remote sensing

MDPI

Article

Assessing the Spatial Similarity of Soil Moisture Patterns and Their Environmental and Observational Drivers from Remote Sensing and Earth System Modeling Across Europe

Thomas Jagdhuber [1,2,*,†], Lisa Jach [1,3,†], Anke Fluhrer [1], David Chaparro [1,4], Florian M. Hellwig [1,2], Gerard Portal [5], Hans-Stefan Bauer [3] and Harald Kunstmann [2,6]

1 Department of Reconnaissance and Security, Microwaves and Radar Institute, German Aerospace Center, 82234 Wessling, Germany; lisa.jach@uni-hohenheim.de (L.J.); anke.fluhrer@dlr.de (A.F.); d.chaparro@creaf.uab.cat (D.C.); florian.hellwig@dlr.de (F.M.H.)
2 Institute of Geography, University of Augsburg, 86159 Augsburg, Germany; harald.kunstmann@geo.uni-augsburg.de
3 Institute of Physics and Meteorology, University of Hohenheim, 70599 Stuttgart, Germany; hans-stefan.bauer@uni-hohenheim.de
4 Centre for Ecological Research and Forestry Applications, 08193 Cerdanyola del Vallès, Spain
5 Institut d'Estudis Espacials de Catalunya IEEC, 08860 Barcelona, Spain; gerard.portal@upc.edu
6 Campus Alpin, Karlsruhe Institute of Technology (KIT), 82467 Garmisch, Germany
* Correspondence: thomas.jagdhuber@dlr.de; Tel.: +49-8153282329
† These authors contributed equally to this work.

Highlights

What are the main findings?

- The spatial similarity of surface soil moisture data between operational passive microwave remote sensing (SMAP DCA) and Earth system modeling (ECMWF IFS) is assessed from a single cell to the continental scale across Europe.
- The environmental and observational drivers behind the spatial distributions of each soil moisture product are investigated to understand their influence on the spatial similarity.

What are the implications of the main findings?

- The newly established spatial pattern correlation metrics (mCD and disagreement) enable the understanding of how spatial patterns of SMAP and ECMWF products can be compared and their spatial correlation can be assessed despite missing spatial congruence.
- This spatial similarity study provides first the insights towards a more seamless fusion of remote sensing products and Earth system simulations as well as towards new strategies (e.g., designing in situ networks) to validate spatial patterns of soil moisture.

Academic Editors: Jaromir Krzyszczak and Anna Siedliska

Received: 16 December 2025
Revised: 5 February 2026
Accepted: 9 February 2026
Published: 15 February 2026

Abstract

Soil moisture is an essential climate variable exhibiting strong spatio-temporal dynamics, especially in the topsoil. Therefore, it is assessed multiple times by sensors within in situ networks, satellites, and by modeling of the Earth system. The resulting soil moisture fields from all methods are individual and non-congruent due to the imperfection of the methods and retrievals. But their spatial patterns have valuable similarities that call for investigation to foster intercomparison or even fusion of soil moisture products. In this research study, the similarity of spatial soil moisture patterns between passive microwave remote sensing products and Earth system modeling is investigated. We configure and apply spatial similarity metrics to enable a spatial comparison of the operational SMAP Dual Channel

Algorithm (DCA) radiometer soil moisture product with the soil moisture output from IFS model runs of the ECMWF. The pattern assessment spans over the whole of Europe and aims to find the drivers behind the spatial soil moisture distributions at scales ranging from single grid cells (minimum) to continental (maximum) spatial scales, and between growing periods of wet (2021) and dry (2022) years. The two specifically configured metrics, total disagreement and mean category distance, showcase the opportunities and challenges when assessing spatial similarity in soil moisture fields across different scales. In addition, the potential drivers of the spatial moisture patterns were screened. Here, soil texture is the most influential single driver of spatial patterns in the IFS soil moisture runs, when analyzed in absolute terms [m^3 m^{-3}]. In relative terms of soil moisture [-] (soil wetness index), precipitation and soil temperature explain most of the variability of the IFS soil moisture for Europe. The SMAP retrievals are predominantly driven by the brightness temperatures, mostly influenced by surface temperature, vegetation water content, and soil roughness. These differences in drivers, as well as in methodology, culminate in an inherent discrepancy between the two soil moisture products. However, the assessment of their spatial patterns reveals the underlying similarity from the local to the continental scale.

Keywords: ECMWF; SMAP; surface soil moisture; spatial variability; spatial correlation metrics; disagreement; Europe

1. Introduction

Land–atmosphere interactions and their feedbacks are still not fully understood due to the manifold and complex interactions between the land surface and the surface proximate atmosphere. The soil–vegetation layer at the land–atmosphere interaction boundary acts as the gatekeeper for mass (water), energy (temperature), and gas (oxygen and carbon dioxide) transfer across this boundary. Especially the moisture content of this soil layer guides the partitioning into latent and sensible heat fluxes, steering evaporation, the provision of plant-available water for uptake, as well as the plant water status and demand for transpiration and photosynthesis. This implicates the potential to transpire and therefore to conduct photosynthesis for plant vitality and growth ([1–4]). Soil moisture is highly variable in time and space and needs a dense monitoring scheme to understand its spatio-temporal variability ([5–7]). Moreover, the spatial distribution of soil moisture across the landscape is an essential asset for reasonable and reliable Earth system modeling, and its role in modeling is highly scale dependent ([8–11]). Hence, spatially resolved information on soil moisture from remote sensing can provide an input (assimilation or direct insertion) or a validation source for Earth system models on a regular temporal and spatially extended basis ([12–16]). In addition, modeled soil moisture fields can provide inputs to initialize and improve soil moisture retrieval from remote sensing (e.g., starting values for a time series retrieval approach [17,18]) or help to assist with certain steps within the remote sensing retrieval process (e.g., permittivity to soil moisture conversion).

However, there is hardly any option to rigorously validate the spatial distribution of soil moisture on larger spatial scales, since observed quantities are mostly available on a point-scale basis from soil moisture networks like the International Soil Moisture Network (ISMN) and the Integrated Carbon Observation System (ICOS) ([19,20]). This is why most studies, which compare different remote sensing retrieved or Earth system modeled soil moisture products, focus on spatio-temporal comparisons with time series data from in situ soil moisture sensors on point locations ([21–26]). For this, a similar area ratio was proposed as a quantitative metric in [27] to select in situ sites with the best spatial

representativeness or even acting as a rationale for spatially deploying new sites for robust validation of satellite products. In addition, it was used in several studies for validating satellite and model-based SM products [28,29].

Furthermore, only a few studies, such as Escorihuela & Quintana-Seguí in 2016 [30], compare operational satellite-based soil moisture products from remote sensing (e.g., Soil Moisture Ocean Salinity (SMOS), Advanced Scatterometer (ASCAT), Advanced Microwave Scanning Radiometer for EOS (AMSR-E)) directly with outputs from a land surface model (e.g., SURFEX—SURFace EXternalisée in French). However, Escorihuela & Quintana-Seguí use only Pearson's correlation as a pixel-based metric, not considering the (local) spatial neighborhood [30]. Moreover, Hain et al. in 2011 added thermal infrared sensing for soil moisture estimation and combined it with microwave-based retrievals [31]. They compared the individual and joint retrievals with soil moisture output from National Oceanic and Atmospheric Administration (NOAA) land surface model runs (version 2.7), but on the basis of a spatial and a time series anomaly analysis. Again, spatial analyses were pixel-based and did not incorporate the local neighborhood. The same applies to the study of Wang et al. in 2024, which is, however, the most comprehensive study found so far [32]. Here, eight operational satellite soil moisture products (SMAP DCA, SMAP-IB, SMAP MTDCA, SMOS-IC, LPRM AMSR2, JAXA AMSR2, FY-3C, and ESA CCI; see the abbreviations list at the end of the paper) and the ECMWF Reanalysis version 5 (ERA5) and Global Land Data Assimilation System (GLDAS) reanalysis soil moisture datasets are used together with in situ soil moisture networks, to understand the performance of three strategies (pixel-based cross validation with model/reanalysis datasets, in situ point comparison and Extended Tripple Collocation) for quality assessment of these soil moisture products at a global scale.

Moreover, Vischel et al. in 2008 [33] compared soil moisture fields, represented by the soil wetness index, from the physics-based hydrological model TOPKAPI (TOPographic Kinematic APproximation and Integration) with remote sensing estimates from European Remote-Sensing Satellite (ERS) scatterometer data. However, the comparison was done on spatial averages of the entire Liebenbergsvlei catchment (4.625 km^2, South Africa) or the ERS footprints (~50 km). Hence, an assessment of the similarity in spatial patterns of soil moisture was not possible.

Hence, there is a lack of continent-wide studies that quantify the similarity between spatial patterns of soil moisture and their spatio-temporal variability beyond simple pixel-based comparisons (rooted in point-scale in situ measurement validation). A quantitative assessment of spatial similarity is crucial when combining, integrating, or fusing two or more spatially resolved soil moisture products, as no method is perfect and can serve as a standard in spatial pattern recognition. This knowledge gap leads to the following research questions:

(a) How similar is the soil moisture spatially distributed in operational remote sensing retrievals and Earth system model outputs across Europe?
(b) Which environmental and observational drivers determine the spatial distributions of each soil moisture product?
(c) How do the spatial patterns of soil moisture products compare on different scales (single cell, local, regional, up to continental) as well as for varying hydrometeorological periods?

This study aims at assessing the similarity of spatial patterns of soil moisture from remote sensing and Earth system modeling across Europe (chosen as a continental domain due to best data availability for ECMWF runs and in situ comparisons) and understanding its environmental and observational drivers for two consecutive years (2021 and 2022) with different hydrometeorological conditions (excluding winter and frost conditions from October to February). The assessment of their spatial patterns explores the underlying

similarity from the local to the continental scale. This could be useful in the future for modelers to further improve the assimilation of both products, even beyond classical stationary (not considering the local neighborhood) data assimilation of satellite data into model runs. Moreover, an important step in assessing the similarity of spatial soil moisture is the establishment of strategies to compare or even validate spatial patterns among different soil moisture products.

2. Experimental Data and Study Region

In the following, we introduce two employed fully operational soil moisture products, one from passive microwave remote sensing within the Soil Moisture Active Passive (SMAP) mission and one from Earth system modeling (Integrated Forecast System—IFS—from the European Center for Medium-Range Weather Forecasts—ECMWF). This choice of datasets ensures the relevance of our study design for data users and practitioners who may aim to jointly use and apply both soil moisture products in the future.

2.1. SMAP Data

The SMAP mission was launched in April 2015 by the National Aeronautics and Space Administration (NASA) and provides a global time series of soil moisture over the land surfaces [34]. Initially, the space-borne instrument combined a radar with a radiometer sensor at L-band (1.4 GHz) for joint active-passive microwave sensing, but the radar sensor ceased its service after three months of operation [35]. Thus, in this study we will only use the soil moisture product of the SMAP microwave radiometer with a nominal footprint resolution of 36 km and a minimum temporal resolution of 2–3 days according to European latitudes. The exact SMAP dataset applied here is the SPL3SMP_E product [35], which includes morning and afternoon satellite overpasses. A 9 km gridding is organized via an enhanced Backhus–Gilbert interpolation of the coarser resolution brightness temperature scenes [36]. The soil moisture retrieval algorithm employed is the dual channel algorithm (DCA), which uses ancillary data, like soil type, soil texture, Normalized Difference Vegetation Index (NDVI), and land surface temperature [37].

The post-processing of the SMAP soil moisture consists of mandatory filtering of cells with non-optimum quality flags and surface properties (e.g., urban areas, water bodies). Afterwards, we re-grid the SMAP product to the IFS resolution (0.09°) using bilinear interpolation. We tested three different approaches for interpolation (bilinear, nearest neighbor, and distance-weighted average) in Appendix C and found very little (Scandinavia) to no (rest of Europe) discrepancies, resulting in an agreement of more than 94% for all of Europe and the entire observation period. Finally, we compute three-day running means of the product to cover all of Europe due to SMAP coverage. In one focus zone of our study region, encompassing Germany, Belgium, the Netherlands, and Luxembourg (Central Europe for brevity; see Section 2.4), the optional filtering strategy, based on vegetation water content, terrain slopes, and peatland coverage, removed 75% of the cells. This largely owes to the intermediate to high vegetation water content in Central and Eastern Europe (cells with a vegetation water content > 5 kg/m^2 were removed) and is excessively strict for the sake of keeping a reasonable number of samples. In an adapted optional filtering strategy, we softened this criterion to 10 kg/m^2—still screening very dense forests—and combined it with orographic (terrain with gentle slopes < 3°) and land surface criteria (peatland fraction below 12.5%). The value of 12.5% is based on peatlands being characterized by, on average, 40% open water areas [37]. A fraction of 40% out of 12.5% results in 5% of open water within a pixel, which is acceptable for a high-quality retrieval [34]. Other peatland fraction thresholds (5% and 15%) were tested visually and considered to be either too loose or too strict, respectively. The optional filtering mask

builds upon a global peatland map from [38]. This adapted strategy filters about 18% of the land cells in Central Europe and 32.8% across all of Europe. However, we will not apply the optional filtering strategy when generating the results to avoid arbitrary filtering biases. Therefore, the spatial analyses are based on the mandatory filtering strategy, meaning all cells with successful retrievals. This allows us to keep spatial consistency, especially for the spatial pattern comparisons (see Section 4.3). In general, highlighting the filtered areas aims to provide insights into the quality of the satellite retrievals to ease the discussion and interpretation of differences in the spatial distribution and to flag the time series from SMAP in case the surface conditions are unfavorable for recommended quality retrievals.

2.2. ECMWF Operational Analysis

The IFS represents the numerical weather prediction system of ECMWF on a global scale. It consists of a spectral atmospheric model with a terrain-following vertical coordinate system applied at a horizontal resolution of 0.09° and 137 vertical levels [39,40]. IFS analyses serve as the initial conditions for the high-resolution 10-day forecasts from ECMWF. The ECMWF operational analyses used here are the output of a 4D-Var data assimilation system. The system merges an earlier forecast of the model with newly arriving observations to provide 6-hourly analysis fields at 00, 06, 12, and 18 UTC daily [40]. Physical processes in land and atmosphere are too small to be explicitly spatially resolved by the model grid [41]. Processes relevant for calculating soil moisture are integrated for the atmosphere in the short- and long-wave radiation schemes [42], whereby convection is parameterized by a modified version of the Tiedtke scheme [43,44]. Moreover, the cloud microphysics module is based on Tiedtke with improvements from [45–47]. Soil moisture processes in the land surface component are embedded in the ecLand module of the IFS [48]. It is a tiled setup within a single pixel, with no sub-tiling of the soil. The top layer is 7 cm, which matches the expected sensing depth of satellite soil moisture (~5 cm, but see also [49]). Three deeper layers are beneath it (21 cm, 72 cm, and 189 cm deep, respectively). The soil is assumed to be uniform across the entire depth. There are only two vegetation tiles (low and high). The others are snow, urban, bare, and water. Therefore, spatial patterns on the ground are only implicit.

2.3. In Situ Data

For comparison with SMAP DCA and IFS model runs, we gathered in situ soil moisture measurements from the ISMN ([19]; https://ismn.earth/en/ (accessed on 22 May 2024)) and ICOS ([50]; https://www.icos-cp.eu/observations/station-network (accessed on 10 March 2023)) networks across Europe. The ISMN database contains harmonized and quality-controlled data from more than 80 networks distributed across the world [19]. In this study, we use data from the networks REMEDHUS (Spain), SMOSMANIA (France), XMS-CAT (Spain), and RSMN (Romania) provided by ISMN. Additionally, ICOS monitors the atmosphere, land, and ocean using primarily tower-based instruments covering areas of 0.01 to 1 km^2. Further, they conduct measurements of air, plant, and soil parameters (including soil moisture) to support studies on factors affecting greenhouse gas fluxes (https://www.icos-cp.eu/ (accessed on 10 March 2023)). Characteristics of the areas where the stations reside are presented in Section 2.4. Of the originally 601 sites in ISMN and 180 sites in ICOS across Europe, a total of 107 sites provide soil moisture measurements in the topsoil with sensing depth between 4 cm and 6 cm for 2021 and 2022. Here, the sensing depth was chosen to ensure consistency with that expected from SMAP (~5 cm) and with the top layer of the ECMWF IFS soil moisture dataset (0 to 7 cm) [49]. In a first step, a quality check was performed for all time series. First, data with quality flags other than "good (G)" in the ISMN were removed. In terms of ICOS, a data set that had already been

quality-checked and contains only good data was used. Second, data with soil moisture above 0.6 [m^3 m^{-3}] were screened out due to soil saturation. Third, data beyond three standard deviations above and below the temporal average of each sensor were removed to avoid extreme outliers. Finally, we manually screened out specific periods when

- A lack of data in one or more sensors affected the site average (e.g., a data gap of some weeks occurring for two or more of the sensors at one measuring station).
- A sensor within one site behaved inconsistently with time trends compared to its nearby sensors and with precipitation patterns.

During the quality assessment, ten stations were completely removed. In the following step, all sites within each pixel were averaged as well. For each resulting time series, a centered 3-day running mean was computed, and data were saved in daily time steps to derive consistency with the satellite and modeling data. The 3-day averages are needed to account for the time SMAP needs to cover all of Europe. Finally, we filtered out pixels with in situ data gaps corresponding to more than 10 days of the full record in each year (this removed 29 stations), and sites where only data for one of the two growing periods were available (this removed 31 stations). This resulted in a final set of 42 stations with continuous, high-quality data records for both growing periods. Their locations are displayed in Figure 1.

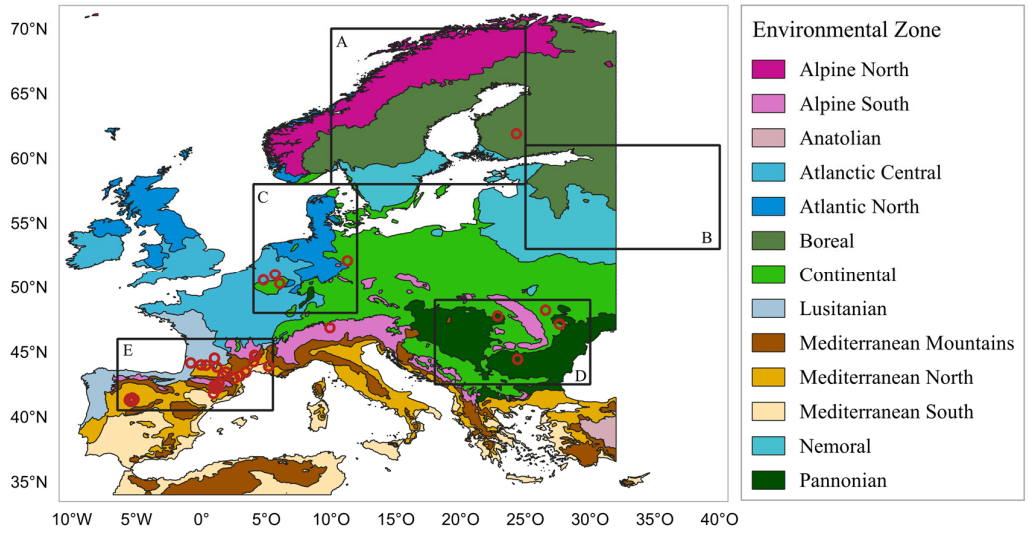

Figure 1. Environmental stratification of ecological zones after [51] overlaid with the five focus regions (black boxes) defined across Europe to cover the main ecological zones in the study. Regions (**A**–**E**) are defined in Table 1. The red rings indicate the in situ networks used at the beginning of the results section for classical in situ to pixel comparison.

Table 1. Introduction of five focus regions following the environmental stratification of ecological zones.

Zone	Name (with Main Countries)	Environmental Zone [51]
A	Scandinavia (Norway, Sweden, Finland)	Alpine North/Boreal
B	Western Russia	Nemoral/Boreal
C	Central Europe (Germany/Denmark)	Atlantic North/Continental
D	Eastern Europe (Hungary/Romania)	Pannonian/Continental
E	Western Europe (France/Spain)	Mediterranean/Lusitanian

2.4. Study Region and Focus Zones

The study region spans the entire European continent from Portugal in the West to the Ural Mountains in the East and from the northern tip of Norway to the southern

islands of Greece. We use the stratification of "environmental zones" from Metzger et al. in 2005 to characterize different homogeneous ecological zones with a spatial resolution of 1 km^2 [51]. This environmental stratification consists of 84 strata in total, which have been aggregated into 13 environmental zones. We have designed five focus regions across Europe for our analyses, from Scandinavia (zone A) to Southern France and Northern Spain (wider Pyrenees region as region E). All focus regions from A to E are presented in Table 1 and visualized in Figure 1. The focus regions span the ecological zones of Europe from North to South (zones A, C, and E) as well as from West to East (zones E, C, D, and B), including the continent's climatological, hydrological, and biological spatial variance and temporal dynamics.

3. Methodology

This section introduces the applied methodology and analysis tools to compare soil moisture from SMAP DCA retrievals, IFS runs, and in situ measurements. The spatial distribution will be investigated by calculating temporal medians for the growing periods 2021 and 2022 (March–September) and their differences across the datasets. Furthermore, differences in the spatial distributions will be analyzed by applying local spatial pattern comparison metrics, introduced and configured in detail in the following subsections. Soil moisture is either represented in absolute terms [m^3 m^{-3}] for quantifying the absolute fit (biases) or in relative terms as a dimensionless soil wetness index [-] for quantifying the relative fit (correlation dynamics). For the min-max normalization of the soil wetness index, the minimum and maximum values of the growing seasons (from March to September of 2021 and 2022) were employed. Please note that this index should not be confused with the SWI based on physical limits (wilting point and field capacity). We subtract from the 3-day running mean soil moisture the minimum of soil moisture, and then divide it by the difference between the maximum and minimum of soil moisture.

3.1. Spatial Pattern Comparisons

The first studies mentioned in Section 1 revealed different local structures in the soil moisture distribution of IFS fields and of SMAP DCA retrievals. Visual comparisons showed that the nature and the strength of the discrepancy in the spatial patterns differed between European regions, suggesting a need for analyses at local and regional scales instead of an immediate continental-scale evaluation of the patterns. Hence, we configure and apply two local and quantitative spatial pattern comparison metrics, the "disagreement" metric from Pontius Jr. and Millones (2011), complemented by the "mean category distance" [52], which is derived based on the fractional skill score [53]. Both metrics are based on categorized variables having inter-comparable unitless values. This enables evaluating whether datasets of variables agree within a certain range of values, and thus allows some degree of uncertainty, meaning discrepancy in quantitative (absolute) values. However, the categorization (see Section 3.1.1) involves a certain loss of detail in spatial accuracy.

For the local spatial pattern comparisons between categorized IFS soil moisture fields and categorized SMAP soil moisture products, we apply the spatial pattern comparison using moving windows of fixed grid size. Please note that in this study, we define the category "local" in a spatial sense, meaning the center cell and its surrounding neighboring cells. A detailed description of the size of the local windows and their derivation follows in Section 3.1.2. This study will be done separately for the years 2021 and 2022.

