Widespread occurrence of anomalous C-band backscatter signals in arid environments caused by subsurface scattering

Wolfgang Wagner, Roland Lindorfer, Thomas Melzer, Sebastian Hahn, Bernhard Bauer-Marschallinger, Keith Morrison, Jean-Christophe Calvet, Stephen Hobbs, Raphael Quast, Isabella Greimeister-Pfeil, et al.

Widespread occurrence of anomalous C-band backscatter signals in arid environments caused by subsurface scattering

Wolfgang Wagner [a,*], Roland Lindorfer [a], Thomas Melzer [a], Sebastian Hahn [a], Bernhard Bauer-Marschallinger [a], Keith Morrison [b], Jean-Christophe Calvet [c], Stephen Hobbs [d], Raphael Quast [a], Isabella Greimeister-Pfeil [a], Mariette Vreugdenhil [a]

[a] *Department of Geodesy and Geoinformation, Technische Universität Wien, Vienna, Austria*
[b] *Department of Meteorology, University of Reading, Reading, United Kingdom*
[c] *CNRM, Université de Toulouse, Meteo-France, CNRS, Toulouse, France*
[d] *Cranfield University, Cranfield, United Kingdom*

ABSTRACT

Backscatter measured by scatterometers and Synthetic Aperture Radars is sensitive to the dielectric properties of the soil and normally increases with increasing soil moisture content. However, when the soil is dry, the radar waves penetrate deeper into the soil, potentially sensing subsurface scatterers such as near-surface rocks and stones. In this paper we propose an exponential model to describe the impact of such subsurface scatterers on C-Band backscatter measurements acquired by the Advanced Scatterometer (ASCAT) on board of the METOP satellites. The model predicts an increase of the subsurface scattering contributions with decreasing soil wetness that may counteract the signal from the soil surface. This may cause anomalous backscatter signals that deteriorate soil moisture retrievals from ASCAT. We test whether this new model is able to explain ASCAT observations better than a bare soil backscatter model without a subsurface scattering term, using k-fold cross validation and the Bayesian Information Criterion for model selection. We find that arid landscapes with Leptosols and Arenosols represent ideal environmental conditions for the occurrence of subsurface scattering. Nonetheless, subsurface scattering may also become important in more humid environments during dry spells. We conclude that subsurface scattering is a widespread phenomenon that (i) needs to be accounted for in active microwave soil moisture retrievals and (ii) has a potential for soil mapping, particularly in arid and semi-arid environments.

1. Introduction

One attractive feature of microwave remote sensing techniques is that their carrier signals are able to penetrate clouds, vegetation, and soils much better than optical and infrared waves can (Ulaby et al., 1981). Therefore, microwave remote sensing can provide information about the physical properties of vegetation and soil undisturbed by atmospheric conditions. In the case of bare soil surfaces, the penetration depth is in the order of millimetres to decimetres depending on the wavelength and wetness conditions of the soils. Particularly in arid regions, penetration depths may be quite large, allowing to map bedrock and gravel surfaces beneath windblown sand several centimetres to possibly meters thick (McCauley et al., 1982; Schaber et al., 1986). Experimental measurements of radar transmission through dry sand

conducted by Williams and Greeley (2001) showed that the signal passes through hyperarid sand (0.3 vol%) up to 50 cm thick with a decrease in signal of less than 6 dB over the frequency range of 0.5–12.6 GHz. But even under moister conditions (4.7 and 10.7 vol%) the penetration was quite large, particularly at the lower frequencies. Nonetheless, when interpreting microwave measurements, the fact that microwave signals may be sensitive to subsurface soil properties is often ignored. Instead, soils are usually treated as pure surface scatterers or emitters. The implicit assumption is that the sensed topmost soil layers can be regarded as a homogeneous dielectric medium. This is for example common practice when retrieving soil moisture and vegetation from both active and passive microwave measurements (Petropoulos et al., 2015; Steele-Dunne et al., 2017).

For the active case, this assumption was driven by the insight that the

---

contribution of scattering from subsurface soil strata to total backscatter is usually small and decreases rapidly with increasing soil moisture content (Schanda, 1986). Therefore, all commonly used bare soil backscatter models ignore subsurface scattering, modelling backscatter just in terms of the geometric and dielectric properties of the soil surface. For characterising the geometric properties of the soil surface, the most commonly used parameters are the root mean square height and auto-correlation length (Verhoest et al., 2008). The soil dielectric properties are a function of water content and soil textural composition (Dobson et al., 1985). All these models have in common that they predict – irrespective of the surface roughness conditions – a monotonic increase of backscatter with increasing surface soil moisture. This is true for theoretical (Bahar, 1981; Fung et al., 2002), semi-empirical (Wegmuller et al., 1994) and empirical models (Dubois et al., 1995; Oh et al., 1992) alike. For the theoretical models this behaviour can be traced back to the Fresnel equations that describe how the reflectivity of a plane interface between two homogeneous media increases with the dielectric contrast between the two media. Guided by the theoretical models and trained by field data sets collected during field experiments (Ulaby et al., 1978; Wegmuller et al., 1994), also all semi-empirical and empirical models predict that the backscattering coefficient as measured by radar instruments increases with soil moisture. This means that, all other things being equal, soil moisture and backscatter are expected to be positively correlated over bare soil surfaces, while a negative correlation would be considered an anomaly.

Surprisingly, anomalous backscattering signals appear to occur worldwide much more often than anticipated. The evidence comes mostly from global studies of C-band backscatter data acquired by the Metop Advanced Scatterometer (ASCAT) and its predecessor, the ERS scatterometer (Wagner et al., 2013). As shown by Fig. 1, ASCAT backscatter exhibits strong negative correlations with modelled soil moisture data over many arid and semi-arid regions. Also global backscatter data sets acquired in L- and Ku-band exhibit spurious backscatter signals over arid environments (Jaruwatanadilok and Stiles, 2014; McColl et al., 2014), suggesting that the problem is not just confined to C-band. Probably an important reason why these anomalies have received little attention so far is that they occur mostly in remote, sparsely vegetated areas, where no or only few in situ data are available. Hence, much of our current knowledge of the phenomenon stems from comparisons with global precipitation and modelled soil moisture data sets (Wagner et al., 2013). Further evidence comes from inter-comparisons with soil moisture data derived from L-band (Kerr et al., 2012; Entekhabi et al., 2010) and higher-frequency (de Jeu et al., 2008) brightness temperature observations. While also brightness temperature data are sensitive to deeper soil layers, they appear to be much less affected by subsurface

effects than the backscatter measurements (Dorigo et al., 2010; Fascetti et al., 2016; Miyaoka et al., 2017). Noting that passive microwave measurements are much less sensitive to the roughness of targets than the highly directional radar measurements (Schanda, 1986), this suggests that roughness effects play an important role in explaining the presence (absence) of anomalies in active (passive) bare soil measurements.

A hypothesis for explaining the anomalies seen in ASCAT data was put forward by Wagner et al. (2013). They speculated that such anomalies may be due to the presence of strong subsurface scatterers that increasingly make up the total backscatter signal when the soil dries. A similar hypothesis was made by McColl et al. (2014) for explaining relatively high L-band backscatter values over desert areas as observed by the Aquarius scatterometer. Such strong scatterers could e.g. be rock surfaces beneath shallow soil layers or small rocks and stones distributed throughout the soil profile. This hypothesis was tested by Morrison and Wagner (2020) in a laboratory experiment that collected high-resolution vertical C-band backscatter profiles of (i) a layer of sand with a flat surface overlying a rough subsurface layer, and (ii) a layer of sand randomly mixed with stones. The results confirmed that soils possessing a distinct, brightly reflecting subsurface layer can produce an anomaly. It is enhanced when the soil surface is smooth and the underlying subsurface rough. Conversely, it is dampened when the soil is rough with stony inclusions. Still, depending on the exact composition of the soil and the characteristics of the radar (frequency, polarisation, incidence angle), anomalies are possible. These findings are in line with the results of an earlier field experiment by Liu et al. (2016) where the authors acquired co- and cross-polarised L-band data over a bare sandy soil in Florida. They found that for the same field plot, the presence of an anomaly is dependent on the roughness of the soil surface, i.e. when the soil surface was rough, surface scattering was always the dominant mechanism. However, when the surface was smooth, a subsurface signal emerged in the co-polarised backscatter data (VV and HH) for soil moisture levels below 7%. For soil moisture levels higher than this threshold, backscatter increased as expected. This resulted in a V-shaped soil moisture-backscatter relationship (see Fig. 4 in Liu et al. (2016)). The cross-polarised VH data showed no anomaly.

The goal of this study is to gain a better understanding of subsurface scattering effects as observed by ASCAT. Where and when does subsurface scattering occur? How strong are the effects and how do they relate to environmental conditions? As study region we chose a 15° by 15° large area covering the Iberian Peninsula and the north-western parts of the African continent (10°W-5°E, 30–45°N) as highlighted by the box in Fig. 1. The European part of the study domain is where Wagner et al. (1999) had first encountered inexplicable backscatter

Fig. 1. Pearson correlation between multi-year ASCAT backscatter time series (at an incidence angle of 40°) and modelled soil moisture data from ERA5-Land (Layer 1) under snow and frost free conditions. Over many arid regions worldwide the correlation is negative, pointing to an anomalous backscattering behaviour. The red box shows our area of investigation. Details about the ASCAT and ERA5-Land data sets are provided in Section 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Pearson correlation coefficient

-1.0   -0.5   0.0   0.5   1.0

anomalies during summer, and where Shamambo et al. (2019) recently noted anomalous backscatter behaviour over karst regions. As the starting point of our analysis, we introduce a backscatter model containing a subsurface scattering term in the next section. By confronting three years of ASCAT backscatter measurements with modelled soil moisture data we test whether the new model is able to explain the ASCAT observations better than a bare soil backscatter model without the subsurface scattering term. We complement this model assessment by an analysis of how the observed spatiotemporal patterns of subsurface scattering correspond to climate, soil properties, and land cover.

2. Theory

Anomalous backscatter signals have so far been mostly noted in arid environments with no or scarce vegetation cover. Nonetheless, for understanding the spatial dimension of this phenomenon, we must also be able to describe the impact of vegetation on the backscatter signature of soils with subsurface scatterers. Therefore, in this section we formulate a subsurface scattering model for soils with and without vegetation cover.

2.1. Bare soil backscatter with subsurface scattering

Let us conceptualise the subsurface scattering problem for a bare soil by assuming that backscatter is the sum of surface contributions and an attenuated subsurface signal:

$$\sigma_{soil}^0 = \sigma_{top}^0 + \Gamma_{soil}^2 \sigma_{sub}^0 \tag{1}$$

where σ_{soil}^0 is the bare soil backscattering coefficient in linear units ($m^2 m^{-2}$), σ_{top}^0 is the backscatter contribution coming directly from the soil surface, Γ_{soil}^2 is the two-way attenuation factor describing how much energy the microwave pulse loses when travelling forth and back through the intermediate soil layer, and σ_{sub}^0 is the backscatter from subsurface targets. Subsurface scatterers may e.g. be buried rock or stones dispersed throughout the soil profile (Fig. 2). For simplicity let us assume that the subsurface scattering cross section is a constant, i.e. $\sigma_{sub}^0 = \psi$.

