Retrieval of sub-kilometric relative surface soil moisture with sentinel-1 utilizing different backscatter normalization factors

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open access

Maslanka, W. ORCID: https://orcid.org/0000-0002-1777-733X, Morrison, K. ORCID: https://orcid.org/0000-0002-8075-0316, White, K., Verhoef, A. ORCID: https://orcid.org/0000-0002-9498-6696 and Clark, J. ORCID: https://orcid.org/0000-0002-0412-8824 (2022) Retrieval of sub-kilometric relative surface soil moisture with sentinel-1 utilizing different backscatter normalization factors. IEEE Transactions on Geoscience and Remote Sensing, 60. 4410613. ISSN 1558-0644 doi: https://doi.org/10.1109/tgrs.2022.3175256 Available at https://centaur.reading.ac.uk/105497/

It is advisable to refer to the publisher’s version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1109/tgrs.2022.3175256

Publisher: Institute of Electrical and Electronics Engineers (IEEE)

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.
www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online
Abstract—Spatiotemporal distribution of soil moisture is important for hydrometeorological and agricultural applications. There is growing interest in monitoring soil moisture in relation to soil- and land-based natural flood management (NFM), to understand the soil's ability, via land-use and management changes, and to delay the arrival of flood peaks in nearby watercourses. This article monitors relative surface soil moisture (rSSM) across the Thames Valley, U.K., using Sentinel-1 data, and the Vienna University of Technology (TU-Wien) Change Detection Algorithm, with a novel exploration of monthly and annual normalization factors and spatial averaging. Two pairs of normalization factors are introduced to remove impacts from varying local incidence angles through direct and multiple regression slopes. The spatiotemporal distribution of rSSM values at various spatial resolutions (1000, 500, 250, and 100 m) is assessed. Comparisons with in situ soil moisture data from the COSMOS-UK network show that, while general temporal trends agree, the difference in effective depth of measurements, coupled with vegetation impacts during the growing season, makes comparison with soil moisture observations difficult. Temporal rSSM trends can be retrieved at spatial resolutions down to 100 m, and the rSSM RMSE was found to decrease as the spatial resolution increases. The vegetation effects upon the rSSM are further explored by comparing the two dominant land cover types: Arable and Horticulture, and Improved Grassland. It was found that, while the rSSM retrieval for these land covers was possible, and the general soil moisture trend is clear, overlying vegetation during the summer artificially increased the rSSM values.

Index Terms—Change detection algorithm, River Thames, Sentinel-1, soil moisture, synthetic aperture radar (SAR).

I. INTRODUCTION

Soil moisture makes up only 0.0012% of the total amount of water in the global water cycle [1], yet it is vitally important for hydrometeorology [2]. Soil moisture plays a pivotal role in a meteorological context: it has a significant impact on the surface energy balance and its interactions and feedbacks with the atmosphere and, hence, on boundary layer development and processes [3]. The ability of weather models to predict soil moisture, and the processes it interacts with, will influence the accuracy and reliability of subseasonal and seasonal forecasting [4]. In the hydrological context, soil moisture impacts infiltration, runoff [5], evapotranspiration, and groundwater recharge [2]; the balance of these fluxes plays an important role in the occurrence of flooding [6] and drought [7]. Finally, from an agricultural point of view, soil moisture is a determining factor in crop water use, growth, development, and productivity [8], [9], which will, in turn, affect the water and energy budgets mentioned above.

Soil and land management will affect the spatiotemporal variation of soil moisture; farmers manage their soil to maximize plant available water and minimize plant water stress. Another more recent example is the soil- and land-use-based natural flood management (NFM). This type of NFM aims to reduce surface run-off by maximizing infiltration into the soil and maximizing storage within the soil [10]. Infiltration crucially depends on near-surface water content; hence, having accurate observations and estimates of global soil moisture is vitally important [11]. Unfortunately, current remote sensing observations of soil moisture are on the kilometric spatial scale or greater [12], [13], whereas some NFM practices (such as land- and soil-based management) occur on sub-kilometric spatial scales (i.e., individual field scale) [14]. In order to better understand the variability of soil moisture across different agricultural practices (such as the comparison of different fields and farms) within the context of NFM practices, finer resolution observations of soil moisture are required.

While in situ point measurements of soil moisture are generally accurate, the highly heterogeneous nature of soils, and their land use, means that in situ networks are too sparsely distributed to reliably capture the spatial variability of soil moisture [15]. Space-based remote sensing remains the only way to get spatially distributed (at catchment-to-global scales) observations of soil moisture. Microwave-based systems are the primary orbital remote sensing platform for soil moisture observations [16] due to their cloud penetrating properties and their ability to make use of the difference in dielectric constant between liquid water and dry soil [17].

There are two categories of remote sensing, which corresponds to the instrumentation type and the way in which the observation takes place: passive and active methods. Passive methods measure the energy that is naturally emitted from the Earth’s surface using radiometers. Both active and
passive remote sensing systems are used in soil moisture remote sensing [18]–[20]. Examples of passive microwave remote sensing systems include the Scanning Multichannel Microwave Radiometer (SMMR [21]), the Special Sensor Microwave Imager (SMMI [22]), the Advanced Microwave Scanning Radiometer (AMSR [23]), the Soil Moisture Active Passive (SMAP [24]) satellite, and the Soil Moisture and Ocean Salinity (SMOS [25]) satellite. Active methods emit a source of radiation and record the strength of the returned signal, often referred to as the radar backscatter ($\sigma^0$). Examples of active microwave platforms include the European Remote Sensing (ERS [26]) satellites, the METOP Advanced Scatterometer (ASCAT [27]), and the Sentinel-1 Constellation [28]. The latter will be used for this study.