3.1.1. Categorization of Soil Moisture Fields

The categorization of the soil moisture fields from IFS and SMAP DCA requires the consideration of (1) how to objectively define the bin edges (category boundaries) of soil moisture, and (2) how many soil moisture categories or bins are needed. Firstly, the bin edges were defined based on fixed boundaries rather than relative values of the soil moisture fields. This fixation of bin ranges (absolute units [m^3 m^{-3}]: 0.05 bin width between 0 and 0.6 = 12 bins; relative units [-]: 0.1 bin width between 0 and 1.0 = 10 bins) has advantages and disadvantages. The benefit is that fixed ranges deliver equally sized and therefore universally comparable bins, which are independent of the value distribution of the soil moisture variables. As a drawback, it may be hindering that the known biases between the datasets (IFS vs. SMAP DCA) cannot be directly considered and accounted for. Moreover, to determine the bin count, we balanced the potential agreement between the datasets, signified by exact agreement in the distribution range between the datasets, and the representation of detail. For the latter, we compare the spatial correlation and the root mean squared error between the original quantitative field and the categorical field (using the threshold value of soil moisture of each category) within each IFS model run or SMAP DCA dataset. A high number of bins (100) would ensure a high accuracy, denoted by high spatial correlations and low RMSE of the quantitative and the categorical fields for all datasets (see Figure 2). However, the disagreement between the fields would increase to almost 100%. On the contrary, a low bin count (2) would allow for agreement between the datasets at the cost of detail accuracy. Figure 2 depicts the disagreement and detail loss statistics on the continental scale (Europe). Here, land mass distribution and the representation of basic features of the soil moisture distribution, such as low soil moisture values in the Sahara Desert, force an elementary agreement in the distributions. This leads to spatial correlation coefficients between 0.7 and 0.77 already with the bin count of 2. Based on this balance, the following analyses were conducted using 12 bins for the absolute soil moisture analysis, and 10 bins for the relative analysis (Soil Wetness Index).

Furthermore, we added a dedicated study with 8, 10, and 12 bins in Appendix D to investigate the sensitivity to the categorization (binning) of the soil moisture in more detail (beyond Figure 2).

3.1.2. Determining the Spatial Window Size for Pattern Analysis

The determination of the window size is based on balancing the degree of heterogeneity of the soil moisture fields with the loss of the dynamic range of soil moisture [54]. This means that the window size needs to be large enough to capture the majority of the heterogeneous land surface elements, but not too large to avoid smoothing out the spatial variability.

The heterogeneity in the soil moisture patterns was determined by applying an adaptation of the local heterogeneity length scale by Tian et al. in 2022 [55] to the 10-bin categorical soil wetness index fields and the 12-bin categorical absolute soil moisture fields of each dataset, respectively (Figure 3). The local heterogeneity length scale is determined in a two-step approach. Firstly, the ratio of the area of the center cell category to the area of the chosen window for all length scales ranging between a single cell and a window size of 41 by 41 cells (maximum shift of 20 cells in each direction) is calculated. Secondly, it determines the last length scale, in which this area ratio of the center cell category is larger than 95%, which suggests nearly homogenous conditions within the window. The heterogeneity length scale corresponds to the maximum window size with homogeneous conditions.

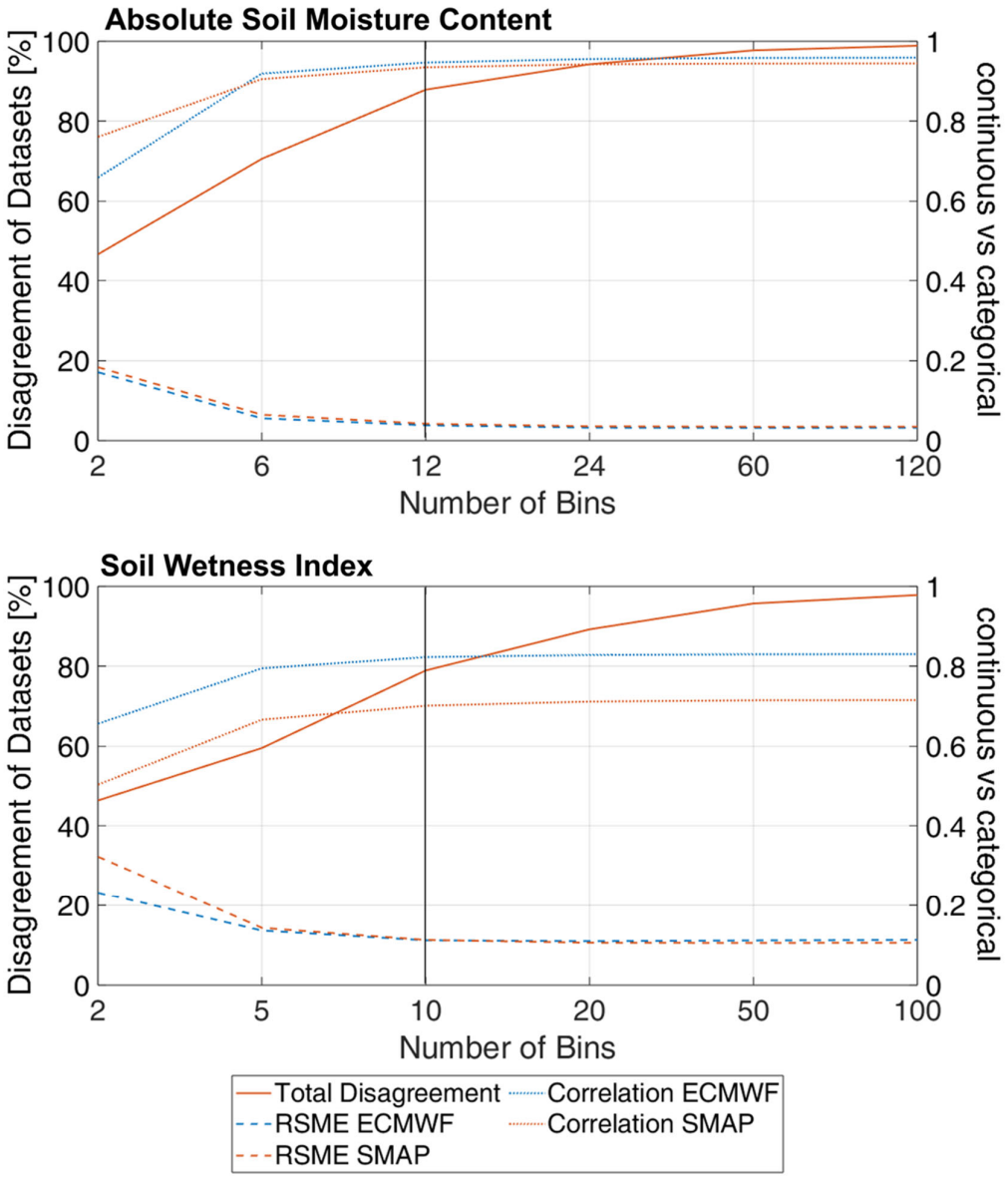

Figure 2. Global total disagreement between the SMAP DCA product and IFS model runs balanced against the spatial loss of details when categorizing the original quantitative soil moisture field with 2, 5, 10, 20, 50, and 100 bins. The loss is represented in terms of the root mean squared error (RMSE) and a spatial correlation between the quantitative and the categorized fields. The vertical black line represents the number of bins finally chosen.

In both the absolute soil moisture content and the soil wetness index, at least 75% of the grid cells have a local heterogeneity length scale below 5 × 5 cells in the investigated datasets (Figure 4). Higher heterogeneity length scales, suggesting more homogeneous conditions, occur primarily in Scandinavia and over southwestern Russia for the absolute soil moisture. The patterns of the soil wetness index show the largest homogenous patches over western Russia in ECMWF analysis and only a few patches with heterogeneity length scales larger than 19 × 19 cells for SMAP DCA.

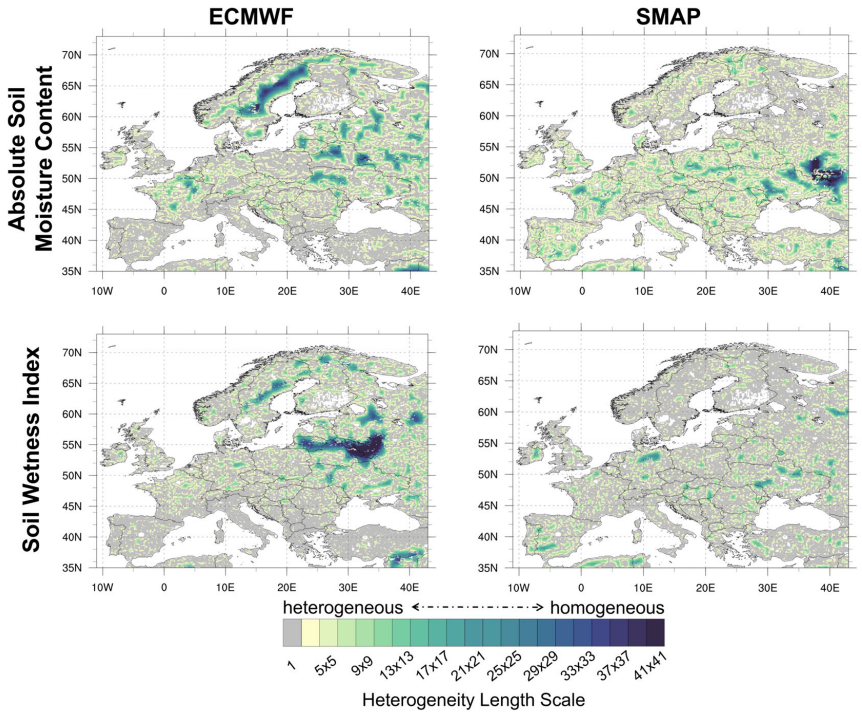

Figure 3. Local heterogeneity length scale across Europe for both investigated datasets (SMAP DCA and IFS), provided as maximum cell edge count with a homogeneity of >95% of cells (41-by-41 cells) in the same category, with one denoting a heterogeneous region around one cell and 41 denoting a highly homogeneous region around one cell.

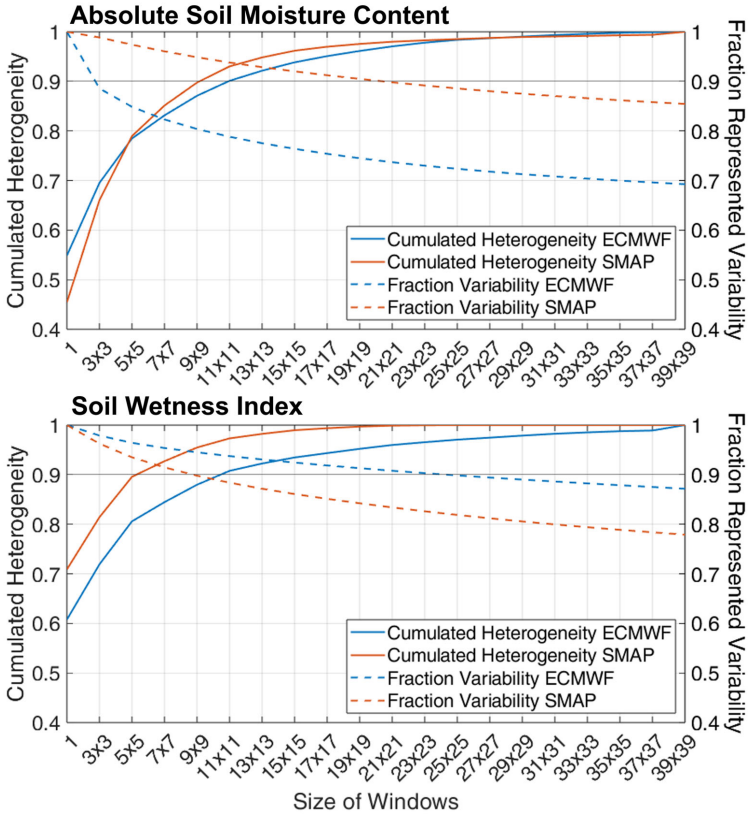

Figure 4. Accumulated heterogeneity plotted against fractional spatial variability in terms of the standard deviation of the smoothed soil moisture field divided by the standard deviation of the categorized unsmoothed soil moisture field for each dataset and algorithm.

For determining the optimum window size, we balanced the heterogeneity length scale with the fraction of represented spatial variability. The latter was derived by calculating the spatial standard deviation of the smoothed, categorized soil moisture and soil wetness index fields. The fields were smoothed by averaging all cells within a moving window. Finally, the spatial standard deviation of the smoothed, categorized fields was divided by the spatial standard deviation of the continuous fields to represent the loss in detail. We aimed for the representation of at least 90% of the original spatial variability while choosing a window that is large enough to encompass most heterogeneous features in the soil moisture and soil wetness index fields. However, for the soil moisture content of the IFS runs in absolute terms, the fraction of represented spatial variability already dropped below the 90% criterion at a window size of 3 × 3 cells (Figure 4). For the rest, we found the best balance at a window size of 5 × 5 cells for both (IFS and SMAP DCA) products. Thus, the size of the moving windows for the comparison analyses was uniformly set to 5 × 5 cells for all local spatial pattern comparisons presented in Section 4 in order to ensure comparability across retrievals.

3.2. The Disagreement Metric as Indicator for Spatial Pattern Similarity

Figure 5 shows the workflow and the methodology of the spatial pattern similarity analysis for an example subregion of 21 × 21 grid cells in Central Europe. In panel 1 of Figure 5, the soil moisture fields of the IFS model run and of the SMAP DCA retrieval are shown with their individual spatial patterns. In panel 2 of Figure 5, the transformation into categorial variables for both products by applying fixed bin ranges for soil moisture is presented, applying a boundary relaxation tolerance window of 2%. The tolerance criterion is set to relax the category allocation of both products at the boundaries of the categorial fields.

Afterwards, the disagreement metric is introduced to compare two categorized soil moisture fields. It was originally developed to compare land cover datasets or assess land cover changes based on a cross-tabulation matrix (Figure 5(3b)), which merges the information about the categories of two datasets under comparison [52]. The cross-tabulation matrix provides the fraction of the cells in the study area being in category i in one dataset and in category j in the other dataset. An agreement is found when the datasets have the same category in a cell ($i = j$). As an example, in the cross-tabulation matrix of Figure 5(3b), a fraction of 0.02 cells is classified as category 2 in the IFS analyses and category 8 in SMAP DCA ($x_{i=2,\,j=8} = 0.02$). Thus, we define the total disagreement D as follows:

$$\text{Total disagreement}: \; D = \sum_{j=1}^{b_{max}} \frac{d_j}{2} \text{with } d_j = \left[\sum_{i=1}^{b_{max}} \left(x_{i,j} + x_{j,i}\right)\right] - 2x_{j,j}, \tag{1}$$

where b_{max} is the number of bins and x are the elements in the cross-tabulation matrix between SMAP DCA and ECMWF IFS (see Figure 5(3b)). The total disagreement metric consists of the quantity and the allocation disagreements (see Figure 5(3c) for calculation and [52]). The quantity disagreement represents differences in the quantity of occurrence of all categories, independent of where they are located within the domain. The allocation disagreement adds whether the category is in the same location within the spatial domain (e.g., neighborhood of 5 × 5 grid cells) or whether it is in different locations for the two datasets.

All components are first assessed for the individual (single) category and then summed over all categories to get the overall disagreement. In Figure 5(3c), we exemplarily assess the disagreement components of category 9. It occurs in a fraction of 0.06 cells in the SMAP DCA product and a fraction of 0.21 cells in the IFS product. Hence, there is a difference $q_j = 9$ of 0.15, which represents the quantity disagreement of category 9. The quantity dis-

agreement over all categories (Q) is then derived from summing the quantity disagreements over all individual categories and dividing them by two (average of both datasets).

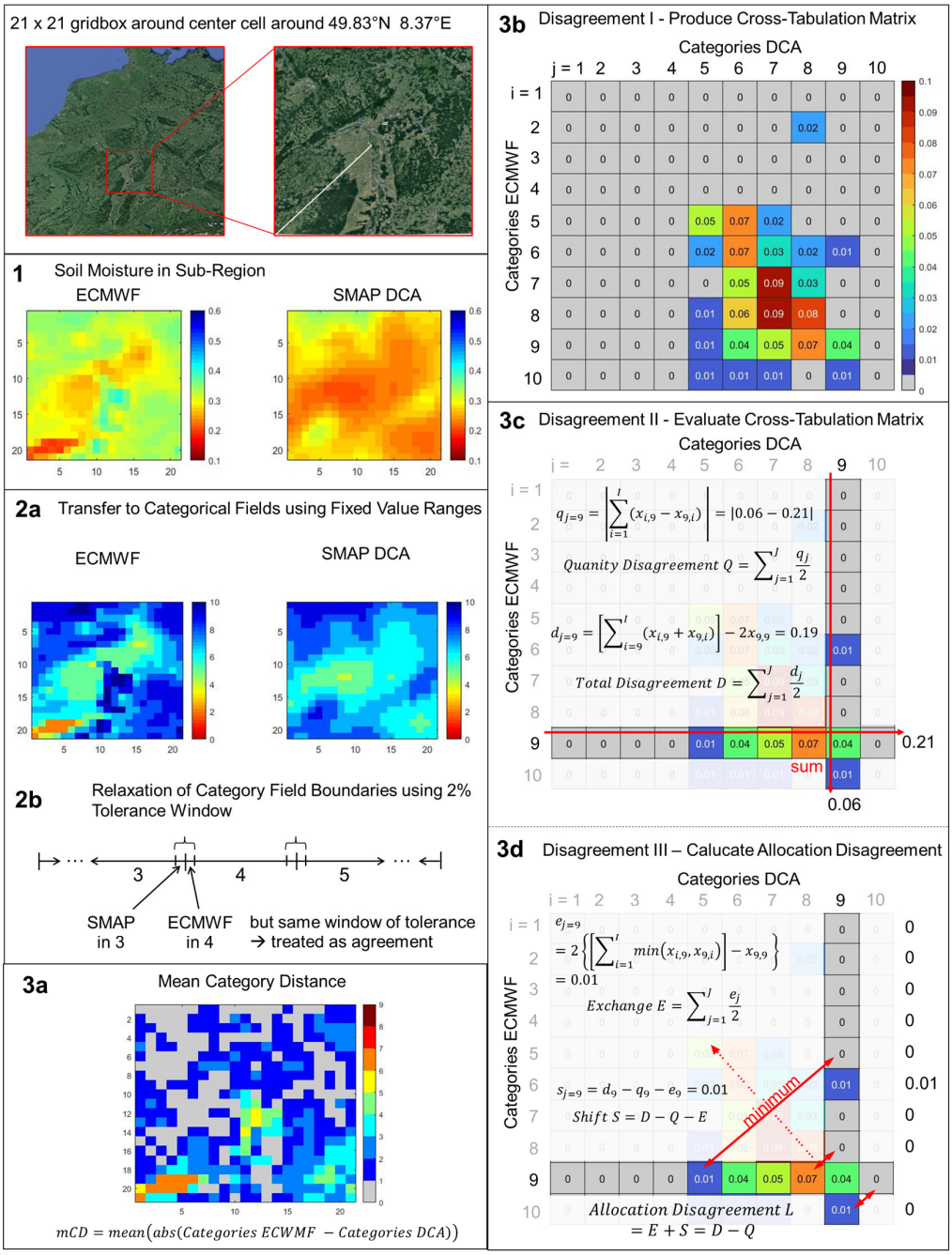

Figure 5. Methodology of the spatial pattern analysis for an example subregion of 21 × 21 grid cells in Central Europe (distance via white scale in gridbox). (**1**): Soil moisture fields of IFS model run and of SMAP DCA retrieval; (**2**): Transformation into categorial variables applying fixed bin ranges and boundary relaxation tolerance window (2%); (**3a**): Result of mean category distance (mCD) metric; (**3b**): Production of cross-tabulation matrix as first step of disagreement metric calculation; (**3c**): Evaluation of cross-tabulation matrix for quantity and total disagreement metrics; (**3d**): Computation of Exchange E and Shift S for calculating allocation disagreement metric.

$$\text{Quantity disagreement}: \; Q = \sum_{j=1}^{b_{max}} \frac{q_j}{2} \text{with } q_j = \left|\sum_{i=1}^{b_{max}} \left(x_{i,j} - x_{j,i}\right)\right|. \tag{2}$$

The allocation disagreement in turn can be distinguished into the components of exchange (E) and shift (S). Exchange occurs in pairs. Hence, the categories occur in the same quantity in the domain, but in different locations. The allocation disagreement is computed by assessing two times the minimum of $x_{i,j}$ and $x_{j,i}$. In the example cross-tabulation matrix of Figure 5(3d), the change between categories 6 and 9, for instance, is 0.01, with $x_{6,9}$ = 0.04 and $x_{9,6}$ = 0.01. Shift summarizes the residual disagreement and occurs between multiple categories. It can involve a local displacement combined with a change in the category, for instance. However, in this study, we always refer to the allocation disagreement without further distinction into its components, exchange or shift, following Pontius and Millones [52,56].

$$\text{Exchange E} = \sum_{j=1}^{b_{max}} \frac{e_j}{2} \text{with } e_j = 2\left\{\left[\sum_{i=1}^{b_{max}} \min\left(x_{i,j}, x_{j,i}\right)\right] - x_{j,j}\right\} \quad (3)$$

$$\text{Shift : S} = \sum_{j=1}^{b_{max}} \frac{s_j}{2} = D - Q - E \; with \; s_j = d_j - q_j - e_j \quad (4)$$

The allocation disagreement L can simply be derived from the sum of E and S, which corresponds to the difference between D and Q:

$$\text{Allocation disagreement : } L = E + S = D - Q. \quad (5)$$

The total disagreement (D) is the sum of all fractions not located on the diagonal of the cross-tabulation matrix. Consequently, the total agreement equals the sum of all fractions on the diagonal. This value was used for determining the number of considered bins for the evaluation (see Section 3.1.1).

3.3. The Mean Category Distance Metric as an Indicator for Spatial Pattern Similarity

The above-mentioned disagreement metric is limited to the scope of the datasets that disagree in the range of their soil moisture value distributions. Since soil moisture is originally a spatially continuous (boundaryless) variable, different from land cover, which has discrete classes and boundaries, the mean category distance (mCD) metric informs about the extent to which the datasets disagree in their value distribution ranges. This metric, thus, builds a bridge between the disagreement metric (see Section 3.2) and the differences in the mean soil moisture of the two products. Therefore, it adds information about differences in the shape of the value distribution of soil moisture.

The mCD is computed by calculating the mean squared errors of the category number of the IFS runs c_i^{IFS} with the category numbers of the SMAP DCA retrievals c_i^{DCA}, normalized by the count of all categories n:

$$mCD = \frac{1}{n}\sum_{i=1}^{n}\left(c_i^{IFS} - c_i^{DCA}\right)^2. \quad (6)$$

The normalization ensures a value range between 0 (total agreement) and 1 (total disagreement) in the spatial window (see Section 3.1.2). The mCD was developed from the fractional skill score [53], which is often applied for validating the locations of precipitation events simulated by Earth system models with in situ or remotely sensed observations. We reformulated it (see Section 3.1) because the original fractional skill score only works when both datasets agree to exceed a certain threshold (e.g., show a precipitation event). The case of not exceeding the threshold in both datasets (e.g., no precipitation) is not defined for the fraction skill score metric. Since soil moisture is a spatially continuous field and a match would be found by exceeding as well as not exceeding the threshold, the original fraction skill score would only detect half of the possible matches compared to the newly

developed mCD. In the following, the spatial similarity metrics (mCD and disagreements) will be applied to the SMAP and IFS datasets.

