salinity (SMOS) is the first satellite mission dedicated to making global observations of soil moisture. It was launched on 2 November 2009. The SMOS satellite carries an L-band radiometer to derive the global soil moisture every three days with a designed accuracy of 0.04 m$^3$ m$^{-3}$ at a spatial resolution of about 50 km (Kerr et al., 2001; Kerr, 2007). Launched in January 2015, Soil Moisture Active Passive (SMAP) mission utilizes L-band radar and radiometer instruments together to monitor surface soil moisture and distinguish frozen from thawed soils.

However, global soil moisture products can also be derived by using higher microwave frequencies (Wagner et al., 2007). For instance, MWRI sensors are carried on three Fengyun satellites (FY-3B/-3C/-3D), and their X-band (10.65 GHz) observations are routinely used to retrieve global surface volumetric soil moisture (VSM) product. Since the launch of the FY-3B satellite in November 2010, MWRI data have been available for a variety of applications. Operational applications and climate change researches can benefit from the VSM products derived from this long-term, continuous, stable X-band dataset (Wang et al., 2010; Yang et al., 2011; Yang J. et al., 2012; Yang Z. D. et al., 2012). However, compared with other satellite VSM products, there is a relatively large error (Parimussa et al., 2014) in the released MWRI VSM product (L2) by the National Satellite Meteorological Center (NSMC) of China Meteorological Administration (CMA).

In general, the observations at X-band are not ideal for soil moisture retrieval, e.g., the estimated accuracy is about 0.06 cm$^3$ cm$^{-3}$ for the Advanced Microwave Scanning Radiometer (AMSR)-E/-X-band observations (Du et al., 2017), because of the reduced penetration in vegetation and the increased effects of various noises from surface thermodynamic temperature, snow cover, topography, surface roughness, soil properties, and the intervening atmosphere (Choudhury et al., 1982; Jackson and Schmugge, 1991; Njoku and Entekhabi, 1996; Wigneron et al., 2003). Therefore, an improvement to VSM retrieval from MWRI is highly designed in accounting for the multivariable nonlinear relationship. The multivariable effects on surface VSM retrieval may be better addressed through a machine learning algorithm such as the random forest (RF). Unlike a traditional linear regression fit technique, the RF method allows many input variables or parameters used for nonlinear regression and can estimate the contributions on VSM estimating from each input parameter.

In this study, in addition to the X-band observations, multifrequency, multipolarization measurements from MWRI, vegetation leaf area index (LAI), surface temperature, topography [height in digital elevation model (DEM)] data, and statistical soil porosity map are considered as the input parameters of the RF training model. The effects from each input parameter are evaluated by the mean square errors (MSEs) and the node impurity derived from the RF model. Hence, the major factors of VSM retrieval for building the RF training model could be determined. The scores of the VSM derived from different RF models with various input parameters are evaluated and the final retrieval accuracy is derived through comparing with various soil moisture products including the in situ data.

2. Datasets

2.1 MWRI products

MWRI L1 products from FY-3C are collected for nearly two years from 1 August 2017 to 31 May 2019 for estimating VSM. Since the vegetation penetration is limited for passive X-band remote sensing, we focused on the MWRI data that have a good dynamic range for measuring the soil moisture. The observations are typically associated with the seven surface types used in this study. The measurements from other surfaces such as water, forest, closed shrublands, permanent wetlands, urban, and snow/ice surfaces are excluded in our training. The International Geosphere–Biosphere Programme (IGBP) data acquired in MWRI L1 dataset are used for surface classification.

MWRI measures TBs at five frequencies (10.65, 18.7, 23.8, 36.5, and 89.0 GHz) with both horizontal and vertical polarizations (TB$^{hv}$ and TB$^{vh}$) at a fixed viewing angle (53.4°). The radio frequency interference (RFI) effects from artificial sources on X-band (10.65 GHz) observations are removed by using the spectral difference method (Li et al., 2004). In addition to the multichannel TBs, the corresponding polarization indices and DEM data in the MWRI L1 products are also used as the RF model input.

Currently, the daily operational product of MWRI L2 VSM is derived by using an algorithm based on the parameterized surface emission model provided by Shi et al. (2006), and the vegetation effects are corrected by using
ancillary vegetation data, e.g., normalized difference vegetation index (NDVI) and vegetation water content (VWC) data. MWRI L2 VSM product is used as a baseline in this study to evaluate the improvement of the retrieval accuracy by using a multivariable method based on the RF training model. Both MWRI L1 and L2 products are resampled to the same 0.25-degree grids for intercomparison using the nearest-neighbor search algorithm. MWRI L1 products are missing for seven days from 25 to 31 March 2018, while MWRI L2 products have more missing data on the dates 24 March and 17 September 2018.

2.2 ECMWF land surface data

In order to analyze the effects on MWRI VSM retrieval from various land surface parameters, we use the low vegetation LAI, dewpoint temperature and air temperature at a height of 2 m, skin temperature, and soil temperature for the top 7 cm derived from ECMWF ERA5 reanalysis dataset at 0000 UTC as the input parameters of the RF training model. Moreover, from the statistical global porosity map (0.25-degree grids) provided by European Space Agency Climate Change Initiative (ESA-CCI) project is also contained in the RF model input to represent the variabilities of soil texture for a global scale. On the other hand, the VSM products for the top 7 cm from ECMWF ERA5 are used to validate the estimated VSM by using the RF model.

2.3 Soil Moisture Operational Products System (SMOPS) and other satellite VSM products

The National Environmental Satellite, Data, and Information Service (NESDIS) SMOPS daily blended satellite VSM products for the top 5 cm are used as the learning objectives in the RF training model. SMOPS merges almost all available VSM products derived from L-, C-, X-band, including from SMAP, SMOS, Global Precipitation Measurement (GPM), Global Change Observation Mission–Water 1 (GCOM-W1), and Meteorological Operational (MetOp)-A/-B satellite missions. SMOPS has much better spatial and temporal coverages than sparse in situ VSM observations for a global scale, and it has been verified by using the in situ VSM observations from various networks (Liu et al., 2016). SMOPS is a very valuable data source for training the MWRI machine learning VSM algorithm. Thus, daily SMOPS VSM products are also collected for validation over the same time period of the MWRI operational products that we used in this study.