Both σ_{top}^0 and Γ_{soil}^2 depend on the soil moisture content θ within the topmost soil layer, but in opposite directions. Assuming exponential relationships, we write

$$\sigma_{soil}^0 = \alpha e^{\beta\theta} + \psi e^{-\xi\theta} \tag{2}$$

where α corresponds to the surface scattering contribution when the soil is dry ($\theta = 0$) and β prescribes the sensitivity of σ_{top}^0 to soil moisture changes. These two parameters depend on the roughness and texture of the soil surface layer. The parameter ξ regulates the strength of the

attenuation of the subsurface scattering signals by the intermediate soil layer and depends on a multitude of factors such as soil texture and the nature and depth of the subsurface scatterer, as does ψ. Note that the logarithmic transform of the term $\alpha e^{\beta\theta}$ corresponds to the widely used linear model for relating σ_{soil}^0 expressed in decibels to θ (Dobson and Ulaby, 1986; Kim and van Zyl, 2009). Furthermore, note that soil moisture is usually expressed in terms of the volumetric soil moisture content in units of $m^3 m^{-3}$, whereas here we prefer to use a relative soil moisture index ranging between 0% (dry soil) and 100% (wet soil), representing the degree of saturation of the remotely sensed topsoil layer.

The scattering behaviour as predicted by Eq. (2) is illustrated by the three graphics shown in Fig. 3. The left graphic (a) represents a land surface with a dominant surface scattering contribution. In this case, backscatter increases monotonically with soil moisture over the complete wetness range, i.e. no anomalies are present. This is the standard scattering regime that is expected to occur over the vast majority of land surface areas. The opposite behaviour occurs when subsurface scattering dominates over the scattering contributions from the soil surface (Fig. 3c). In this case, σ_{soil}^0 decreases monotonically with soil moisture, i.e. an anomaly is always present.

Between these two limit cases, Fig. 3b illustrates a mixed scattering regime where the subsurface scattering signal shapes the backscatter curve under dry conditions, while under wet conditions the surface signal dominates. This results in a kind of U-shaped curve comparable to the more pronounced V-shaped relationship observed by Liu et al. (2016) in their local-scale field measurements. In such a situation, there exists a turning point $\theta_{turn} \in [0,1]$ where σ_{soil}^0 is minimal. It can be found by setting $\partial\sigma_{soil}^0/\partial\theta$ to zero:

$$\theta_{turn} = \frac{1}{\beta + \xi} ln \frac{\psi\xi}{\alpha\beta} \tag{3}$$

For soil moisture values smaller than θ_{turn} one will observe backscatter anomalies, while for values $\theta > \theta_{turn}$ backscatter is positively correlated with soil moisture. Such a mixed scattering regime becomes apparent when $\psi\xi \geq \alpha\beta$ and $(\beta + \xi) \geq ln\frac{\psi\xi}{\alpha\beta}$, i.e. subsurface scattering must be sufficiently strong to lead to an initial decrease of total backscatter at least under dry conditions, but not too strong to still dominate under wet conditions. Note that such a mixed scattering regime is obviously problematic from the point of view of soil moisture retrievals as there is no unique mapping of one backscatter value to one soil moisture value. Instead, depending on the wetness conditions, a backscatter increase could signify both a wetting (at the wet edge) or drying (at the dry edge) of the soil surface. To resolve this ambiguity one would either need additional radar observations (multiple polarisations, incidence angles, or frequencies) (Morrison and Wagner, 2020) or other
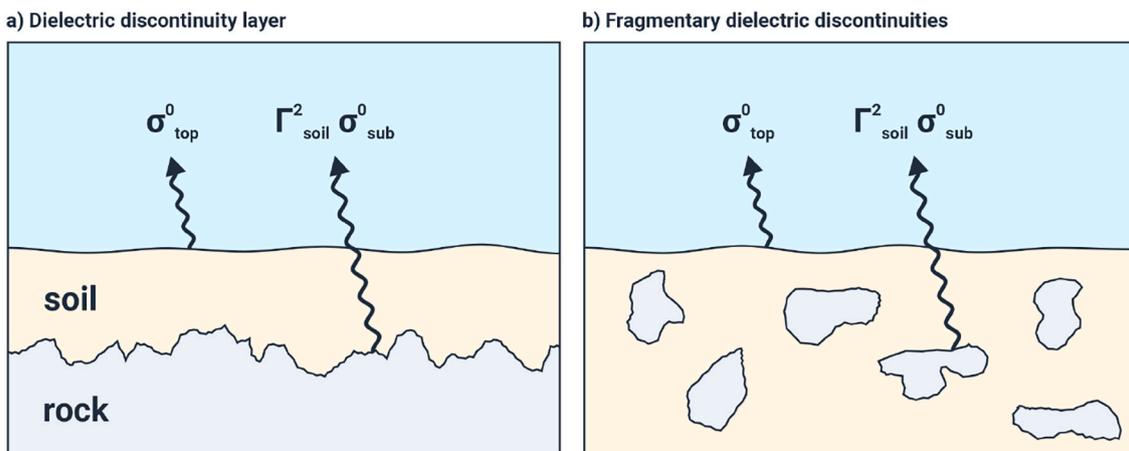

Fig. 2. Bare soil backscatter with contributions from the surface (σ_{top}^0) and subsurface dielectric discontinuities (σ_{sub}^0) dampened by the soil medium (Γ_{soil}^2).
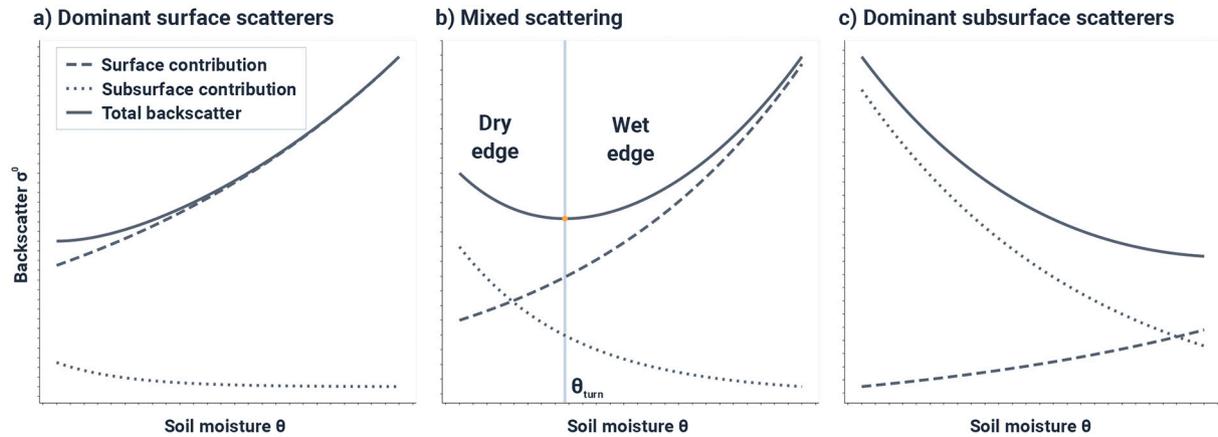
Fig. 3. Backscattering behaviour of a bare soil with (a) dominant surface, (b) mixed, and (c) dominant subsurface scatterers.

contextual information.

2.2. Subsurface scattering signals in vegetated regions

To account for vegetation, we use the widely used Water Cloud model introduced by Attema and Ulaby (1978) which models backscatter from vegetation by assuming that the vegetation canopy can be conceptualised as a layer of identical and randomly distributed water droplets. This model represents a zeroth-order solution of the radiative transfer equation applied to a tenuous distribution of particulate media overlying a rough surface. According to this model, the principal effects of vegetation are to dampen the bare soil backscatter signal and to add a contribution from the vegetation canopy:

$$\sigma^0 = \Gamma_{veg}^2 \sigma_{soil}^0 + \sigma_{veg}^0 \qquad (4)$$

where σ^0 is the backscatter coefficient measured by a radar sensor, Γ_{veg}^2 is the two-way attenuation factor describing the two-way loss of energy through the vegetation, and σ_{veg}^0 is the volume scattering contribution from the canopy. Substituting Eq. (2) yields

$$\sigma^0 = \widehat{\alpha} e^{\beta\theta} + \widehat{\psi} e^{-\xi\theta} + \sigma_{veg}^0 \qquad (5)$$

where $\widehat{\alpha}$ and $\widehat{\psi}$ are the surface- and subsurface scattering terms dampened by the vegetation layer, i.e. $\widehat{\alpha} = \Gamma_{veg}^2 \alpha$ and $\widehat{\psi} = \Gamma_{veg}^2 \psi$. Other than α and ψ, $\widehat{\alpha}$ and $\widehat{\psi}$ vary over the year reflecting the seasonal vegetation development. This is also true for σ_{veg}^0. However, the two exponential coefficients β and ξ are not affected by vegetation, nor is the ratio $\frac{\widehat{\psi}\xi}{\widehat{\alpha}\beta}$. Hence, the general U-shape and in particular the position of the turning point θ_{turn} should remain unaltered by the presence of vegetation and not change over the year. Of course, over dense vegetation Γ_{veg}^2 becomes so small that changes in σ_{soil}^0 cannot be reliably discerned any longer. Therefore, over dense forests and shrub land it becomes impossible to discern subsurface scattering effects.

Finally, let us specify the signal range of the two scattering terms from completely dry ($\theta = 0\%$) to wet ($\theta = 100\%$) conditions with

$$\mathscr{S}_{top} = \widehat{\alpha}\left(e^{\beta} - 1\right) \qquad (6)$$

$$\mathscr{S}_{sub} = \widehat{\psi}\left(1 - e^{-\xi}\right) \qquad (7)$$

where $\mathscr{S}_{top}$ is the signal range of the surface contributions and $\mathscr{S}_{sub}$ to the subsurface scatterers. Both quantities do not just reflect soil properties but also the optical thickness of the vegetation (via Γ_{veg}^2), and inform us about how sensitive the two scattering terms are to soil moisture changes. Note that $\mathscr{S}_{top}$ applies also to the backscatter model without the subsurface scattering term.

3. Study region and data

Our study region spans the longitudes from $10°$W to $5°$E and the latitudes from $30°$N to $45°$N, covering mainland Portugal and Spain, south-western France and the northern parts of Morocco and Algeria (Fig. 4). It was chosen to cover a wide range of environmental conditions governing subsurface scattering, from hyperarid regions in the Sahara with strong and permanent subsurface signals, to semi-arid conditions in central and south-eastern Spain with weaker, intermittent anomalies (Wagner et al., 1999; Wanders et al., 2012; Escorihuela and Quintana-Seguí, 2016) and more humid regions in the northern and north-westerly parts of the study domain, where subsurface scattering is unlikely to occur. The gradient from arid to more humid conditions is reflected by land cover patterns (from bare to densely vegetated areas) and soil properties (from sandy and shallow soils to well developed fertile soils). To analyse the impacts of climate, soil type, and land cover on subsurface scattering, we use the Köppen-Geiger climate classification by Beck et al. (2018), the soil groups from the ISRIC SoilGrids250m map (Hengl et al., 2017), and the land cover map produced within the framework of the European Space Agency's Climate Change Initiative (ESA CCI) (ESA, 2017). Furthermore, we included the World Karst Aquifer map (Chen et al., 2017) in our analysis given that rough karstic rocks may cause strong backscatter returns and have been found to cause unexpected backscatter behaviour (Shamambo et al., 2019). Fig. 4 shows for each ASCAT grid point the dominant classes for each of these four maps. The dominant classes were determined by calculating the class frequency within each ASCAT pixel (squares of 12.5 × 12.5 km) and selecting for each grid point the most frequently occurring class.

