



**Statistical approaches to assimilate ASCAT
soil moisture information: Uncertainty
assessment, NN localisation, and evaluation**

Filipe Aires, Patricia De Rosnay, Peter Weston, and
David Fairbairn

EUMETSAT activity ID: **H_AS21_01**

November 30th, 2021

Abstract

Land surfaces are characterised by strong heterogeneities of soil texture, orography, land cover, soil moisture, snow and other variables. These are very challenging to represent accurately in radiative transfer models which have a limited reliability over land, especially for observations such as given. This has resulted in difficulties for assimilation of land surface related satellite observations in numerical weather prediction models. Simple statistical relationships between the satellite observations and surface variables have therefore been considered in the last 20 years. In this study, we propose to compare two such approaches: the CDF-matching (used as a normalisation and an inversion technique) and the Neural Network (NN) methods. The CDF-matching finds a simple mono-variate relationship at the local scale and is very dependent on the LSM on which it is calibrated. NNs are global multivariate models able to exploit auxiliary information and the synergy of multiple instruments, but the solution is global and no local characteristics are constraining the solution, leading sometimes to local biases. One of these two methods can be better suited, depending on the application, in particular the simplicity of the satellite/variable relationship. We illustrate here these concepts using the ASCAT (Advanced SCATerometer) observations for the SM retrieval in an assimilation context. The two approaches are compared and combined. We also compare the more traditional inversion scheme and the forward modelling that could be attractive for assimilation purposes. It is shown that in the context of SCAT, the inversion approach seems better suited than the forward modelling. It is also shown that it is possible to combine the global model obtained using the NN and the localised information of the LSM offered by the CDF-matching.

In a second step, we propose to adjust the NN approach by using a localization strategy. Localization is defined here as the process to help the NN model to be more specific for local conditions. Several strategies can be employed for localization, we tested here the training of individual NNs in each pixel. Tests over USA are conducted and show that localization results in better retrievals of soil moisture. This is true when compared to the target (the reanalysis) but also when compared to a large network of in situ measurements over the USA. Localization appears to be an important tool that could facilitate the assimilation of satellite data (in forward or inverse mode) into NWP data assimilation frameworks. Another innovation is the estimation of the NN uncertainties. We proposed here to estimate the uncertainties based on a sliding time window, for each pixel, that provides

an empirical errors statistics. Then a NN is requested to estimated these uncertainty statistics (i.e. error STD) based on the same inputs as those for the original NN retrieval model. With this study, we move forward in our understanding of the NN methodology and we proposed new approach to better assimilate satellite observations into NWP centres.

Contents

1	Introduction	3
2	Neural network for satellite remote sensing	7
3	Datasets	10
3.1	ERA5 database	10
3.1.1	Soil Temperature (ST)	10
3.1.2	Leaf Area Index (LAI)	10
3.1.3	Amplitude of the diurnal cycle of surface temperature ΔT	12
3.1.4	Soil Moisture	12
3.2	ASCAT information	12
3.3	Preliminary analysis	15
4	Methodology for the exploitation of soil moisture information	17
4.1	CDF-matching	17
4.1.1	Methodology	17
4.1.2	Rescaling tool	19
4.1.3	Retrieval scheme	22
4.2	Neural Networks	24
4.2.1	Model	24
4.2.2	NN database	25
4.3	Quality criteria	26
5	Experiments and results	29
5.1	Retrieval schemes	31
5.2	Forward modeling	33
5.3	Comparison of the CDF-matching techniques	35
5.4	Combining algorithms	35

6	Evaluation	38
6.1	Localization of the NN model	38
6.1.1	Dataset	40
6.1.2	SM retrieval results over USA	43
6.2	NN uncertainty estimation	44
6.2.1	Using the learning database \mathcal{B} as a reference	44
6.2.2	Estimation of the “empirical” uncertainty	45
6.2.3	Uncertainty NN model	45
6.2.4	Uncertainty estimation results	45
6.3	Validation using in situ Measurements	48
6.3.1	In situ soil moisture database	48
6.3.2	Evaluation	48
7	Conclusion and perspectives	51
8	Annexe	55

Chapter 1

Introduction

Satellite observations are an important source of information for assimilation into Numerical Weather Prediction (NWP) models [54, 16, 41, 65]. Two different approaches to the assimilation of satellite observations can be considered: the assimilation of the raw radiances R or the assimilation of the retrieved products P . The assimilation process weights the different sources of information (i.e. satellite and model) based on their respective uncertainties [56]. An uncertainty estimate for the satellite information is therefore necessary on R or P .