2.4 In situ soil moisture observations

The daily mean in situ VSM observations at 5-cm depth from 12 stations in Natural Resources Conservation Service–Soil Climate Analysis Network (NRCS–SCAN) over the American continent (Schaefer and Paetzold, 2000) are used to validate the estimated VSM using RF models from 1 August 2017 to 31 May 2019.

3. Methodology

3.1 Measurement physics

The TB\(^p\) measured by a spaceborne microwave radiometer is given by Schmugge and Jackson (1994):

\[
TB^p = TB_{\text{atm\_up}} + I_{\text{atm}}\left(r_{\text{surface}} \cdot TB_{\text{atm\_down}} + TB_{\text{surface}}\right),
\]

where the superscript \(p\) denotes polarization (V or H); \(TB_{\text{atm\_up}}\) is the upwelling TB from the atmosphere; \(I_{\text{atm}}\) is the atmospheric transmittance from surface to the top of atmosphere, which is related to the atmospheric opacity along the viewing path; \(TB_{\text{atm\_down}}\) represents the downwelling TB from the atmosphere, which is partly reflected depending on the surface reflectivity \(r_{\text{surface}}\); \(TB_{\text{surface}}\) is the TB from the surface emission. At the MWRI 10.65-GHz frequency, the atmospheric contributions to the measured \(TB^p\) at clear-sky conditions are often negligible, i.e., \(TB^p \approx TB_{\text{surface}}\) but at higher frequencies, and for precipitating clouds, the effects of atmospheric emission and scattering are required to separate \(TB_{\text{surface}}\) from the observed \(TB^p\).

For a homogenous vegetated surface, the vegetation layer attenuates the radiation emitted from soil and emits the radiation at its own emission. In general,

\[
TB_{\text{surface}}^p = e_{\text{soil}}^p \cdot I_{\text{veg}}^p \cdot TB_{\text{soil}}^p + \left(1 - \omega_{\text{veg}}^p\right)\left(1 - I_{\text{veg}}^p\right)\left[1 + r_{\text{soil}}^p \cdot I_{\text{veg}}^p\right] TB_{\text{veg}}^p,
\]

where the first term is the soil emission attenuated by vegetation layer: \(e_{\text{soil}}^p\) is the soil emissivity, \(I_{\text{veg}}^p\) is the vegetation transmissivity related to the vegetation optical thickness, and \(TB_{\text{soil}}^p\) is the soil TB; the second term represents the sum of the direct vegetation emission and the reflected vegetation emission by soil surface: \(\omega_{\text{veg}}^p\) is the single scattering albedo of the vegetation surface, \(r_{\text{soil}}^p\) is the soil reflectivity, and \(TB_{\text{veg}}^p\) is the vegetation temperature. Approximately, \(TB_{\text{soil}}^p\) and \(TB_{\text{veg}}^p\) are assumed equal, and \(e_{\text{soil}}^p = 1 - r_{\text{soil}}^p\).

Equation (2) has been used for low vegetation surfaces at frequencies up to X-band, in which vegetation transmissivity can be approximated as:

\[
I_{\text{veg}}^p = \exp\left(-b^p \cdot \frac{VWC}{\cos\theta}\right),
\]

where VWC is the vegetation water content (kg m\(^{-2}\)), \(\theta\) is the incident angle, and coefficient \(b^p\) depends on the fre-
frequency, vegetation type, VWC, polarization, $\theta$, crop phonology, and especially the vertical structure of the canopies (Ulaby and Wilson, 1985; Jackson and Schmugge, 1991; Wigneron et al., 2003).

In addition to the complexity of the vegetation effects, the soil reflectivity also varies with many factors. For a smooth soil, $\frac{\rho_p}{\rho_{\text{soil}}}$ can be calculated from the soil permittivity related to soil moisture and the incidence angle ($\theta$) by using Fresnel equations. But in general, the effects of surface roughness and soil texture should also be taken into account (Wang and Choudhury, 1981; Wang, 1983). It is difficult to build a physical parameterized model to cover all the possible situations over the global lands. For example, although the roughness and vegetation effects are considered in the parametric model used for deriving MWRI L2 VSM products, the retrieval results are still suboptimal for many conditions. This may be due to using only X-band MWRI observations. Hence, the corresponding factors related to vegetation, surface temperature, roughness, and soil properties are all required in the RF model input to deal with such a complex multivariable issue for estimating VSM from MWRI data.

3.2 RF training model input

Because multifrequency and multipolarization TB measurements have different sensitivities to vegetation and roughness than soil moisture, it is possible to correct the vegetation and roughness effects on soil moisture retrieval by combining various frequency and polarization TB observations together (Njoku and Entekhabi, 1996). There is a large polarization difference (PD = $TB_V - TB_H$) from bare soil emission ($TB_H < TB_V$), but the difference is reduced with increasing vegetation biomass and/or vegetation cover density. This property has been commonly used in the vegetation correction of the soil moisture retrieval from passive microwave radiometers (Owe et al., 2001; Paloscia et al., 2001). The most common index is the TB polarization ratio (PR), defined as:

$$PR = \frac{(TB_V - TB_H)}{(TB_V + TB_H)}.$$  (4)

Since the PR index is less sensitive to the effects of surface temperature than the PD index, it is adopted in this study. PR is less affected by soil moisture and more affected by vegetation and roughness with increasing frequency. This is the physical basis for a multivariable approach by combining multifrequency TBs and PRs as the input of the RF training model to distinguish soil moisture contribution from that of surface temperature and vegetation. Since the DEM data are provided in the MWRI L1 dataset and the statistical soil porosity map is available, we can simply use them as the input parameters to represent the changes in topography and soil texture for a global scale. Moreover, the VSM retrieval derived from MWRI is very robust and fairly independent by using TBs, PRs, DEM, and soil porosity as the RF training model input.