The Sentinel-1 constellation is currently made up of one satellite (Sentinel-1A), launched in April 2014. A second satellite (Sentinel-1B) was launched in April 2016 and ceased operations in December 2021. A third satellite (Sentinel-1C) is planned to be launched in 2023. Each satellite observes the Earth with a fixed 12-day repeating orbit. The Sentinel-1 satellites each house a C-band synthetic aperture radar (CSAR), operating at a central frequency of 5.405 GHz, with both single- and dual-polarization modes and four separate acquisition modes. The interferometric wide (IW) swath mode is the main mode over nonpolar land and has a 250-km swath.

This article uses the Vienna University of Technology (TU-Wien) change detection algorithm [29], adapted for use with the Sentinel-1 constellation [30], to estimate relative Surface Soil Moisture (rSSM) in the top few centimeters of the surface, across the Thames Valley, in Southern England, to assess the capability of broad-scale rSSM monitoring and parameters of the rSSM retrieval algorithm in context of the processing steps, the Sentinel-1 normalization factor may be needed to account for the seasonal variations in vegetation cover and surface roughness, as crops are tilled, grown, and harvested. It is well known ([31]–[33]) that vegetation affects radar backscatter, yet the standard normalization factor of Sentinel-1 assumes a static annual value. Potential within-year variations in this normalization factor may be needed to account for the seasonal changes, yet few have considered a temporal variation to the normalization factor for use with Sentinel-1.

Section II reviews the TU-Wien Change Detection Algorithm used, along with the Sentinel-1 processing steps. As part of the processing steps, the Sentinel-1 normalization factor will be discussed. Section III details the area of interest (AOI) and parameters of the rSSM retrieval algorithm in context with the AOI. Section IV compares the retrieved rSSM at various spatial scales with different normalization factors with a preexisting in situ soil moisture dataset, while Section V closes with a comparison of the retrieved rSSM time series for two dominant land use types with the AOI.

II. TU-WIEN CHANGE DETECTION ALGORITHM

The Sentinel-1 rSSM retrieval algorithm presented here is based on the TU-Wien Change Detection Algorithm [29], [33] (referred to as TWCDCA), which derives rSSM directly from a time series of observed radar backscatter $\sigma^0$. The TWCDCA was originally designed for use with the ERS platform and has been used in conjunction with both ASCAT [27] and Sentinel-1 [30], [34] previously.

In this model, changes in backscatter are accredited as changes in soil moisture, while other properties that can influence the observed backscatter (such as surface roughness and geometry) are assumed to be temporally constant. For the rSSM estimation, the model uses actual backscatter $\sigma^0(\theta, t)$, taken at time $t$ and local incidence angle (LIA) $\theta$, and normalized to a reference angle $\Theta$, before linearly scaling the normalized backscatter between the wet and dry reference values, as shown in the following:

$$rSSM(t) = \frac{\sigma^0(\theta, t) - \sigma^0(\Theta)}{\sigma^0(\Theta) - \sigma^0(\theta)} \times 100\%$$

where $rSSM(t)$ is the rSSM at a time $t$ (%), $\Theta$ is a reference angle ($^\circ$), $\sigma^0(\Theta)$ is a dry threshold radar backscatter value (dB), and $\sigma^0(\theta)$ is a wet threshold radar backscatter value (dB). Derivation of these thresholds is described in Section II-E.

A. Preprocessing the Raw Sentinel-1 Data

Sentinel-1 data must first be geocoded and radiometrically corrected before the TWCDCA can be used to estimate rSSM. The geocoding and radiometric calibration is completed using ESA's Sentinel Application Platform (SNAP1), which allows for the production and completion of workflows. The workflow used for the geocoding and radiometric calibration is very similar to that laid out by [35]. As an initial step, the workflow processor, as an input, takes the VV-polarized IW Ground Range Detected High Resolution (GDRH) Sentinel-1 data and applies orbital corrections, thermal and border noise removals, and radiometric calibrations. A “Refined Lee” speckle filter and terrain correction are applied, using the Shuttle Radar Topography Mission (SRTM) 3Sec data. The resulting radar backscatter is then subset to the required AOI and then exported for individual frames. If there are multiple subsequent frames that fall over the AOI (i.e., the AOI is split in two), then the frames are stitched together to create a single radar file per scan day. The subset-stitching is also conducted on exported LIA ($\theta$) data to be used in the normalization process.

B. Backscatter Normalization

As $\sigma^0$ will be observed across a range of LIAs over a number of different orbits, $\sigma^0$ must be normalized to a common angle for comparisons across the time series to be made. This is to ensure that LIA dependencies have been removed from the $\sigma^0$ time series, and thus, variations in $\sigma^0$ are only a result of changing ground conditions (i.e., variations in soil moisture). To access and conduct the normalization of LIA, a normalization parameter, $\beta$, is calculated at the pixel scale (units of dB/\(^\circ\)). Using this normalization factor, the observed radar backscatter can be normalized to the reference angle, as shown in the following equation:

$$\sigma^0(\theta, t) = \sigma^0(\theta, t) - \beta(\theta - \Theta) \ [\text{dB}].$$

1SNAP-ESA Sentinel Application Platform v6.0.0. https://step.esa.int/
For this study, a reference incidence angle, $\Theta$, of 40° was used. For the TWCDA, two current methods of LIA normalization can be calculated. The first, simpler, method is to model the normalization factor as an indirect linear relationship between $\theta$ and $\sigma^{\theta}$ [36] and is known as the direct regression slope, $\beta_d$. This direct regression slope estimation, however, is only reliable for areas of sufficient repeating observations and a large enough range of LIA values. If the area does not have sufficient repeating observations, or if the LIA values are too similar, the normalization factor can be calculated using a multiple linear regression model. As the correlation between normalization factor $\beta$, the non-normalization sensitivity between wet and dry thresholds, $S$, and mean backscatter ($\langle \sigma \rangle$) are found to be generally high [33], the multiple regression slope, $\beta$, can be modeled via the following equation:

$$\beta = aS + b\sigma + c \quad [\text{dB}/\circ].$$ \hspace{1cm} (3)

where $a$, $b$, and $c$ are constants calculated in Table I.