3.4. Design of Analyses for Assessing the Environmental and Observational Drivers of Soil Moisture Patterns

In Section 4.2, the environmental and observational drivers for the soil moisture patterns from the SMAP DCA retrievals and those from the IFS runs of ECMWF will be analyzed. For this analysis, we apply a multi-linear regression model via the Matlab©-function *fitlm*. At first, we set up a baseline model, which includes available driving variables and a few carefully selected interactions for the respective datasets. For the SMAP retrieval, the base model in Wilkonson formulation is described as follows:

$$SM \sim VWC + clay + T_{sfc} + TB_v + TB_h + rough + VWC{:}T_{sfc} + VWC{:}clay + rough{:}TB_v + rough{:}TB_h + Lat{*}Lon$$

The v- and h-polarized emissivity, as well as the bulk density, were excluded to reduce multicollinearity between the driving variables. The model formulation for the IFS data looks as follows:

$$SM \sim LAI + T_{sfc} + precip + Veg_L + Veg_H + SoilText + LAI{:}T_{sfc} + LAI{:}SoilText + SoilText{:}Precip + T_{sfc}{:}Precip + Lat{*}Lon$$

With LAI being the sum of the LAIs from low and high vegetation, Tsfc being the surface temperature from the model, precip being the total precipitation, VegL and VegH being the low and high vegetation types in the model, and SoilText being the prescribed soil texture types in the model system. Please note that the total precipitation fields stem from the 24-hour forecasts initiated from the 00 UTC operational analysis.

Secondly, since all driving variables and the soil moisture fields are spatially structured (see Table A3, Moran's I), we must acknowledge that spatial autocorrelation causes an overestimation of the probability statistics (*p*-values) and the t-statistics in a linear regression model. However, we assume that the relative importance of the respective variables in X remains robust. In order to test this hypothesis, we performed sensitivity tests using different special autoregressive models [57] to analyze whether the relative magnitudes of the beta-coefficients for the individual drivers remain similar under consideration of spatial autocorrelation or not. In Appendix E, we explain the theory and the results of this analysis in detail. The spatial autoregressive models either consider spatial autocorrelation in the dependent variable by including a spatial lag (SAR), spatial autocorrelation in the error term (SEM), or they consider both (SAC) [58]. This analysis showed that the signs of the major drivers, as well as their magnitudes relative to each other, remain similar. Hence, a ranking of the relative importance of the different variables in creating the spatial patterns can be based on a multi-linear regression model, although, e.g., *p*-values are likely overestimated and should be interpreted with caution.

Finally, we applied the linear regression model to investigate the relative importance of the drivers for the spatial patterns in soil moisture by fitting reduced models and taking out one driver and its respective interactions at a time. We then compared the coefficients of determination from the base model with those from the reduced models to understand how much additional information the driver and its respective interactions provide compared to a model without the respective driver. In Section 4.2. we will show this discrepancy as a loss of explained variability, calculated from the difference in the coefficient of determination from the base model with the one from the reduced model, where one driver and its respective interactions were discarded.

4. Results and Discussion

This section starts in Section 4.1 with investigating the deviations of IFS modeled and SMAP DCA satellite-derived soil moisture for the growing seasons (March–September) in 2021 and 2022 over all of European continent. In Section 4.2, SMAP and IFS products are investigated individually to assess the environmental and observational drivers causing the spatial patterns. Afterwards, in Section 4.3, the similarity in spatial pattern distribution is investigated and evaluated from single-cell (minimum) to continental (maximum) scales, applying the metrics introduced in Section 3. Please note beforehand that the monthly averages of total soil moisture, as well as the 5th and 95th percentiles, revealed very similar patterns in the differences between both datasets (IFS and SMAP). Hence, the following analysis focuses primarily on the means of the growing seasons (March–September). Here, we show soil moisture in absolute terms [m^3 m^{-3}], and in relative terms [-], meaning as a soil wetness index.

4.1. Spatial Distribution of Mean Total Soil Moisture Content

The mean total soil moisture distribution of IFS model runs and SMAP DCA retrievals over Europe is investigated in the following. We examine similarities and differences in the amount of soil moisture in both datasets, including the five different focus regions (Figure 1) across Europe, and evaluate these against in situ measurements (see Section 2.3) to provide the classical (point-to-pixel) evaluation reference and a basis to assess the absolute soil moisture patterns. The total soil moisture of IFS model runs, and SMAP DCA retrievals averaged over the entire vegetation period from March to September 2021, is depicted in Figure 6 in absolute terms [m^3 m^{-3}] and in Figure 7 in relative terms [-] (soil wetness index), including zooms into the five focus regions. The circles denote the mean total soil moisture of the topsoil layer measured at in situ stations of the ICOS and ISMNs.

In Figure 6, the mean soil moisture from SMAP DCA retrievals is significantly drier across Europe than that of the IFS model runs, which is consistent with the literature [59–63]. Both SMAP DCA and the IFS soil moisture products exhibit a north–south gradient in top-layer soil moisture across Europe, which follows the environmental stratification of ecological zones (Figure 1) driven by climate and weather. However, this gradient is clearly more pronounced in the SMAP DCA retrievals and extends from the dry Mediterranean to the very wet Scandinavia, same as with height from sea level to the Alps.

In the IFS model runs, Scandinavia shows drier topsoil than Central Europe and the Eastern European plain. In more detail, topographical features such as the Alps, the Pyrenees, the Carpathian and Scandinavian mountains are clearly reflected in moister top soils compared to surrounding regions, while valleys have less topsoil moisture. IFS model runs have higher soil moisture in regions where orographically triggered precipitation (e.g., western coastlines and the western edges of mountain ranges) can be expected. Although SMAP comes from a coarser original resolution (36 km), the dependence on the orography is more clearly visible in the SMAP products than in IFS model runs.

Additionally, the Finnish wetlands and western Russian plains (regions A and B) show high soil moisture and significantly different patterns between the SMAP DCA retrievals and IFS model runs. Moreover, the differences in spatial patterns are especially visible when comparing the five focus regions for the two retrievals, where the northern focus regions A-C reveal these stronger pattern differences than the southern one (region E). When compared to Figure 7, where the soil moisture is illustrated in relative terms, the latitudinal spatial pattern does not change severely at a continental scale.

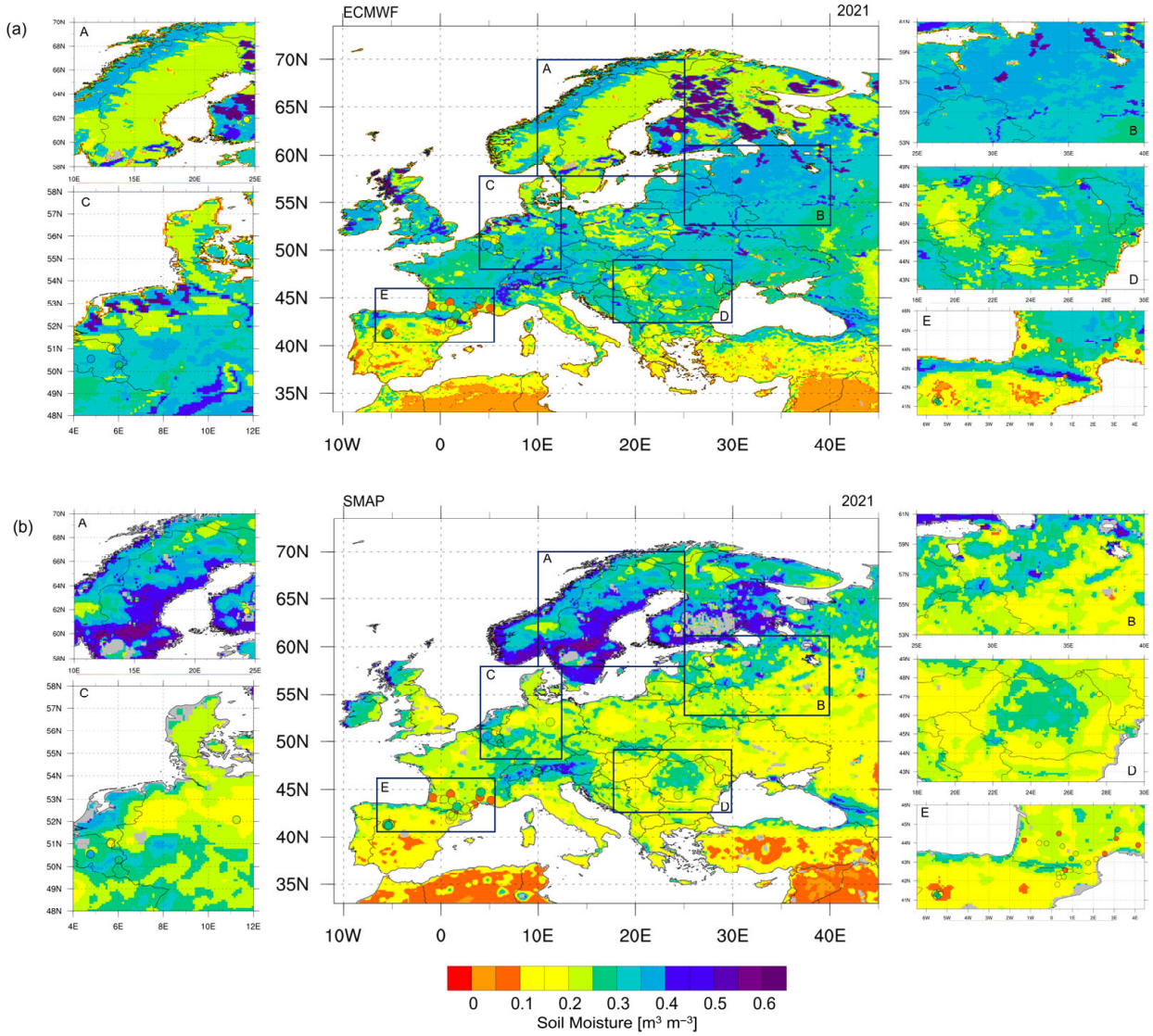

Figure 6. Total soil moisture amount [m^3 m^{-3}] averaged over the vegetation period 2021 (March–September) for (**a**) IFS model runs and (**b**) SMAP DCA retrievals. Both subplots include zooms on the five focus regions: (**A**) Scandinavia, (**B**) western Russia, (**C**) Central Europe, (**D**) Eastern Europe, and (**E**) Western Europe. The location and mean soil moisture of the ICOS and ISMN station data are represented by the circles (color-filled gray circles), where coloring follows the legend.

Nevertheless, at regional and local scales, the spatial patterns are distinctively different from the ones in Figure 6 (soil moisture in absolute terms). Here, the strongest discrepancy appears in Scandinavia and western Russia (including regions A and B) for both products.

Focusing on the in situ measurements (circles in Figures 6 and 7) for comparison, especially the Western European zone (E) stands out where the measured soil moisture at the single locations is significantly lower (below -0.1 [m^3 m^{-3}]) than the estimated soil moisture from SMAP DCA (higher than 0.15 [m^3 m^{-3}]) and the simulated one from IFS model runs (higher than 20 vol.%). This may indicate that, especially for drier soil conditions, the discrepancy from point in situ measurements to pixel-size averages (IFS or SMAP DCA) of soil moisture is larger due to an increased spatial heterogeneity in the drying process (Figure 3). A detailed assessment of all stations for both products is not provided here, but in a companion study focusing on the temporal dynamics between IFS and SMAP DCA [64], since the focus of this study is not on the point-to-pixel comparison

between in situ measurements and gridded products (IFS, SMAP DCA), but on the spatial pattern assessment.

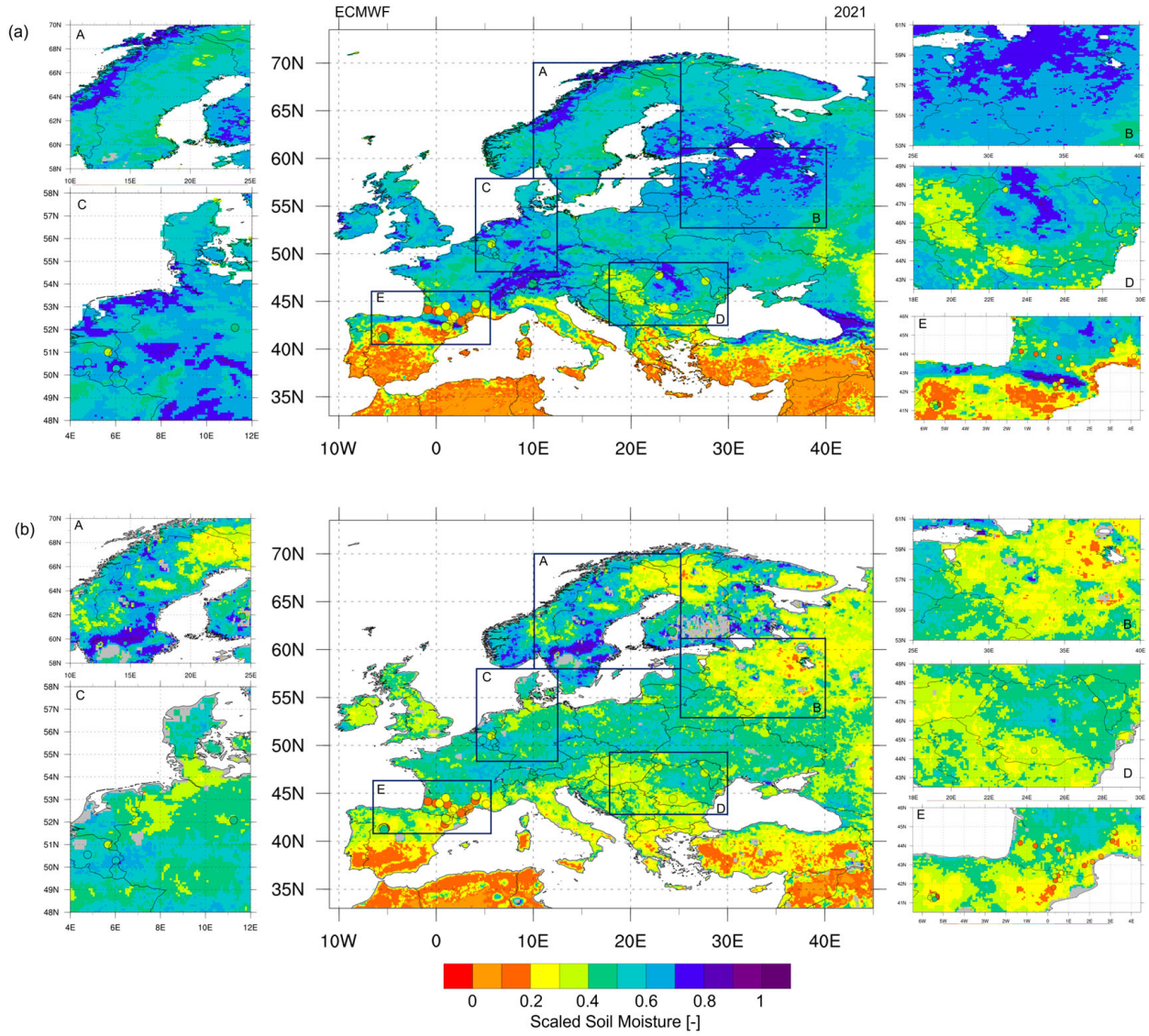

Figure 7. Relative soil moisture (soil wetness index) [-] averaged over the vegetation period 2021 (March–September) for (**a**) IFS model runs and (**b**) SMAP DCA retrievals. Both subplots include zooms on the focus regions: (**A**) Scandinavia, (**B**) western Russia, (**C**) Central Europe, (**D**) Eastern Europe, and (**E**) Western Europe. The location and mean soil moisture of the ICOS and ISMN station data (color-filled gray circles), where coloring follows the legend.

Furthermore, the differences in soil moisture patterns of both datasets between the two vegetation periods under investigation (2021 and 2022) are shown in Figure 8. These patterns are spatially similar in absolute as well as in relative terms (rows in Figure 8). However, they are noticeably different between IFS model runs and SMAP DCA retrievals (columns in Figure 8), suggesting a strong component in methodological (remote sensing retrieval vs. Earth system modeling) differences between the datasets. Nonetheless, the interannual differences in soil moisture between the years 2021 and 2022 shown in Figure 8 still indicate that both data products reflect climatic influences on soil moisture in a reasonable manner. Both show clearly lower soil moisture in Western Europe and North of the Black Sea in 2022, but slightly higher moisture (in the range of 0.02 to 0.04 [m^3 m^{-3}]) in the Mediterranean and the Eastern European Plain. The strongest difference in Figure 8,

indicating wetter soils in 2021, occurs in Central Europe (Zone C). Here, severe flooding events happened in 2021, leading to soil moisture values well beyond field capacity [65].

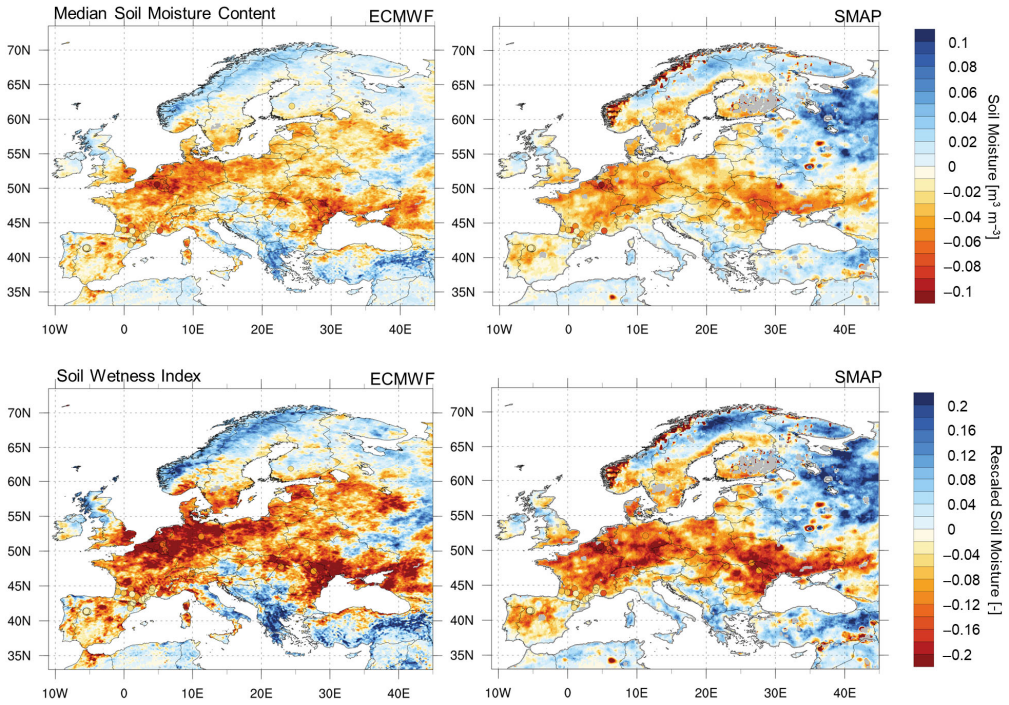

Figure 8. Temporal differences (2022–2021) in soil moisture as mean over the vegetation period (March–September) ((**top row**): absolute terms [m^3 m^{-3}], (**bottom row**): relative terms [-] = soil wetness index) for IFS model runs (**left column**) and SMAP DCA retrievals (**right column**). The maps are overplotted with colored circles indicating the soil moisture as the mean over the vegetation period (March–September) at all stations with consistent data records in 2021 and 2022. Reddish colors indicate that 2021 was drier than 2022, while blueish colors indicate that 2022 was drier than 2021.

4.2. Environmental and Observational Drivers of Soil Moisture Patterns

In the following, we investigate the underlying drivers influencing the spatial distributions in the soil moisture datasets of IFS and SMAP DCA products. This means exploring whether soil moisture patterns follow inputs for their product generation, for instance, soil properties, vegetation cover, air temperature, precipitation, and other representative environmental factors, as well as observational factors, like brightness temperature or emissivity, in the SMAP retrieval. Hence, the set of investigated drivers is specific to each method (IFS model run vs. SMAP DCA retrieval) and assessed across Europe. Furthermore, Peng et al. studied global soil moisture trends and their driving mechanisms, providing a larger perspective than our study [66]. They report greater influence of precipitation and vegetation than temperature on a global scale. Moreover, in the long-term study by Li et al., they use global GLDAS-Noah SM data from 1948 to 2024 to decompose the total soil moisture variability into two major temporal dynamics: long-term trends and inter-annual variability, going far beyond the time scale of this study, providing the longer-term perspective [67].

For indicating the explained variability by each driver or driver combination, we calculate the individual loss of explained variability when the respective driver or driver combination is left out from the analysis (see calculus and statistical details in Section 3.4 and Appendix E). In this way, the influence of different environmental and observational drivers on the respective soil moisture product can be understood.

4.2.1. IFS Model Runs

In Earth system models, such as the IFS of ECMWF, the soil moisture amount is driven by vertical inflows (precipitation) and outflows (upward: evapotranspiration; downward: infiltration and deep drainage), and the water holding capacity of the soil is determined from the parameterized soil properties (porosity, matrix potential, and hydraulic conductivity). Moreover, biases of precipitation, variations in vegetation parameters (traits and phenology), as well as evapotranspiration fluctuating with surface temperature dynamics, also directly impact soil moisture content and its spatio-temporal variation. The IFS system applies an operational data assimilation scheme for SMOS and ASCAT satellite remote sensing data, and thus constantly updates the soil moisture fields in the direction of observed soil moisture [40]. The influence of data assimilation and the implemented model processes for soil infiltration are not considered in the multivariate regression analyses, which probably explains a certain amount of the missing variability in Figure 9 (see variabilities of the base model). Please note that precipitation fields are not directly available from the IFS analyses themselves. They come from the ECMWF forecasts starting at 00:00 UTC from the analyses. Hence, the regressions with precipitation need to be treated with caution due to this dependency, and they should be seen only as an indication for regional differences in the precipitation fields.