On the other hand, since the proposed multivariable method (RF model) in this study is expected to assess the vegetation and surface temperature effects on soil moisture retrieval, LAI and soil/air/skin temperature reanalysis from the ECMWF ERA5 are added to the model input together with the independent input parameters to perform dependent experiments. The VSM estimates derived from the dependent experiments are compared with those from the independent experiments.

3.3 RF training model experiments

First, a representative training dataset is needed to build the RF training model for VSM estimating. Here, MWRI observations from four random days (31 January, 30 April, 31 July, and 31 October 2018) are used to represent the seasonal differences. In addition, to account for the spatial heterogeneity of vegetation covers and soil properties at a global scale, we also select the training samples from MWRI 0.25-degree grid global data to cover the spatial variations. Totally, 642,408 samples are used as our training dataset.

Four RF models are trained through the R package “RandomForest” (Liaw and Wiener, 2002), using the same training dataset and the same number of decision trees (480) but various model inputs. EXP1 is a dependent experiment containing all available 22 parameters: 17 independent parameters from MWRI (multifrequency TBs, PRs, and DEM) and the statistical soil porosity map, and 5 dependent parameters from ECMWF ERA5 reanalysis (low vegetation LAI, dewpoint temperature and air temperature at a height of 2 m, skin temperature, and soil temperature for the top 7 cm). Different from EXP1, EXP2 is an independent experiment with only the 17 independent input parameters in EXP1. From our analysis (see Section 4.1), we remove the less important input parameters from MWRI (PRs from 36.5 and 89.0 GHz) in EXP1 and EXP2 to perform EXP3 and EXP4, respectively. In EXP3, we only use LAI and soil temperature from ECMWF to represent the variations in vegetation and surface temperature, excluding the dew-point, air, and skin temperatures from ECMWF in EXP1 input. Finally, EXP3 is a dependent experiment with 17 input parameters, while EXP4 is an independent one with 15 input parameters. The detailed flowchart for the optimal RF model decision is shown in Fig. 1.
4. Results

4.1 Performance of the MWRI L2 VSM products

Since the performance of the MWRI L2 VSM products compared with other satellite VSM products is rarely presented in the past, an intuitive comparison is presented first in Fig. 2 between MWRI L2 and four daily gridded VSM products derived from SMAP, SMOS, AMSR2 on GCOM-W1, and GPM Microwave Imager (GMI), respectively, for the focused land covers from 1 August 2017 to 31 May 2019. The number of daily available VSM estimates from each satellite is different but with a similar trend across time.

The daily coefficient of determination ($R^2$) scores between MWRI L2 and SMAP are much higher than the scores by comparing MWRI L2 with the other satellite products. This is positive to use MWRI observations to estimate VSM. But it is noticed that the range of the $R^2$ scores between MWRI L2 and SMAP is large from 0.1 to 0.6, and the average value of the $R^2$ scores is 0.44 over the whole time period. Moreover, the average values of the $R^2$ scores are only 0.25, 0.09, and 0.09 for MWRI L2 compared with SMOS, GMI, and AMSR2, respectively. On the other hand, the unbiased root mean square difference (ubRMSD) scores are relatively large. During the whole time period, the smallest average value of ubRMSD is 0.10 m$^3$ m$^{-3}$ from the comparison between MWRI L2 VSM and SMAP, while the average ubRMSD values are all around 0.13 m$^3$ m$^{-3}$ for MWRI L2 VSM compared with the other three satellite VSM products. Therefore, we need to increase the stability of the MWRI VSM retrieval and reduce the values of the ubRMSD scores.

4.2 RF training dataset and models

The training dataset used in this work is simply selected from four random days to represent the four seasons in a year, but it includes the variations in global soil texture and land cover using a high-density spatial sampling (0.25 degree). Figure 3 presents the statistical distributions of three variables in the training dataset, including land cover type, SMOPS product (learning objective), and TB observation from the 10.65 GHz horizontal channel (TB10H) typically used for VSM retrieval [e.g., National Snow and Ice Data Center (NSIDC) AMSR-E/-2 VSM product]. In Fig. 3, 91% of the VSM values range between 0.10 and 0.35 m$^3$ m$^{-3}$, and 89% of the TB10H values range from 220 to 280 K; while the proportion of extreme values is small in both SMOPS and TB10H. It should be noted that the variation of the land cover types significantly affects the VSM estimation. The maximum
percentage (30%) is from open shrublands. Hence, the retrieval accuracy from open shrublands contributes a lot to the global average value of the retrieval accuracy. The proportions from woody savannas, grasslands, croplands, and barrens are similar, from 13% to 16%, while the percentages of savannas and croplands/natural vegetation mosaic lands are small (9% and 4%). Although the statistical distributions in Fig. 3 are only from 4-day observations, it approximately represents the annual distributions for a global scale.

Based on the above training dataset, EXP1 and EXP2 models are trained by using the RF method with different input parameters. The contribution of each input in each model is evaluated by two factors: the percentage for increasing the MSE and the node impurity. The higher percentage for increasing MSE or the larger node impurity value (meaning impurity), the larger contribution of the input parameter on VSM retrieval. Although DEM and soil porosity are ancillary data, not observed by MWRI, both of them are very important for increasing MSE and node impurity shown in Fig. 4. In EXP1 with all available input parameters, soil porosity is the most important parameter for increasing node impurity and the second important parameter for increasing MSE, while DEM is the top one for increasing MSE and the sixth for increasing node impurity. The following are vegetation and temperature from ECMWF and low frequency TBs and PRs from MWRI. It is similar in EXP2, without the contributions from ECMWF reanalysis data, DEM rises to the third one for increasing node impurity. From both EXP1 and EXP2 models, we consider that PRs from 36.5 and 89.0 GHz are the two least important input parameters in increasing both MSE and node impurity and can be removed from the input parameters. This is similar to that only low frequency PRs (from 6.9, 10.6, and 18.9 GHz of AMSR-E) are used in Njoku et al. (2003). Here, PR at 18.7 GHz is less important than that at 23.8 GHz in both EXP1 and EXP2 models.