3.1. ERA5-land soil moisture

As relative soil moisture index, θ, we use the 0–7 cm soil moisture simulations of ERA5-Land (Muñoz-Sabater et al., 2021) and scale them between minimum and maximum values for each pixel. ERA5-Land is the off-line and independently operated land model component of the 5th generation of the European ReAnalysis (ERA5) system (Hersbach et al., 2020) operated by the European Centre for Medium-Range Weather Forecasts (ECMWF). It uses ERA5 for atmospheric forcing and the Carbon Hydrology-Tiled ECMWF Scheme for Surface Exchanges over Land (CHTESSEL) model to simulate soil moisture and many other land surface variables on a 9 km grid with hourly time steps. Other than ERA5, which assimilates ASCAT soil moisture data, ERA5-Land does not assimilate land surface observations, i.e. the soil moisture data from ERA5-Land and the ASCAT backscatter measurements can be treated as independent variables. Also important for this study is that CHTESSEL simulates soil moisture dynamics much better than the older TESSEL scheme used in past re-analysis. The mean value and standard deviation
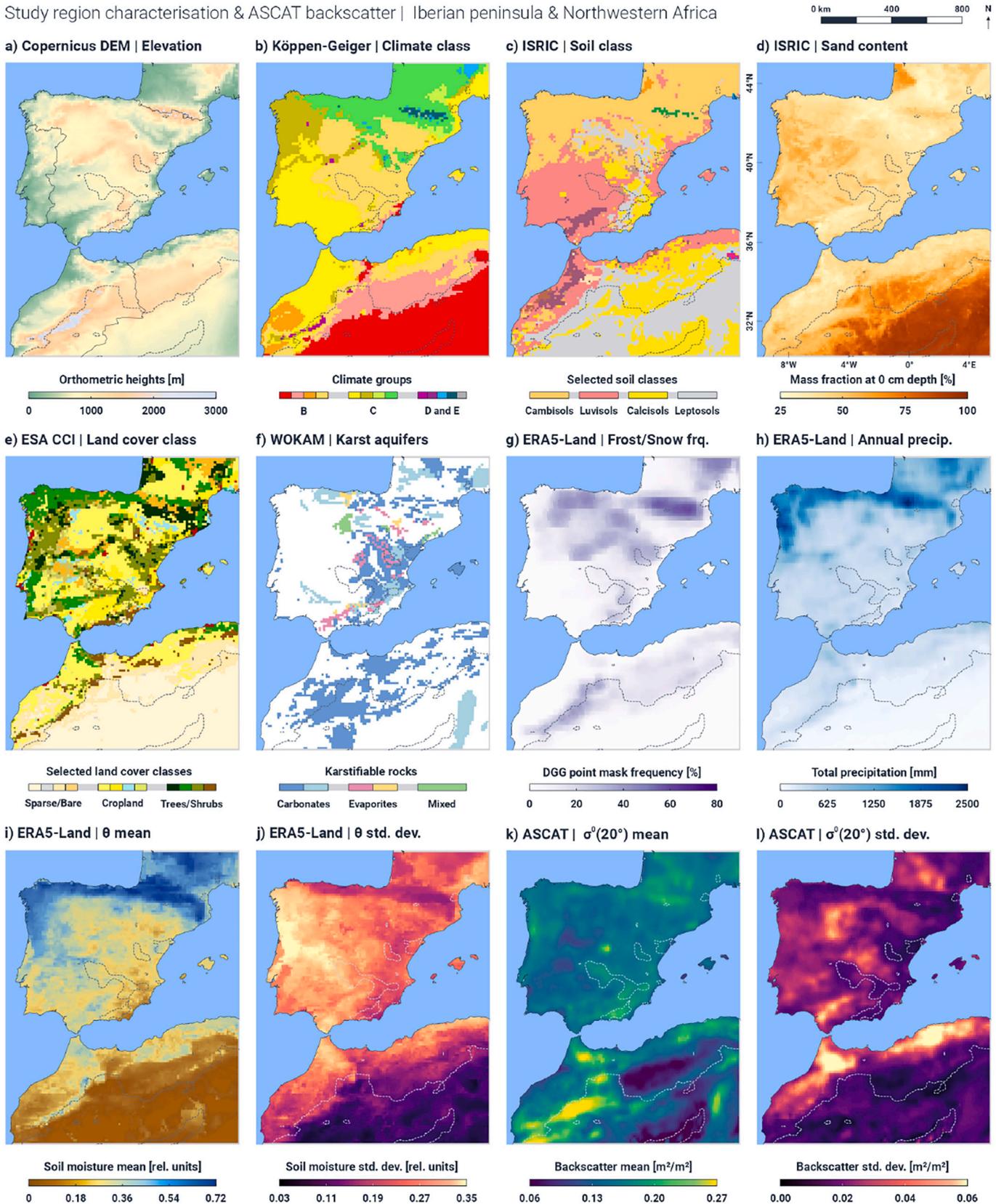
Fig. 4. Study region characteristics and ASCAT backscatter: a) elevation and country boundaries (solid lines), b) dominant Köppen-Geiger climate classes per ASCAT pixel, c) dominant soil groups, d) sand fraction, e) dominant ESA CCI land cover, f) karst acquifer map, g) number of ASCAT measurements masked due to frost or snow cover (in %), h) mean annual precipitation (mm) from ERA5-Land, i) mean of ERA5-Land soil moisture (relative units), j) standard deviation of ERA5-Land soil moisture (relative units), k) mean of ASCAT backscatter at 20° incidence angle over all snow- and frost free days (in m^2m^{-2}), l) standard deviation of ASCAT backscatter (m^2m^{-2}). The dotted contour line shows the $R = 0$ isoline from the Pearson correlation plot between ASCAT backscatter and ERA5-Land soil moisture shown in Fig. 1.

of θ, calculated for the years 2015–2017, are shown in Fig. 4. One can observe the expected soil moisture gradients with low and rather constant θ values in the desert regions, and higher and more variable θ values in the more humid parts of the study domain.

3.2. ASCAT backscatter

The Advanced Scatterometer (ASCAT) is a C-band (5.255 GHz) fan beam scatterometer that uses six antennas to acquire backscatter triplets in VV polarisation along two 550 km wide swaths (Figa-Saldaña et al., 2002). For the two mid beam antennas, the incidence angle varies over each swath from 25° in near range to 55° in far range. For the fore and aft beam antennas, the range is 34–65°. Thanks to its multiple-viewing capability, it is possible to characterise the dynamic relationship between the ASCAT backscatter measurements and the incidence angle (Hahn et al., 2017), which in turn allows to extrapolate the ASCAT triplets to any chosen reference angle. While the standard reference angle used in the EUMETSAT H SAF processing chain to retrieve surface soil moisture from ASCAT is 40° (Wagner et al., 2013), in this study we extrapolated the data to 20° in order to minimise the effects of seasonal vegetation development on σ^0 (Wagner et al., 1999; Hahn et al., 2021). Azimuthal effects are corrected using the methods introduced by Bartalis et al. (2006), which is important for reducing the noise of the extrapolated backscatter measurements particularly in sloping to steep terrain.

We constructed $\sigma^0(20)$ time series for the years 2015 to 2017 by combining data from the two ASCAT instruments flown on board of METOP-A (2006–2021) and METOP-B (launch 2012). This is possible thanks to the excellent cross-calibration of the two sensors (Anderson et al., 2017). The data have a spatial resolution of 25 km and are sampled to a 12.5 km grid. In total, there are 9746 ASCAT land pixels in our study domain. Backscatter values affected by frost or the presence of snow were masked using ERA5-Land soil temperature ($\leq 0°C$) and snow depth data (> 0 mm). The percentage of masked ASCAT measurements (in %) is shown in Fig. 4g. One can see that particularly over the high-altitude regions (Pyrenees, Spain's high plateaus, and Morocco's Atlas mountains) measurements are masked. The mean and standard deviation of the ASCAT backscatter measurements over all three years (after masking and normalisation to 20°) are shown in Fig. 4 as well. Their spatial patterns reflect land cover and surface roughness properties, with Mean(σ^0) being lowest over sand dunes and highest over bare rocks and urban areas. The standard deviation StDev(σ^0) is highest over agricultural and grassland regions with high soil moisture variability, while it is low over dense vegetation and regions with low soil moisture variability.

4. Methods

We investigate our subsurface scattering theory as introduced in Section 2 by assessing under which environmental conditions the backscatter model with a subsurface scattering term is able to describe the relationship between ASCAT backscatter and ERA5-Land soil moisture better than the model without a subsurface scattering term. Furthermore, we examine whether the observed model behaviour may plausibly be explained by the presence of buried rocks, stones and pebbles within a sandy soil horizon or other subsurface scatterers such as e.g. described by Schaber et al. (1986). Throughout this text we use the least squares criterion to quantify the goodness of fit of the models to the observations. Let us denote our land surface backscatter model without the subsurface scattering terms with $\mathscr{M}_0$, and the corresponding one with the subsurface scattering term with $\mathscr{M}_1$:

$$\mathscr{M}_0 : \sigma^0 = f_0(\theta) = c_\sigma + \widehat{\alpha} e^{\beta\theta}$$
$$\mathscr{M}_1 : \sigma^0 = f_1(\theta) = c_\sigma + \widehat{\alpha} e^{\beta\theta} + \widehat{\psi} e^{-\xi\theta} \tag{8}$$

where σ^0 is the backscattering coefficient measured by ASCAT in m²m⁻² and θ is the relative soil moisture content in % from ERA5-Land. In these

two models we have replaced the vegetation term σ^0_{veg} from Eq. (5) with a more generic constant backscatter term c_σ to account for the fact that ASCAT, due to its coarse resolution, inevitably senses not just soils and vegetation but also many other land features that have an impact on the overall magnitude of the ASCAT signal. Notably, urban areas and exposed rocks enhance backscatter strongly, whilst inland water bodies decrease backscatter (Wagner et al., 1999). Recall from Section 2.2 that the terms $\widehat{\alpha}$ and $\widehat{\psi}$ vary over the seasons reflecting vegetation development. Because time-dependency would complicate the tasks of model fitting and model selection quite significantly, we work with ASCAT backscatter measurements standardised to an incidence angle of 20°. This minimises seasonal vegetation effects and allows us to treat $\widehat{\alpha}$ and $\widehat{\psi}$ in a first approximation as constants. From now on, unless otherwise noted, we use the symbol σ^0 to represent ASCAT backscatter measurements extrapolated to 20°.

To obtain collocated ASCAT and ERA5-Land data pairs we used the nearest neighbour method. In space, we selected for each ASCAT grid point the closest ERA5-Land point. Temporal matching was done by choosing from the hourly ERA5-Land soil moisture time series the simulation closest to the ASCAT overflight, which is around ~9:30 local time for the descending pass and ~21:30 for the ascending pass.