Traditionally, the normalization factors, using either the direct or the multiple regression slope, have been calculated to not have a temporal component (i.e., a single value for normalization parameterization is used) [30], [36]–[40]. This study, however, adopts a novel approach, considering the normalization parameterization (both the direct and multiple regression slopes) as a monthly varying normalization factor by applying the direct and multiple regression slope methodology to Sentinel-1 orbits occurring on individual months. Table I also shows the multiple regression model constants $a$, $b$, and $c$ from (3) for the individual months.

When calculating the monthly normalization parameter (either via the direct or multiple regression slope method), an interesting seasonal cycle can be produced. Fig. 1 shows the (traditionally used) annual and monthly $\beta_d$ and $\beta$ values. The observed seasonal pattern exhibits a maximum value during the summer months (JJA) and a minimum value over the winter months (DJF). This seasonal oscillation is most likely present due to combined contributions from changes in vegetation structure and surface soil moisture; soil moisture levels tend to be higher during the winter than during summer; and inversely, vegetation growth tends to be greater during the summer than the winter. Compared to mean monthly NDVI (obtained from the Copernicus Global Land Service2 based on data acquired by PROBA-V) across the AOI (see Section III), the oscillation in NDVI generally correlates to that of the monthly $\beta_d$ with a minimum monthly $\beta_d$ and minimum NDVI value during February, and a noticeable dip in both NDVI and monthly $\beta_d$ in August; both are potentially due to the senescing and crop harvest period in the AOI. The lag of a couple of months between the peak of NDVI and that of monthly $\beta_d$ and $\beta$, may be explained by the growth and development of crops across the entire AOI. As the crops develop and grow during the spring and early summer, they would develop green leafy shoots and stalks; this would affect the NDVI values. However, later in the season, flowers and crop products appear, generally with a nongreen or less green. An example is the yellow flowering of *Brassica napus* or Rapeseed. However, once the flowers develop into green seedpods, as well as after these have been harvested, the yellow of the flower would be replaced with the green of the developed stalks, thus resulting in the decrease-then-increase in August.

Normalization factor oscillations have been found previously with the TWCDA [41]; however, these normalization factors cannot be used for Sentinel-1 application due to satellite platform differences (both ERS and METOP satellites had multiple antennae, while Sentinel-1 has a single antenna).

### C. Dynamic Backscatter Masking

Very high and very low backscatter measurements are unlikely to hold any soil moisture signal and, thus, can be discarded from the analysis. To dynamically remove the effects of these very high and very low backscatter values, a pair of thresholds were applied to each individual Sentinel-1 SAR image. The threshold limits for this project were set to $-5$ and $-22$ dB with the upper limit derived from qualitative comparisons with urban and suburban areas from the Centre of Ecology and Hydrology 2015 Land Cover Model (CEH LCM 2015 [42]; see Fig. 2) and the lower limit taken from the sensor’s noise equal sigma zero. In addition to the very high and very low backscatter measurement mask, areas allocated

---

2https://land.copernicus.eu/global/products/ndvi/, accessed May 10, 2021.
D. Spatial Averaging

In the IW mode, Sentinel-1 has a single look spatial resolution of $5 \times 20$ m, with a resulting multilook spatial resolution of $20 \times 22$ m. Soil moisture estimation at this spatial resolution is not advisable, as the backscatter signal is influenced by many different, highly variable features. These features include vegetation water content, crop row orientation, crop bending (e.g., lodging for grain crops, due to inadequate standing power of the crop and adverse weather conditions), underlying surface roughness, and soil moisture content. To mitigate the unwanted influences from many of the heterogeneous variables, while conserving the influence of soil moisture variations, the backscatter data can be upscaled to coarser resolutions. The pixel aggregation was completed via the calculation of the arithmetic mean. Spatial resolutions used as part of this study include 1 km, 500 m, 250 m, and 100 m. Spatial averaging affects how the data product can be used to evaluate different types of land use for NFM. These coarse resolutions have been shown to have a high correlation ($\geq 0.7$) with in situ measurements [39], [43], [44].

E. Wet and Dry Threshold

To use the TWGDA, the wet and dry thresholds ($\sigma^0_w(\Theta)$ and $\sigma^0_d(\Theta)$, respectively) must be calculated to indicate the upper and lower limits of the backscatter time series. Ideally, the radar backscatter would extend over long time periods (i.e., several years) to capture occasions where the soil is completely dry and completely saturated so that the wettest and driest events can correspond to the largest and smallest magnitudes of normalized $\sigma^0$, respectively [29].

In order to remove outliers (potential peaks in $\sigma^0$, which do not correspond to extreme wet or dry conditions due to signal interference, frozen soil conditions, and so on [36]) and address the relatively short time series length, estimated threshold values can be statistically calculated at the pixel scale, as detailed by [34]. By using the 10% and 90% values of normalized radar backscatter time series and, assuming that these correlate with rSSM values of 10% and 90%, respectively, radar backscatter values associated with 0% and 100% (completely dry and saturated) can be statistically calculated.

As the wet and dry threshold parameters are statistically determined, it is possible for extreme events to cause rSSM values to be out of the 0%–100% range. rSSM values that are $\pm 20\%$ off-limit are set to 0% or 100%, respectively, with values outside the 20% buffer being set to No Data.