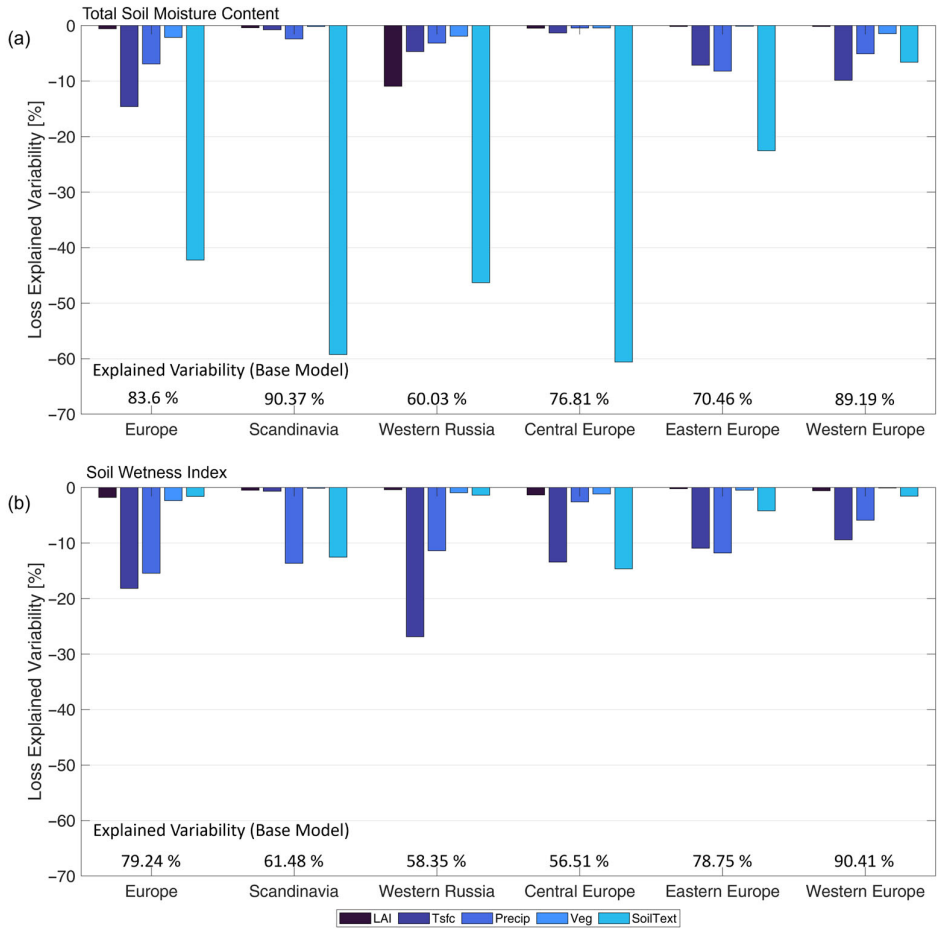

Figure 9. Loss of explained variability (coefficients of determinations) for taking single or combined variables out of the multi-variable regressions of potential influential factors of the IFS model runs with the mean top layer soil moisture over the vegetation period 2021 for all of Europe and the focus regions (A–E); (**a**) in absolute terms [m^3 m^{-3}], (**b**) in relative terms [-]; Tsfc = air temperature in 2 m height, Precip = accumulated precipitation of the day, SoilText = Soil texture, Veg = vegetation.

Soil texture is clearly the most influential single driver of spatial patterns in the IFS soil moisture product for all of Europe and for four out of five focus regions, when analyzed in absolute terms [m^3 m^{-3}]. It explains about 43% of the variability in the patterns in all of Europe and up to 59% of the patterns in Scandinavia (Zone A). Surface features of extra-tropical organic soils seem to cause very high soil moisture values in some areas of Finland (Figure 9, top row), and the riverbeds of the East European plain are reflected in fine soils and coarse soil texture patches. They appear as drier patches in Northern Germany, Poland, and parts of the Scandinavian lowlands (Figure 6). Only the soil texture classes medium and medium-fine have a considerable overlap in their value ranges. The strong dependence on soil texture leads to a stratification of the data points, with cells with coarse soil texture clustering at lower soil moisture values, intermediate fine soils clustering at intermediate soil texture, and extra-tropical organic soils around very high soil moisture contents (e.g., Scandinavia in Zone A).

In relative terms of soil moisture [-], the wetness index in Figure 9 (bottom row) reveals a different dependency pattern compared to the soil moisture in absolute terms (Figure 9, top row). Here, soil texture has a minor influence (<15%) due to normalization, which largely removes the dependency on the soil type due to the focus on relative dynamics. In contrast, precipitation (18%) and soil temperature (15%) explain much of the variability for most of Europe. For the focus regions, the situation is diverse in terms of drivers.

The soil wetness index [-] is driven mainly by precipitation and soil texture in Scandinavia (zone A) and by temperature and soil texture in Central Europe (zone C), while the index is guided in Western Europe (zone D) and Eastern Europe (zone E) by precipitation and temperature as their climate regimes are more water-limited by temperature (zone D) and by continentality (zone E). Western Russia (zone B) is mainly dominated by temperature as the main driver for soil moisture due to its colder climate and absence of maritime influences (oceanic humidity).

Analyzing the soil moisture in absolute terms [m^3 m^{-3}] and focusing on the individual focus regions paints a different picture. In Western Europe, temperature, precipitation, and soil texture are the main environmental drivers, whereas in Eastern Europe, soil texture plays a minor role as the third environmental influencing factor. For Central Europe, Scandinavia, and western Russia, soil texture is by far the most influential and dominant factor, which is similar to the driver allocation for the soil wetness index for the first two focus regions (Central Europe and Scandinavia). However, for the latter, temperature (Central Europe) and precipitation (Scandinavia) are equally important.

Interestingly, vegetation as an environmental driver (see Figure 9) is in none of the focus regions an influential factor for the IFS model runs. This is the case, no matter if soil moisture is analyzed in absolute or relative terms. This may come from the way vegetation is represented in the IFS model architecture and could have two reasons. For one, leaf area index as the proxy of plant phenological dynamics through the season is represented as a seasonally invariant climatology [41]. In addition, the atmosphere module regulates the land surface fluxes in case of inconsistency between the soil and the vegetation compartments [41]. This may result in temperature and precipitation having a stronger influence on soil moisture in IFS runs than vegetation for the focus regions.

4.2.2. SMAP DCA Retrievals

The environmental drivers for the SMAP DCA retrievals are different from the ones for the IFS model runs due to the specific remote sensing-based retrieval inputs and retrieval architecture (see Section 2.1). This is why, for instance, the input signals (brightness temperatures) to the retrieval are also investigated as observational drivers besides all other environmental drivers.

A minor influence on the soil moisture distributions of the remote sensing retrievals is found for the vegetation distribution, the precipitation, and the surface temperature. One could argue that the vegetation removes water from deeper layers by plant water uptake. Moreover, spatial patterns of soil moisture in the topsoil are fairly short-lived due to interception in the vegetation canopies. The underlying driver is precipitation, which is mostly at a larger scale than vegetation variability, which is not resolved at the satellite level in Europe due to landscape heterogeneity with minor forest (grown-up vegetation) and major agricultural and grassland (non-grown-up vegetation) proportions.

Moreover, the atmosphere drives high soil moisture amounts at the west coasts as well as the western edges of mountain ranges (Scandinavian mountains, Scottish west coast, Pyrenees, and Alps), leading to the typical biomes in these regions with respective topography and mountainous soil types. The strongest influence of the vegetation and climatic factors can be observed over Western Europe (zone E), where they both explain about 60% of the variability. The stronger influence of temperature appears to be linked to the moisture limitation over the Iberian Peninsula, which has a semi-arid climate, wherefore a stronger relation between temperature (or climate factors in general) and soil moisture is to be expected but is modulated by vegetation accessing deeper water resources. The focus on Western Europe (zone E) is the region with the most explainable variability within the domain (above the *x*-axis in Figure 10). The least explainable variability is present in western Russia (zone B) for the soil wetness index as well as for the soil moisture in absolute terms. In Central Europe (zone C), most of the patterns arise from the soil roughness for the soil wetness index. Climatic influences and vegetation cover explain a smaller percentage of the pattern distribution and add about the same amount of explained variability to the variability explained by soil roughness.

In contrast to IFS model runs, soil moisture from SMAP DCA is not a conserved quantity, as the satellite retrievals provide regular snapshots (at the time of overpass) of the soil moisture distribution with a frequency of twice a day in Scandinavia (zone A) and up to three days in the lower (equatorial) latitudes. SMAP patterns do not have one single major environmental driver (see Figure 10), compared to IFS model runs, where soil texture types as categorical classes dominate the spatial moisture patterns in absolute terms in most cases. Over Europe, the retrievals are predominantly driven by a combination of brightness temperature (light green bars in Figure 10) (up to −29% loss of explained variability when taken out; all numbers for soil moisture in absolute terms), which is followed by the vegetation water content (up to −14%; ancillary from MODIS NDVI), the soil roughness (up to −8%; static ancillary from SMAP mission defaults) and the surface temperature (up to −4%; dynamic ancillary from GEOS-FP modeling system). Hence, soil characteristics still play a role as a driver of the SMAP patterns.

Furthermore, among the influential factors of the SMAP acquisition system, H-polarization fields are less influential than V-polarization fields of the brightness temperatures on the DCA patterns, especially for soil moisture in absolute terms [m^3 m^{-3}]. This is unexpected from soil moisture retrieval theory with passive microwaves, where both polarizations have distinct and similar soil moisture sensitivity (see Figure 12-3 in [68]). Another factor influencing the transformation from permittivity to volumetric soil moisture with the Mironov model used in L-band retrievals is the clay fraction. It has only a second-order influence (up to −2%) on the soil moisture patterns, since its analytical conversion does not affect the spatial pattern directly [37]. An overlooked input to the SMAP DCA soil moisture retrieval is the roughness coefficient of the soil surface. It is a temporally static ancillary field, which is used during the soil moisture retrieval from the brightness temperatures measured by the satellite radiometer. It shows a considerable effect

(up to −8% of variability loss when left out) on the soil moisture estimates for absolute and relative terms.

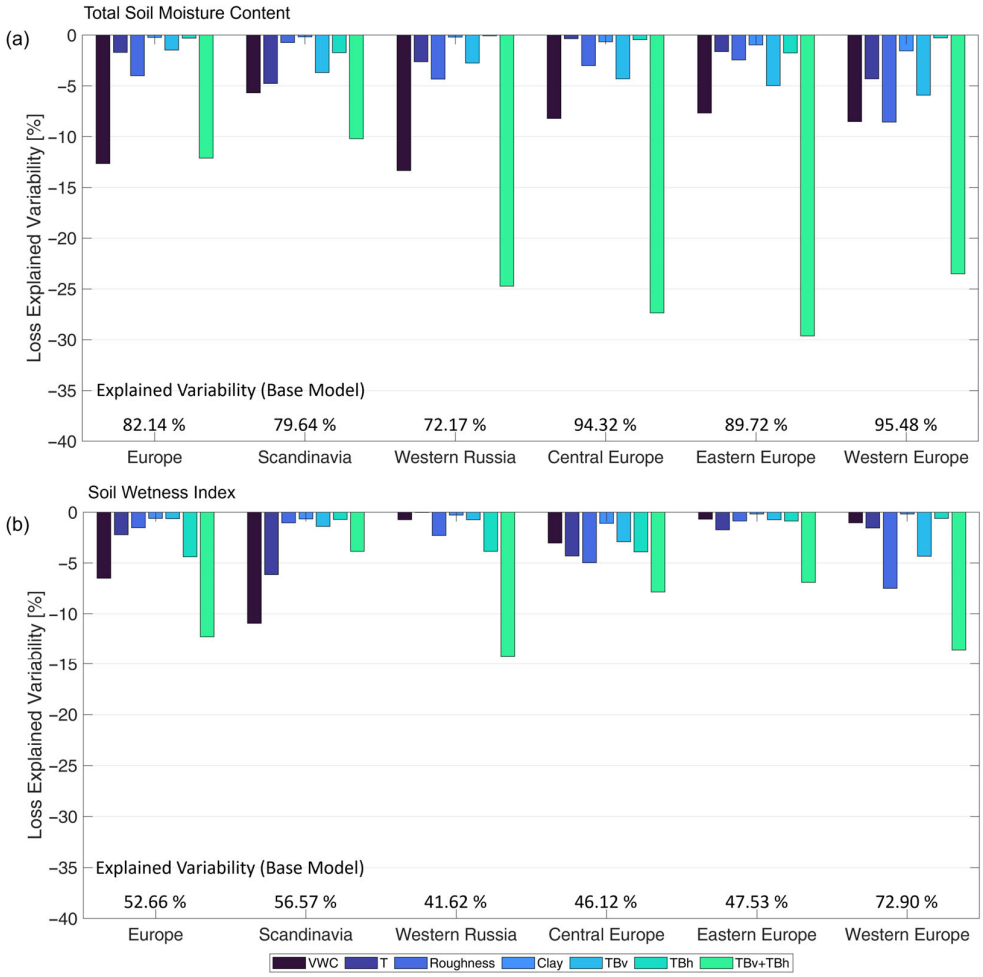

Figure 10. Loss of explained variability (coefficients of determinations) from taking single or combined variables out of the multi-variable regressions of potential influential factors of the SMAP DCA retrievals over the vegetation period 2021 for all of Europe and the focus regions (A–E); (**a**) in absolute terms [m^3 m^{-3}], (**b**) in relative terms [-]; VWC = vegetation water content, T = temperature, TBh = horizontally polarized brightness temperature, TBv = vertically polarized brightness temperature, Rough = soil roughness, Clay = clay content of the soil.

Nonetheless, all ancillary variables fed into the multivariate regression models explain a large range of explainable variability within Europe for soil moisture in absolute terms (80–95%). Only Scandinavia (zone A) lacks about 3% of explainable variability, which leads to a share of explainable variability of about 82% for all of Europe.

Though all ancillary variables have an influence on the patterns, there is some regional difference in the weighting of the importance of each influential factor for the retrieval. Similar to IFS model runs, western Russia (zone B) has the highest share of explained variability by a single variable (VWC) for soil moisture in absolute terms, whereas in Western Europe (zone E) and Scandinavia (zone A), it appears to be a combination of variables. Bulk density, being a direct influential factor for water tension capabilities, has a strong influence in Western Europe and the Mediterranean. There, the vegetation water content has a stronger influence, keeping in mind that vegetation is connected and influenced by the soil properties and dynamics. The positive correlation with vegetation water content could be rooted in the passive microwave acquisition nature of the retrieval, since the SMAP radiometer sensor is also sensitive to the moisture in the vegetation, besides

the moisture underneath the vegetation in the soil [68] p. 566f). This is why the SMAP DCA soil moisture retrieval uses a radiative transfer-based method to model soil as well as vegetation emissions as contributions to the brightness temperature signal [37].

When comparing soil moisture in absolute terms [m^3 m^{-3}] to relative terms (soil wetness index), the all-variable explainability (numbers above *x*-axis in Figures 9 and 10) is significantly different, with values well above 80% in absolute soil moisture values and values between 42% and 73% for the soil wetness index. One reason for this could be that normalization is done on a limited amount of data (two seasons spanning 2021 and 2022) for finding the extremes to calculate the index.

For the focus regions in Figure 10, the situation is a bit more diverse. Soil moisture in absolute terms is explained between 6% and 13% (absolute values of explainability loss) by the vegetation water content (VWC), which is a direct follow-on from the imperfect characterization of the vegetation in the retrieval model that is based on zeroth-order radiative transfer theory, neglecting higher-order vegetation–wave interactions [69]. Temperature plays a significant role with a variability loss of above −4% to −5% at focus regions A (Scandinavia) and E (Western Europe).

The explained variability of the soil wetness index shows, in general, significantly lower values compared to the absolute soil moisture values. When the single focus regions are investigated for the different drivers, the vertically and horizontally polarized brightness temperature (sum is represented as bright green bars in Figure 10) as input signals to the algorithm stand out in explainability for all of Europe, western Russia (zone B), Eastern (zone D), and Western (zone E) Europe. In addition, the vegetation water content showed a significant influence on the explained variability for Scandinavia (zone A) and Europe.

4.3. Similarity in Spatial Pattern Distribution from Cell to Continental Scales

This section aims to quantify the similarity in spatial pattern distribution of soil moisture from IFS and from SMAP DCA at varying spatial scales, from single cell, local, as well as at regional, up to continental scale.

4.3.1. Single-Cell Scale

At the scale of single cells, the spatial pattern comparison is a pixel-to-pixel comparison of both products. Figure 11 compares the differences between the IFS model-simulated and SMAP DCA-retrieved topsoil moisture for the vegetation period of 2021 and 2022 in absolute [m^3 m^{-3}] and in relative [-] terms. The hatching in the SMAP DCA subplots indicates areas with potential issues in the quality of the retrieval (e.g., the surface properties entered in the retrieval process) from SMAP data quality flags. We implemented two quality levels as introduced in Section 2: Dashed hatching indicates recommended quality after the optional SMAP flagging strategy. Dotted hatching indicates an even softer quality standard, allowing also a vegetation water content of 10 kg m^{-2} instead of 5 kg m^{-2}.

Interestingly, the difference between the two spatial datasets in Figure 11 reveals significant deviations in spatial patterns when shown in absolute terms for the median soil moisture compared to relative terms for the soil wetness index. This stems from the removal of the soil porosity impact when calculating the soil wetness index (no absolute values). For the median soil moisture content, the largest negative differences occur for the peatland areas in northern Scandinavia and western Russia (dark red regions in Figure 11, upper left plot). Moreover, the largest differences are found in southern Sweden and the coastal areas of Norway. This pattern of extremes stays constant over both years 2021 and 2022 (upper row of Figure 11), indicating their origin to be less influenced by meteorology but more by inputs like soil characterization to the SMAP retrieval and the IFS modeling.

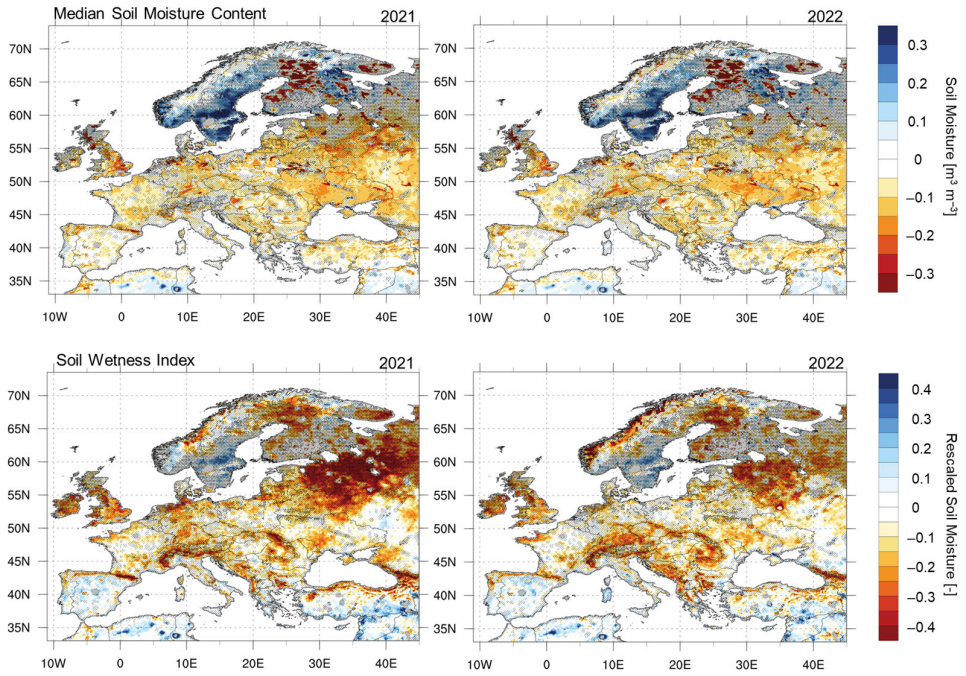

Figure 11. Differences between IFS model simulated and SMAP DCA retrieved top layer soil moisture in absolute [m^3 m^{-3}] (**top row**) and in relative [-] (**bottom row**) terms for the vegetation period 2021 (**left column**) and 2022 (**right column**). The dashed hatching indicates areas where the retrievals are not flagged as recommended quality in the retrieval metadata. The dotted hatching areas denote regions with a vegetation water content higher than 10 kg m^{-2}, mountain ranges with a slope larger than 3°, as well as more than 12.5% peatland.

The hatched zones are also predominantly found in these regions with a strong underlying variability in their topography, dominant vegetation, or significant wetness (including open water). The differences in relative terms (soil wetness index) are most significant in the regions of western Russia, northern Sweden, the United Kingdom, and in regions of high altitude (e.g., the Alps and the Pyrenees) in Atlantic Central, Atlantic North up to the Boreal climate regions, except Finland (Figure 1). Here, IFS-based soil moisture shows significantly lower moisture levels (Figure 7) compared to the SMAP-based soil moisture.

The deviations between the two products are stronger for the soil wetness index in 2021 than in 2022. This might be due to the wetter weather conditions in 2021 compared to 2022. However, overall continent-wide patterns stay similar between the years.

4.3.2. Local Scale

Here, we compare the spatial patterns of soil moisture at the local scale. Previous analyses at the cell scale, investigating single cells and their surrounding neighboring cells, have shown that the patterns do not locally (in a context of 5 × 5 cells, see Figure 4 for window size selection) agree between the two soil moisture products (SMAP and IFS). Therefore, the spatial pattern comparison metrics (disagreements and mean category distance) will be applied based on 5 × 5 pixel moving windows to quantify the divergence in quantity and allocation of the soil moisture patterns within this window. The mCD refines the quantity disagreement by stating how far the category deviates.

In Figures 12 and 13, the spatial pattern comparison metrics are shown for both years and compared in terms of absolute values (median soil moisture in [m^3 m^{-3}]) and in terms of relative values (soil wetness index [-]). Overall, the mCD complements the quantity disagreement metric.

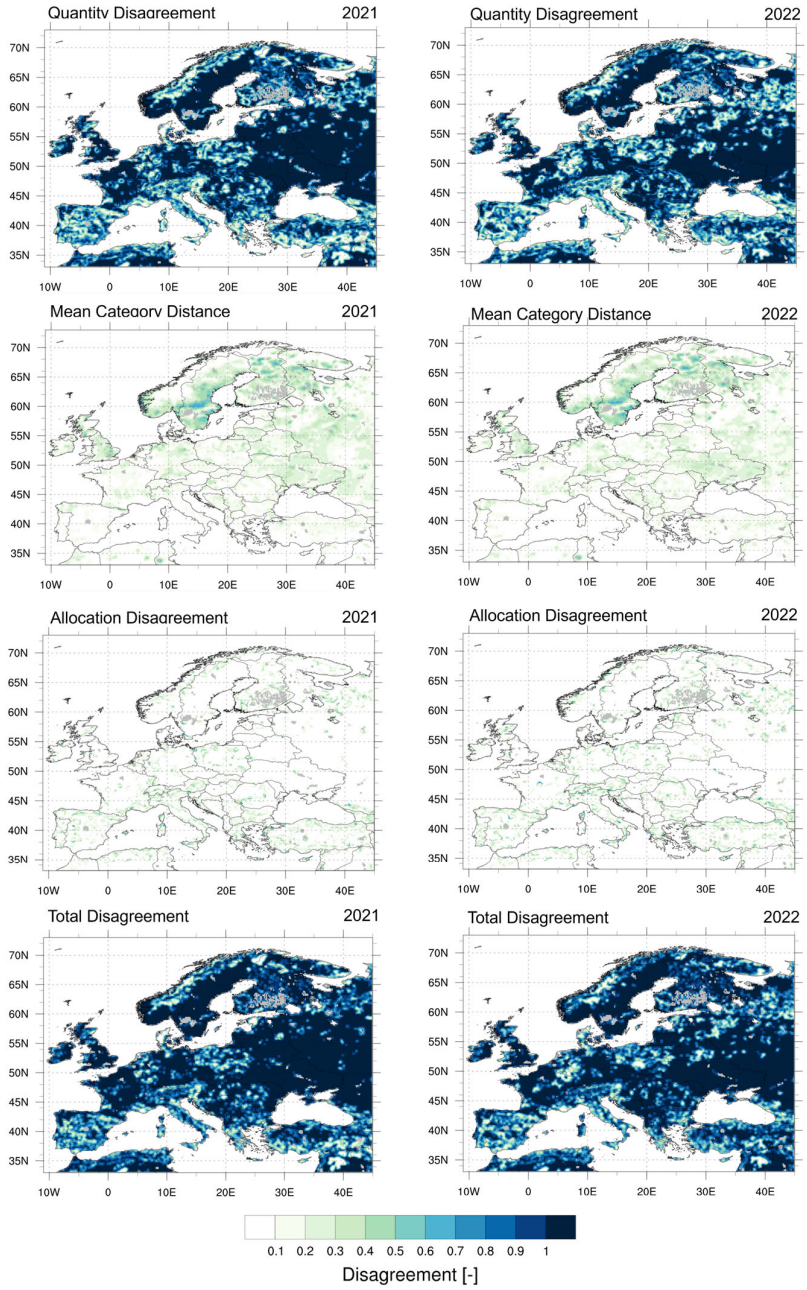

Figure 12. Comparison of the spatial patterns of the median soil moisture content (in absolute terms [m^3 m^{-3}]) during the vegetation periods 2021 and 2022; window size for pattern comparison 5 × 5 pixels; boundary tolerance 2%. Focus regions are detailed in Figure A2 (Appendix B).