4.1. Model fitting

We fitted the two models $\mathscr{M}_0$ and $\mathscr{M}_1$ to three years (2015–2017) of collocated ASCAT σ^0 (in m²m⁻²) and ERA5-Land θ data pairs for each ASCAT land pixel. Before fitting the non-linear model terms, we determined the constant c_σ by calculating the mean σ^0 value for ten soil moisture intervals and selecting the minimum from the resulting ten values. For a grid point without subsurface scattering, c_σ will normally represent the mean σ^0 value within the [0,0.1[soil moisture interval, while for a point with subsurface scattering it is supposed to represent a mean σ^0 value near the turning point θ_{turn} of the U-shaped curve. To illustrate this, Fig. 5 shows a scatter plot of σ^0 vs. θ for a region around the city of Ciudad Real in Central Spain. Over this region the Mean(σ^0)

Fig. 5. Scatter plot of ASCAT backscatter measurements at 20° versus relative soil moisture data from ERA5-Land over a region in central Spain (pixel centre: 39.059°N, 4.044°W). The red dashed lines show the mean σ^0 values for each of the ten intervals and the two solid lines the fitted models $\mathscr{M}_0$ (green line) and $\mathscr{M}_1$ (orange). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

values initially decreases with increasing soil moisture, with the minimum being reached in the [0.2,0.3[soil moisture interval. Thereafter, Mean(σ^0) increases. Thus, c_σ is set equal to the Mean(σ^0) value from the [0.2,0.3[interval. Note that c_σ estimated in this way is a surrogate parameter reflecting land cover patterns, whereas high values indicate urban areas, open rock and dense vegetation, low to medium values grassland and agricultural areas, and very low values sandy deserts.

After fixing a value for c_σ, only two respectively four free parameters are left in models $\mathcal{M}_0$ and $\mathcal{M}_1$. The non-linear model fits are computed using the Python/scipy-implementation of the *trust region reflective* (TRF) algorithm (Branch et al., 1999), with start values estimated from the distribution of the observed backscatter values. Furthermore, we restrained the range of model parameters to physically meaningful ranges, most importantly the condition that $\widehat{\alpha}$ must be >0, and β, $\widehat{\psi}$, and ξ must be ≥ 0. The two model fits for our ASCAT pixel in Central Spain are also shown in Fig. 5. It can be seen that over this area only model $\mathcal{M}_1$ is able to reproduce the U-shaped relationship between σ^0 and θ, while model $\mathcal{M}_0$ increases monotonically. Compared to the Mean(σ^0) values for the ten wetness intervals (red dotted lines in Fig. 5), the fitted instance of the $\mathcal{M}_1$ exhibits a quicker decrease at low soil moisture values, causing the turning point θ_{turn} of $\mathcal{M}_1$ (blue point) to be situated in the [0.1,0.2[interval instead of the [0.2,0.3[interval.

4.2. Model selection

Model selection refers to the task of choosing from several possible candidate model classes $\mathcal{M}_0$, $\mathcal{M}_1$, ..., $\mathcal{M}_p$ the one which explains best the observations $\mathcal{O}$ (Hastie et al., 2009), and combines classification, i.e., the selection of the correct model (class) $\mathcal{M}_j$, with parameter estimation for the model instances: for each model class $\mathcal{M}_j$, a model instance $\widehat{f}_j$ is fitted to the observations – the training set $\mathcal{O} = (\sigma_i^0, \theta_i), 1 \leq i \leq N$ – resulting in a training error expressed in terms of the mean squared error (MSE):

$$\text{MSE}_j = \frac{1}{N} \sum_{i=1}^{N} \left(\sigma_i^0 - \widehat{f}_j(\theta_i) \right)^2. \tag{9}$$

However, one cannot simply choose the model $\mathcal{M}_c$ with the smallest training error $c = \arg \min \text{MSE}_j$. This is because the results may differ by less than what is a physically meaningful threshold considering the uncertainty of the data, in which case the simpler model is to be chosen over the more complex model. Moreover, complex models will typically lead to a smaller MSE on the training set given their higher flexibility compared to more parsimonious models. Nonetheless, beyond a given complexity, they will incur an increase in the MSE for non-training data, a phenomenon known as overfitting.

There are two principal avenues to model selection. The first, k-fold cross validation, is very general and can be used for any kind of model, but requires repeated estimation of the model parameters. The second are approaches that inflate the training error based on the effective degrees of freedom of the model, but these are limited (at least, in theory) to the class of linear models. We shall consider each one in turn.

4.2.1. k-fold cross validation

The *k*-fold cross validation (CV) method is based on the concept of empirical cross validation, i.e., the use of a separate validation data set (apart from the training set) to determine how compatible a particular model class $\mathcal{M}_j$ (more precisely, a fitted instance of this class) is with the observations. It is implemented by subdividing the training set in k equally large parts (the folds) of size $N_s = N/k$, and using each of these folds as validation set for a model instance trained on the remaining $k - 1$ folds: thus, for each model class $\mathcal{M}_j$, k model instances $\widehat{f}_{ji}, 1 \leq i \leq k$ are trained on a reduced training set, with the removed part of the training set with indices $i_1, ..., i_{N_s}$ used for validation, giving

$$CV_{ji} = \frac{1}{N_s} \sum_{l=1}^{N_s} \left(\sigma_{i_l}^0 - \widehat{f}_{ji}(\theta_{i_l}) \right)^2. \tag{10}$$

The goodness of model $\mathcal{M}_j$ is obtained as average of the k validation errors, $CV_j = \sum_{i=1}^{k} CV_{ji}/k$. The simplest strategy is to choose the model class c with the smallest average validation error $c = \arg \min CV_j$, although more refined approaches are possible. Here, we adopt the *one standard error* rule (Hastie et al. (2009), p. 216), which states that the simplest model $\mathcal{M}_j$ with CV_j within one standard error ε of CV_c should be chosen; this means specifically that the complex model accounting for subsurface effects will only be chosen if it decreases the *CV* by more than ε.

4.2.2. Methods based on the modification of the training error

These approaches work by inflating the training error of a fitted model instance by adding a term reflecting the complexity of the model class (e.g., two vs. one exponential basis functions); the complexity is quantified by the number of effective parameters d, which, in our case, is identical with the number of model parameters. The goal here is not to infer an estimate of the true test error, but to compute a quantity that reflects the relative performance of the different models on new data.

The Akaike Information Criterion (AIC) is probably the best known representative of this group

$$\text{AIC}_j = MSE_j + 2 \frac{d}{N} \widehat{s}^2 \tag{11}$$

where d is the number of parameters, *N* is the number of training points, and $\widehat{s}^2$ an estimate of the noise variance (which can, for example, be derived from the RMSE of a complex, i.e. low bias, model fit). More complex models (with more parameters) are more heavily penalised.

In our experiments, we do not use the AIC but the Bayesian Information Criterion (BIC) because of its tendency to select simpler models:

$$\text{BIC}_j = \frac{N}{\widehat{s}^2} \left(MSE_j + logN \frac{d}{N} \widehat{s}^2. \right) \tag{12}$$

It can be seen that it is basically proportional to the AIC, with the factor 2 replaced by log*N*. The motivation and derivation of the BIC, however, is rooted in Bayesian Statistics and quite different from the AIC; all the more remarkable that they are so similar. For a full derivation and more detailed discussion, the reader is referred to Hastie et al. (2009).

5. Results

In this section we present the results of fitting the two backscatter models without ($\mathcal{M}_0$) and with ($\mathcal{M}_1$) subsurface scattering term to the ASCAT backscatter and ERA5-Land soil moisture data, and selecting the best fitting model using both the CV and BIC methods.

5.1. Model fitting

As described in Section 4.1, the first step of the fitting procedure was to estimate the scattering constant c_σ by finding the minimum Mean(σ^0) value from ten soil moisture intervals. The resulting c_σ map (not shown) mirrors the spatial patterns of the Mean(σ^0) map displayed in Fig. 4k, but at reduced backscatter levels. The c_σ values range from minimum values around 0.04 m²m⁻² (−14 dB) over sand desert areas in Algeria to maximum values of around 0.3 m²m⁻² (−5.2 dB) over the Atlas mountains in Morocco.

The results of the non-linear model fitting process are summarised by Fig. 6 for both $\mathcal{M}_0$ and $\mathcal{M}_1$. The figure shows all non-linear model parameters along with the RMSE values and the derived parameters, namely the signal ranges $\mathcal{S}_{top}$ and $\mathcal{S}_{sub}$, and the soil moisture value at the turning point θ_{turn}. One can observe that in the more humid parts of the study domain, all parameters related to the surface scattering
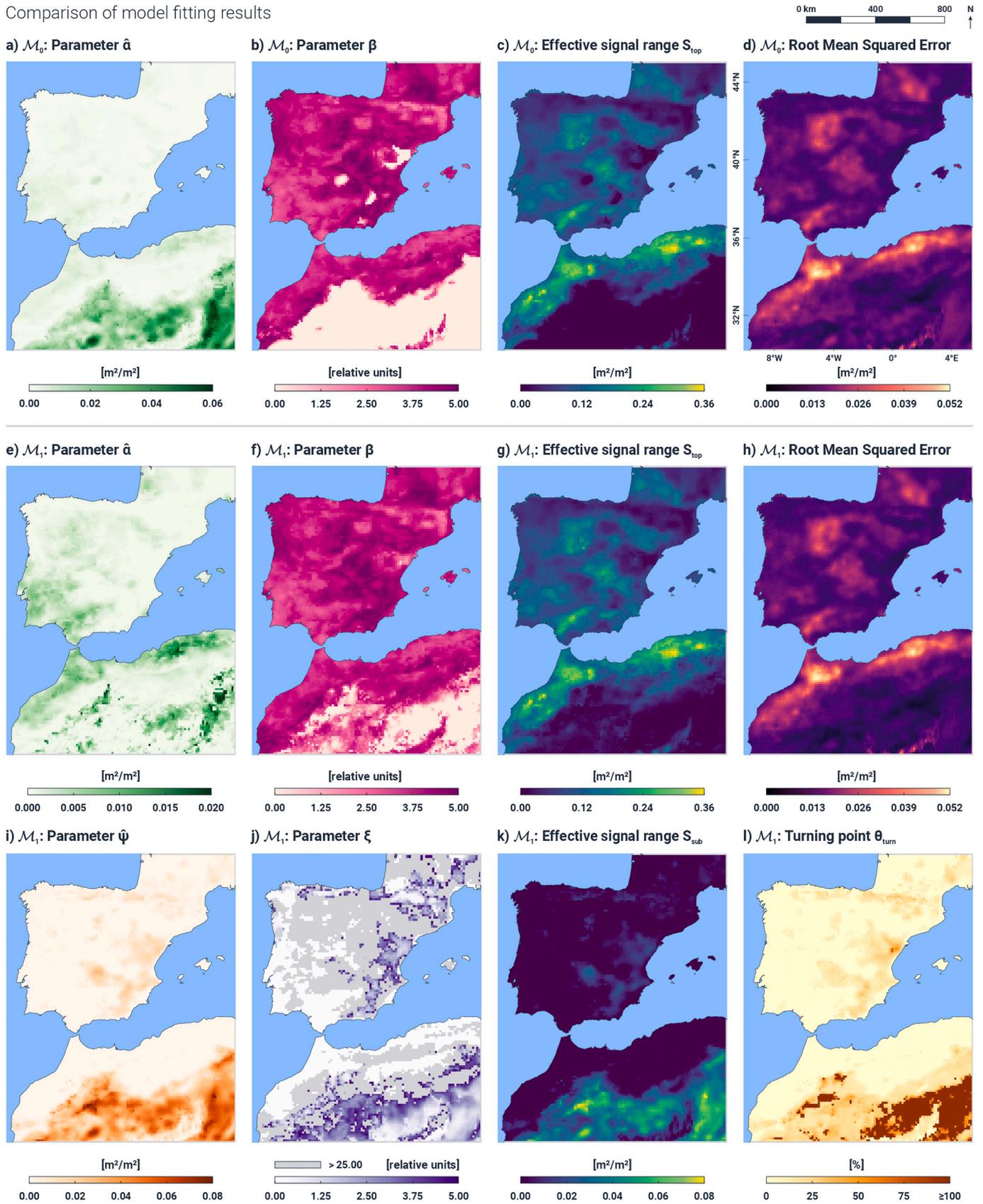
Comparison of model fitting results