III. AREA OF INTEREST AND DATASET

A. Area of Interest

The River Thames Valley (containing 18 different tributary catchments) covers approximately 16 200 km$^2$ in Southeast England. The upstream area to the west is predominately rural, comprising a mix of agricultural land and woodland over rolling hills on chalk and limestone geology, with flatter areas on clays (see Fig. 2). Toward the center and the east of the basin, the land becomes increasingly urbanized, as the River Thames flows through Reading, Slough, and into London. The source of the River Thames is in the west (elevation 350-m asl, in Kemble, Gloucestershire), with the fluvial endpoint being Teddington Lock in London, over 230-km downstream [45]. To remove large parts of urban London from the study, the AOI is bounded between 51.1501 N–52.0622 N and 2.22475 W–0.5708 W. The land use within the AOI comprises agricultural land (42.0%), grassland (36.2%, comprised of 33.2% “Improved Grassland” and 3.0% “Unimproved Grassland”), woodland (10.6%), freshwater (0.7%), and urban (9.9%) areas, as characterized by the CEH LCM 2015 model [42]. Table II shows the aggregation of CEH LCM 2015 land uses to the land uses presented in Fig. 2.

The climate of the Thames Valley is categorized by the Köppen climate classification as Cfb (Temperate Oceanic) and received an average of 747-mm precipitation annually, in the 1981–2010 time period [46]. Higher monthly precipitation values generally occur over the autumn and winter months, from cyclonic frontal systems, although flashier, intense thunderstorms have produced some large rainfall totals, especially over the summer months. Both frontal and convective rainfall events have led to a number of fluvial flooding events in recent years, across all reaches of the Thames Valley [47], [48].

B. Dataset

Sentinel-1 IWGRDH data spanning six years (October 2015–September 2021) were extracted over the AOI and processed using the preprocessing steps, normalization, masking, spatial averaging, and TWGDA detailed in Section II. In total, 608 individual ascending orbits, from relative orbits
Fig. 3. TWCDA parameters for the AOI between October 2015 and September 2021 using the monthly multiple regression model normalization factor and a spatial resolution of 100 m. (a) Wet threshold \(\sigma_\text{w}(40)\). (b) Dry threshold \(\sigma_\text{d}(40)\). (c) Backscatter sensitivity \(S(40) = \sigma_\text{w}(40) - \sigma_\text{d}(40)\).

### TABLE II

| Abbrev. | Long Name | Land Uses in LCM 2015 [42] |
|---------|-----------|---------------------------|
| AR      | Artificial| Urban Suburban             |
| AH      | Arable and Horticulture | Arable and Horticulture |
| BW      | Broadleaf Woodland       | Broadleaf Woodland         |
| CW      | Coniferous Woodland      | Coniferous Woodland        |
| FW      | Freshwater               | Freshwater                 |
| IG      | Improved Grassland       | Improved Grassland         |
| IG      | Unimproved Grassland     | Improved Grassland         |
| HG      | Acid Grassland           | Acid Grassland             |
| LG      | Calcareous Grassland     | Calcareous Grassland       |
| HE      | Heather                 | Heather                    |
| OT      | Other                   | Fur, Marsh, or Swamp Island Rock |

Using the 608 ascending orbits, the wet [see Fig. 3(a)] and dry [see Fig. 3(b)] thresholds, as calculated in Section II-E, were derived at spatial resolutions of 1000, 500, 250, and 100 m, using the four normalization parameters. For each individual observation, data collected over the Artificial and Freshwater land uses (as defined in Table II and Fig. 2) were discarded. The wet and dry threshold values for the AOI are displayed in Fig. 3(a) and (b), respectively, with the difference between the wet and dry thresholds, known as the sensitivity, \(S(\theta)\), as shown in Fig. 3(c).

Compared to the land cover map in Fig. 2, it can be seen that the land use defined as “Arable and Horticulture” has the greatest values of sensitivity, followed by “Improved Grassland.” This suggests that the algorithm is better at detecting rSSM changes across these two land uses due to the larger difference between the wet and dry thresholds. The sensitivity range calculated in Fig. 3(c) ranges between 0 and 14 dB, which is a similar range to previous findings in the literature [36], [39]. The land cover types with the lowest sensitivity (as aggregated by Table II) were the “Broadleaf Woodland” and “Coniferous Woodland.” This is presumably due to the lack of change throughout the year compared to the change seen by the Arable and Horticulture land covers, and the impact upon backscatter. In the case of Broadleaf Woodland, the dominant difference throughout the time series would be the growth and decay (senescence and subsequent leaf fall) of the leaves in the canopy, and the resultant change in leaf area index within the canopy would introduce additional volume scattering components to the observed backscatter signal [49]. While backscatter is sensitive to the dynamics of the surface beneath the forest canopy, extracting the surface contribution (and, thus, the rSSM contribution) from the numerous volume scattering present within the forest (i.e., from the trunk, branches, and leaves) is extremely difficult. It is for this reason that soil moisture retrieval in forested areas is not feasible [34], [50].

#### C. Rainfall Event

Using the algorithm detailed in Section II, along with the model parameters, as shown in Fig. 3, it is possible to examine the effect of different meteorological events within the timeframe, upon rSSM. Fig. 4 shows two different localized precipitation events within the AOI using data recorded by the U.K. Met Office NIMROD System [51]. The 3-h precipitation accumulations for September 11, 2018, and September 6, 2019 (Fig. 4(a) and (c), respectively) are plotted alongside the corresponding rSSM estimations (Fig. 4(b) and (d), respectively). Only precipitation accumulations greater than 0.25 mm are shown in order to ignore very light precipitation events (i.e., drizzle). The spatial patterns of both precipitation events correspond well to the spatial distribution of increases in rSSM (> 50%).
Fig. 4. Spatial analysis of precipitation and rSSM patterns. (a) 3-h precipitation between 15:00 and 18:00 UTC on September 09, 2018, across the AOI. (b) rSSM across the AOI at 100-m resolution on September 09, 2018 (orbit overpass at approximately 18:00 UTC). (c) 3-h precipitation between 15:00 and 18:00 UTC on September 06, 2019, across the AOI. (d) rSSM across the AOI at 100-m resolution on September 06, 2019 (orbit overpass at approximately 18:00 UTC).