Forward model - Since reliable uncertainties on retrieved products are difficult to obtain (i.e. they are state-dependent and sensitive to several factors), the assimilation of raw radiances R instead of retrieved products P has been preferred in the past. The assimilation of R is possible only when the Radiative Transfer (RT) to simulate R is reliable enough. For instance, this has been the case for atmospheric applications, e.g. sounding of temperature and humidity atmospheric profiles. For many years the assimilation of radiances was restricted to clear skies, with careful screening for cloud and precipitation. However, a number of NWP centres have recently made significant progress in assimilating radiances in the presence of cloud and precipitation. This was made possible in part by improvements to scattering parameterisations in the RT models [28]. The RT calculations are less accurate for surface simulations because required auxiliary parameters (soil texture, surface emissivities, vegetation state, *etc.*) are often difficult to obtain and have large uncertainties at the global scale. Furthermore, it is often assumed that observation errors can be represented by the sum of (1) a constant instrument error, and (2) a constant RT error, whatever the surface or atmospheric conditions. This simple assumption is satisfactory in many cases, but it can also be misleading since instrument and RT errors might be state-dependent [14] and very difficult to model (e.g., [8] for an investi-

gation of more realistic assumptions for cloudy situations with microwave observations).

Inverse model - Since RT simulations at the surface are difficult, the assimilation of retrieved products P is a viable alternative to direct assimilation of raw radiances R . For instance, the assimilation of retrieved SM from satellite observations can be more useful [2] than the assimilation of the raw satellite data R . This has been shown for instance for SMOS (Soil Moisture and Ocean Salinity) observations [57]. The assimilation of retrieved products P requires some estimation of the uncertainties, but they are often considered to be constant. In fact, estimating retrieval uncertainties is a difficult task but it should be noted that *a priori* errors (i.e., the error of the *a priori* information before the satellite observations are used in the assimilation) need to be characterised too, and this was not an easy task for the NWP community: uncertainty estimates were global at first, then latitudinally-dependent, and they became state-dependent (see for instance [36]) and this is still a subject of research. Such an effort can be performed too for retrieval errors. For instance, when retrieved by a classical optimal interpolation technique, a state-dependent error is provided by the theory [56]. The same process can be performed with any statistical inversion scheme, including for the NN retrievals [7].

One important aspect of assimilation is that assimilated data need to be coherent enough with the numerical model. Otherwise, jumps will be generated in the time series of the model variables, with values alternating between the satellite observations and the model climatology. Therefore, an often adopted strategy is to bias-correct the satellite observations using globally available predictors, in order to better match the model at the pixel level. One example for atmospheric variables is variational bias-correction [20]. Another form of bias-correction often used for surface applications is the CDF-matching (Cumulative Distribution Function) [23, 12, 19, 47]. With this approach, the CDF of the assimilated satellite information is first transformed towards the CDF of the corresponding variable in the model (both CDFs are assumed to be Gaussian), before the assimilation is done. This means that averages, ranges of values, and even physical units of the variable are harmonised. This process has some advantages and some drawbacks, which will be discussed below. The relationship between observation (backscatter) and variable (soil moisture) is not simple. However, thanks to many simplifications and assumptions leading to a pixel by pixel temporal change detection retrieval for ASCAT, the relationship between satellite derived soil moisture and model soil moisture is represented by a simple relationship. The CDF-matching represents a simple scaling between the satellite measurement and the variable to retrieve, enabling it to be used as

a local retrieval scheme [67].

NN are statistical models that allow for a link to be built between satellite observations and the geophysical variables to retrieve [24, 5, 49]. They have mostly been used as a retrieval tool, in inversion mode. They can also be used as forward models to simulate satellite observations [15]. This approach hasn't been used as frequently, in part because RT can be very complex and necessitates many inputs which are not always available. NN are designed to represent relatively simple, statistical relationships at the global scale, and very important physical dependencies might not be attainable only from a data-driven approach. A compromise could be found in the form of an hybrid model where inputs of the NN are chosen to simplify the forward model (as in the RTTOV model [61, 46, 62]) and physical relationships could be coded in the NN, the training would then estimate the appropriate coefficients of such relationships. Furthermore, the use of a NN direct model in an assimilation scheme would mean that errors are present in the forward (the NN) and the inverse mode (inside the assimilation), where a NN retrieval scheme would include only inversion errors.

Therefore, several strategies can be used to exploit satellite observations into land surface models. In this study, several questions will be addressed in the context of the ASCAT observations: (1) should we develop an inverse (for the assimilation of retrieved products P) or a forward (for the assimilation of raw observations R) statistical model? (2) What is the best matching/inversion technique, NN or CDF-matching? While these approaches can be implemented separately, their combination will also be considered. The NN can bring its global perspective and its exploitation of auxiliary/ancillary information, while the CDF-matching can help improve the local matching of the satellite information with the LSM if this is required to facilitate the assimilation. (3) The principle of localization will be introduced and tested for the retrieval of soil moisture. This approach to obtain a NN with more local properties can be a way to avoid using CDF-matching and also to improve the results of a standard NN retrieval. (4) We will introduce also a methodology to estimate the NN uncertainties, i.e. the uncertainties made by the NN when estimating the soil moisture. (5) Finally, evaluation of this strategy will be made over the USA using in situ soil moisture measurements.