In regions where the quantity disagreement saturates (dark blue regions with values of one in Figures 12 and 13), the mCD is capable of resolving and quantifying this divergence. It indicates the number of cells in the moving (5 × 5) window for which the category completely diverges. Moreover, this category would also not show up in another pixel of the moving window. The strongest deviations are seen in Scandinavia (Sweden) and in western Russia, which is logically consistent, but not collocated, with what is shown in Figure 11 due to soil type divergence.

Moreover, it can be seen in both Figures 12 and 13 that the allocation disagreement is more pronounced in regions where the quantity disagreement and the mCD-values are rather small (0.1) to medium (~0.6) levels. At high disagreement levels (e.g., quantity disagreement > 0.9), the allocation disagreement plunges close to zero, which is expected and consistent with the logic of the disagreement metric. Looking at the sum of both

disagreements expressed in the total disagreement, it is well understood that the quantity disagreement is dominating the dynamics of the total disagreement, indicating the minor influence of the allocation disagreement for both years and both terms (absolute and relative). In the end, the overall spatial patterns of all metrics on the local scale stay similar between the two years, 2021 and 2022, showing no large-scale year-to-year dynamics across Europe for the two years. This means that the local patterns of soil moisture are less dependent on the year-to-year change in the hydrological situation and more on the drivers, like soil and vegetation characteristics, as presented in Section 4.2.

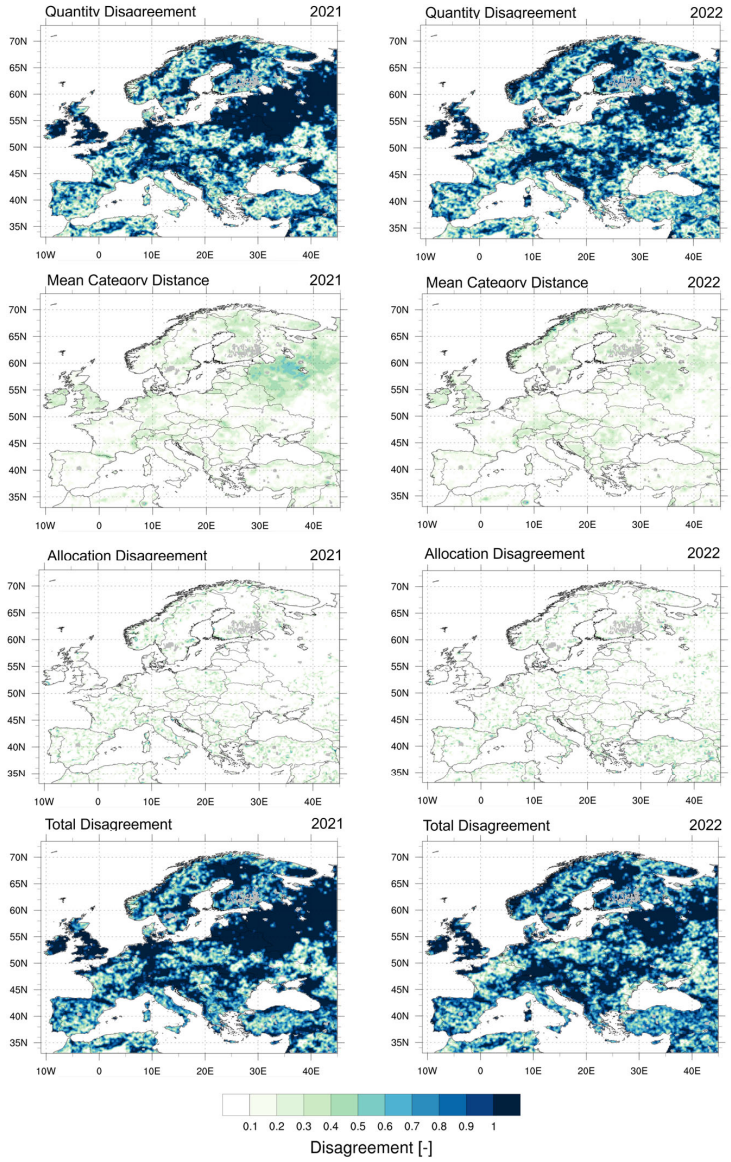

Figure 13. Comparison of the spatial patterns of the soil wetness index (in relative terms [-]) during the vegetation periods 2021 and 2022; window size for pattern comparison 5 × 5 pixels; boundary tolerance 2%. Focus regions are detailed in Figure A1 (Appendix B).

However, the spatial patterns are significantly different for the soil wetness index compared to the absolute soil moisture values, meaning in terms of units ([-] vs. [m^3 m^{-3}]). This is especially apparent for the mCD, comparing Figure 12 with Figure 13. Here, the regions containing Sweden and Finland (both in zone A) and parts of north–western Russia show the highest values (up to 0.5 to 0.6) of mCD for the absolute values of soil moisture in Figure 12, while the soil wetness index reveals the biggest differences in

mCD for western Russia (zone B) in Figure 13. This underlines the different drivers behind the local patterns of soil moisture, presented in Figures 9 and 10, and how they are pronounced when using absolute or relative units of soil moisture. The strength of the different spatial pattern correlation metrics for the individual focus zones (A–E) is detailed in Figure A1 of Appendix B.

For spatial patterns comparisons in absolute terms in Scandinavia, the main drivers are vegetation water dynamics, soil roughness, and soil temperature for the SMAP DCA retrievals, while for the IFS model runs, the main driver is the soil texture input. For the spatial patterns comparison in relative terms in western Russia, the main drivers are bulk density for the SMAP DCA retrievals, whereas the IFS model runs are mainly driven by air and soil temperature. This means the drivers of the soil moisture fields are especially different in these regions, leading to the highest mCD values in Figures 12 and 13.

4.3.3. Regional up to Continental Scale

Figure 14 presents the spatial pattern comparison metrics for both years (2021 and 2022), either for absolute values [m^3 m^{-3}] (top row) or for the soil wetness index [-] (bottom row) over the entire domain (Europe). Tables with statistics are included in Appendix B.

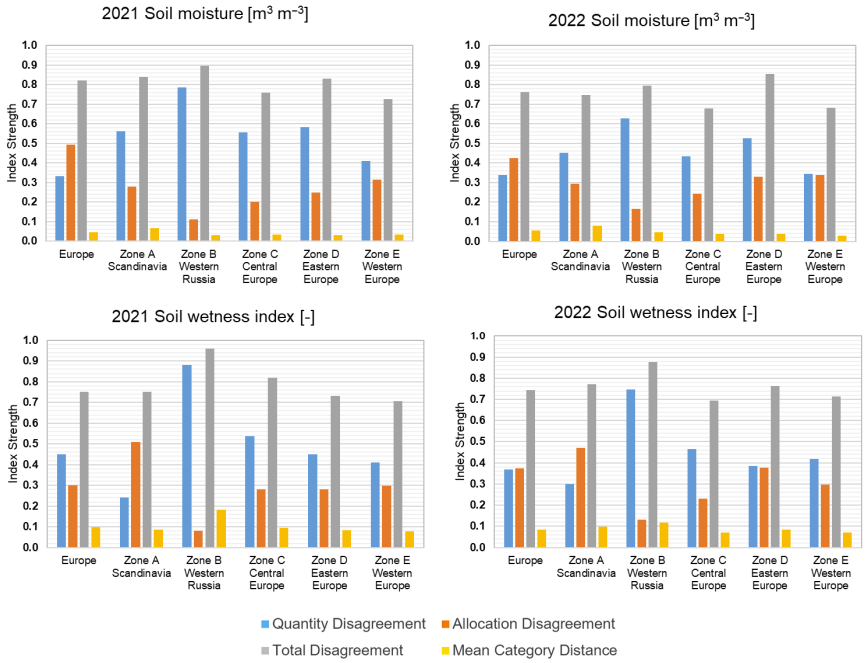

Figure 14. Spatial pattern comparison metrics of the entire domain (Europe) and of all five focus regions A–E for both years (2021 and 2022) (columns) as well as for absolute values [m^3 m^{-3}] (**top row**) and for the soil wetness index (relative values) [-] (**bottom row**).

For absolute values [m^3 m^{-3}], the mCD is largest in Scandinavia (zone A) for 2021 and 2022, indicating the strong soil type influence in this zone. In 2021 and 2022, the smallest total disagreement is seen in Western Europe (zone E). In 2021, the total disagreement is largest for western Russia (zone B), and in 2022 it is largest for Eastern Europe (zone D), both times due to the large allocation disagreement. For western Russia (zone B), the quantity disagreement is at its maximum, and the allocation disagreement is at its minimum compared to all focus regions in 2021 and 2022, meaning that the location of patterns in space is similar, but the quantities of the soil moisture values are significantly different. This fits well with Figures 11 and 12, indicating this already at the cell scale and even for mCD at the local scale, respectively. Reasons here are the different environmental

and observational drivers (see Section 4.2) influencing the IFS model runs and the SMAP DCA retrievals to lead to similar spatial patterns but diverging soil moisture values.

For the soil wetness index (relative values) [-], the mean category distance is largest in western Russia (zone B) for 2021 and 2022. In 2021, the smallest total disagreement is seen in Western Europe (Zone E), and in 2022 in Central Europe (Zone C). In 2021 and 2022, the total and quantity disagreement is largest in western Russia (zone B). However, the allocation disagreement is largest in Scandinavia (zone A) in 2021 and 2022, which fits the organic soils in these zones. In contrast, the allocation disagreement is smallest in western Russia (zone B) due to a calmer and colder continental climate, leading to higher spatial similarity due to homogeneous moisture input (precipitation).

5. Conclusions and Outlook

This study investigates and discusses the discrepancy of spatial soil moisture patterns between operational products from passive remote sensing and Earth system modeling. We configure and apply spatial similarity metrics to enable spatial comparison of the SMAP radiometer soil moisture product with the soil moisture output from IFS model runs available from the ECMWF. For this, a set of spatial pattern correlation metrics is used: allocation disagreement, quantity disagreement, total disagreement, and mean category distance (mCD). The spatial similarity is compared on different spatial scales from the single cell up to the European continent. Key findings are that soil moisture patterns diverge distinctively between a state-of-the-art model-based (ECWMF-IFS) soil moisture product and a state-of-the-art satellite-based (SMAP DCA) remote sensing product. This is shown across Europe in Figures 6–8 for the mean soil moisture fields and on a single-cell scale in Figure 11. Figures 12–14 indicate that the spatial pattern divergence (non-similarity) occurs from the local scale up to the continental scale. However, the novel spatial pattern correlation metrics help to understand how spatial patterns can still be compared and their spatial correlation assessed across different spatial scales.

We find that the spatial similarity is essential to enable pattern comparison due to significant differences between the initial products. The spatial similarity varies across scales from single cells to continents and over the two years. The two established metrics, disagreement, containing total, quantity, and allocation disagreement, and mCD, showcase the potential and challenges when assessing spatial similarity in soil moisture fields across scales. Essential steps are the categorization of the two soil moisture input products, as well as the adjustment of the spatial window size and the boundary relaxation criteria.

Moreover, we analyze the environmental and observational drivers behind these soil moisture patterns for both products and find that they are individual for each product and diverse across focus regions. The analysis shows that soil texture is clearly the most influential single driver of spatial patterns in the ECMWF IFS soil moisture product, when analyzed in absolute terms [m^3 m^{-3}]. In relative terms of soil moisture [-] (soil wetness index), soil texture has the least influence, whereas precipitation and soil temperature explain most of the variability for all of Europe. The environmental drivers, but also algorithmic input factors, called observational drivers, for the SMAP DCA retrievals are different from the ones for the IFS model runs due to the specific remote sensing-based retrieval inputs and retrieval architecture (see Section 2.1). The vegetation distribution, the precipitation, and the surface temperature have a minor influence on the soil moisture distributions of the remote sensing retrievals. Over Europe, the retrievals are predominantly driven by a combination of brightness temperatures, which in turn are influenced by the vegetation water content (ancillary from MODIS NDVI), the soil roughness (static ancillary from SMAP mission defaults), and the surface temperature (dynamic ancillary from GEOS-FP modeling system). These differences in drivers, as well as in the methodology and

algorithm of both products, culminate in an inherent difference in soil moisture from ECWMF IFS runs and SMAP DCA retrievals. Hence, they cannot be directly used jointly for large-scale mapping or as a spatial reference for validation, e.g., other remote sensing soil moisture products (SMOSs). However, the assessment of their spatial patterns reveals the underlying similarity from the local to the continental scale. This could be tested in the future to further improve assimilation schemes for both products beyond classical data assimilation of satellite data into ECWMF model runs. One option could be to use the spatial pattern comparison metrics as an indicator where the assimilation could be done classically due to high allocation agreement, and novel machine learning (ML)-based options could be derived and tested where the allocation agreement is rather low.

Furthermore, another important step for assessing the similarity of spatial soil moisture is the establishment of strategies (up to in situ network strategies and proximal remote sensing) to validate spatial patterns of different soil moisture products. Here, ML could be a contemporary asset to be implemented and used for spatial similarity analyses. Data assimilation is already part of the operational practice within ECMWF forecasts with the assimilation of remote sensing-based soil moisture, e.g., from SMOS [70], but not the focus of this research study. Moreover, 90% of the spatial patterns in ECWMF IFS runs appear to be explainable according to the multi-linear regression analyses through a combination of static fields and meteorological variables. This suggests that the assimilation of satellite data, like from ASCAT, may only have a reduced influence on the spatial distribution of ECWMF-based soil moisture, since the spatial distribution is also derived through the data assimilation of screen-level temperature and relative humidity of the atmosphere at the pixel scale.

In the end, assessing the spatial similarity of soil moisture patterns across Europe by applying novel spatial comparison metrics from local to continental scale and considering the underlying drivers leads the way towards a more seamless fusion of remote sensing products and Earth system model runs.

Author Contributions: Conceptualization, L.J. and T.J.; methodology, L.J., T.J., and A.F.; software, L.J.; validation, L.J. and T.J.; formal analysis, L.J. and T.J.; investigation, L.J., T.J., and A.F.; resources, T.J. and H.K.; data curation, L.J., T.J., H.-S.B., D.C., A.F., F.M.H., and G.P.; writing—original draft preparation, T.J., L.J., and A.F.; writing—review and editing, T.J., L.J., H.-S.B., D.C., A.F., F.M.H., G.P., and T.J.; visualization, L.J., T.J., and A.F.; supervision, H.K. All authors have read and agreed to the published version of the manuscript.

Funding: L. Jach acknowledges funding from the Anton and Petra Ehrmann-Stiftung Research Training Group "Water–People–Agriculture" for the guest research stay at DLR Oberpfaffenhofen. F. Hellwig is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—514721519 in the project "Remote sensing of vegetation canopy properties: States and spatio-temporal dynamics" as part of the DFG Research Unit 5639: Land Atmosphere Feedback Initiative (LAFI). D. Chaparro has been funded by the projects 'la Caixa' Junior Leader Fellowship LCF/BQ/PI25/12100008 (lead: D. Chaparro) and LCF/BQ/PI23/11970013 (lead: O. Binks) and by the project H2020 FORGENIUS (Improving access to FORest GENetic resources Information and services for end-USers) #862221. This research study is also part of the scientific investigations within the DFG research unit 5639: Land Atmosphere Feedback Initiative - LAFI, where F.M.H. and L.J. are funded from.

Data Availability Statement: The in situ measurements and SMAP data are publicly available via the in-text provided references. ECMWF operational analyses are available upon request from ECMWF.

Acknowledgments: The authors want to warmly acknowledge the dear support of Christoph Rüdiger, team leader of the Land Modeling team within the Research Department (Earth System Modeling Section) of ECMWF. His deep insights and very valuable comments have improved the manuscript substantially. We are very thankful for our close collaboration throughout the years.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

4D-Var	Four-dimensional Variation Assimilation
AMSR-E	Advanced Microwave Scanning Radiometer
AMSR2	Advanced Microwave Scanning Radiometer 2
ASCAT	Advanced Scatterometer
CCI	Climate Change Initiative
CESBIO	Centre d'Etudes Spatiales de la Biosphère
DCA	Dual Channel Algorithm
ECMWF	European Center for Medium-Range Weather Forecasts
ERS	European Remote Sensing Satellite
ESA	European Space Agency
FY-3C	FengYun-3C satellite
ICOS	Integrated Carbon Observation System
IFS	Integrated Forecast System
INRAE	Institut National de la Recherche pour l'agriculture, l'alimentation et 'environment
ISMN	International Soil Moisture Network
JAXA	Japanese Aerospace Exploration Agency
LPRM	Land Parameter Retrieval Model
mCD	Mean Category Distance
NASA	National Aeronautics and Space Administration
NDVI	Normalized Differential Vegetation Index
NOAA	National Oceanic and Atmospheric Administration
RMSE	Root Mean Square Error
SMAP	Soil Moisture Active Passive
SMAP-IB	SMAP-INRAE Bordeaux
SMAP MTDCA	SMAP Multi-Temporal Dual Channel Algorithm
SMOS	Soil Moisture Ocean Salinity
SMOS-IC	SMOS-INRAE CESBIO
SURFEX	SURFace EXternalisée
TOPKAPI	TOPographic Kinematic APproximation and Integration
V	Vertical
VWC	Vegetation Water Content

Appendix A. Spatial Patterns Comparison Metrics

In this appendix, the tables of the spatial patterns comparison metrics for all of Europe and the five focus regions (A–E) of Section 4.3 are collected for documentation of the statistics.

Table A1. Spatial patterns comparison metrics for all of Europe and the five focus regions (A–E) regarding years 2021 (**top**) and 2022 (**bottom**); calculus: soil moisture input was in absolute values [m^3 m^{-3}] and with fixed bins in categorization (including a tolerance window of 2%).

Soil Moisture in Absolute Values [m^3 m^{-3}]						
Year 2021	**Europe**	**Scandinavia Region A**	**Western Russia Region B**	**Central Europe Region C**	**Eastern Europe Region D**	**Western Europe Region E**
Quantity Disagreement	0.331	0.561	0.787	0.556	0.582	0.410
Allocation Disagreement	0.492	0.279	0.110	0.202	0.248	0.315
Total Disagreement	0.823	0.840	0.897	0.759	0.831	0.725

Table A1. *Cont.*

Soil Moisture in Absolute Values [m^3 m^{-3}]						
Mean Category Distance	0.046	0.067	0.031	0.032	0.030	0.033
Year 2022	**Europe**	**Scandinavia Region A**	**Western Russia Region B**	**Central Europe Region C**	**Eastern Europe Region D**	**Western Europe Region E**
Quantity Disagreement	0.337	0.453	0.628	0.435	0.525	0.344
Allocation Disagreement	0.426	0.293	0.167	0.242	0.329	0.338
Total Disagreement	0.763	0.746	0.795	0.677	0.854	0.681
Mean Category Distance	0.054	0.078	0.045	0.037	0.037	0.030

Table A2. Spatial patterns comparison metrics for all of Europe and the five focus regions (A–E) for soil wetness index [-] regarding years 2021 (**top**) and 2022 (**bottom**); calculus: soil moisture input was in relative values [-] (soil wetness index) and with fixed bins in categorization (including tolerance window of 2%).

Soil Wetness Index [-]						
Year 2021	**Europe**	**Scandinavia Region A**	**Western Russia Region B**	**Central Europe Region C**	**Eastern Europe Region D**	**Western Europe Region E**
Quantity Disagreement	0.451	0.242	0.880	0.538	0.451	0.409
Allocation Disagreement	0.301	0.509	0.081	0.281	0.279	0.298
Total Disagreement	0.752	0.751	0.961	0.819	0.730	0.707
Mean Category Distance	0.098	0.087	0.182	0.095	0.084	0.077
Year 2022	**Europe**	**Scandinavia Region A**	**Western Russia Region B**	**Central Europe Region C**	**Eastern Europe Region D**	**Western Europe Region E**
Quantity Disagreement	0.368	0.301	0.748	0.464	0.386	0.419
Allocation Disagreement	0.375	0.470	0.130	0.231	0.378	0.296
Total Disagreement	0.744	0.771	0.878	0.695	0.764	0.715
Mean Category Distance	0.085	0.097	0.119	0.072	0.084	0.070

Appendix B. Relative Frequency Distribution of Spatial Pattern Correlation Metrics on Local Scale

In Section 4.3, the spatial pattern correlation metrics are applied on a local scale for a moving window of 5 × 5 pixels with a boundary tolerance of 2%. Figures 12 and 13 show the comparisons of the spatial patterns of the median soil moisture content during the vegetation periods 2021 and 2022 in absolute terms [m^3 m^{-3}] and in relative terms [-] (soil wetness index), respectively. In the following, the relative frequency distribution of the spatial pattern correlation metrics is shown for the local scale (5 × 5 cells) for the different focus regions A–E.