IV. COMPARISONS WITH IN SITU DATA FROM COSMOS-UK

In situ soil moisture data were taken from the Cosmic-ray Soil Moisture Observing System United Kingdom (COSMOS-UK) network [52]–[54] in order to evaluate the retrieval of the rSSM time series. The COSMOS-UK network has been producing observations of meteorological and soil moisture variables since 2013 [55]. Both the depth
and footprint diameter of the COSMOS-UK cosmic-ray neutron (CNR) sensor vary with the underlying soil moisture values [56] with both footprint size and penetration depth being inversely proportional to soil moisture. For this project, volumetric water content (VWC) data from three sites that fall within the AOI (Chimney Meadows, Sheepdrove, and Waddesdon, soil, and land properties displayed in Table III), taken between January 2016 and December 2019, were used to compare and evaluate the rSSM algorithm and time series.

Across this period, the effective measurement depth of the VWC measurements via CRNs from these three sites ranged from 12.3 to 25.2 cm with a mean effective depth of 17.5 cm. This measurement depth is deeper than that of the rSSM time series, which, by its nature, is only of the top few centimeters of the surface [29]. The footprint radius of the CRNs was assumed to be 200 m, as changes in soil moisture (and, thus, observed cosmic rays) outside of 200 m have previously been found to make negligible contributions [56] to the return of the COSMOS-UK CRNs.

In order to make appropriate comparisons between the rSSM and the VWC time series, the VWC data were transformed into a normalized dimensionless index, similar in method to the rSSM time series, as shown in the following equation:

\[ \text{VWC}_1(t) = \frac{\text{VWC}(t) - \text{VWC}_{\text{min}}}{\text{VWC}_{\text{max}} - \text{VWC}_{\text{min}}} \times 100\% \]  

(4)

where \( \text{VWC}_1(t) \) is the VWC dimensionless index at time \( t \), \( \text{VWC}(t) \) is the VWC at time \( t \), and \( \text{VWC}_{\text{max}} \) and \( \text{VWC}_{\text{min}} \) are the minimum and maximum VWC values across the time series.

The VWC time series consists of data taken on the same day as the Sentinel-1 orbits so that the VWC1 and rSSM time series can be directly comparable. Then, a 14-orbit moving average is calculated across both the VWC1 and rSSM data in order to focus more on the general temporal trends of soil moisture.

Fig. 5 shows the smoothed VWC1 and rSSM time series (using the monthly multiple regression normalization factors, spatially averaged to a spatial resolution of 100 m), for the three COSMOS-UK sites. To aid with the comparison of the two soil moisture products, daily precipitation totals (measured using a Pluviometer rain gauge at the COSMOS-UK sites) are shown.

In general, the VWC1 and rSSM time series capture similar temporal trends, with particularly good agreement over the October–May period at Chimney Meadows (\( r^2 = 0.53 \) and RMSE = 7.0%) and some agreement over the same time period at Waddesdon (\( r^2 = 0.41 \) and RMSE = 12.4%). At Chimney Meadows, there appears to be an overestimation over the summer/early autumn months, where the VWC1 values decrease, yet the rSSM values increase, before decreasing approximately in October. This deviation is clear over the summer of 2018, where the Thames catchment (and the rest of Southern England) suffered a period of unseasonably hot and dry weather. This deviation may be due, in part, to vegetation effects on the Sentinel-1 observations, as there would be more vegetation biomass over the surface during the summer months. It should be noted that, currently, the TWCDA does not have an appropriate way to explicitly take vegetation dynamics (such as biomass or vegetation water content) into account. At Chimney Meadows, the hay meadow grows during the summer, before being harvested in the autumn. This over-estimation can also be seen in the summers in the Waddesdon data; however, there is a general overestimation of rSSM in the 2018 time period.

The VWC1 data at Sheepdrove have a low agreement with the rSSM time series (\( r^2 = 0.35 \) and RMSE = 9.5) and with some degree of the temporal pattern being exhibited (increase during the winter months, decrease in the summer months, with a larger decrease being seen in 2018); however, the rSSM values often underestimate the soil moisture compared to the VWC1. One potential reason behind the overarching underestimation could be the land use type, as both Chimney Meadows and Waddesdon have been denoted as “Improved Grassland,” whereas Sheepdrove is denoted as being “Arable and Horticulture.”

Comparisons were also made between the COSMOS-UK VWC1 and rSSM estimations using spatial resolutions of 1000, 500, and 250 m in addition to 100 m, as well as using the two pairs of different normalization factors (direct and multiple regression slopes, calculated at both the traditional annual and novel monthly timescales, as detailed in Section II-B), using a 14-orbit moving average, for all three COSMOS-UK sites, as shown in Table IV.

Across all COSMOS-UK sites, regardless of the normalization factor used, it can be seen that RMSE values decrease as the spatial resolution increases (for example, Chimney Meadows, using the monthly regression normalization factor, the RMSEs are 11.9, 11.5, 7.8, and 6.7 for the 1000-m, 500-mm, 250-m, and 100-m spatial averagings, respectively). This is to be expected, as the CRNs’ footprint radius of 200 m is smaller than the spatial resolution of the larger grid sizes, meaning that the rSSM values using the larger grid sizes would contain rSSM information from outside the CRNs’ footprint, whereas the smaller grid sizes only use the relevant pixels that fall within the CRNs’ footprint. Unlike the RMSE values, the \( r^2 \) values increase with an increasing spatial resolution for only Chimney Meadows. For Sheepdrove and Waddesdon, \( r^2 \) decreases slightly with increasing spatial resolution. However, the decrease in \( r^2 \) is small (by approximately 0.25 for Sheepdrove and approximately 0.2 for