Fig. 6. Results of fitting the models without ($\mathcal{M}_0$) and with ($\mathcal{M}_1$) a subsurface scattering term to ASCAT backscatter and ERA5-Land soil moisture data. The top row shows from left to right the model parameters $\hat{\alpha}$ and β, the signal ranges $\mathscr{S}_{top}$, and the RMSE of $\mathcal{M}_0$. The second row shows the corresponding maps for $\mathcal{M}_1$. The last row shows the subsurface scattering parameters $\hat{\psi}$, ξ, $\mathscr{S}_{sub}$, and θ_{turn}.

contribution, i.e. $\widehat{\alpha}$, β, and $\mathscr{S}_{top}$, are in general quite similar for both models. This speaks for the robustness of our fitting procedure and implies that subsurface scattering is either weak or non-existent in these regions. Differences between the surface terms of $\mathscr{M}_0$ and $\mathscr{M}_1$ are most pronounced in the arid parts of the study area, and best visible when comparing the two β maps: While β of $\mathscr{M}_1$ has values equal to zero only in the Sahara, β of $\mathscr{M}_0$ also vanishes in some regions in Spain (white areas in Fig. 6b). This means that without allowing for an initial decrease of backscatter, it would appear that backscatter exhibits no sensitivity to soil moisture in these regions. Given that this behaviour cannot be explained by dense vegetation cover that fully blocks the microwave signals from the ground, this is a first indication that $\mathscr{M}_1$ is able to describe the physical reality better than $\mathscr{M}_0$. Overall, it appears that β and $\mathscr{S}_{top}$ of $\mathscr{M}_1$ reflect vegetation patterns quite well. When looking at $\mathscr{S}_{top}$, it is highest over grassland and agricultural regions with values in the range from 0.15 to 0.3 m^2m^{-2}. Over forests and shrubland $\mathscr{S}_{top}$ is typically below 0.1 m^2m^{-2}.

The subsurface scattering parameters $\widehat{\psi}$, ξ, and $\mathscr{S}_{sub}$ become important mostly over some of the arid regions of our study domain. The values of $\mathscr{S}_{sub}$ are highest in the Sahara, with values of up to 0.08 m^2m^{-2}. But also over the European parts of the study domain $\mathscr{S}_{sub}$ can be significant with values exceeding 0.02 m^2m^{-2} in some areas in Spain. In those areas where $\mathscr{S}_{sub} > 0$, one also finds valid solutions for the turning point ($\theta_{turn} \in [0, 100]$), with values of θ_{turn} varying mostly in the range from 10 to 30% (cf. Fig. 6k and l). Nonetheless, in some parts of the Sahara our data only depict decreasing backscatter values with increasing soil moisture, implying that in these areas subsurface scattering dominates under all conditions and there is no physically valid solution for θ_{turn}.

The RMSE maps of $\mathscr{M}_0$ and $\mathscr{M}_1$ look at first sight very similar to each other, with most values being small in the range from 0.005 to 0.015 m^2m^{-2}. Substantially higher RMSE values in the range from 0.02 to 0.05 m^2m^{-2} can be found in some of the agricultural and grassland regions. However, this does not mean that the fit was less successful, but merely reflects the fact that the natural variability of the ASCAT backscatter measurements is highest in these regions. Indeed, when comparing the RMSE maps with the StDev(σ^0) map shown in Fig. 4l one finds very

similar spatial patterns, only the magnitude is different with RMSE values being $\sim \frac{2}{3}$StDev(σ^0). Subtle differences between the two RMSE maps are mostly apparent in the arid regions. These differences are the basis for the model selection as discussed in the next section.

5.2. Model selection

The results of model selection using the *k*-fold cross validation (CV) and Bayesian Information Criterion (BIC) methods are shown in Fig. 7. For the CV method we used a constant ε value equal to 10^{-3} m^2m^{-2} (−30 dB), which is at least a factor of 10 smaller than any signal variation we can reasonably hope to capture given the uncertainty of our observations $\mathscr{O} = (\sigma_i^0, \theta_i)$. The areas shown in orange colour are those areas where $\mathscr{M}_1$ yields a better fit to the observations than $\mathscr{M}_0$ does. One can see that the CV method favours $\mathscr{M}_1$ over $\mathscr{M}_0$ over more and larger regions than the BIC method. Nonetheless, both methods yield consistent results, selecting $\mathscr{M}_1$ over $\mathscr{M}_0$ in those regions where $\mathscr{S}_{sub}$ and θ_{turn} take on non-zero values. The majority of these subsurface scattering areas are situated in arid climates. Nonetheless, some of them can even be found in more humid climatic zones, e.g. in the north of Portugal and Spain and even some areas in France.

To obtain a more quantitative understanding of where subsurface scattering occurs, we determined for each class $\mathscr{C}$ of the climate, soil, and land cover maps shown in Fig. 4 the number of ASCAT pixels affected (n_1) and unaffected (n_0) by subsurface scattering. From this we calculated the likelihood $\mathscr{P}(\mathscr{M}_1 | \mathscr{C}) = n_1/(n_0 + n_1)$ for each class per map and for both the CV and BIC methods. The resulting probabilities (in %) are shown in Table 1. The results confirm our hypothesis that the aridity of the environment in combination with Leptosols (very shallow soils over hard rock or highly calcareous material, but also deeper soils that are extremely gravelly and/or stony) or Arenosols (unconsolidated sand deposits) represent ideal environmental conditions for the occurrence of subsurface scattering. Bare soils, sparse vegetation cover, and shrubland are also indicative, but this does not mean that there are no subsurface scatteres in more densely vegetated areas. For example, in some forest-dominated areas inland of the Spanish Mediterranean coastline, subsurface scattering is detected in spite of a rather small
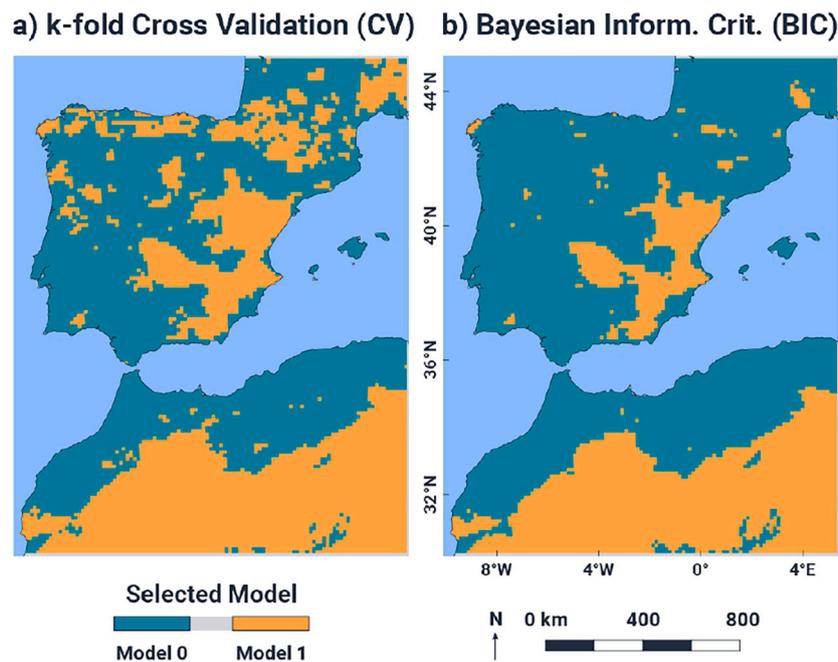

Fig. 7. Models selected according to the (a) *k*-fold cross validation (CV) method and (b) Bayesian Information Criterion (BIC) method. Green shows where $\mathscr{M}_0$ (no subsurface scattering) was selected, orange $\mathscr{M}_1$ (with subsurface scattering). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Likelihood (%) of subsurface scattering occurring for different classification schemes and classes. The second column shows the number of ASCAT pixels covered by the respective class. The third and forth columns show the likelihood calculated based on the cross validation (CV) and Bayesian Information Criterion (BIC) criteria respectively. Classes with $N_{ASCAT} < 100$ pixels are not shown.

Class	N_{ASCAT}	CV (%)	BIC (%)
Köppen-Geiger Climate Classification			
BWh - Arid desert hot	2203	92.6	91.7
BWk - Arid desert cold	995	47.9	45.4
BSh - Arid steppe hot	264	9.8	9.1
BSk - Arid steppe cold	2137	33.5	31.3
Csa - Temperate dry hot summer	2138	6.6	6.0
Csb - Temperate dry warm summer	685	11.2	9.3
Cfa - Temperate no dry season hot summer	106	5.7	2.8
Cfb - Temperate no dry season warm summer	1073	9.6	7.0
ISRIC Soil Groups			
15–20 - Arenosols	639	86.9	84.8
21–24 - Calcisols	1369	27.0	25.3
25–35 - Cambisols	2242	7.1	4.8
68–72 - Leptosols	2762	80.0	78.9
76–84 - Luvisols	2225	12.9	11.7
115–118 - Vertisols	267	0.0	0.0
ESA CCI Land Cover Classification			
10 - Cropland rainfed	600	7.2	6.0
11 - Cropland rainfed, herbaceous cover	1612	8.9	7.9
12 - Cropland rainfed, tree or shrub cover	417	12.9	10.3
20 - Cropland irrigated or post-flooding	166	2.4	1.8
40 - Mosaic natural vegetation (>50%)/cropland (<50%)	234	7.7	7.7
60 - Tree cover broadleaved deciduous, closed to open (>15%)	648	3.5	1.5
70 - Tree cover needleleaved evergreen, closed to open (>15%)	416	13.7	9.9
100 - Mosaic tree and shrub (>50%)/herbaceous cover (<50%)	901	30.3	28.6
120 - Shrubland	326	37.7	33.1
130 - Grassland	304	6.9	5.9
150 - Sparse vegetation (<15%)	210	38.6	36.2
153 - Sparse herbaceous cover (<15%)	813	39.1	38.5
200 - Bare areas	2913	84.2	82.9

backscatter variability that indicates a dense vegetation canopy with few gaps through which subsurface scattering effects can be sensed. The karst map shown in Fig. 4f helps little in understanding subsurface scattering in the African part of our study area, but nonetheless gives some clues to where subsurface scattering may become detectable in the European part. This is because karstified carbonate rocks and evaporites are extremely rough targets for a radar, potentially yielding very strong subsurface scattering signals.