4.3 Soil moisture retrieval

Figure 5 displays an example of the estimated VSM from the independent EXP4 using the least input para-
meters on 15 March 2018, which is compared with SMOPS and MWRI L2 VSM products. The retrievals from other experiments are similar and not shown here. Clearly, the VSM retrievals from EXP4 are closer to
SMOPS products than MWRI L2 products. Due to the MWRI data gap and the limited land covers for estimating MWRI VSM, the number of the available EXP4 estimates is smaller than that of SMOPS, but the number of the public MWRI L2 VSM products is even less, especially over the Tibetan Plateau and Northern Hemisphere high latitudes. The spatial distribution of MWRI soil moisture estimates is much improved in EXP4.

For an intercomparison, SMOPS and ECMWF ERA5 are used as the benchmark, respectively, for the whole time period from 1 August 2017 to 31 May 2019. Daily global scores of the MWRI VSM retrievals from EXP1, EXP2, EXP3, EXP4, and MWRI L2 are calculated only using the data where the VSM estimates from all data sources are available. Hence, the data size is the same for the four experiments and MWRI L2 product. Because of the missing data in SMOPS, the amount of SMOPS data is slightly smaller than that of ECMWF ERA5 data. We can see the dynamic change of the available data size in Fig. 6. The amount of data is much larger in summer and smaller in winter. Due to the soil freezing and snow/ice cover problems in MWRI L2 VSM products, missing data are much more common in winter.

From the daily mean bias scores between various MWRI VSM and SMOPS in Fig. 7, all of the four experiments are close and around zero, while for MWRI L2 the daily mean bias ranges between $-0.025$ and $0.025$ m$^3$ m$^{-3}$. Since SMOPS is the learning objective in the RF models, the bias between SMOPS and the model derived value is very small. But note that only 4-day SMOPS data are used in the training data, and most of the model estimates (656 of 660 days) are independent. This means that our multivariable method can produce MWRI VSM estimates with a really small bias related to SMOPS. However, based on the independent data source, ECMWF ERA5, the mean bias values of our four experiments are all below zero, displaying a negative bias, as low as $-0.025$ m$^3$ m$^{-3}$. The mean bias scores of the MWRI L2

![Fig. 5. Surface VSM distributions from (a) SMOPS products, (b) EXP4 VSM estimates, and (c) MWRI L2 VSM products on 15 March 2018.](image)

![Fig. 6. Time series of the number of daily global VSM estimates used in this study, from SMOPS (red line) and ECMWF ERA5 (green line) from 1 August 2017 to 31 May 2019.](image)
products are also worse, with a larger variation from −0.050 to 0.025 m$^3$ m$^{-3}$.

In Fig. 8, the VSM retrievals from all four experiments generally present a good agreement with SMOPS: $R^2$ is not lower than 0.55, ubRMSD does not exceed 0.05 m$^3$ m$^{-3}$, and p-values are all smaller than 0.01. A significant improvement is found in the retrievals derived from our proposed multivariable method, compared with the MWRI L2 VSM products ($R^2$ below 0.45 and ubRMSD around 0.11 m$^3$ m$^{-3}$). Among the four experiments, with the help of the ancillary LAI and temperature data, $R^2$ scores in EXP1 and EXP3 are relatively higher than those in EXP2 and EXP4: the average value of $R^2$ scores is 0.66 (0.65) in EXP1 (EXP3) and 0.63 in both EXP2 and EXP4. It presents the enhancement in estimating MWRI VSM with vegetation and temperature information. Moreover, the difference between EXP2 and EXP4 is neglected and thus this proves the 15 input parameters in EXP4 enough for training the independent RF model. For dependent experiments, EXP1 is slightly better than EXP3. There is no doubt that the best $R^2$ and ubRMSD scores are from the four random days (31 January, 30 April, 31 July, and 31 October 2018) used for constructing the training dataset.

We also compare the MWRI retrievals in this study and the MWRI L2 products with the independent ECMWF ERA5 reanalysis. An improvement for estimating MWRI VSM using our method is also found in $R^2$ scores but not significant in ubRMSD scores. The average value of $R^2$ scores during the whole time period from 1 August 2017 to 31 May 2019 is 0.68 (0.64) in EXP1 (EXP3) and 0.62 in both EXP2 and EXP4, while for MWRI L2, it is only 0.42. Moreover, the difference between EXP1 and EXP2/EXP4 is obvious. But note that the ancillary data used in EXP1 are also from ECMWF ERA5, which are not independent and might increase the effects of vegetation and temperature on VSM retrieval. On the other hand, the average value of ubRMSD scores is around 0.11 m$^3$ m$^{-3}$ for all MWRI experiments and around 0.12 m$^3$ m$^{-3}$ for the MWRI L2 products, which is very close and much larger than that based on SMOPS. Figure 8 shows a slight improvement in MWRI experiments in spring and summer.

The detail scores for various land covers are shown in Tables 1 and 2 using SMOPS and ECMWF ERA5 as the benchmark, respectively. The largest percentage of land covers is from open shrublands (IGBP 7). Moreover, all scores ($R^2$, bias, and ubRMSD) are good in open shrublands (Table 1). Besides, $R^2$ scores in grasslands (IGBP 10) and ubRMSD scores in barrens (IGBP 16) are the best in all land covers. While the lowest $R^2$ values and the largest bias values are both from croplands/natural vegetation mosaic lands (IGBP 14), and the largest ubRMSD values are from woody savannas (IGBP 8). But
note the number of the daily estimates in croplands/natural vegetation mosaic lands (IGBP 14) is relatively small due to the limited global coverage. In Table 2, taking ECMWF ERA5 as the benchmark, $R^2$ and bias scores in woody savannas (IGBP 8), croplands/natural vegetation mosaic lands (IGBP 14), and barrens (IGBP 16) are relatively low in all land covers, but ubRMSD scores in barrens (IGBP 16) are the best (around 0.05 m$^3$ m$^{-3}$) and much better than others (from 0.08 to 0.11 m$^3$ m$^{-3}$ in our experiments). From the above results, we conclude the retrieval accuracies for croplands/natural vegetation mosaic lands and woody savannas are the worst, and the re-