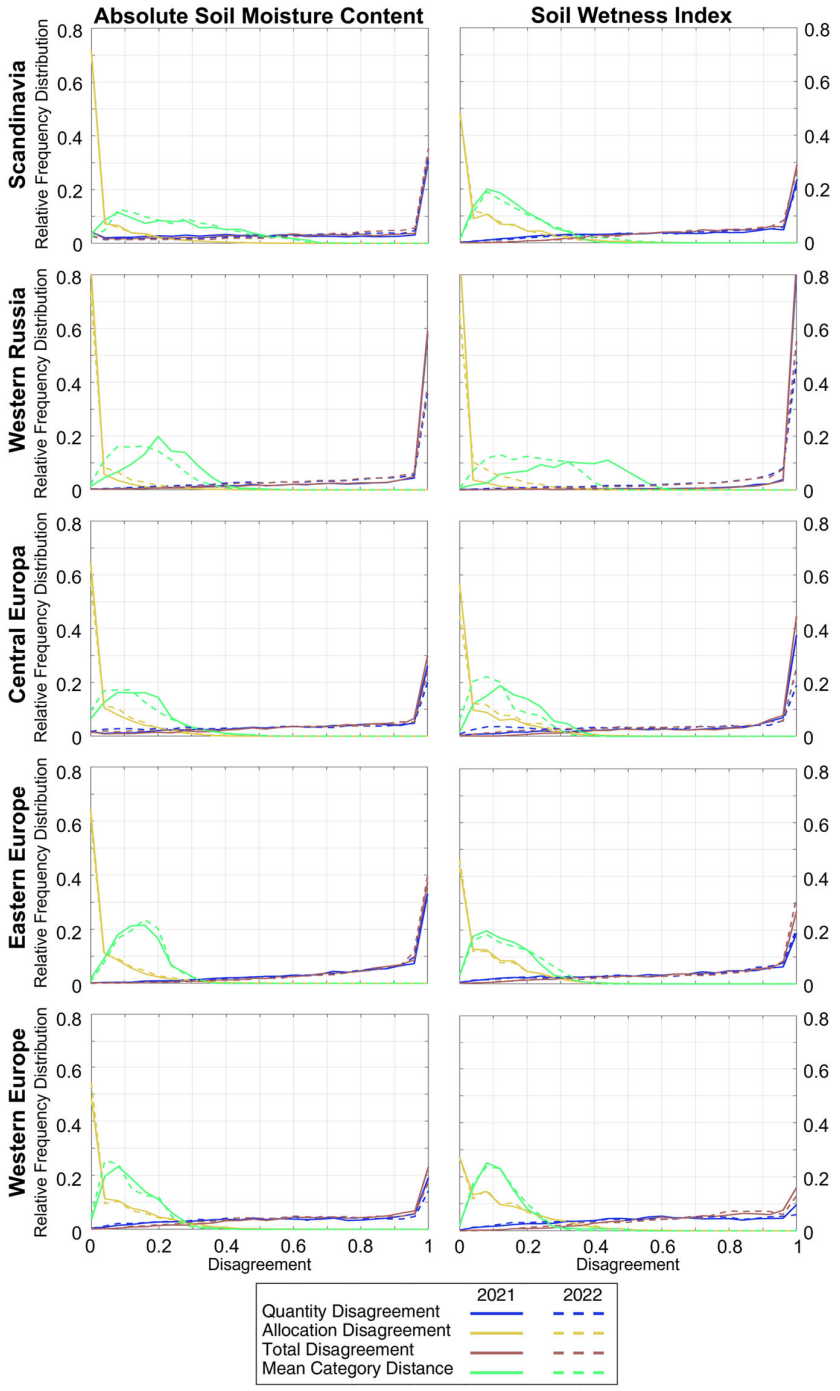

Figure A1. Relative Frequency Distribution of Spatial Pattern Correlation metrics on local scale (5 × 5) for the different focus regions A–E.

Appendix C. Comparison of Soil Moisture Interpolation Techniques

We conducted a resampling test with three different interpolation techniques (nearest neighbor interpolation, bilinear interpolation, and distance-weighted averaging) to understand the sensitivity to resampling of the SMAP DCA soil moisture product to match the resolution of the IFS grid.

In Figure A2, we show the difference in soil moisture [m^3 m^{-3}] for 2021 (top row) and for 2022 (bottom row), comparing the three different interpolation techniques across Europe. Figure A3 confirms that the different interpolation techniques have only very little (Scandinavia) to no (rest of Europe) influence on the resampled result.

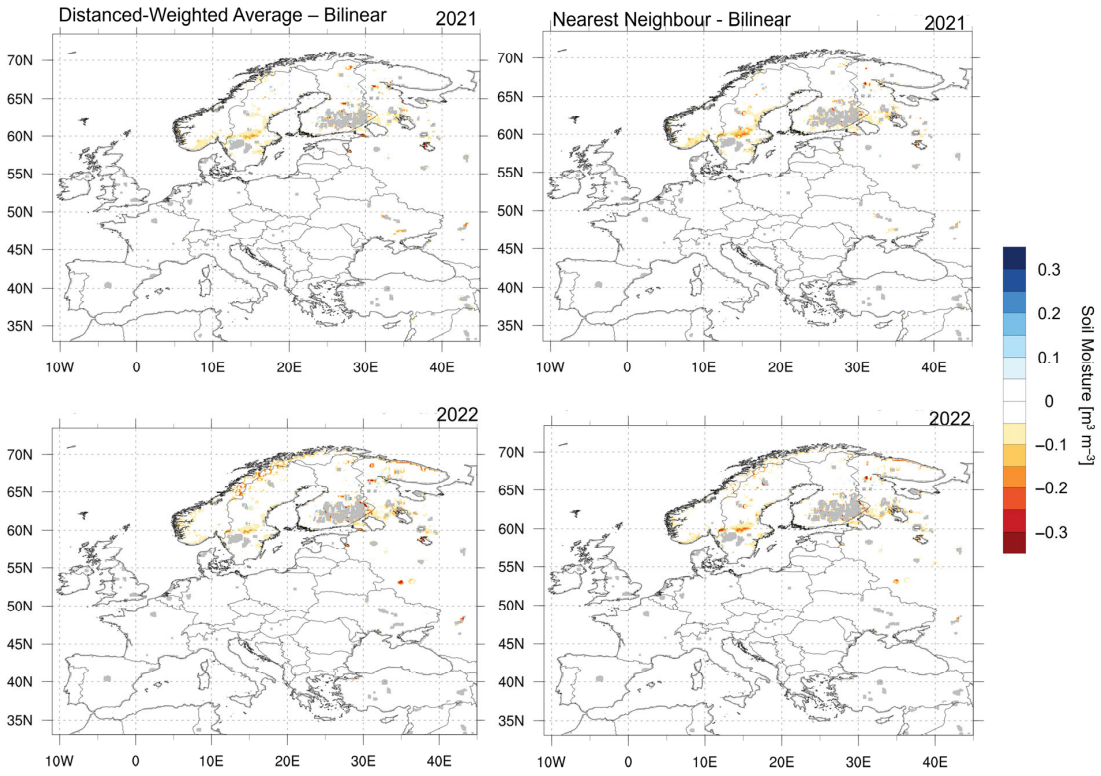

Figure A2. Difference in soil moisture [m^3 m^{-3}] for 2021 (**top row**) and for 2022 (**bottom row**), comparing three different interpolation techniques: Distance-weighted average, bilinear interpolation, and nearest neighbor interpolation. Left column: Difference in Distance-Weighted Average with Bilinear interpolation; Right column: Difference between Nearest Neighbor and Bilinear Interpolation. Gray areas are masked according to mandatory SMAP filtering.

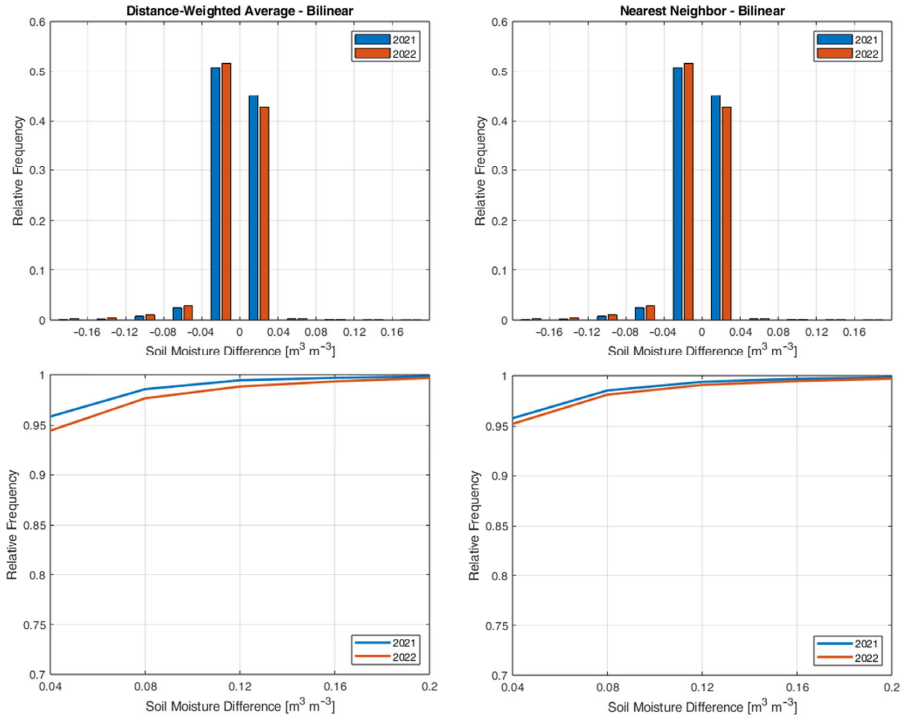

Figure A3. Statistical measures showing the soil moisture difference [m^3 m^{-3}] of the three tested interpolation methods in 2021 and 2022: Histograms (**top row**) and cumulated histograms (**bottom row**).

In Figure A3, the respective statistics are presented as histograms (top row) and cumulated histograms (bottom row) of the difference in soil moisture [m^3 m^{-3}] for couples of interpolation techniques and both years. More than 94% of the differences for all years and for all of Europe are below the official SMAP target accuracy of 0.04 [m^3 m^{-3}].

Appendix D. Sensitivity of Spatial Similarity Metrics to Binning

Beyond the descriptions and derivations of the final binning numbers in Section 3.1.1, we add here a binning test (8/10/12 bins) to show the sensitivity of the four similarity metrics, calculated according to Section 3 in the manuscript, to the binning. Figures A4 and A5 depict the results of the binning test as histograms of the four similarity metrics and both years (2021 and 2022) for soil moisture in absolute terms [m^3 m^{-3}] and in relative terms [-], respectively.

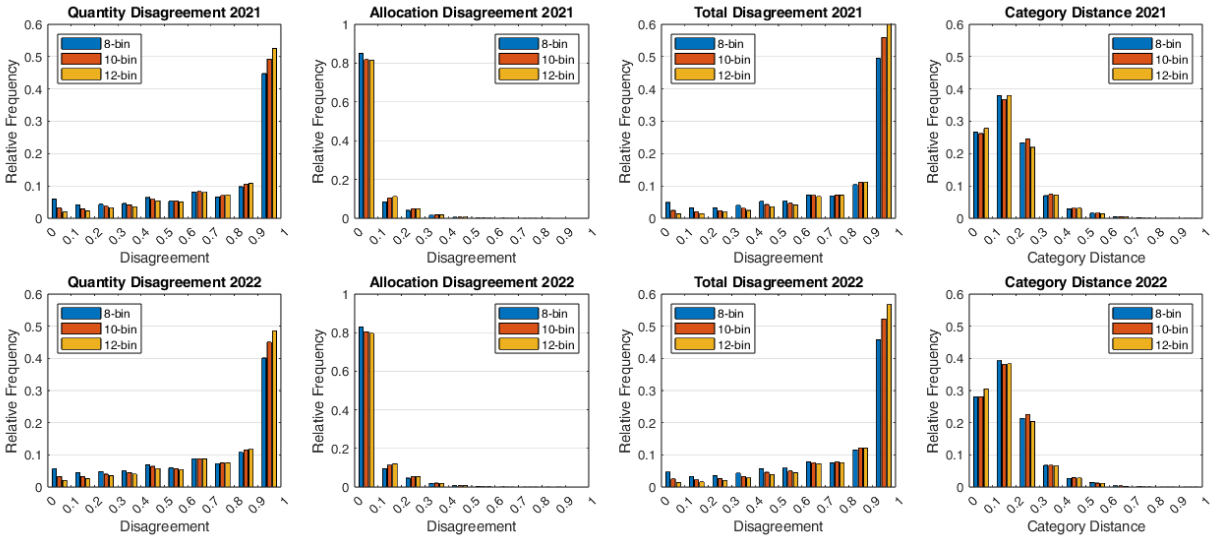

Figure A4. Histograms of the four similarity metrics (columns) and both years (rows) for soil moisture in absolute terms [m^3 m^{-3}] calculated with different binning (8, 10 or 12 bins).

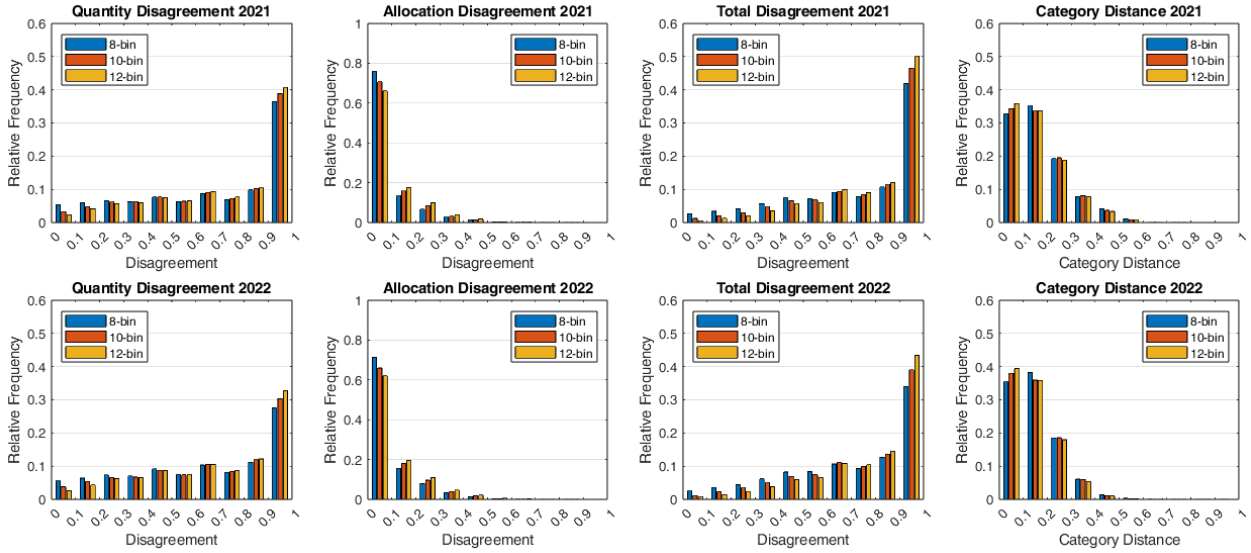

Figure A5. Histograms of the four similarity metrics (columns) and both years (rows) based for soil moisture in relative terms (soil wetness index) [-] calculated with different binning (8, 10 or 12 bins).

In both Figures, the overall patterns of the histograms of all metrics stay the same. There is not a single change in pattern. This is the most important observation, as we are not evaluating and discussing these metrics below the 0.1-digit level at all. Hence, the message and the findings we address are fully valid, no matter which of the three binning sizes would be used.

Moreover, we show in Figures A6–A9 the spatial patterns (maps) of the four similarity metrics for all years and soil moisture in absolute terms [m^3 m^{-3}] (Figures A6 and A7) and in relative terms [-] (Figures A8 and A9), respectively. These four figures should provide insights into the spatial distributions of the four similarity metrics. They indicate that the spatial distributions across Europe and for larger domains (e.g., focus zones of the study) stay stable, and the message and findings are fully valid.

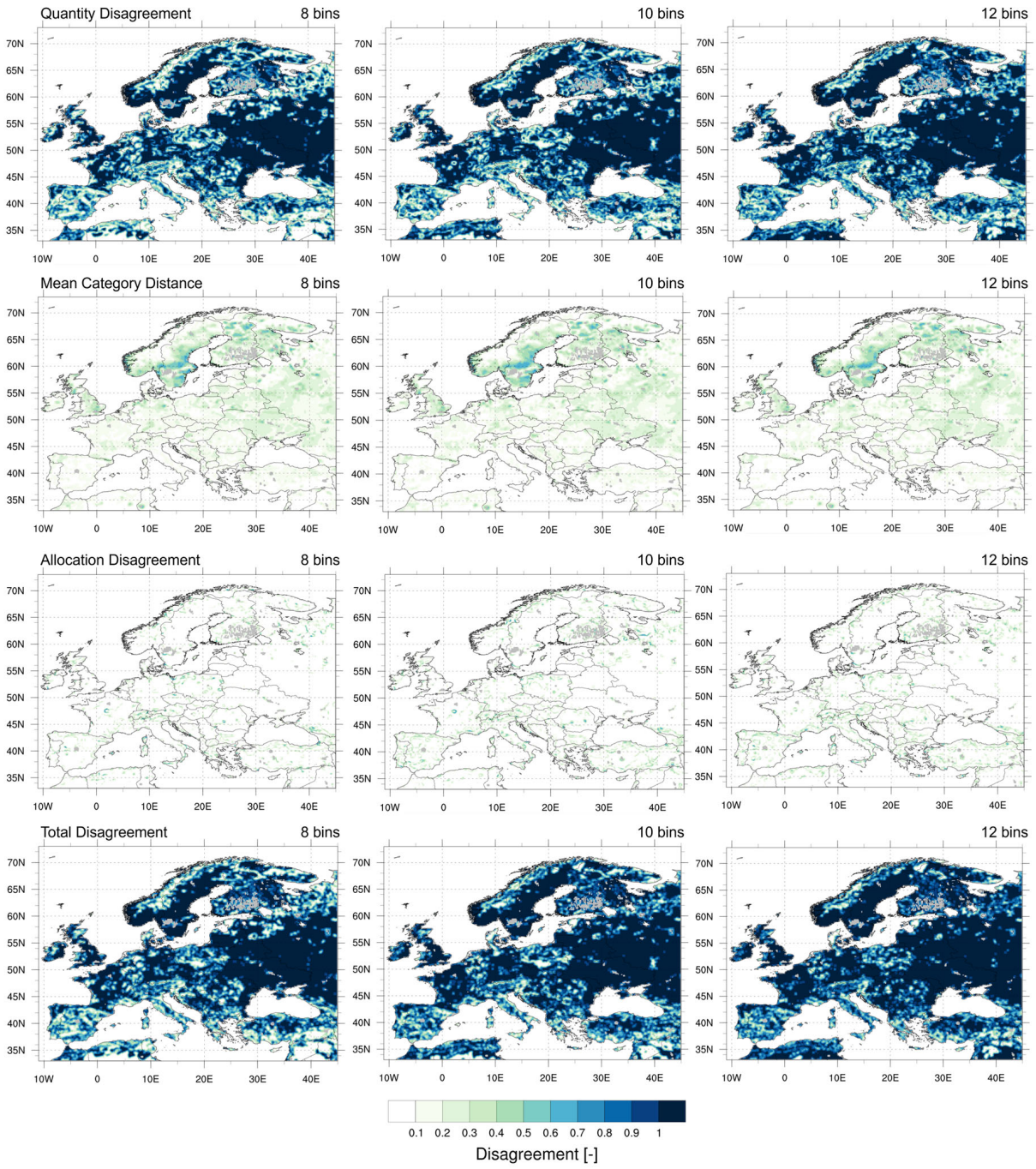

Figure A6. Maps of the four similarity metrics (rows) in 2021 for soil moisture in absolute terms [m^3 m^{-3}] calculated with different binning (8, 10, or 12 bins) (columns).

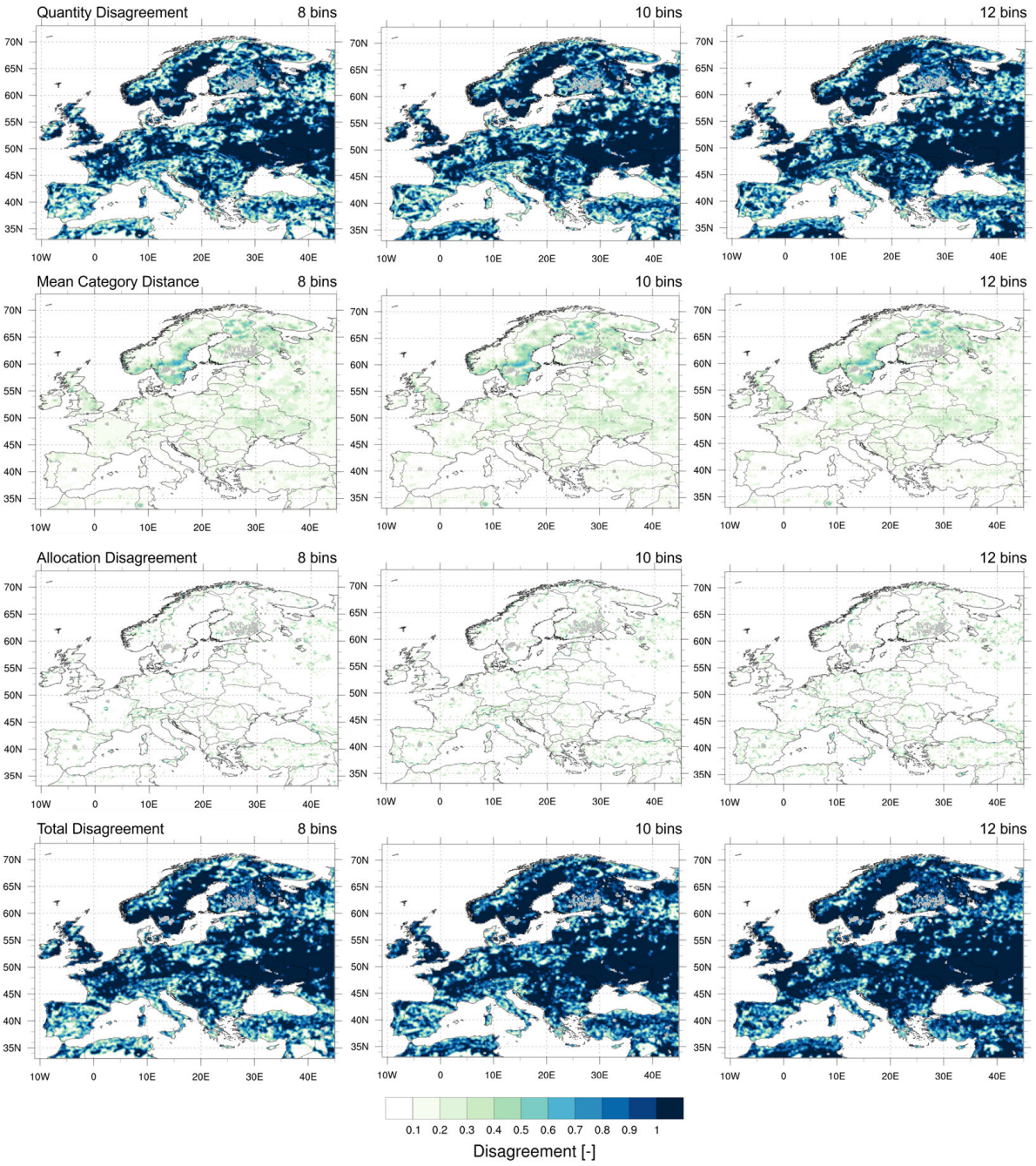

Figure A7. Maps of the four similarity metrics (rows) in 2022 for soil moisture in absolute terms [m^3 m^{-3}] calculated with different binning (8, 10, or 12 bins) (columns).