### Table III

| Site                  | Chimney Meadows | Sheepdrove | Waddesdon Manor |
|-----------------------|-----------------|------------|-----------------|
| Latitude (°N)         | 51.7080         | 51.5303    | 51.8395         |
| Longitude (°W)        | 1.4788          | 1.4819     | 0.9484          |
| Elevation (m)         | 65 m            | 183 m      | 90 m            |
| Soil Type             | Deep Clay to Sandy Loam | Shallow choky, silty loam with flints | Deep Clay to Loam |
| Land Cover in Footprint | Improved Grassland | Arable and Horticulture | Improved Grassland |

**Comparisons** between the COSMOS-UK VWC1 and rSSM estimations using spatial resolutions of 1000, 500, and 250 m in addition to 100 m, as well as using the two pairs of different normalization factors (direct and multiple regression slopes, calculated at both the traditional annual and novel monthly timescales, as detailed in Section II-B), using a 14-orbit moving average, for all three COSMOS-UK sites, as shown in Table IV.

Across all COSMOS-UK sites, regardless of the normalization factor used, it can be seen that RMSE values decrease as the spatial resolution increases (for example, Chimney Meadows, using the monthly regression normalization factor, the RMSEs are 11.9, 11.5, 7.8, and 6.7 for the 1000-m, 500-mm, 250-m, and 100-m spatial averagings, respectively). This is to be expected, as the CRNs’ footprint radius of 200 m is smaller than the spatial resolution of the larger grid sizes, meaning that the rSSM values using the larger grid sizes would contain rSSM information from outside the CRNs’ footprint, whereas the smaller grid sizes only use the relevant pixels that fall within the CRNs’ footprint. Unlike the RMSE values, the \( r^2 \) values increase with an increasing spatial resolution for only Chimney Meadows. For Sheepdrove and Waddesdon, \( r^2 \) decreases slightly with increasing spatial resolution. However, the decrease in \( r^2 \) is small (by approximately 0.25 for Sheepdrove and approximately 0.2 for
Fig. 5. VWC$_1$ and rSSM time series, using the monthly multiple regression normalization factor, 100-m spatial resolution, and 14-orbit moving average for (a) and (b) Chimney Meadows, (c) and (d) Sheepdrove, and (e) and (f) Waddesdon. For subplots (a), (c), and (e), daily Pluviometer rain accumulations are plotted in blue, with dates where erroneous data are returned being plotted in gray. For subplots (b), (d), and (f), the VWC$_1$ and rSSM are directly compared, with the color denoting the month of observation.

Waddesdon). Together, the RMSE and $r^2$ patterns suggest that, while the rSSM time series cannot be used to estimate VWC$_1$ values (something that is to be expected, given the differences in measurement depths, methodologies, and assumptions made as part of the estimation), the general temporal trends are captured.
When looking at the four normalization factors used, it can be seen in Table IV that the RMSE values are not primarily impacted, with similar values being found for each normalization factor used, regardless of spatial resolution. The largest range of RMSE was found for the 250-m grid box size at Chimney Meadows, being just 1.1. The $r^2$ values, however, have a larger, spread of values for a given resolution. Indeed, the $r^2$ values also have a distinct difference between the monthly and annual normalization factors, with the annual normalization factors having a greater $r^2$ value for Chimney Meadows and Sheepdrove. For Waddesdon, this distinction between annual and monthly normalization factors cannot be seen.

It can also be seen, when looking at the values of RMSE and $r^2$ in Table IV, that using the monthly slopes slightly decreases the performance of the TWCDA compared to using the annual slopes. This slight performance drop may be due to both the direct and multiple monthly regression slopes, as well as the TWCDA, not explicitly taking the vegetation dynamics into account, as shown by the time lag between the peak NDVI value (taking place in May) and the peak value of monthly regression slopes (taking place in July), as shown in Fig. 1.

Using Fig. 5 and Table IV, it is clear that, while the TWCDA may not be appropriate for use of determining the rSSM of an exact 100-m pixel, the RMSE and $r^2$ values suggest that it would be appropriate for determining the general temporal trends at different spatial scales, whether that be catchment-wide or, to a lesser extent, in-field scales. As such, it shows that the TWCDA would be of use for land- and soil-based NFM measures, providing a general temporal pattern for areas of different (nonforest-based) NFM measures.

### V. Evaluation of rSSM Values Across Dominant Land Cover Types

Using Fig. 2, the rSSM time series across the entire timespan were produced for the two dominant land cover types across the AOI: Arable and Horticulture and Improved Grassland, as shown in Fig. 6, using the monthly multiple regression normalization factors, and a spatial resolution of 100 m. The raw temporal time series, similar to the analysis in Section IV, is subject to a 14-orbit moving average window in order to reflect the general trend in rSSM across both land uses.

It can be seen that both the Arable and Horticulture and the Improved Grassland time series follow a similar general temporal trend, with larger rSSM values during the later autumn, winter, and early spring (October–March), and with a rapid decrease in rSSM over March and April, before slowly increasing in rSSM until October. Looking at the time series in full, the Arable and Horticulture rSSM data trend to be larger over the wetter winter period with the Improved Grassland rSSM values being larger over the drier summer months; as shown in Fig. 7, displaying the monthly averages of rSSM.

Using the analysis of Fig. 6, an increase in rSSM estimations is present over the summer months. This increase is unexpected, as the summer months have reduced precipitation totals, as well as an increase in evapotranspiration, due to the water temperatures; both of which should lead to a decrease, not an increase, in rSSM values. This perceived increase in the rSSM signal is most likely, like that of Fig. 5, due to an increase in vegetation biomass above the surface. This increase in vegetation canopy would, thus, produce a contributing backscatter influence, which would artificially increase the rSSM value calculated by the TWCDA. This vegetation impact is more noticeable when looking at the mean annual cycle for both the Arable and Horticulture (and to a lesser extent) Improved Grassland land uses (see Fig. 7). The sudden decrease in rSSM between March and April coincides with both a reduction of rainfall, following the seasonal pattern of precipitation, and (in the case of Arable and Horticultural land use) the planting and early growth of vegetation. This increase in rSSM between May and July (especially with the Arable and Horticulture land cover), however, is counterintuitive to the expectation that soil moisture content would be lowest during the summer months due to the seasonality of both soil moisture and precipitation.