Fig. 8. Daily scores of (a, b) $R^2$ and (c, d) the ubRMSD of the soil moisture estimates from MWRI EXP1 (red line), EXP2 (blue line), EXP3 (orange line), EXP4 (black line), and L2 (green line) based on the validations from (a, c) SMOPS and (b, d) ECMWF ERA5, respectively, during the time period from 1 August 2017 to 31 May 2019.
ults for the savannas and croplands are in the middle, while those for barrens, grasslands, and open shrublands are relatively better for estimating MWRI VSM. Since the scores between MWRI L2 VSM and other satellites VSM are shown in Fig. 2, we also present the scores between EXP4 VSM (the independent estimates) and the same benchmarks in Fig. 9. The average ubRMSD score between MWRI VSM and AMSR2 is significantly improved from 0.13 (L2) to 0.07 m$^3$ m$^{-3}$ (EXP4), while for comparing with SMAP, GMI, and SMOS, the average ubRMSD values are also reduced to 0.08, 0.10, and 0.11 m$^3$ m$^{-3}$, respectively. An improvement is also found in the $R^2$ scores. The highest one (0.51) is from the comparison between EXP4 and SMAP.

### Table 1
Comparisons of VSM between MWRI estimates (from EXP1, EXP2, EXP3, EXP4, and L2) and SMOPS products at a depth of 5 cm from 1 August 2017 to 31 May 2019. Mean $R^2$, bias, and ubRMSD scores are given for each land cover type. IGBP numbers 7, 8, 9, 10, 12, 14, and 16 are related with open shrublands, woody savannas, savannas, grasslands, croplands, croplands/natural vegetation mosaic lands, and barrens, respectively; $N$ is the number of the samples.

| IGBP type | $N$ | $R^2$ between SMOPS and EXP1 | Bias (m$^3$ m$^{-3}$) between SMOPS and EXP1 | ubRMSD (m$^3$ m$^{-3}$) between SMOPS and EXP1 |
|-----------|-----|-----------------------------|---------------------------------------------|-----------------------------------------------|
| All       | 72645 | 0.66 0.63 0.65 0.63 0.31 | 0.0013 0.0021 0.0005 0.0020 0.0032 | 0.0423 0.0442 0.0429 0.0440 0.1079 |
| IGBP 7    | 14849 | 0.57 0.53 0.55 0.53 0.26 | 0.0008 0.0019 0.0014 0.0018 0.0250 | 0.0386 0.0403 0.0392 0.0402 0.0801 |
| IGBP 8    | 10876 | 0.51 0.45 0.48 0.46 0.10 | 0.0012 0.0033 0.0019 0.0030 0.0521 | 0.0478 0.0505 0.0490 0.0501 0.1283 |
| IGBP 9    | 9689  | 0.53 0.46 0.50 0.47 0.21 | 0.0014 0.0017 0.0014 0.0016 0.0410 | 0.0452 0.0480 0.0465 0.0477 0.1173 |
| IGBP 10   | 10616 | 0.58 0.56 0.57 0.56 0.26 | 0.0030 0.0025 0.0027 0.0023 0.0125 | 0.0425 0.0437 0.0430 0.0435 0.1013 |
| IGBP 12   | 10425 | 0.49 0.46 0.49 0.46 0.09 | 0.0042 0.0047 0.0046 0.0046 0.0139 | 0.0465 0.0478 0.0466 0.0476 0.1155 |
| IGBP 14   | 2776  | 0.43 0.39 0.42 0.39 0.07 | 0.0085 0.0086 0.0070 0.0084 0.0478 | 0.0456 0.0472 0.0461 0.0469 0.1240 |
| IGBP 16   | 12406 | 0.48 0.42 0.48 0.42 0.10 | 0.0025 0.0036 0.0021 0.0035 0.0419 | 0.0305 0.0326 0.0306 0.0326 0.0472 |

### Table 2
As in Table 1, but for comparisons of VSM between MWRI estimates and the top 7-cm VSM reanalysis from ECMWF ERA5.

| IGBP type | $N$ | $R^2$ between ECMWF and EXP1 | Bias (m$^3$ m$^{-3}$) between ECMWF and EXP1 | ubRMSD (m$^3$ m$^{-3}$) between ECMWF and EXP1 |
|-----------|-----|-----------------------------|---------------------------------------------|-----------------------------------------------|
| All       | 74028 | 0.68 0.62 0.64 0.62 0.42 | 0.0151 0.0143 0.0149 0.0144 0.0130 | 0.1111 0.1137 0.1128 0.1138 0.1206 |
| IGBP 7    | 15073 | 0.56 0.50 0.53 0.49 0.35 | 0.0313 0.0324 0.0320 0.0323 0.0057 | 0.0952 0.0974 0.0967 0.0976 0.1041 |
| IGBP 8    | 11040 | 0.39 0.30 0.30 0.29 0.15 | 0.0850 0.0829 0.0843 0.0831 0.0338 | 0.1025 0.1068 0.1064 0.1069 0.1420 |
| IGBP 9    | 9864  | 0.55 0.46 0.48 0.45 0.31 | 0.0649 0.0646 0.0649 0.0648 0.0250 | 0.1041 0.1081 0.1070 0.1081 0.1247 |
| IGBP 10   | 10753 | 0.53 0.47 0.49 0.46 0.31 | 0.0320 0.0326 0.0324 0.0328 0.0475 | 0.0960 0.0984 0.0976 0.0986 0.1147 |
| IGBP 12   | 10550 | 0.41 0.36 0.37 0.36 0.11 | 0.0528 0.0524 0.0524 0.0524 0.0423 | 0.0815 0.0839 0.0832 0.0840 0.1274 |
| IGBP 14   | 2823  | 0.28 0.23 0.22 0.21 0.06 | 0.0081 0.0082 0.0079 0.0075 0.0392 | 0.0875 0.0906 0.0900 0.0906 0.1404 |
| IGBP 16   | 12895 | 0.41 0.34 0.38 0.34 0.26 | 0.0942 0.0954 0.0938 0.0952 0.0500 | 0.0505 0.0527 0.0514 0.0529 0.0555 |

### Fig. 9
Daily scores of (a) $R^2$ and (b) ubRMSD from EXP4 VSM compared with AMSR2 VSM (green line), GMI VSM (blue line), SMOS VSM (red line), and SMAP VSM (black line), respectively, during the time period from 1 August 2017 to 31 May 2019.
the second one (0.29) is from the comparison with SMOS, while for comparing with AMSR2 and GMI, $R^2$ scores are 0.17 and 0.14.