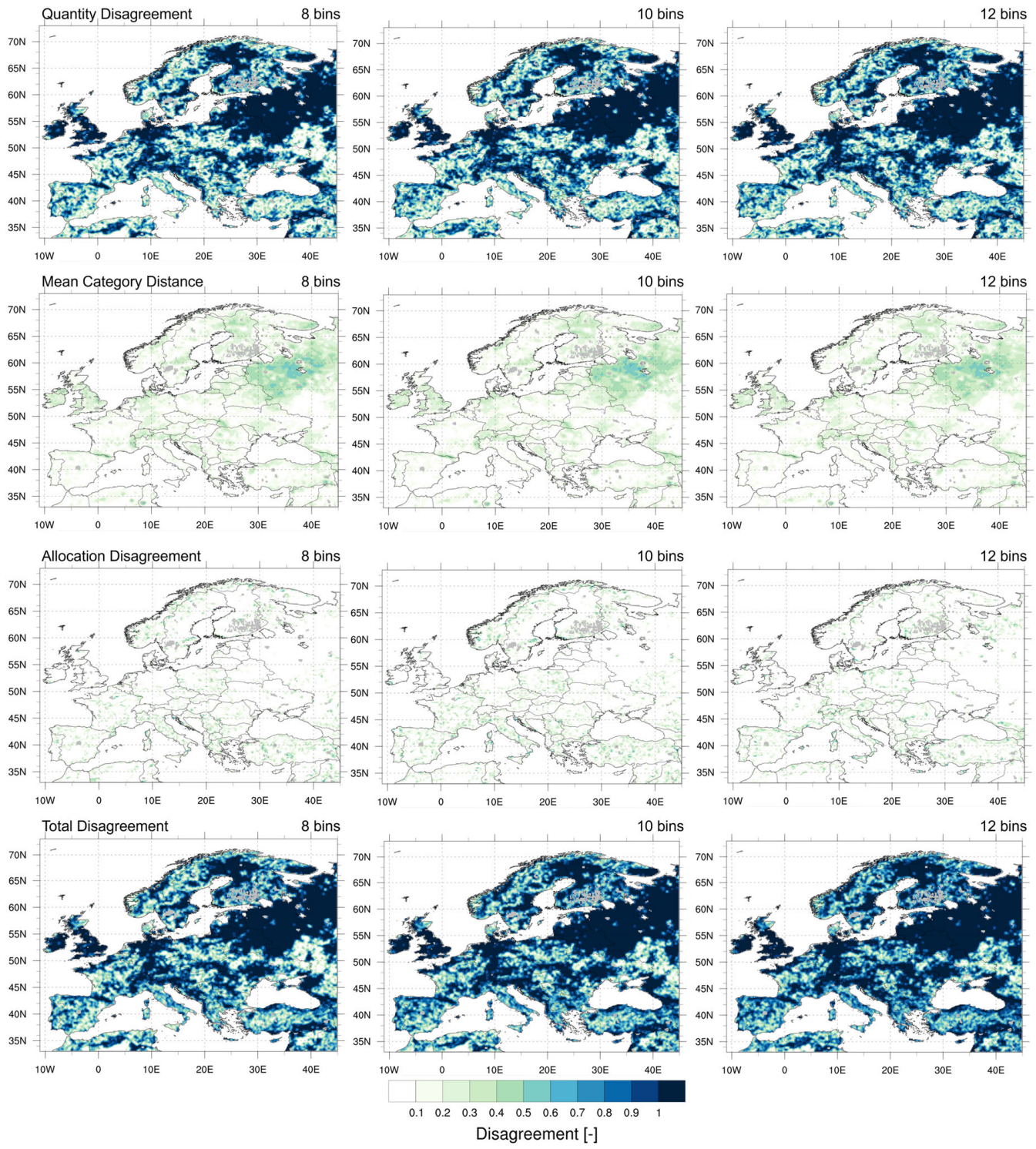

Figure A8. Maps of the four similarity metrics (rows) in 2021 for soil moisture in relative terms (soil wetness index) [-] calculated with different binning (8, 10, or 12 bins) (columns).

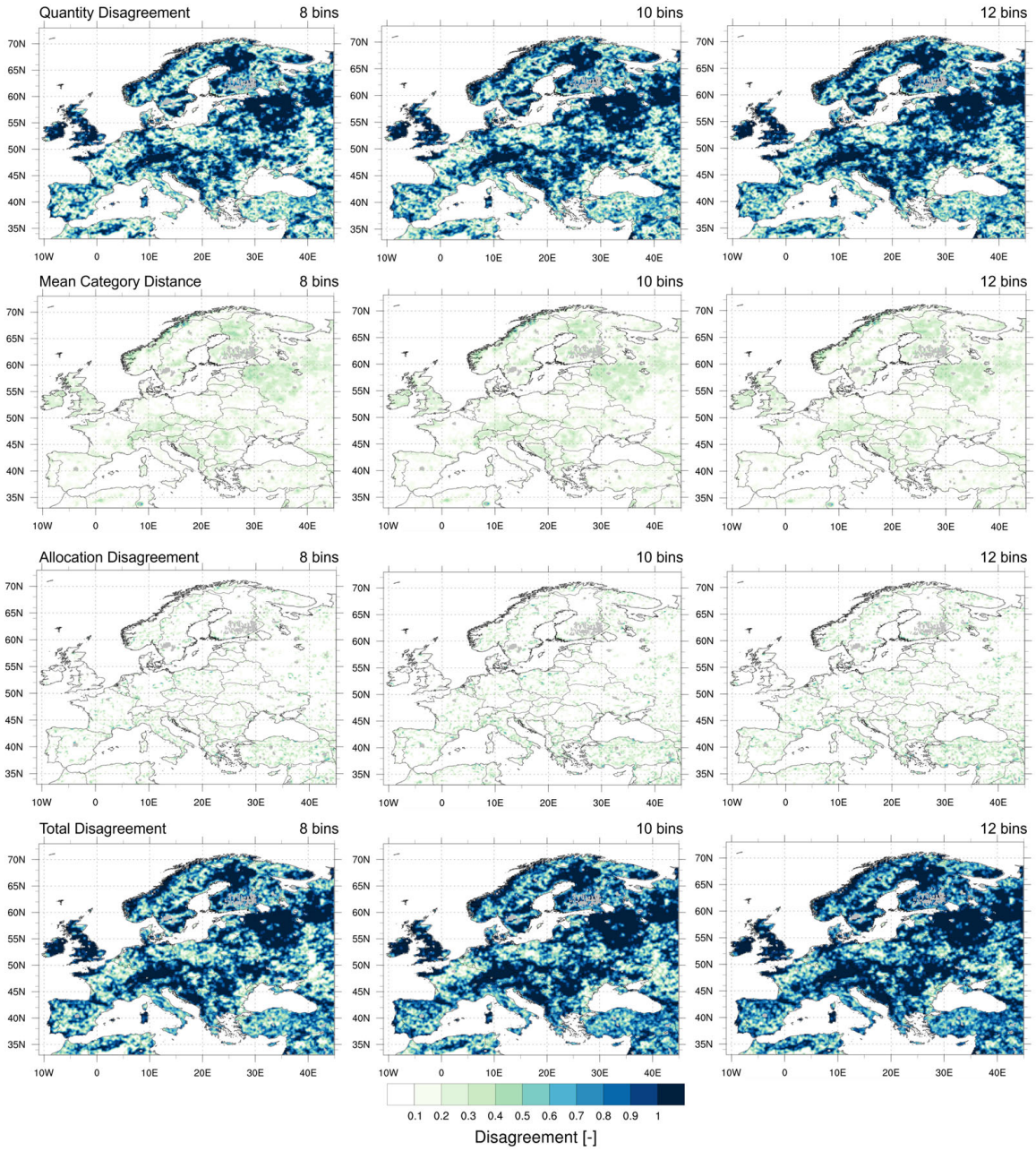

Figure A9. Maps of the four similarity metrics (rows) in 2022 for soil moisture in relative terms (soil wetness index) [-] calculated with different binning (8, 10, or 12 bins) (columns).

Appendix E. Evaluation of Statistical Models for Driver Analysis of Soil Moisture Patterns

In this appendix, we introduce the sensitivity analysis of the beta-coefficients to the choice of models considering spatial autocorrelation in comparison to the multi-linear regression model for analyzing the drivers of soil moisture. This is meant to support the results presented in Section 4.2. We acknowledge here that the multivariate linear regression analysis (see Section 3.4 for details) is influenced by the spatial autocorrelation within the patterns of the environmental and observational drivers.

To test the spatial autocorrelation-induced uncertainty in the driver ranking, we tested several designs of spatial autoregressive models, which directly account for spatial autocorrelation. Applying the same model formulation as for the linear model, several designs of the spatial autoregressive models, such as the spatial autoregressive lag model

(SAR), the spatial autoregressive combined model (SAC), and the spatial error model (SEM), were investigated with differently designed and sized neighborhoods [58]:

$$\text{LM}: \ y = \beta X + \epsilon \tag{A1}$$

with slope β and intercept ϵ.

$$\text{SAR}: \ \text{spatial autoregressive lag model}: \ y = \rho W y + \beta X + \epsilon \tag{A2}$$

where $\rho W y$ is a term representing spatial autocorrelation in the dependent variable y, and includes "spillover" effects from surroundings on the center cell with ρ being the fitted parameter and W a weight matrix.

$$\text{SEM}: \ \text{spatial error model}: \ y = \beta X + \lambda W u + \epsilon \tag{A3}$$

where $\lambda W u$ represents the autocorrelation in the error term (unrepresented spatial processes, such as topography in the SMAP DCA case)

$$\text{SAC -} > \ \text{spatial autoregressive combined model}: \ y = \rho W y + \beta X + \lambda W u + \epsilon \tag{A4}$$

SAC: represents both the spatial autocorrelation in the dependent variable and in the error term.

All models require a weight matrix as input, which defines the considered neighborhood [57,58]. We set up the neighborhoods using different weight matrices to be included in the autocorrelation analysis. We evaluated different configurations for the weight matrix, such as distance weighted influence that is based on k-NN neighbor search, as well as for a pixel-based neighborhood. The pixel-based weight matrix was set up for (1) direct neighbors, (2) 3 × 3, (3) 5 × 5 pixels, and the kNN-based weight matrix, equivalently, with (1) 4 neighbors (called direct), (2) 8 neighbors, and (3) 24 neighbors. Differences between those methods mainly occur in coastal areas, where a center cell is located next to water cells, which are excluded from the analyses. Motivation behind the size of the neighborhood is that the autocorrelation length in the soil moisture fields themselves for the entire continent is larger than 5°. Hence, the size of clusters is big and not captured by including direct neighbors and large-scale latitudinal and longitudinal gradients only. The latter are considered in all models by including latitude and longitude coordinates as variables in the statistical model. Despite the possibility to represent larger patterns, larger neighborhoods reduce the coefficients of determination of the full model and deprive the impact of the driver variables in the weight matrix. Therefore, larger neighborhoods with 9 × 9 and 15 × 15 cells or 80 as well as 224 neighbors were investigated but not included here. The final choice of the models for which we compared the beta-coefficients is based on the following list:

1. Reduction in the autocorrelation in the residuals of the spatial model against the LM represented by the Global Moran's I. value range: −1 representing a perfect checkerboard pattern, 0 exhibiting random spatial distribution (no autocorrelation), 1 showing highly clustered values.
2. No considerable loss in explainable variability (coefficient of determination)
3. Log-likelihood of the model maximized within the respective model category

Table A3 shows the different criteria for the choice of the spatial autocorrelation models and highlights (in bold) the models that were used in this study. The best configuration could be achieved with SACs, considering spatial autocorrelation in the dependent variable and the error term, while SARs brought significant rho coefficients, but could not reduce

Moran's I. The table does not include results for SEM, as none of the configurations were appropriate for the comparison.

Table A3. Different criteria for the choice of the spatial autocorrelation models (SAR, SEM, SAC) as well as LM (reference) and for the four different soil moisture products (SMAP DCA absolute values [m^3 m^{-3}], SMAP DCA relative values [-], IFS ECMWF absolute values [m^3 m^{-3}], IFS ECMWF relative values [-]), Bold font shows the models directly applied in the study.

			Global Moran's I		**r2**		**Likelihood**		**Rho**		**Lambda**	
			kNN	**Pixel**	**kNN**	**Pixel**	**kNN**	**Pixel**	**kNN**	**Pixel**	**kNN**	**Pixel**
SMAP DCA absolute	LM	direct	0.889	0.895	0.787		-8.748×10^4					
		3 × 3	0.866	0.842								
		5 × 5	0.814	0.796								
	SAR	**direct**	**0.920**	**0.925**	**0.799**	**0.799**	$\mathbf{1.268 \times 10^5}$	$\mathbf{1.308 \times 10^5}$	**0.942**	**0.943**		
		3 × 3	**0.912**	0.900	**0.767**	0.722	$\mathbf{1.203 \times 10^5}$	8.782×10^4	**0.949**	0.892		
		5 × 5	0.938	0.924	0.465	0.543	1.049×10^5	7.887×10^4	0.966	0.920		
	SAC	**direct**	**−0.034**	**−0.005**	**0.986**	**0.988**	$\mathbf{1.266 \times 10^5}$	$\mathbf{1.309 \times 10^5}$	**0.804**	**0.953**	**0.555**	**−0.059**
		3 × 3	0.004	−0.010	0.987	0.983	1.310×10^5	1.168×10^5	−0.093	0.053	1.015	1.027
		5 × 5	0.033	0.039	0.983	0.981	1.242×10^5	1.159×10^5	0.570	0.591	1.026	1.034
SMAP relative	LM	direct	0.871	0.876	0.527		-1.425×10^5					
		3 × 3	0.845	0.833								
		5 × 5	0.784	0.777								
	SAR	**direct**	**0.881**	**0.884**	**0.523**	**0.526**	$\mathbf{3.704 \times 10^4}$	$\mathbf{3.924 \times 10^4}$	**0.942**	**0.942**		
		3 × 3	0.862	**0.890**	**0.502**	**0.350**	$\mathbf{3.308 \times 10^4}$	2.516×10^4	**0.952**	0.950		
		5 × 5	0.952	**0.994**	−1.109	**−2.437**	2.245×10^4	1.673×10^4	0.989	0.989		
	SAC	**direct**	**0.004**	**0.003**	**0.952**	**0.955**	$\mathbf{3.720 \times 10^4}$	$\mathbf{4.012 \times 10^4}$	**0.952**	**0.961**	**−0.119**	**−0.195**
		3 × 3	−0.003	−0.025	0.942	0.932	3.309×10^4	2.592×10^4	0.946	0.820	0.081	0.531
		5 × 5	0.019	0.013	0.929	0.927	2.774×10^4	2.507×10^4	0.747	0.636	0.855	0.954
IFS absolute	LM	direct	0.627	0.629	0.836		-7.056×10^4					
		3 × 3	0.606	0.597								
		5 × 5	0.552	0.546								
	SAR	**direct**	**0.765**	**0.771**	**0.838**	**0.838**	$\mathbf{5.687 \times 10^4}$	$\mathbf{5.819 \times 10^4}$	**0.742**	**0.748**		
		3 × 3	0.713	0.706	0.836	0.828	4.499×10^4	3.929×10^4	0.698	0.709		
		5 × 5	0.649	0.639	0.825	0.829	3.219×10^4	2.950×10^4	0.662	0.659		
	SAC	**direct**	**−0.037**	**−0.048**	**0.963**	**0.962**	$\mathbf{6.568 \times 10^4}$	$\mathbf{6.559 \times 10^4}$	**0.616**	**0.622**	**0.535**	**0.494**
		3 × 3	**−0.039**	−0.067	0.956	0.954	5.741×10^4	5.253×10^4	**0.437**	0.226	**0.744**	0.906
		5 × 5	−0.001	−0.006	0.952	0.949	5.431×10^4	5.105×10^4	0.450	0.388	0.921	0.952
IFS relative	LM	direct	0.867	0.868	0.792		-8.651×10^4					
		3 × 3	0.838	0.827								
		5 × 5	0.776	0.767								
	SAR	**direct**	**0.925**	**0.921**	**0.723**	**0.743**	$\mathbf{1.167 \times 10^5}$	$\mathbf{1.167 \times 10^5}$	**0.958**	**0.953**		
		3 × 3	0.910	0.877	0.710	0.728	1.054×10^5	7.979×10^4	0.950	0.912		
		5 × 5	0.913	0.866	0.542	0.662	8.336×10^4	6.427×10^4	0.953	0.914		
	SAC	**direct**	**−0.019**	**−0.021**	**0.985**	**0.985**	$\mathbf{1.189 \times 10^5}$	$\mathbf{1.182 \times 10^5}$	**0.885**	**0.891**	**0.416**	**0.368**
		3 × 3	**−0.001**	−0.006	**0.983**	0.981	$\mathbf{1.165 \times 10^5}$	1.076×10^5	**0.157**	0.056	**0.972**	1.031
		5 × 5	0.021	0.028	0.977	0.974	1.029×10^5	9.670×10^4	0.482	0.441	1.033	1.041

Afterwards, the beta coefficients of the LM are compared with those from these spatial models that fulfill the criteria in the best way.

The beta coefficients of all chosen models were normalized by dividing by the absolute value of the largest beta coefficient of the respective model. Thus, the beta coefficients are scaled between −1 and 1 to compare the relative magnitude and the signs of the coefficients. We plot all normalized beta coefficients for the four cases (SMAP DCA in absolute values and relative values (Figure A10), as well as IFS model runs of ECMWF in absolute values and relative values (Figure A11). In all plots, the LM can be compared with the four to five additional spatial autocorrelation models, and we see that the normalized beta coefficients have the same sign and a similar magnitude, except for the drivers that show between +/−0.2 in the normalized beta coefficient. Here, the correlation is anyway low and of minor importance for driving the soil moisture products and their patterns. Since the spatial autocorrelation models state similar importance of individual drivers as the LM, our LM-based driver analysis is at least representative for the order of priority of the environmental and observational drivers influencing the soil moisture products (SMAP DCA and IFS ECMWF).

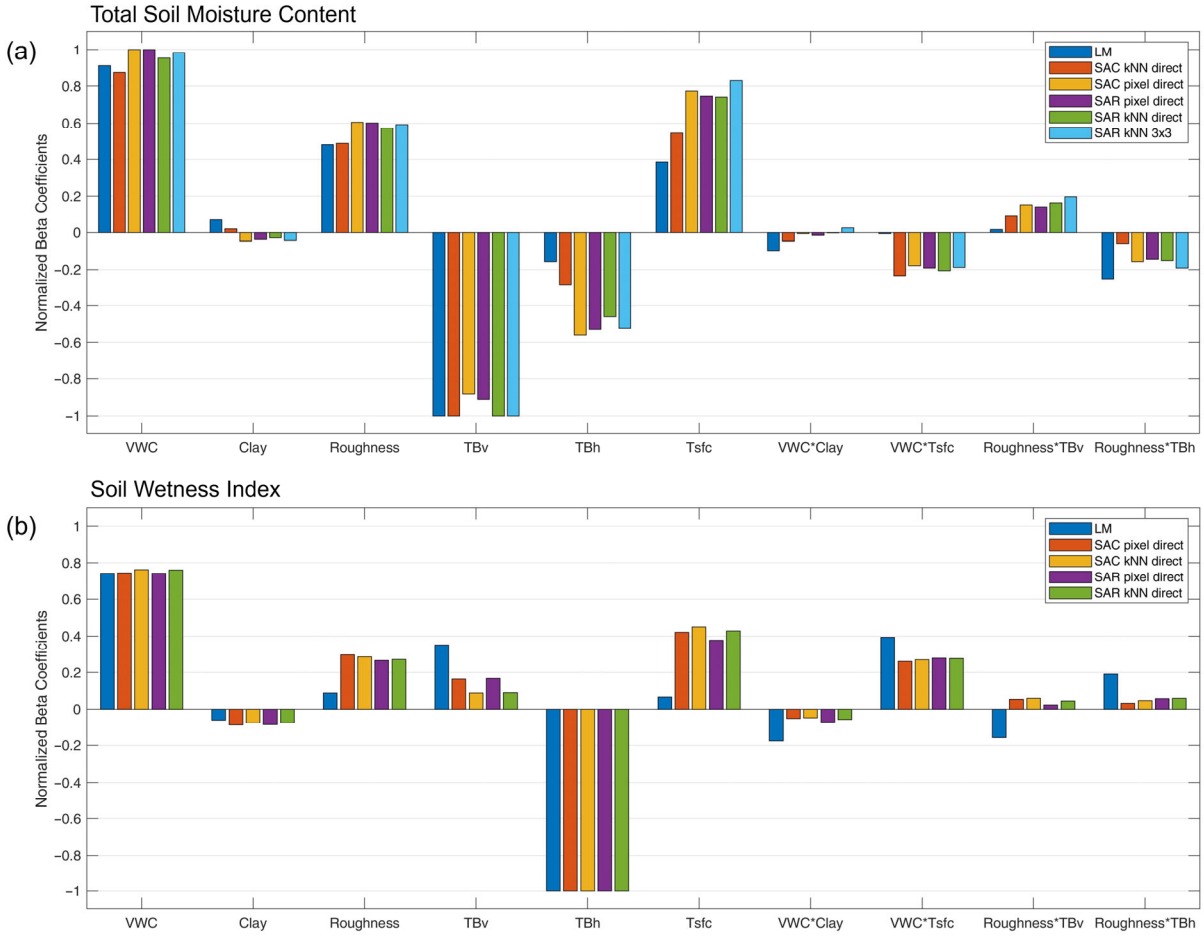

Figure A10. Comparison of the different pixel-based and k-NN neighbor-based statistical models (LM, SAC, and SAR) for their normalized beta coefficient along the individual drivers (environmental and observational) and their combinations (*) regarding the SMAP DCA soil moisture in absolute terms (**a**) and in relative terms (**b**).