6. Discussion

Our method for detecting subsurface scattering in ASCAT backscatter measurements rests on several simplifying assumptions. Firstly, the models $\mathcal{M}_0$ and $\mathcal{M}_1$ only account for zeroth-order radiative effects in the vegetation layer, neglecting high-order soil-vegetation interactions that are typically weak but may nonetheless be important for modelling ASCAT backscatter measurements over low to medium vegetation cover (Crow et al., 2010; Quast et al., 2019). However, these interactions are not expected to impact the analysis of backscatter anomalies as they always enhance total backscatter with increasing ground scattering, i.e. they do not alter the general direction of signals from the ground surface. Another simplification is that by fitting the models $\mathcal{M}_0$ and $\mathcal{M}_1$ to the complete three-years long $\sigma^0(20)$ time series, seasonal vegetation effects are neglected. While it is clear that backscatter shows the highest sensitivity to the ground at steep incidence angles (Magagi and Kerr,

1997; Hahn et al., 2021), this assumption introduces some noise in the $\sigma^0 - \theta$ scatterplots as shown in Fig. 5. This noise comes on top of the measurement noise, uncertainties related to the extrapolation of the backscatter measurements to 20°, and uncertainties in the modelled soil moisture data. Nonetheless, the functional relationship between σ^0 and θ could be clearly discerned over all pixels of our study domain, resulting in satisfactory model fits and model selections.

Our results demonstrate that subsurface scattering occurs mostly in arid environments, but is not necessarily confined to them. Apparently, when there are strong subsurface scatterers (e.g. karstic rock), especially when they are near the soil surface, they may also become detectable in more temperate climatic regions during dry spells. To illustrate the resulting scattering behaviour in different parts of the study area, Fig. 8 shows selected $\sigma^0 - \theta$ scatterplots along with the fitted instances of $\mathcal{M}_1$. Over the three barren landscape sites in the Sahara (DZ-1 and -2, MA-2), the subsurface scattering term $\widehat{\psi}e^{-\xi\theta}$ clearly dominates over the surface term $\widehat{\alpha}e^{\beta\theta}$, resulting in an initially fast decrease of σ^0 with θ, and more or less constant backscatter for wetter conditions. Interestingly, such a behaviour may emerge for very different backscatter backgrounds, with c_σ being as small as 0.055 m^2m^{-2} over the sand desert at point DZ-1, and as high as 0.27 m^2m^{-2} over the Atlas mountain range (MA-2). Over all other sites σ^0 either decreases slightly or is more or less constant for dry to medium wet soil conditions. Thereafter, it increases at a rate that depends mostly on vegetation density, from very low values over densely vegetated areas (e.g. ES-3 and FR-2) to high values over agricultural and sparsely vegetated regions (e.g. MA-1, and ES-1 and -2). This behaviour has two important consequences: Firstly, dependent on the subsurface scattering strength, one should observe intermittent backscatter anomalies during dry conditions. Secondly, subsurface scattering may substantially reduce the sensitivity of σ^0 to θ under dry conditions, even in situations when the subsurface scatterers are not strong enough to enforce an initial decline at the dry edge (cf. Fig. 3). The implications of these points are discussed below.

6.1. Occurrence of intermittent backscatter anomalies

From the functional relationship between σ^0 and θ over mixed surface-subsurface scattering areas, we expect to observe negative correlations between these two variables under dry conditions and positive correlations during wet conditions. Moreover, we expect that the sensitivity $\mathscr{S}_{top}$ to subsurface scatterers is an important control as to how often intermittent backscatter anomalies occur. To verify these expectations, we calculated the Spearman rank correlation ρ between ASCAT backscatter and ERA5-Land soil moisture for each day using a sliding window of one month (31 days) and determined the number of days for which $\rho < -0.4$, adopting the methods first tested by Frießenbichler (2020) over South Africa. Dividing this number by the total number of days yields an estimate of the probability of occurrence of backscatter anomalies, $\mathscr{P}_{ano}$, shown in Fig. 9a. A comparison to Fig. 4 shows that the $\mathscr{P}_{ano}$ maps nicely reflect mean annual rainfall and soil moisture patterns, with $\mathscr{P}_{ano} < 10\%$ over regions with mean annual rainfall >800 mm and reaching values of up to 80% over the desert regions. As expected, it is even more closely related to $\mathscr{S}_{sub}$, as the scatterplot between these two variables in Fig. 9b demonstrates: at first there is a rapid increase of $\mathscr{P}_{ano}$ with increasing $\mathscr{S}_{sub}$ values, and then a flattening out of the curve as $\mathscr{P}_{ano}$ saturates. The rank correlation between these two variables is 0.71. This confirms that the anomalous ASCAT backscatter behaviour in dry regions, as observed in many soil moisture validation studies (Fascetti et al., 2016; Mousa and Shu, 2020; Zhang et al., 2021), is caused by subsurface scattering. It may also explain why rainfall variability better explains temporal variability of soil moisture retrieval errors than leaf area index (LAI), as recently found by Wu et al. (2021).
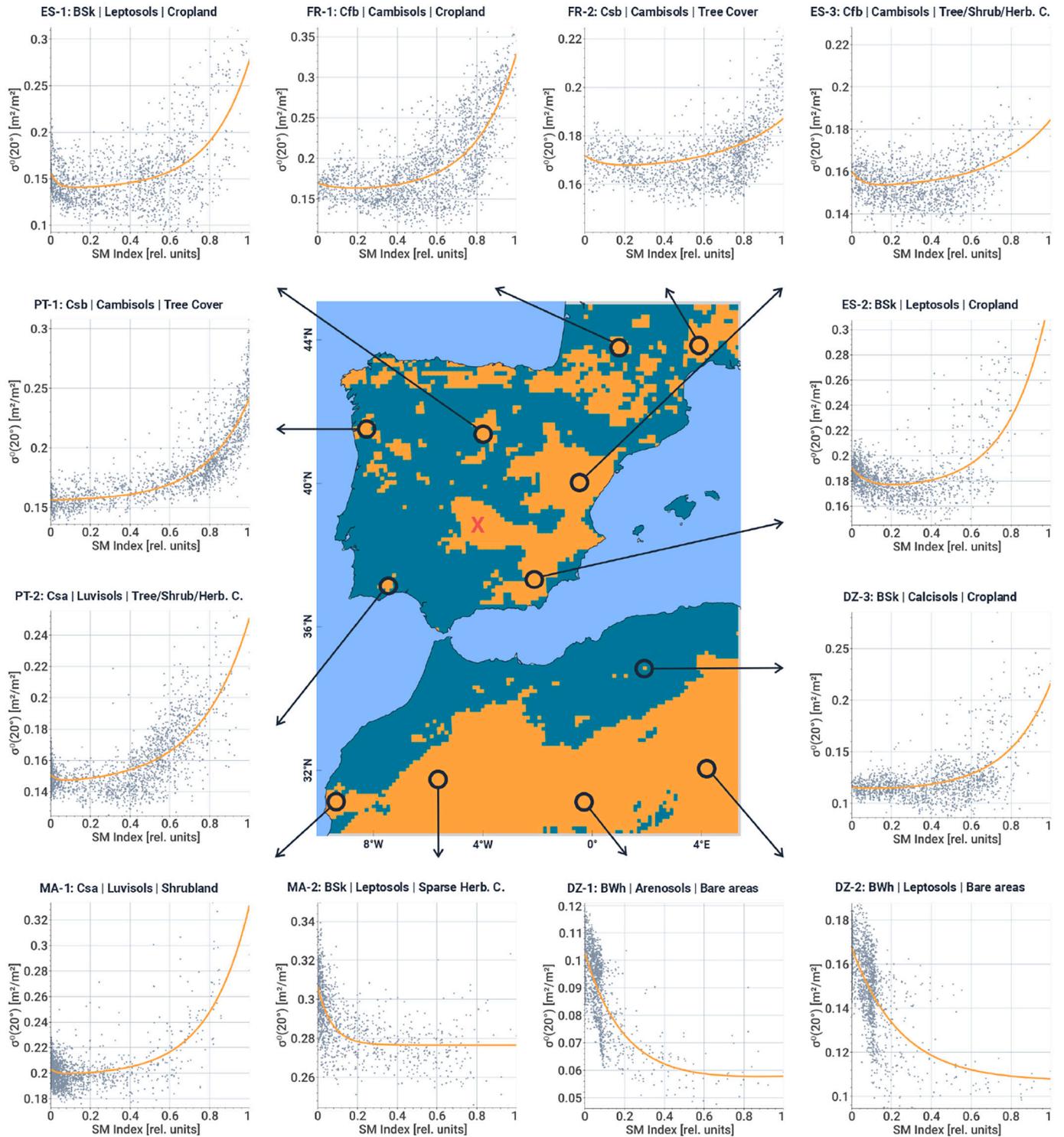
Fig. 8. Relationship between ASCAT backscatter σ^0 and ERA5-Land soil moisture θ over a number of locations in Algeria (DZ), France (FR), Morocco (MA), Spain (ES) and Portugal (PT), where the CV method selected the model $\mathcal{M}_1$ with the subsurface scattering term (the map in the middle is identical to the one shown in Fig. 7). The X-symbol indicates the location of the data from Fig. 5.

6.2. Reduced sensitivity of backscatter to soil moisture

On top of causing intermittent backscatter anomalies, subsurface scattering reduces the sensitivity of backscatter to soil moisture under dry soil conditions. From a modelling point of view this is not such a big problem given that also a model $\mathcal{M}_0$ without the subsurface scattering terms is normally able to describe the functional relationship between σ^0 and θ in weak subsurface scattering areas (e.g. PT-1 und −2, DZ-3).

Nonetheless, it is of course problematic from a soil moisture retrieval point of view, as noise in the backscatter data translates into higher soil moisture retrieval errors. Even in regions with a dominant surface signal this may have implications given that subsurface scattering reduces the signal range of backscatter to soil moisture, such that $\mathscr{S}_{top}$ of a model instance of $\mathcal{M}_0$ is in reality $\mathscr{S}_{top} - \mathscr{S}_{sub}$ of an unobservable model with subsurface scattering $\mathcal{M}_1$ (cf. Fig. 3a).