In situ VSM observations (daily mean values) at 5 cm from the NRCS–SCAN are used to further evaluate the performance of the EXP4 VSM estimates derived by our RF model and the improvement compared with the MWRI L2 VSM products. Since the VSM estimates are made at a 0.25-degree grid resolution, not all in situ stations are suitable for validating. The VSM observations from 12 stations from 1 August 2017 to 31 May 2019 are selected to assess the correlations with the SMOPS products, MWRI L2 VSM products, and EXP4 estimates, respectively. The scores from the 12 stations are shown in Table 3. The correlation coefficient ($R$) scores between scaled EXP4 estimates and scaled in situ observations range from 0.4 to 0.7, which are significantly better than the $R$ scores (from −0.2 to 0.5) of scaled MWRI L2 VSM products, and slightly worse than the $R$ scores (from 0.5 to 0.8) of scaled SMOPS products. The mean values of the ubRMSD scores from the 12 stations are 0.87, 0.98, and 1.35 m$^3$ m$^{-3}$ for scaled SMOPS products, scaled EXP4 estimates, and scaled MWRI L2 VSM products, respectively. Similar to $R$ scores, the performance of ubRMSD scores of EXP4 estimates is much better than that of MWRI L2 products but slightly worse than that of SMOPS products.

Scaled EXP4 VSM estimates from three typical stations in Table 3 are compared in Fig. 10 with three scaled VSM time series: in situ VSM observations at a depth of 5 cm, MWRI L2 VSM products, and SMOPS products. EXP4 estimates are significantly closer to the in situ observations than the MWRI L2 VSM products. Moreover, the available MWRI L2 VSM products ($N$ values in Table 3) are obviously less than the EXP4 estimates. Most of the missing data in MWRI L2 VSM are much more common in winter, shown in Fig. 10. It proves that more MWRI VSM estimates can be obtained and better performance of MWRI VSM can be reached by using our RF model.

5. Discussion

In using machine learning methods, it is very important to make sure the spatial–temporal representation of training sample space. According to the characteristics of the surface soil moisture, we emphasized the spatial sampling at 0.25-degree grids but the temporal sampling was sparse, i.e., only four days in a year were used. Therefore, we evaluate whether the sample space is big enough by comparing the difference among 4, 8, and 12 random days used in a whole year with the same input parameters in EXP4. The statistical scores are shown in Fig. 11. The 4-day experiment is the same as EXP4, and the 4 random days are 31 January, 30 April, 31 July, and 31 October 2018; for the 8-day experiment, the 8 random days are 15 January, 31 January, 15 April, 30 April, 15 July, 31 July, 15 October, and 31 October 2018; while for the 12-day experiment, the 12 random days are 31 January, 28 February, 24 March, 30 April, 31 May, 30 June, 31 July, 30 August, 30 September, 31 October, 30 November, and 31 December 2018. Since MWRI L2 VSM products are missing on 24 March, the VSM estimates on this day are not shown in Fig. 11.

Compared with SMOPS products, the statistical scores are slightly better by using more random days: the average $R^2 = 0.67, 0.65,$ and 0.63, and the average ubRMSD $= 0.042, 0.043,$ and 0.044 m$^3$ m$^{-3}$ for 12, 8, and 4 days, respectively, while there is no improvement in scores compared with ECMWF ERA5 products. It is clearly shown that more random days cannot significantly improve the scores. However, using more days of MWRI data in the training can consume huge computing re-

### Table 3. Comparisons of VSM between scaled in situ VSM observations at 5 cm from 12 stations in NRCS–SCAN and three scaled VSM time series (EXP4 VSM estimates, MWRI L2 VSM products, and SMOPS products) during the same time period from 1 August 2017 to 31 May 2019. The number of samples ($N$), correlation coefficient ($R$), and ubRMSD scores are given.

| SCAN station | SMOPS vs. in situ | EXP4 vs. in situ | MWRI L2 vs. in situ |
|--------------|------------------|-----------------|-------------------|
|              | $N$   | $R$   | ubRMSD (m$^3$ m$^{-3}$) | $N$   | $R$   | ubRMSD (m$^3$ m$^{-3}$) | $N$   | $R$   | ubRMSD (m$^3$ m$^{-3}$) |
| Tok          | 637   | 0.62  | 0.870            | 637   | 0.74  | 0.712            | 163   | −0.22 | 1.690            |
| Ku-Nesa      | 618   | 0.46  | 1.051            | 618   | 0.36  | 1.132            | 190   | −0.19 | 1.459            |
| Conrad Ag Re | 642   | 0.69  | 0.790            | 642   | 0.43  | 1.064            | 362   | 0.48  | 1.072            |
| Fort Assiniboine | 653   | 0.48  | 1.020            | 653   | 0.53  | 0.964            | 346   | −0.18 | 1.587            |
| Moccasin     | 635   | 0.57  | 0.929            | 635   | 0.47  | 1.030            | 355   | 0.42  | 1.150            |
| Violett      | 654   | 0.76  | 0.699            | 654   | 0.51  | 0.994            | 345   | 0.27  | 1.301            |
| Sheldr        | 648   | 0.53  | 0.963            | 648   | 0.38  | 1.110            | 363   | 0.20  | 1.207            |
| Chicken      | 621   | 0.58  | 0.906            | 621   | 0.53  | 0.966            | −     | −     | −                |
| Enterprise   | 630   | 0.53  | 0.973            | 630   | 0.45  | 1.049            | 455   | 0.06  | 1.341            |
| Ephraim      | 642   | 0.71  | 0.762            | 642   | 0.63  | 0.862            | 361   | 0.14  | 1.227            |
| Grouse       | 655   | 0.77  | 0.685            | 655   | 0.59  | 0.910            | 4     | 0.47  | 1.287            |
| Park         | 653   | 0.70  | 0.770            | 653   | 0.55  | 0.945            | 212   | −0.19 | 1.546            |
6. Conclusions and future perspectives