In addition to the driving factor of additional contribution to the backscatter from overlying vegetation, early harvesting at some locations may lead to a small increase in rSSM due to a reduction in transpiration from the now bare soil. Vegetation
Figs. 6 and 7. Mean rSSM time series using the monthly multiple regression normalization factors and spatial resolution of 100 m with a 14-orbit moving average for Arable and Horticulture (black) and Improved Grassland (red) across the AOI.

roots would stop absorbing moisture from the soil, as well as there being a reduced amount of interception from the now missing vegetation. The early harvesting of some Arable and Horticulture land, but not of Improved Grassland, goes some way to explain why there is a smaller increase in rSSM in the Improved Grassland than that of Arable and Horticulture in Fig. 7. The Improved Grassland areas may experience some grazing or mowing; however, grow-back would occur quite rapidly, thus reducing the impact of the loss of transpiration and interception of precipitation to the soil.

VI. CONCLUSION

In this study, Sentinel-1 SAR backscatter data have been used, in conjunction with the TWCDA, in order to assess temporal patterns of rSSM across the study area. In addition to using the TWCDA, numerous different normalization factors were applied, a direct linear relationship known as the direct regression slope and, in addition, a more complex multiple regression slope. Both the direct and the multiple regression slopes were calculated as a single annual value and at a novel monthly resolution. Wet and dry algorithm thresholds are calculated, and the sensitivity of the algorithm is briefly discussed with respect to land cover.

The spatial pattern of rSSM on two different orbits was compared against two localized precipitation events, showing that the spatial distribution of high and low rSSM values closely matches the spatial distribution of precipitation.

The rSSM time series were also compared against in situ data from three different sites within the COSMOS-UK network (Chimney Meadows, Sheepdrove, and Waddesdon) using four different spatial resolutions (1000, 500, 250, and 100 m), as well as four combinations of normalization factors (direct regression slope and multiple regression slope, at a traditional annual, and a novel monthly temporal resolution). Across all three sites, using any spatial resolution or normalization factor, the rSSM time series captures the general trend exhibited within the VWC$_1$ data although a difference in measurement depth does introduce variability into the observations, making day-to-day direct comparison of soil moisture values problematic.

It was shown that the two annual normalization factors yielded slightly smaller RMSE and a higher $r^2$ value for two of the COSMOS-UK sites than using the two monthly normalization factors. For the third site, Waddesdon, the differences in RMSE and $r^2$ values were negligible. In addition, the impact of using the direct regression slope or multiple regression slope
normalization factors was minimal, with similar RMSE and $r^2$ values across all three COSMOS-UK sites and across all four spatial resolutions.

The presence of a vegetation effect, not accounted for by the TWCD, was noted in the rSSM time series over the growing season (May–October), with an artificial increase in rSSM being present due to an added contribution to the backscatter signal by the growing vegetation. This vegetation impact could be seen in the two dominant land use types: Arable and Horticulure, and Improved Grassland. The vegetation impact was shown to have a greater artificial rise in rSSM in the Arable and Horticulure time series although it was still significant within the Improved Grassland time series. It should be noted that early harvesting in summer (July–August) at some locations in the study area could lead to an increase in rSSM due to reductions in evaportranspiration and interception from the nonexisting vegetation and bare soil in the Arable and Horticulure signal.

The retrieved rSSM time series indicates that the TWCD is capable of recording and measuring rSSM trends over time. The comparison with VWC data indicates that rSSM trends can be recorded at sub-kilometer scales; however, additional analysis using a larger sample of the COSMOS-UK could give further evidence to this.

Sub-kilometer rSSM data retrieval can be used to investigate soil moisture changes and trends across different spatial scales, from large river basins to smaller catchments and subcatchments, and interfield and intrafield scales. For NFM, this could enable exploring the role of different land uses on soil water storage, including looking at the impact of soil types. Beyond NFM, there are wider applications to hydrological modeling, where soil moisture is needed as an input or as verification model, there are wider applications to hydrological modeling, including looking at the impact of soil types. Beyond NFM, there are wider applications to hydrological modeling, where soil moisture is needed as an input or as verification.