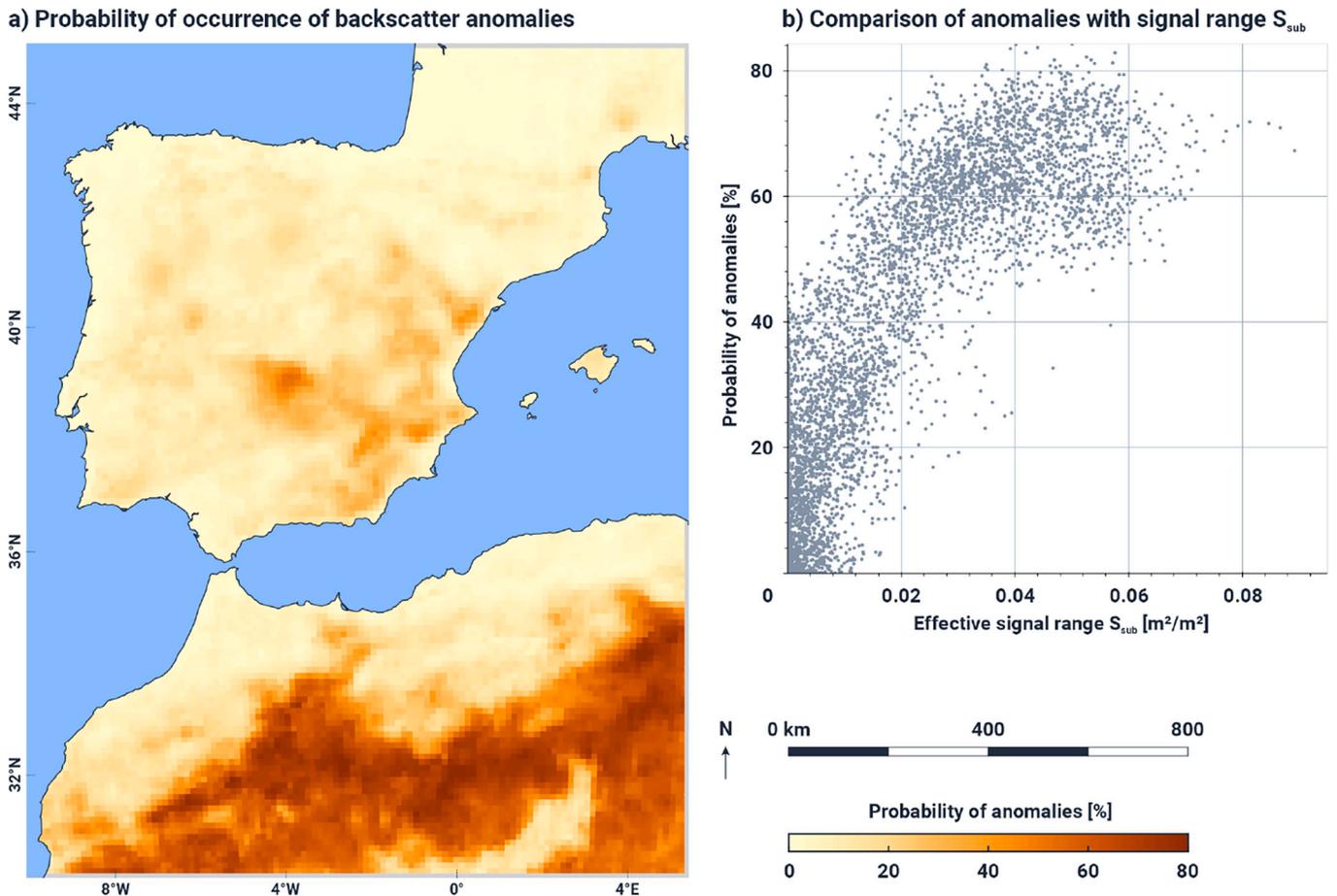It also has implications as regards to the choice of backscatter models

Fig. 9. Probability of the occurrence of ASCAT backscatter anomalies (a), and its relationship to the effective signal range $\mathscr{S}_{sub}$ of the subsurface scattering term (b).

to be used for soil moisture retrievals from ASCAT and other C-band radars. In the TU Wien change detection algorithms for ASCAT (Bartalis et al., 2007) and Sentinel-1 (Bauer-Marschallinger et al., 2019), which are used in the EUMETSAT H SAF and Copernicus Land Monitoring Services (CLMS), we assume that backscatter expressed in decibels is linearly related to soil moisture. Unfortunately, even when converting the σ^0 data as shown in Fig. 8 to decibels, the curves retain a slightly convex shape. Therefore, the H SAF ASCAT and CLMS Sentinel-1 soil moisture retrievals tend to overestimate soil moisture in case of subsurface scattering. This problem would be even more pronounced when using bare soil backscatter models based on the Fresnel equations, such as Fung's integrated equation model (IEM) or the new reflectivity index recently published by Zribi et al., 2021, given that these have a concave shape due to the saturation of the Fresnel reflectivity for higher soil moisture values. Interestingly, when Zribi et al. (2014) considered a heterogeneous soil moisture profile in their IEM simulations, they could establish a linear relationship between their multilayer backscatter model and soil moisture. Therefore, it seems overall quite important to account for subsurface effects in bare soil backscatter models.

7. Conclusion

In this study we showed that a subsurface scattering term of the form $\psi e^{-\xi\theta}$ is key for explaining the behaviour of ASCAT backscatter measurements in arid environments. Moreover, weak subsurface scattering signals may also occur in more temperate climates during dry spells. Therefore, it appears that subsurface scattering is a widespread phenomenon that has so far not received enough attention by the microwave remote sensing community. Much more work will be needed to

understand the adverse impacts of subsurface scattering on soil moisture and vegetation retrievals, and to come up with improved backscatter models and retrieval approaches. This entails the need to improve our understanding of how subsurface scattering changes with incidence angle, polarisation, and frequency. Thankfully, the increasing availability of radar observations at multiple polarisations, incidence angles, or even frequencies will help to study these dependencies and come up with means to differentiate the different scattering mechanisms acting at the soil surface and within (Morrison and Wagner, 2020). For ASCAT, the possibility to use the so-called slope and curvature along with backscatter to characterise the land surface (Steele-Dunne et al., 2019) may be an opportunity to discern subsurface scattering effects. For the successor instrument of ASCAT, the SCA instrument to be flown on the second generation METOP-SG B satellites, the two additional polarisations acquired by the mid-beam antennas (VH and HH) represent another big opportunity to improve scatterometer soil moisture retrievals (Stoffelen et al., 2017). Much will also be learned from the analysis of Synthetic Aperture Radar (SAR) backscatter time series given that the much higher spatial resolution of the data will allow to directly relate the radar signals to field observations and small-scale geomorphological patterns (Ullmann et al., 2019). Furthermore, the phase as measured by SAR sensors is sensitive to the scattering mechanisms within the soil profile in low moisture regimes (Zwieback et al., 2015), which will not only help to investigate soil moisture processes in arid environments but also to estimate the depth of subsurface scatterers (Morrison and Wagner, 2022). Therefore we see a large potential of scatterometer and SAR techniques for characterising soil properties, particularly in arid and semi-arid environments.

Credit author statement

Conceptualization: Wagner; Methodology: Wagner, Melzer; Data Curation: Hahn, Lindorfer, Greimeister-Pfeil; Software: Hahn, Quast, Melzer, Lindorfer; Formal analysis: Lindorfer, Wagner; Investigation: Morrison, Calvet, Hobbs, Quast, Vreugdenhil; Visualisation: Lindorfer, Bauer-Marschallinger; Funding acquisition: Wagner, Vreugdenhil; Writing – original draft: Wagner, Melzer; Writing – review … editing: All authors.

Declaration of Competing Interest

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

Acknowledgements

References

Anderson, C., Figa-Saldana, J., Wilson, J.J.W., Ticconi, F., 2017. Validation and cross-validation methods for ASCAT. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 2232–2239. URL.

Attema, E.P.W., Ulaby, F.T., 1978. Vegetation modeled as a water cloud. Radio Sci. 13, 357–364. https://doi.org/10.1029/RS013i002p00357.

Bahar, E., 1981. Scattering cross sections for composite random surfaces: full wave analysis. Radio Sci. 16, 1327–1335. https://doi.org/10.1029/RS016i006p01327.

Bartalis, Z., Scipal, K., Wagner, W., 2006. Azimuthal anisotropy of scatterometer measurements over land. IEEE Trans. Geosci. Remote Sens. 44, 2083–2092. URL. http://ieeexplore.ieee.org/document/1661797/. https://doi.org/10.1109/TGRS.2006.872084.

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J., Anderson, C., 2007. Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). Geophys. Res. Lett. 34, L20401. https://doi.org/10.1029/2007GL031088.

Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., Modanesi, S., Massari, C., Ciabatta, L., Brocca, L., Wagner, W., 2019. Toward global soil moisture monitoring with Sentinel-1: harnessing assets and overcoming obstacles. IEEE Trans. Geosci. Remote Sens. 57, 520–539. URL.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci. Data 5, 180214. URL. http://www.nature.com/articles/sdata2018214. https://doi.org/10.1038/sdata.2018.214.

Branch, M.A., Coleman, T.F., Li, Y., 1999. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. SIAM J. Sci. Comput. 21, 1–23. https://doi.org/10.1137/S1064827595289108.

Chen, Z., Auler, A.S., Bakalowicz, M., Drew, D., Griger, F., Hartmann, J., Jiang, G., Moosdorf, N., Richts, A., Stevanovic, Z., Veni, G., Goldscheider, N., 2017. The world karst aquifer mapping project: concept, mapping procedure and map of Europe. Hydrogeol. J. 25, 771–785. https://doi.org/10.1007/s10040-016-1519-3.

Crow, W.T., Wagner, W., Naeimi, V., 2010. The impact of radar incidence angle on soil-moisture-retrieval skill. IEEE Geosci. Remote Sens. Lett. 7, 501–505. https://doi.org/10.1109/LGRS.2010.2040134.

de Jeu, R.A.M., Wagner, W., Holmes, T.R.H., Dolman, A.J., van de Giesen, N.C., Friesen, J., 2008. Global soil moisture patterns observed by space borne microwave radiometers and scatterometers. Surv. Geophys. 29, 399–420. https://doi.org/10.1007/s10712-008-9044-0.

Dobson, M., Ulaby, F., 1986. Active microwave soil moisture research. IEEE Trans. Geosci. Remote Sens. GE-24, 23–36. URL.

Dobson, M., Ulaby, F., Hallikainen, M., El-rayes, M., 1985. Microwave dielectric behavior of wet soil-part II: dielectric mixing models. IEEE Trans. Geosci. Remote Sens. GE-23, 35–46. URL.

Dorigo, W.A., Scipal, K., Parinussa, R.M., Liu, Y.Y., Wagner, W., de Jeu, R.A.M., Naeimi, V., 2010. Error characterisation of global active and passive microwave soil moisture datasets. Hydrol. Earth Syst. Sci. 14, 2605–2616. URL.

Dubois, P., van Zyl, J., Engman, T., 1995. Measuring soil moisture with imaging radars. IEEE Trans. Geosci. Remote Sens. 33, 915–926. URL. http://ieeexplore.ieee.org/document/406677/. https://doi.org/10.1109/36.406677.

Entekhabi, D., Njoku, E.G., O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, K.C., Moghaddam, M., Moran, S., Reichle, R.,

Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Van Zyl, J., 2010. The soil moisture active passive (SMAP) mission. Proc. IEEE 98, 704–716. URL.

ESA, 2017. ESA Land Cover CCI Product User Guide - v2.0 - CCI-LC-PUGv2. URL. https://www.esa-landcover-cci.org/?q=webfm_send/112.

Escorihuela, M.J., Quintana-Seguí, P., 2016. Comparison of remote sensing and simulated soil moisture datasets in Mediterranean landscapes. Remote Sens. Environ. 180, 99–114. URL. https://linkinghub.elsevier.com/retrieve/pii/S0034425716300748. https://doi.org/10.1016/j.rse.2016.02.046.

Fascetti, F., Pierdicca, N., Pulvirenti, L., Crapolicchio, R., Muñoz-Sabater, J., 2016. A comparison of ASCAT and SMOS soil moisture retrievals over Europe and Northern Africa from 2010 to 2013. Int. J. Appl. Earth Obs. Geoinf. 45, 135–142. URL. https://linkinghub.elsevier.com/retrieve/pii/S0303243415300313. https://doi.org/10.1016/j.jag.2015.09.008.