Based on the MWRI measurement physics, a multivariable approach using the RF machine learning technique is proposed to estimate MWRI VSM for improving retrieval accuracy. Moreover, the VSM estimates derived from different RF models are compared with the validations from SMOPS or ECMWF ERA5 VSM products during the same time period from 1 August 2017 to 31 May 2019, to evaluate the differences in retrieval accuracy due to the various model input parameters. From the MWRI experiments, the independent input parameters from all 10-channel TBs and three PRs at 10.65, 18.7, and 23.8 GHz, DEM, and soil porosity are determined as the significant important ones. Based on these model input parameters, the independent MWRI retrievals are in good agreement with the SMOPS products: $R^2 = 0.63$, ubRMSD = 0.044 m$^{-3}$, and mean bias = 0.002 m$^{-3}$ for a global scale. This positive result shows the possibility of using the RF training model to estimate MWRI VSM independently.

It should be noted that there is a risk of data gap from the present to 2025, for continuing the low frequency microwave observations, i.e., from GCOM-W1 AMSR2 and GMI on GPM. Meanwhile, MWRI observations from FY-3 series satellites can be a beneficial supplement to ensure data integrity and increase the data density, but more work should be investigated to make an operational RF model for releasing daily MWRI VSM products with high accuracy. This is useful to promote the reasonable use of global satellite resources and to speed up the development of national satellite product applications.

Fig. 10. Four scaled VSM time series from 1 August 2017 to 31 May 2019: daily mean in situ VSM observations (black line), SMOPS products (grey line), MWRI L2 VSM products (green line), and EXP4 VSM estimates (red line) from three NRCS–SCAN stations: (a) Tok, (b) Violett, and (c) Ephraim.
First, vegetation and land surface temperature are useful for MWRI VSM retrieval, but we need to find good proxies of them from Fengyun satellite observations or other satellites with the least spatial and temporal influences. Second, the amount of the missing data is significantly large in MWRI L2 VSM products. It is needed to assess the possibility of increasing the available VSM estimates based on the multivariable machine learning technique. Once the training model is constructed, we can obtain the MWRI VSM estimates using the MWRI observations, but we need to validate the retrieval accuracy especially over the many places where the soil exper-

Fig. 11. Daily scores of (a, b) $R^2$ and (c, d) the ubRMSD of the soil moisture estimates from MWRI EXP4 derived from sampling space with 12 (red line), 8 (blue line), and 4 random days (black line), and MWRI L2 products (green line) based on the validations from (a, c) SMOPS and (b, d) ECMWF ERA5, respectively, from 1 August 2017 to 31 May 2019.
ences freezing and thawing. We will further explore the ways of selecting the training data samples with the best representation of the MWRI observations and the VSM observations.

Acknowledgments. We thank Jean-Christophe Calvet at Météo France for his suggestions on this work. We thank Yang Liu at Qian Xuesen Laboratory of Space Technology for her revision on this paper.

REFERENCES

Albergel, C., P. de Rosnay, C. Gruhier, et al., 2012: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations. Remote Sens. Environ., 118, 215–226, doi: 10.1016/j.rse.2011.11.017.

Cao, B., J. F. Mao, S., W. Zhang, et al., 2018: Assessing the performance of separate bias Kalman filter in correcting the model bias for estimation of soil moisture profiles. J. Meteor. Res., 33, 519–527, doi: 10.1007/s13351-019-8057-6.

Choudhry, B., J. T. Schmugge, and T. M., 1982: A parameterization of effective soil temperature for microwave emission. J. Geophys. Res. Oceans, 87, 1301–1304, doi: 10.1029/JC087iC02p01301.

Du, J. Y., J. S. Kimball, L. A. Jones, et al., 2017: A global satellite environmental data record derived from AMSR-E and AMSR2 microwave Earth observations. Earth Syst. Sci. Data, 9, 791–808, doi: 10.5194/essd-9-791-2017.

Green, J. K., S. I. Seneviratne, A. M. Berg, et al., 2019: Large influence of soil moisture on long-term terrestrial carbon uptake. Nature, 565, 476–479, doi: 10.1038/s41586-018-0848-x.

Jackson, T. J., and T. J. Schmugge, 1991: Vegetation effects on the microwave emission of soils. Remote Sens. Environ., 36, 203–212, doi: 10.1016/0034-4257(91)90057-D.

Kerr, Y. H., 2007: Soil moisture from space: Where are we? Hydrogeol. J., 15, 117–120, doi: 10.1007/s10040-006-0095-3.

Kerr, Y. H., P. Waldteufel, J.-P. Wigneron, et al., 2001: Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. IEEE Trans. Geosci. Remote Sens., 39, 1729–1735, doi: 10.1109/36.942551.

Li, L., E. G. Njoku, E. Im, et al., 2004: A preliminary survey of radio-frequency interference over the U.S. in Aqua AMSR-E data. IEEE Trans. Geosci. Remote Sens., 42, 380–390, doi: 10.1109/TGRS.2003.817195.

Liao, W. L., D. G. Wang, G. L. Wang, et al., 2019: Quality control and evaluation of the observed daily data in the North American Soil Moisture Database. J. Meteor. Res., 33, 501–518, doi: 10.1007/s13351-019-8121-2.

Liaw, A., and M. Wiener, 2002: Classification and regression by randomForest. R News, 2, 18–22.

Liu, J. C., X. W. Zhan, C. Hain, et al., 2016: NOAA Soil Moisture Operational Product System (SMOPS) and its validations. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, Beijing, China, 3477–3480, doi: 10.1109/IGARSS.2016.7729899.