REFERENCES

[1] V. Chow, D. Maidment, and L. Mays, Applied Hydrology. New York, NY, USA: McGraw-Hill, 1988.
[2] D. R. Legates et al., “Soil moisture: A central and unifying theme in physical geography,” Prog. Phys. Geog., vol. 35, no. 1, pp. 65–86, Nov. 2011, doi: 10.1177/0309133310386514.
[3] K. L. Findell and E. A. B. Eltahir, “Atmospheric controls on soil moisture-boundary layer interactions. Part I: Framework development,” J. Hydrometeorol., vol. 4, no. 3, pp. 552–569, Jun. 2003, doi: 10.1175/1525-7541(2003)004<0552:ACOSM2.0.CO;2.
[4] H. Zhang, J. Liu, H. Li, X. Meng, and A. Ablikim, “The impacts of soil moisture initialization on the forecasts of weather research and forecasting model: A case study in Xinjiang, China,” Water, vol. 12, no. 7, p. 1892, Jul. 2020, doi: 10.3390/w12071892.
[5] J. Liu, B. A. Engel, Y. Wang, Y. Wu, Z. Zhang, and M. Zhang. “Runoff response to soil moisture and micro-topographic structure on the plot scale,” Sci. Rep., vol. 9, no. 1, p. 2532, Dec. 2019, doi: 10.1038/s41598-019-39409-6.
[6] M. G. Grillakis, A. G. Koutroulis, J. Komma, I. K. Tzanis, W. Wagner, and G. Blöschl, “Initial soil moisture effects on flash flood generation—A comparison between basins of contrasting hydro-climatic conditions,” J. Hydrol., vol. 541, pp. 206–217, Oct. 2016, doi: 10.1016/j.jhydrol.2016.03.007.
[7] M. Zeri et al., “Importance of including soil moisture in drought monitoring over the Brazilian semiarid region: An evaluation using the Jules model, in situ observations, and remote sensing,” Climate Resilience Sustainability, vol. 1, no. 1, pp. 1–18, Feb. 2021, doi: 10.1002/2021.0001.
[8] S. I. Seneviratne et al., “Investigating soil moisture-climate interactions in a changing climate: A review,” Earth Sci. Rev., vol. 99, nos. 3–4, pp. 125–161, 2010, doi: 10.1016/j.earscirev.2010.02.004.
William Maslanka received the M.Met. degree in meteorology and the Ph.D. degree in remote sensing and radiative modeling of snow from the University of Reading, Reading, U.K., in 2013 and 2017, respectively. From 2017 to 2019, he was a Flood Modeller and a Forecaster for the Environment Agency, before working as a Post-Doctoral Researcher with the LANDWISE Project, focusing on the remote sensing of soil moisture, for use with natural flood management. His research interests include microwave remote sensing of soil moisture, remote sensing of snow, and the application of drones to remote sensing techniques.

Keith Morrison (Member, IEEE) received the B.Sc. degree in physics with astrophysics from the University of Leicester, Leicester, U.K., in 1983, and the Ph.D. degree in astronomy and astrophysics from the University of St Andrews, St Andrews, U.K., in 1987. He has carried out extensive fieldwork campaigns across Europe, synergistically with indoor laboratory studies. He was with British Antarctic Survey, Cambridge, U.K., from 1987 to 1994. The University of Sheffield, Sheffield, U.K., from 1994 to 2000, the Defence Evaluation and Research Agency (DERA)/QinetiQ, Malvern, U.K., from 2000 to 2002, and Cranfield University, Cranfield, U.K., from 2002 to 2016. Since 2016, he has been a Professor with the Meteorology Department, University of Reading, Reading, U.K. His main research interests are the development and application of ground- and drone-based synthetic aperture radar (SAR) imaging and signal processing techniques applied to environmental remote sensing, namely, snow, soil, crops, and forestry. Prof. Morrison was a recipient of the NASA's Group Achievement Award in 1998 and the Annual Best Publication Award in 2001 from Taylor & Francis Publishing for one of the first demonstrations of 3-D SAR tomography. He held a Joint-Chair at the National Physical Laboratory from 2018 to 2021.
Kevin White received the Ph.D. degree from the University of Reading, Reading, U.K., in 1990. He subsequently worked at the University of Oxford, Oxford, U.K., and the University of Canterbury, Christchurch, New Zealand, before returning to his current post at the University of Reading. He has over 30 years of experience in conducting remote sensing research, mostly focused on soils, geology, and vegetation in dryland regions. He is currently an Associate Professor of environmental remote sensing with the School of Archaeology, Geography and Environmental Science, University of Reading. He is currently working on remote sensing of cyanobacterial soil crusts, monitoring dryland sediment transport processes, and unmanned aerial vehicle (UAV) survey methods for site characterization in geomorphology, vegetation studies, archeology, and agronomy.

Anne Verhoef received the combined B.Sc./M.Sc. degree in agricultural and environmental sciences and the Ph.D. degree in micrometeorology from Wageningen University & Research, Wageningen, The Netherlands, in 1990 and 1995, respectively. She has extensive expertise in land surface modeling (LSM) and experimental campaigns relating to heat, water, and gas transfer in soils, as well as to the land surface energy, water, and carbon balance. She has a strong interest in the synergistic use of LSM, field observations, and remote sensing data. Over the years, she has built up a large portfolio of research grants mostly relating to applied and theoretical environmental science, broadly in the context of water, food, and energy security, and related impacts of weather extremes (flooding, droughts, and heatwaves). She is currently a Professor of soil physics and micrometeorology with the University of Reading, Reading, U.K.

Dr. Verhoef has been an Executive Board Member of the International Soil Modelling Consortium since 2017 and co-chairing the GEWEX GLASS Panel since 2020, where she guides international developments relating to land surface modeling (e.g., in the context of Intergovernmental Panel on Climate Change (IPCC) model intercomparison projects).

Joanna Clark received the B.Sc. degree in geography from Durham University, Durham, U.K., in 1999, the M.Sc. degree in monitoring, modeling, and management of environmental change from King’s College London, London, U.K., in 2000, and the Ph.D. degree in physical geography from the University of Leeds, Leeds, U.K., in 2005. She was a Post-Doctoral Researcher with the University of Leeds from 2005 to 2009 on biogeochemistry of soils and catchments, looking at the impact of declining acid rain and climate change on dissolved organic carbon release from organic soils to streams and lakes. She spent a year working on a knowledge exchange project on climate change and the British Uplands at Bangor University, Bangor, U.K., from 2008 to 2009, including coordinating workshops, journal special issues, model intercomparison, and developing bioclimatic models for blanket peatland. In 2010, she started a fellowship at the Grantham Institute for Climate Change, Imperial College on Peatlands and Climate Change. She moved on to a Lectureship in soil carbon at the University of Reading, Reading, U.K., later in 2010, where she was promoted to an Associate Professor of environmental science in 2016 and a Professor of environmental science in 2019. She works on carbon and water cycles from soil to catchment scales, particularly the impacts on carbon sequestration, natural flood management, and drinking water quality.

Dr. Clark is also a member of the British Hydrological Society, the British Ecological Society, and the British Society of Soil Science. She is also a Principal Investigator of the NERC Landwise Project.