Figa-Saldaña, J., Wilson, J.J.W., Attema, E.P.W., Gelsthorpe, R., Drinkwater, M.R., Stoffelen, A., 2002. The advanced scatterometer (ASCAT) on the meteorological operational (MetOp) platform: a follow on for European wind scatterometers. Can. J. Remote. Sens. 28, 404–412. https://doi.org/10.5589/m02-035.

Frießenbichler, M., 2020. An Analysis of the Radar Subsurface Scattering Phenomenon Observed in the Metop ASCAT Soil Moisture Product over South Africa. Ph.D. thesis. Technische Universität Wien, Vienna, Austria.

Fung, A., Liu, W., Chen, K., Tsay, M., 2002. An improved IEM model for bistatic scattering from rough surfaces. J. Electromagnet. Wave Appl. 16, 689–702. https://doi.org/10.1163/156939302X01119.

Hahn, S., Reimer, C., Vreugdenhil, M., Melzer, T., Wagner, W., 2017. Dynamic characterization of the incidence angle dependence of backscatter using Metop ASCAT. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 2348–2359. URL.

Hahn, S., Wagner, W., Steele-Dunne, S.C., Vreugdenhil, M., Melzer, T., 2021. Improving ASCAT soil moisture retrievals with an enhanced spatially variable vegetation parameterization. IEEE Trans. Geosci. Remote Sens. 59, 8241–8256. URL.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning. Springer Series in Statistics. Springer New York, New York, NY. https://doi.org/10.1007/978-0-387-84858-7.

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., 2017. SoilGrids250m: global gridded soil information based on machine learning. PLoS One 12, e0169748. https://doi.org/10.1371/journal.pone.0169748.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146, 1999–2049. https://doi.org/10.1002/qj.3803.

Jaruwatanadilok, S., Stiles, B.W., 2014. Trends and variation in Ku-band backscatter of natural targets on land observed in QuikSCAT data. IEEE Trans. Geosci. Remote Sens. 52, 4383–4390. URL.

Kerr, Y.H., Waldteufel, P., Richaume, P., Wigneron, J.P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S.E., Leroux, D., Mialon, A., Delwart, S., 2012. The SMOS soil moisture retrieval algorithm. IEEE Trans. Geosci. Remote Sens. 50, 1384–1403. URL.

Kim, Y., van Zyl, J., 2009. A time-series approach to estimate soil moisture using polarimetric radar data. IEEE Trans. Geosci. Remote Sens. 47, 2519–2527. URL.

Liu, P.W., Judge, J., DeRoo, R.D., England, A.W., Bongiovanni, T., Luke, A., 2016. Dominant backscattering mechanisms at L-band during dynamic soil moisture conditions for sandy soils. Remote Sens. Environ. 178, 104–112. URL. https://linkinghub.elsevier.com/retrieve/pii/S0034425716300918. https://doi.org/10.1016/j.rse.2016.02.062.

Magagi, R.D., Kerr, Y.H., 1997. Retrieval of soil moisture and vegetation characteristics by use of ERS-1 wind scatterometer over arid and semi-arid areas. J. Hydrol. 188-189, 361–384. URL. https://www.sciencedirect.com/science/article/pii/S0022169496031666. https://doi.org/10.1016/S0022-1694(96)03166-6.

McCauley, J.F., Schaber, G.G., Breed, C.S., Grolier, M.J., Haynes, C.V., Issawi, B., Elachi, C., Blom, R., 1982. Subsurface valleys and geoarcheology of the eastern Sahara revealed by shuttle radar. Science 218, 1004–1020. https://doi.org/10.1126/science.218.4576.1004.

McColl, K.A., Entekhabi, D., Piles, M., 2014. Uncertainty analysis of soil moisture and vegetation indices using aquarius scatterometer observations. IEEE Trans. Geosci. Remote Sens. 52, 4259–4272. URL.

Miyaoka, K., Gruber, A., Ticconi, F., Hahn, S., Wagner, W., Figa-Saldaña, J., Anderson, C., 2017. Triple collocation analysis of soil moisture from Metop-a ASCAT and SMOS against JRA-55 and ERA-interim. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 2274–2284. URL.

Morrison, K., Wagner, W., 2020. Explaining anomalies in SAR and scatterometer soil moisture retrievals from dry soils with subsurface scattering. IEEE Trans. Geosci. Remote Sens. 58, 2190–2197. URL.

Morrison, K., Wagner, W., 2022. A novel DInSAR algorithm for the retrieval of soil moisture and soil depth over arid regions of the world. Can. J. Remote. Sens. submitted.

Mousa, B.G., Shu, H., 2020. Spatial evaluation and assimilation of SMAP, SMOS, and ASCAT satellite soil moisture products over Africa using statistical techniques. Earth Space Sci. 7 https://doi.org/10.1029/2019EA000841.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G.,

Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.N., 2021. ERA5-land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13, 4349–4383. URL.

Oh, Y., Sarabandi, K., Ulaby, F., 1992. An empirical model and an inversion technique for radar scattering from bare soil surfaces. IEEE Trans. Geosci. Remote Sens. 30, 370–381. URL. http://ieeexplore.ieee.org/document/134086/. https://doi.org/10.1109/36.134086.

Petropoulos, G.P., Ireland, G., Barrett, B., 2015. Surface soil moisture retrievals from remote sensing: current status, products & future trends. Phys. Chem. Earth Part A/B/C 83-84, 36–56. URL.

Quast, R., Albergel, C., Calvet, J.C., Wagner, W., 2019. A generic first-order radiative transfer modelling approach for the inversion of soil and vegetation parameters from scatterometer observations. Remote Sens. 11 https://doi.org/10.3390/rs11030285.

Schaber, G., McCauley, J., Breed, C., Olhoeft, G., 1986. Shuttle imaging radar: physical controls on signal penetration and subsurface scattenng in the eastern Sahara. IEEE Trans. Geosci. Remote Sens. GE-24, 603–623. URL.

Schanda, E., 1986. Physical Fundamentals of Remote Sensing. Springer Berlin Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-48733-0.

Shamambo, D., Bonan, B., Calvet, J.C., Albergel, C., Hahn, S., 2019. Interpretation of ASCAT radar scatterometer observations over land: a case study over southwestern France. Remote Sens. 11, 2842. URL.

Steele-Dunne, S.C., McNairn, H., Monsivais-Huertero, A., Judge, J., Liu, P.W., Papathanassiou, K., 2017. Radar remote sensing of agricultural canopies: a review. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 2249–2273. URL.

Steele-Dunne, S.C., Hahn, S., Wagner, W., Vreugdenhil, M., 2019. Investigating vegetation water dynamics and drought using Metop ASCAT over the North American Grasslands. Remote Sens. Environ. 224, 219–235. URL. https://www.sciencedirect.com/science/article/pii/S0034425719300045. https://doi.org/10.1016/j.rse.2019.01.004.

Stoffelen, A., Aaboe, S., Calvet, J.C., Cotton, J., De Chiara, G., Saldana, J.F., Mouche, A.A., Portabella, M., Scipal, K., Wagner, W., 2017. Scientific developments and the EPS-SG scatterometer. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 2086–2097. URL.

Ulaby, F.T., Batlivala, P.P., Dobson, M.C., 1978. Microwave backscatter dependence on surface roughness, soil moisture, and soil texture: part I-bare soil. IEEE Trans. Geosci. Electron. 16, 286–295. URL.

Ulaby, F.T., Moore, R.K., Fung, A.K., 1981. Microwave Remote Sensing: Active and Passive. Volume 1 Microwave Remote Sensing Fundamentals and Radiometry. Artech House Inc.

Ullmann, T., Serfas, K., Büdel, C., Padashi, M., Baumhauer, R., 2019. Data processing, feature extraction, and time-series analysis of Sentinel-1 synthetic aperture radar (SAR) imagery: examples from Damghan and Bajestan Playa (Iran). Z. Geomorphol.

9–39. https://doi.org/10.1127/zfg_suppl/2019/0524. Supplementary Issues 62. place: Stuttgart, Germany Publisher: Schweizerbart Science Publishers.

Verhoest, N., Lievens, H., Wagner, W., Álvarez Mozos, J., Moran, M., Mattia, F., 2008. On the soil roughness parameterization problem in soil moisture retrieval of bare surfaces from synthetic aperture radar. Sensors 8, 4213–4248. URL.

Wagner, W., Lemoine, G., Borgeaud, M., Rott, H., 1999. A study of vegetation cover effects on ERS scatterometer data. IEEE Trans. Geosci. Remote Sens. 37, 938–948. URL. http://ieeexplore.ieee.org/document/752212/. https://doi.org/10.1109/36.752212.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Blöschl, G., Eitzinger, J., Steinnocher, K., 2013. The ASCAT soil moisture product: a review of its specifications, validation results, and emerging applications. Meteorol. Z. 22, 5–33. URL.

Wanders, N., Karssenberg, D., Bierkens, M., Parinussa, R., de Jeu, R., van Dam, J., de Jong, S., 2012. Observation uncertainty of satellite soil moisture products determined with physically-based modeling. Remote Sens. Environ. 127, 341–356. URL. https://linkinghub.elsevier.com/retrieve/pii/S0034425712003574. https://doi.org/10.1016/j.rse.2012.09.004.

Wegmuller, U., Matzler, C., Huppi, R., Schanda, E., 1994. Active and passive microwave signature catalog on bare soil (2-12 GHz). IEEE Trans. Geosci. Remote Sens. 32, 698–702. URL. http://ieeexplore.ieee.org/document/297987/. https://doi.org/10.1109/36.297987.

Williams, K., Greeley, R., 2001. Radar attenuation by sand: laboratory measurements of radar transmission. IEEE Trans. Geosci. Remote Sens. 39, 2521–2526. URL. http://ieeexplore.ieee.org/document/964990/. https://doi.org/10.1109/36.964990.

Wu, K., Ryu, D., Nie, L., Shu, H., 2021. Time-variant error characterization of SMAP and ASCAT soil moisture using triple collocation analysis. Remote Sens. Environ. 256, 112324. URL.

Zhang, R., Kim, S., Sharma, A., Lakshmi, V., 2021. Identifying relative strengths of SMAP, SMOS-IC, and ASCAT to capture temporal variability. Remote Sens. Environ. 252, 112126. URL. https://linkinghub.elsevier.com/retrieve/pii/S0034425720304995. https://doi.org/10.1016/j.rse.2020.112126.

Zribi, M., Gorrab, A., Baghdadi, N., Lili-Chabaane, Z., Mougenot, B., 2014. Inuence of radar frequency on the relationship between bare surface soil moisture vertical profile and radar backscatter. IEEE Geosci. Remote Sens. Lett. 11, 848–852. URL.

Zribi, M., Foucras, M., Baghdadi, N., Demarty, J., Muddu, S., 2021. A new reflectivity index for the retrieval of surface soil moisture from radar data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 818–826. URL.

Zwieback, S., Hensley, S., Hajnsek, I., 2015. A polarimetric first-order model of soil moisture effects on the DInSAR coherence. Remote Sens. 7, 7571–7596. URL.