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

Njoku, E. G., T. J. Jackson, V. Lakshmi, et al., 2003: Soil moisture retrieval from AMSR-E. IEEE Trans. Geosci. Remote Sens., 41, 215–229, doi: 10.1109/TGRS.2002.808243.

Owe, M., R. de Jeu, and J. Walker, 2001: A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index. IEEE Trans. Geosci. Remote Sens., 39, 1643–1654, doi: 10.1109/36.942542.

Paloscia, S., G. Macelloni, E. Santi, et al., 2001: A multifrequency algorithm for the retrieval of soil moisture on a large scale using microwave data from SMMR and SSM/I satellites. IEEE Trans. Geosci. Remote Sens., 39, 1655–1661, doi: 10.1109/36.942543.

Pangaluru, K., I. Velicogna, G. A., et al., 2019: Soil moisture variability in India: Relationship of land surface–atmosphere fields using maximum covariance analysis. Remote Sens., 11, 335, doi: 10.3390/rs11030335.

Parinussa, R. M., G. Wang, T. R. H. Holmes, et al., 2014: Global surface soil moisture from the Microwave Radiation Imager onboard the Fengyun-3B satellite. Int. J. Remote Sens., 35, 7007–7029, doi: 10.1080/01431161.2014.960622.

Parrens, M., J.-C. Calvet, P. de Rosnay, et al., 2014: Benchmarking of L-band soil microwave emission models. Remote Sens. Environ., 140, 407–419, doi: 10.1016/j.rse.2013.09.017.

Rudd, A. C., A. L. Kay, and V. A. Bell, 2019: National-scale analysis of future river flow and soil moisture droughts: Potential changes in drought characteristics. Climatic Change, 156, 323–340, doi: 10.1007/s10584-019-02528-0.

Ruosteenmaa, K., T. Markkanen, A. Venäläinen, et al., 2018: Seasonal soil moisture and drought occurrence in Europe in CMIP5 projections for the 21st century. Climate Dyn., 50, 1177–1192, doi: 10.1007/s00382-017-3671-4.

Schaefer, G. L., and R. F. Paetzold, 2000: SNOTEL (SNOWpack TELEmetry) and SCAN (Soil Climate Analysis Network). Presented at the Automated Weather Stations for Applications in Agriculture and Water Resources Management: Current Use and Future Perspectives, USDA-NRCS, Lincoln, NE.

Schmugge, T., and T. J. Jackson, 1994: Mapping surface soil moisture with microwave radiometers. Meteor. Atmos. Phys., 54, 213–223, doi: 10.1007/BF01030061.

Seager, R., J. Nakamura, and M. F. Ting, 2019: Mechanisms of seasonal soil moisture drought onset and termination in the southern Great Plains. J. Hydrometeor., 20, 751–771, doi: 10.1175/JHM-D-18-01911.1.

Shi, J., L. Jiang, L. Zhang, et al., 2006: Physically based estimation of bare-surface soil moisture with the passive radiometers. IEEE Trans. Geosci. Remote Sens., 44, 3145–3153, doi: 10.1109/TGRS.2006.876706.

Ulaby, F. T., and E. A. Wilson, 1985: Microwave attenuation properties of vegetation canopies. IEEE Trans. Geosci. Remote Sens., GE-23, 746–753, doi: 10.1109/TGRS.1985.289393.

Wagner, W., G. Blöschl, P. Pampaloni, et al., 2007: Operational readiness of microwave remote sensing of soil moisture for hydrologic applications. Hydroil. Res., 38, 1–20, doi: 10.2166/hd.2007.029.

Wang, J. R., 1983: Passive microwave sensing of soil moisture content: The effects of soil bulk density and surface roughness. Remote Sens. Environ., 13, 329–344, doi: 10.1016/
Wang, J. R., and B. J. Choudhury, 1981: Remote sensing of soil moisture content, over bare field at 1.4 GHz frequency. *J. Geophys. Res. Oceans*, **86**, 5277–5282, doi: 10.1029/JC086iC06p05277.

Wang, J. S., J. S. Zhang, and Y. Q. Tang, 2010: Fengyun satellites: Achievements and future. *Chinese J. Space Sci.*, **30**, 468–473.

Wigneron, J.-P., J.-C. Calvet, T. Pellarin, et al., 2003: Retrieving near-surface soil moisture from microwave radiometric observations: Current status and future plans. *Remote Sens. Environ.*, **85**, 489–506, doi: 10.1016/S0034-4257(03)00051-8.

Xu, J. W., W. C. Zhang, Z. Y. Zheng, et al., 2012: Establishment of a hybrid rainfall-runoff model for use in the Noah LSM. *Acta Meteor. Sinica*, **26**, 85–92, doi: 10.1007/s13351-012-0108-1.

Yang, H., F. Z. Weng, L. Q. Lv, et al., 2011: The FengYun-3 Microwave Radiation Imager on-orbit verification. *IEEE Trans. Geosci. Remote Sens.*, **49**, 4552–4560, doi: 10.1109/TGRS.2011.2148200.

Yang, J., P. Zhang, N. M. Lu, et al., 2012: Improvements on global meteorological observations from the current Fengyun 3 satellites and beyond. *Int. J. Digi. Earth*, **5**, 251–265, doi: 10.1080/17538947.2012.658666.

Yang, Z. D., N. M. Lu, J. M. Shi, et al., 2012: Overview of FY-3 payload and ground application system. *IEEE Trans. Geosci. Remote Sens.*, **50**, 4846–4853, doi: 10.1109/TGRS.2012.2197826.

Zhang, F. M., Z. X. Pu, and C. H. Wang, 2019: Impacts of soil moisture on the numerical simulation of a post-landfall storm. *J. Meteor. Res.*, **33**, 206–218, doi: 10.1007/s13351-019-8002-8.

Zhang, S. W., Y. H. Liu, and W. D. Zhang, 2013: Ensemble square root filter assimilation of near-surface soil moisture and reference-level observations into a coupled land surface-boundary layer model. *Acta Meteor. Sinica*, **27**, 541–555, doi: 10.1007/s13351-013-0402-6.