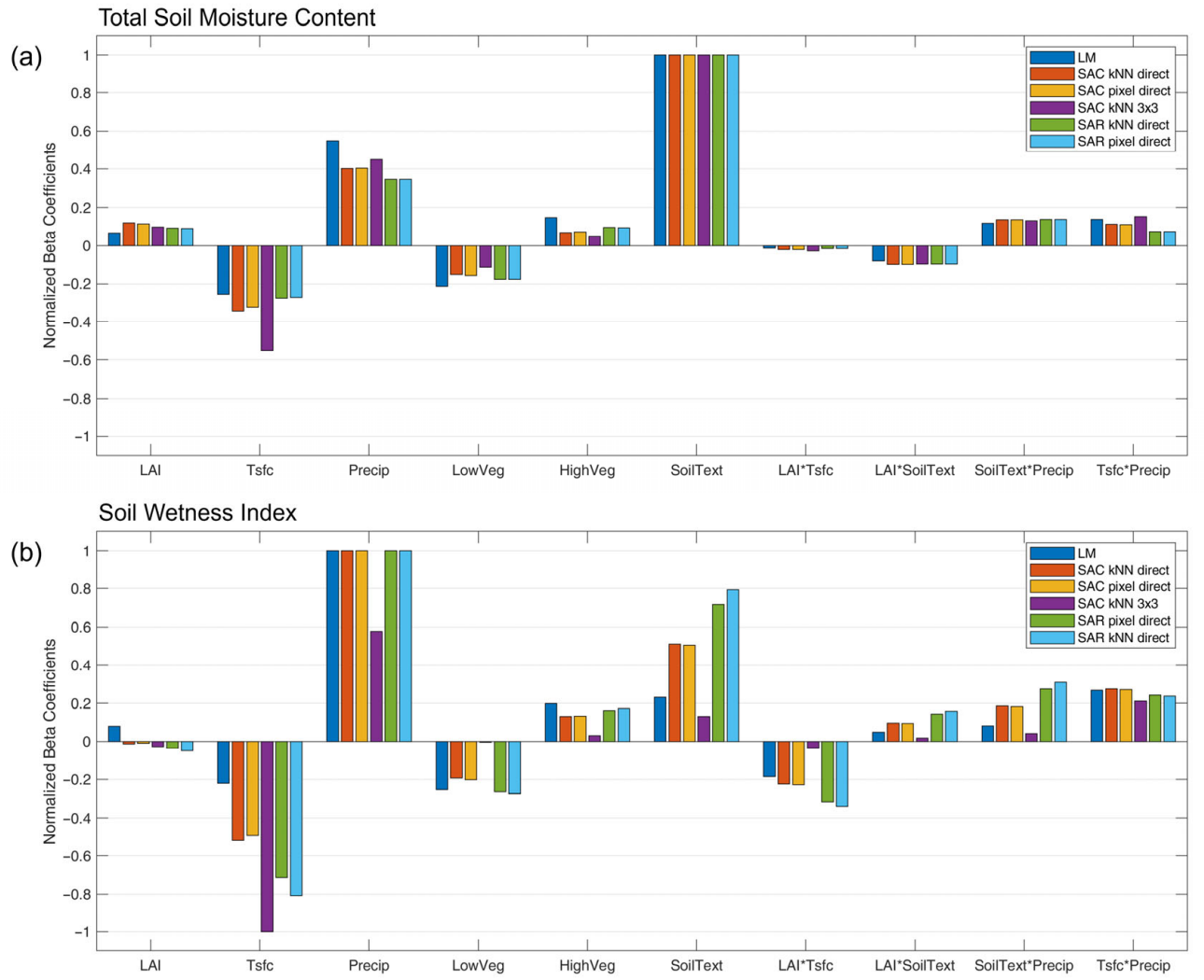

Figure A11. Comparison of the different pixel-based and k-NN neighbor-based statistical models (LM, SAC, and SAR) for their normalized beta coefficient along the individual drivers (environmental and observational) and their combinations (*) regarding the IFS model runs of soil moisture in absolute terms (**a**) and in relative terms (**b**).

References

1. Molz, F.J.; Remson, I.; Fungaroli, A.A.; Drake, R.L. Soil Moisture Availability for Transpiration. *Water Resour. Res.* **1968**, *4*, 1161–1169. [CrossRef]
2. Pospíšilová, J.; Nobel, P.S. Physicochemical & Environmental Plant Physiology. *Biol. Plant.* **1999**, *42*, 498. [CrossRef]
3. Lal, P.; Shekhar, A.; Gharun, M.; Das, N.N. Spatiotemporal Evolution of Global Long-Term Patterns of Soil Moisture. *Sci. Total Environ.* **2023**, *867*, 161470. [CrossRef] [PubMed]
4. Pablos, M.; Portal, G.; Camps, A.; Vall-llossera, M.; López-Martínez, C. Advances in Soil Moisture Mapping Techniques and Disaggregation Algorithms for Environmental and Agricultural Applications. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2025**, *XLVIII-G-2025*, 1139–1146. [CrossRef]
5. Western, A.W.; Blöschl, G. On the Spatial Scaling of Soil Moisture. *J. Hydrol.* **1999**, *217*, 203–224. [CrossRef]
6. Bronstert, A.; Bárdossy, A. The Role of Spatial Variability of Soil Moisture for Modelling Surface Runoff Generation at the Small Catchment Scale. *Hydrol. Earth Syst. Sci.* **1999**, *3*, 505–516. [CrossRef]
7. Bronstert, A.; Creutzfeldt, B.; Graeff, T.; Hajnsek, I.; Heistermann, M.; Itzerott, S.; Jagdhuber, T.; Kneis, D.; Lück, E.; Reusser, D.; et al. Potentials and Constraints of Different Types of Soil Moisture Observations for Flood Simulations in Headwater Catchments. *Nat. Hazards* **2012**, *60*, 879–914. [CrossRef]
8. Pauwels, V.R.N.; Hoeben, R.; Verhoest, N.E.C.; De Troch, F.P. The Importance of the Spatial Patterns of Remotely Sensed Soil Moisture in the Improvement of Discharge Predictions for Small-Scale Basins through Data Assimilation. *J. Hydrol.* **2001**, *251*, 88–102. [CrossRef]

9. Crosson, W.L.; Laymon, C.A.; Inguva, R.; Schamschula, M.P. Assimilating Remote Sensing Data in a Surface Flux–Soil Moisture Model. *Hydrol. Process.* **2002**, *16*, 1645–1662. [CrossRef]
10. Western, A.W.; Zhou, S.-L.; Grayson, R.B.; McMahon, T.A.; Blöschl, G.; Wilson, D.J. Spatial Correlation of Soil Moisture in Small Catchments and Its Relationship to Dominant Spatial Hydrological Processes. *J. Hydrol.* **2004**, *286*, 113–134. [CrossRef]
11. Hirschi, M.; Mueller, B.; Dorigo, W.; Seneviratne, S.I. Using Remotely Sensed Soil Moisture for Land–Atmosphere Coupling Diagnostics: The Role of Surface vs. Root-Zone Soil Moisture Variability. *Remote Sens. Environ.* **2014**, *154*, 246–252. [CrossRef]
12. Jackson, T.J.; Schmugge, J.; Engman, E.T. Remote Sensing Applications to Hydrology: Soil Moisture. *Hydrol. Sci. J.* **1996**, *41*, 517–530. [CrossRef]
13. van der Kwast, J. *Quantification of Top Soil Moisture Patterns: Evaluation of Field Methods, Process-Based Modelling, Remote Sensing and an Integrated Approach*; Nederlandse Geografische Studies/Netherlands Geographical Studies; Koninklijk Nederlands Aardrijkskundig Genootschap: Faculteit Geowetenschappen Universiteit Utrecht: Utrecht, The Netherlands, 2009.
14. Mohanty, B.P.; Cosh, M.H.; Lakshmi, V.; Montzka, C. Soil Moisture Remote Sensing: State-of-the-Science. *Vadose Zone J.* **2017**, *16*, 1–9. [CrossRef]
15. Fairbairn, D.; De Rosnay, P.; Weston, P. Evaluation of an Adaptive Soil Moisture Bias Correction Approach in the ECMWF Land Data Assimilation System. *Remote Sens.* **2024**, *16*, 493. [CrossRef]
16. Gomis-Cebolla, J.; Garcia-Arias, A.; Perpinyà-Vallès, M.; Francés, F. Evaluation of Sentinel-1, SMAP and SMOS Surface Soil Moisture Products for Distributed Eco-Hydrological Modelling in Mediterranean Forest Basins. *J. Hydrol.* **2022**, *608*, 127569. [CrossRef]
17. Balenzano, A.; Mattia, F.; Satalino, G.; Davidson, M.W.J. Dense Temporal Series of C- and L-Band SAR Data for Soil Moisture Retrieval Over Agricultural Crops. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2011**, *4*, 439–450. [CrossRef]
18. Mengen, D.; Jagdhuber, T.; Balenzano, A.; Mattia, F.; Vereecken, H.; Montzka, C. High Spatial and Temporal Soil Moisture Retrieval in Agricultural Areas Using Multi-Orbit and Vegetation Adapted Sentinel-1 SAR Time Series. *Remote Sens.* **2023**, *15*, 2282. [CrossRef]
19. Dorigo, W.; Himmelbauer, I.; Aberer, D.; Schremmer, L.; Petrakovic, I.; Zappa, L.; Preimesberger, W.; Xaver, A.; Annor, F.; Ardö, J.; et al. The International Soil Moisture Network: Serving Earth System Science for over a Decade. *Hydrol. Earth Syst. Sci.* **2021**, *25*, 5749–5804. [CrossRef]
20. Schröter, I.; Paasche, H.; Doktor, D.; Xu, X.; Dietrich, P.; Wollschläger, U. Estimating Soil Moisture Patterns with Remote Sensing and Terrain Data at the Small Catchment Scale. *Vadose Zone J.* **2017**, *16*, 1–21. [CrossRef]
21. Wagner, W.; Scipal, K.; Pathe, C.; Gerten, D.; Lucht, W.; Rudolf, B. Evaluation of the Agreement between the First Global Remotely Sensed Soil Moisture Data with Model and Precipitation Data. *J. Geophys. Res.* **2003**, *108*, 2003JD003663. [CrossRef]
22. Albergel, C.; De Rosnay, P.; Gruhier, C.; Muñoz-Sabater, J.; Hasenauer, S.; Isaksen, L.; Kerr, Y.; Wagner, W. Evaluation of Remotely Sensed and Modelled Soil Moisture Products Using Global Ground-Based in Situ Observations. *Remote Sens. Environ.* **2012**, *118*, 215–226. [CrossRef]
23. Al-Shrafany, D.; Rico-Ramirez, M.A.; Han, D.; Bray, M. Comparative Assessment of Soil Moisture Estimation from Land Surface Model and Satellite Remote Sensing Based on Catchment Water Balance: Land Surface Model Coupled to Numerical Weather Model. *Met. Apps* **2014**, *21*, 521–534. [CrossRef]
24. Fang, L.; Hain, C.R.; Zhan, X.; Anderson, M.C. An Inter-Comparison of Soil Moisture Data Products from Satellite Remote Sensing and a Land Surface Model. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *48*, 37–50. [CrossRef]
25. Babaeian, E.; Sadeghi, M.; Jones, S.B.; Montzka, C.; Vereecken, H.; Tuller, M. Ground, Proximal, and Satellite Remote Sensing of Soil Moisture. *Rev. Geophys.* **2019**, *57*, 530–616. [CrossRef]
26. Ford, T.W.; Quiring, S.M. Comparison of Contemporary In Situ, Model, and Satellite Remote Sensing Soil Moisture With a Focus on Drought Monitoring. *Water Resour. Res.* **2019**, *55*, 1565–1582. [CrossRef]
27. Peng, C.; Zeng, J.; Chen, K.-S.; Ma, H.; Letu, H.; Zhang, X.; Shi, P.; Bi, H. Spatial Representativeness of Soil Moisture Stations and Its Influential Factors at a Global Scale. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 1–15. [CrossRef]
28. Wang, Q.; Lu, C.; Zheng, Y.; Wang, Z.; Li, J.; Tan, Y. In-Situ and Triple Collocation-Based Evaluations of Tianmu-1 Global Soil Moisture Products. *Remote Sens. Environ.* **2025**, *328*, 114892. [CrossRef]
29. Zhu, Z.; Feng, M.; Cornelis, W.; Miralles, D.G.; Zeng, Z.; Chen, Y.; Yang, Z.; Zou, S.; Liu, Y.; De Maeyer, P.; et al. Global Soil Moisture Dynamics since 1980: Datasets Biases, Trends, and Science-Informed Selection. *Sci. Bull.* **2025**, *70*, 4253–4262. [CrossRef]
30. Escorihuela, M.J.; Quintana-Seguí, P. Comparison of Remote Sensing and Simulated Soil Moisture Datasets in Mediterranean Landscapes. *Remote Sens. Environ.* **2016**, *180*, 99–114. [CrossRef]
31. Hain, C.R.; Crow, W.T.; Mecikalski, J.R.; Anderson, M.C.; Holmes, T. An Intercomparison of Available Soil Moisture Estimates from Thermal Infrared and Passive Microwave Remote Sensing and Land Surface Modeling. *J. Geophys. Res.* **2011**, *116*, D15107. [CrossRef]
32. Wang, P.; Zeng, J.; Chen, K.-S.; Ma, H.; Zhang, X.; Shi, P.; Peng, C.; Bi, H. Global-Scale Assessment of Multiple Recently Developed/Reprocessed Remotely Sensed Soil Moisture Datasets. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 1–18. [CrossRef]

33. Vischel, T.; Pegram, G.G.S.; Sinclair, S.; Wagner, W.; Bartsch, A. Comparison of Soil Moisture Fields Estimated by Catchment Modelling and Remote Sensing: A Case Study in South Africa. *Hydrol. Earth Syst. Sci.* **2008**, *12*, 751–767. [CrossRef]
34. Entekhabi, D.; Yueh, S.; O'Neill, P.E.; Kellogg, K. *SMAP Handbook*; JPL: Pasadena, CA, USA, 2014; pp. 400–1567.
35. ONeill, P.; Chan, S.; Njoku, E.; Jackson, T.; Bindlish, R.; Chaubell, J.; Colliander, A. SMAP Enhanced L3 Radiometer Global and Polar Grid Daily 9 Km EASE-Grid Soil Moisture, Version 6. 2023. Available online: https://nsidc.org/data/spl3smp_e/versions/6 (accessed on 10 March 2023).
36. Chaubell, J. *SMAP L1B Enhancement Radiometer Brightness Temperature Data Product*; California Institute of Technology. U.S. Government sponsorship acknowledged. Soil Moisture Active Passive (SMAP) Project Algorithm Theoretical Basis Document SMAP L1B Enhancement Radiometer Brightness Temperature Data Product Julian Chaubell Jet Propulsion Laboratory; California Institute of Technology: Pasadena, CA, USA, 2016; p. 24.
37. O'Neill, P.; Bindlish, R.; Chan, S.; Chaubell, J.; Colliander, A.; Njoku, E.; Jackson, T. *Soil Moisture Active Passive (SMAP) Algorithm Theoretical Basis Document Level 2 & 3 Soil Moisture (Passive) Data Products—Revision G*; JPL: La Cañada Flintridge, CA, USA, 2021.
38. Melton, J.R.; Chan, E.; Millard, K.; Fortier, M.; Winton, R.S.; Martín-López, J.M.; Cadillo-Quiroz, H.; Kidd, D.; Verchot, L.V. A Map of Global Peatland Extent Created Using Machine Learning (Peat-ML). *Geosci. Model Dev.* **2022**, *15*, 4709–4738. [CrossRef]
39. ECMWF. *IFS Documentation CY47R3—Part I: Observations*; ECMWF: Reading, UK, 2021. [CrossRef]
40. ECMWF. *IFS Documentation CY47R3—Part II: Data Assimilation*; ECMWF: Reading, UK, 2021. [CrossRef]
41. ECMWF. *IFS Documentation CY47R3—Part IV Physical Processes*; ECMWF: Reading, UK, 2021. [CrossRef]
42. Hogan, R.J.; Bozzo, A. A Flexible and Efficient Radiation Scheme for the ECMWF Model. *J. Adv. Model. Earth Syst.* **2018**, *10*, 1990–2008. [CrossRef]
43. Becker, T.; Bechtold, P.; Sandu, I. Characteristics of Convective Precipitation over Tropical Africa in Storm-resolving Global Simulations. *Quart. J. R. Meteoro Soc.* **2021**, *147*, 4388–4407. [CrossRef]
44. Tiedtke, M. A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models. *Mon. Wea. Rev.* **1989**, *117*, 1779–1800. [CrossRef]
45. Tiedtke, M. Representation of Clouds in Large-Scale Models. *Mon. Wea. Rev.* **1993**, *121*, 3040–3061. [CrossRef]
46. Forbes, R.; Tompkins, A. An Improved Representation of Cloud and Precipitation. *ECMWF Newsl.* **2011**, *129*, 13–18. [CrossRef]
47. Tompkins, A.M.; Gierens, K.; Rädel, G. Ice Supersaturation in the ECMWF Integrated Forecast System. *Quart. J. R. Meteoro Soc.* **2007**, *133*, 53–63. [CrossRef]
48. Boussetta, S.; Balsamo, G.; Arduini, G.; Dutra, E.; McNorton, J.; Choulga, M.; Agustí-Panareda, A.; Beljaars, A.; Wedi, N.; Munõz-Sabater, J.; et al. ECLand: The ECMWF Land Surface Modelling System. *Atmosphere* **2021**, *12*, 723. [CrossRef]
49. Feldman, A.F.; Short Gianotti, D.J.; Dong, J.; Akbar, R.; Crow, W.T.; McColl, K.A.; Konings, A.G.; Nippert, J.B.; Tumber-Dávila, S.J.; Holbrook, N.M.; et al. Remotely Sensed Soil Moisture Can Capture Dynamics Relevant to Plant Water Uptake. *Water Resour. Res.* **2023**, *59*, e2022WR033814. [CrossRef]
50. Heiskanen, J.; Brümmer, C.; Buchmann, N.; Calfapietra, C.; Chen, H.; Gielen, B.; Gkritzalis, T.; Hammer, S.; Hartman, S.; Herbst, M.; et al. The Integrated Carbon Observation System in Europe. *Bull. Am. Meteorol. Soc.* **2022**, *103*, E855–E872. [CrossRef]
51. Metzger, M.J.; Bunce, R.G.H.; Jongman, R.H.G.; Mücher, C.A.; Watkins, J.W. A Climatic Stratification of the Environment of Europe. *Glob. Ecol. Biogeogr.* **2005**, *14*, 549–563. [CrossRef]
52. Pontius, R.G.; Millones, M. Death to Kappa: Birth of Quantity Disagreement and Allocation Disagreement for Accuracy Assessment. *Int. J. Remote Sens.* **2011**, *32*, 4407–4429. [CrossRef]
53. Roberts, N.M.; Lean, H.W. Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Mon. Weather Rev.* **2008**, *136*, 78–97. [CrossRef]
54. Portal, G.; Vall-llossera, M.; Piles, M.; Camps, A.; Chaparro, D.; Pablos, M.; Rossato, L. A Spatially Consistent Downscaling Approach for SMOS Using an Adaptive Moving Window. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 1883–1894. [CrossRef]
55. Tian, J.; Zhang, Y.; Klein, S.A.; Öktem, R.; Wang, L. How Does Land Cover and Its Heterogeneity Length Scales Affect the Formation of Summertime Shallow Cumulus Clouds in Observations From the US Southern Great Plains? *Geophys. Res. Lett.* **2022**, *49*, e2021GL097070. [CrossRef]
56. Pontius, R.G. Component Intensities to Relate Difference by Category with Difference Overall. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *77*, 94–99. [CrossRef]
57. LeSage, J.P. Applied Econometrics Using Matlab. University of Toledo, Department of Economics, October 1999. Available online: https://www.spatial-econometrics.com/html/mbook.pdf (accessed on 22 May 2024).
58. Dormann, C.F.; McPherson, J.M.; Araújo, M.B.; Bivand, R.; Bolliger, J.; Carl, G.; Davies, R.G.; Hirzel, A.; Jetz, W.; Daniel Kissling, W.; et al. Methods to Account for Spatial Autocorrelation in the Analysis of Species Distributional Data: A Review. *Ecography* **2007**, *30*, 609–628. [CrossRef]

59. Al-Yaari, A.; Wigneron, J.-P.; Kerr, Y.; Rodriguez-Fernandez, N.; O'Neill, P.E.; Jackson, T.J.; De Lannoy, G.J.M.; Al Bitar, A.; Mialon, A.; Richaume, P.; et al. Evaluating Soil Moisture Retrievals from ESA's SMOS and NASA's SMAP Brightness Temperature Datasets. *Remote Sens. Environ.* **2017**, *193*, 257–273. [CrossRef]
60. Zwieback, S.; Colliander, A.; Cosh, M.H.; Martínez-Fernández, J.; McNairn, H.; Starks, P.J.; Thibeault, M.; Berg, A. Estimating Time-Dependent Vegetation Biases in the SMAP Soil Moisture Product. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 4473–4489. [CrossRef]
61. Fan, X.; Liu, Y.; Gan, G.; Wu, G. SMAP Underestimates Soil Moisture in Vegetation-Disturbed Areas Primarily as a Result of Biased Surface Temperature Data. *Remote Sens. Environ.* **2020**, *247*, 111914. [CrossRef]
62. Colliander, A.; Reichle, R.H.; Crow, W.T.; Cosh, M.H.; Chen, F.; Chan, S.; Das, N.N.; Bindlish, R.; Chaubell, J.; Kim, S.; et al. Validation of Soil Moisture Data Products From the NASA SMAP Mission. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 364–392. [CrossRef]
63. Portal, G.; Jagdhuber, T.; Vall-llossera, M.; Camps, A.; Pablos, M.; Entekhabi, D.; Piles, M. Assessment of Multi-Scale SMOS and SMAP Soil Moisture Products across the Iberian Peninsula. *Remote Sens.* **2020**, *12*, 570. [CrossRef]
64. Jach, L.; Bauer, H.-S.; Chaparro, D.; Fluhrer, A.; Hellwig, F.M.; Portal, G.; Jagdhuber, T. Comparing Temporal Dynamics of Soil Moisture from Remote Sensing, Modeling and Field Observations across Europe. *Remote Sens.* **2026**, *18*, 445. [CrossRef]
65. Tradowsky, J.S.; Philip, S.Y.; Kreienkamp, F.; Kew, S.F.; Lorenz, P.; Arrighi, J.; Bettmann, T.; Caluwaerts, S.; Chan, S.C.; De Cruz, L.; et al. Attribution of the Heavy Rainfall Events Leading to Severe Flooding in Western Europe during July 2021. *Clim. Change* **2023**, *176*, 90. [CrossRef]
66. Peng, C.; Zeng, J.; Chen, K.-S.; Li, Z.; Ma, H.; Zhang, X.; Shi, P.; Wang, T.; Yi, L.; Bi, H. Global Spatiotemporal Trend of Satellite-Based Soil Moisture and Its Influencing Factors in the Early 21st Century. *Remote Sens. Environ.* **2023**, *291*, 113569. [CrossRef]
67. Li, Y.; Bao, L.; He, P.; Wang, C.; Zhu, S.; Ma, J.; Liang, Y.; Xiao, J. Global Soil Moisture Dynamics: Attribution of Contributions and Their Association with GPP. *Front. Plant Sci.* **2025**, *16*, 1691082. [CrossRef]
68. Ulaby, F.T.; Long, D.G. *Microwave Radar and Radiometric Remote Sensing*; University of Michigan Press: Ann Arbor, MI, USA, 2014; ISBN 978-0-472-11935-6.
69. Wigneron, J.-P.; Jackson, T.J.; O'Neill, P.; De Lannoy, G.; De Rosnay, P.; Walker, J.P.; Ferrazzoli, P.; Mironov, V.; Bircher, S.; Grant, J.P.; et al. Modelling the Passive Microwave Signature from Land Surfaces: A Review of Recent Results and Application to the L-Band SMOS & SMAP Soil Moisture Retrieval Algorithms. *Remote Sens. Environ.* **2017**, *192*, 238–262. [CrossRef]
70. Muñoz-Sabater, J.; Lawrence, H.; Albergel, C.; Rosnay, P.; Isaksen, L.; Mecklenburg, S.; Kerr, Y.; Drusch, M. Assimilation of SMOS Brightness Temperatures in the ECMWF Integrated Forecasting System. *Quart. J. R. Meteoro Soc.* **2019**, *145*, 2524–2548. [CrossRef]