The ASCAT (Advanced SCATerometer) instrument [26] is an active microwave scatterometer acting in the C-band (5.3 GHz). The instrument is flown aboard the MetOp satellites with a revisit time of 1-2 days. It provides a backscatter value at a 25 km spatial resolution. The ASCAT soil penetration depth is typically around 1-2 cm, but can be significantly larger for very dry soils [52]. A global dataset of surface SM has been built from ASCAT backscatter observations using the change detection technique [67]. The EU-

METSAT Satellite Applications Facility for Hydrology (H SAF) has adopted this approach for the development and production of level 2 ASCAT-derived surface SM products. The 25 km sampling H SAF product [30, 32] is assimilated operationally in the UK Met Office and the ECMWF forecasting systems [21, 63, 18, 25], and it is used for research developments in other NWP centres [12], but prior to assimilation the SM is bias-corrected via the CDF matching technique. The same ASCAT backscatter observations will be considered here for the retrieval of surface SM, but instead using the NN approach. In the past, ASCAT observations have been combined *a posteriori* with passive microwave observations [45] or combined as inputs to a retrieval process, e.g. [43]. It is foreseen that the techniques presented in this paper could be extended to passive microwave observations such as those derived from the SMOS mission [57, 59].

Section 2 gives the historical background of neural networks for satellite remote sensing. Section 3 introduces the datasets used in this study. In Section 4, the NN and CDF-matching are presented, including new CDF-matching approaches; and the diagnostics used in this study are proposed in the context of the assimilation framework. Experiments based on SM retrieval from ASCAT data are presented in Section 5. Evaluation of the localization over the USA is conducted in Section 6, together with NN uncertainty estimation and validation using in situ measurements. Finally, Section 7 concludes this study and provides some practical perspectives.

Chapter 2

Neural network for satellite remote sensing

Neural Networks (NN) have been widely used in the field of satellite remote sensing, for more than 20 years, for the atmosphere [24, 5], or the land [5] and ocean [49] surfaces. Two key limitations have driven the need to develop NN approaches: Firstly that no physical RT model is accurate enough for surface applications, and secondly, that no empirical SM database is large or representative enough to calibrate a statistical retrieval scheme. Therefore, [2] proposed a NN approach based on four innovations:

- (1) *Use a NN for the retrieval of SM* - Although NN methods were not easily accepted initially, because they were considered as “black-box” approaches, their relevance to better understand physical problems is now widely acknowledged. NNs can extract in a synthetic way physical relationships buried in a large database. Hence, NN methods for satellite retrievals has grown tremendously in popularity in the past decade [42, 58, 44], with also recent enthusiasm on big data such as deep learning [11] and convolutional neural networks [64].
- (2) *Train this NN using model SM outputs* - The use of model SM to train the NN was strongly rejected in the early days of NN developments. It was thought at that time that a NN trained in this way would then only reproduce the spatial and temporal behaviour of the model, and consequently would be useless. This is wrong because the NN is trained at the pixel and instantaneous scales, with no information on location or time. Spatial or temporal patterns from the model are not imposed on the NN: Once trained, the NN temporal and spatial dynamics is driven by its inputs (i.e., the satellite observations) and not by the model. The NN can even be trained with an imperfect database provided that

this database includes the general sensitivities required for the retrieval. The machine learning approach will find these general sensitivities and will remove any spurious behaviour not coherent with the input data. During the learning phase, the NN will not take into account these errors if they are not coherent with the satellite observations. [2] illustrates this principle and show how to use it to identify suspect regions in the model SM, and [39] checks how a synthetic error can be corrected by the NN. Therefore, an imperfect dataset can be used to calibrate a NN retrieval scheme. There are still controversial aspects regarding model-trained NN retrievals. For example, a NN retrieval was trained using SMOS SM retrievals performed using a physical inversion scheme [59] instead of using the SM model outputs, with the aim of speeding up the physical inversion scheme. This is a valid approach, since the NN is designed to reproduce the physical inversion scheme and it might correct some of its errors. However, the training using LSM SM outputs will continue to be preferable because they often provide better SM estimates at the global scale than the dedicated satellite datasets. Furthermore, there is no “incestuous” process prohibiting the use of a model-trained NN in the assimilation. The NN is trained using model data, just as the CDF-matching is calibrated using model outputs. Either way, this should not invalidate the assimilation of the retrieved SM observations.

- (3) *Facilitate the exploitation of SM satellite information by assimilating the NN-retrieved SM instead of using a “physical” forward model* - The assimilation community generally prefers to assimilate raw satellite observations instead of retrieved products. One of the reasons is that it is necessary to specify uncertainties of the data assimilated into the model. It is considered that errors on the satellite observations are easier to quantify than errors in retrieved products. However, errors in satellite observations are not trivial, they must include instrument errors (often considered constant in the microwave) but also RT errors and other errors of representation which are much more complex [8, 38]. Furthermore, “physical” forward models largely rely on empirical parametrizations that would be more accurate if they were based on NN approaches rather than simple linear regressions. For observations related to the atmosphere, RT errors are easier to specify (at least in clear sky), but when dealing with surface parameters, RT errors are large, highly state-dependent, and difficult to estimate [19]. A characterisation of retrieved parameters is comparatively not so difficult because they are statistical in nature. LSM assimilation of retrieved

surface products instead of the raw observations has now started [57] and it is a good alternative to direct assimilation of the radiances while no reliable-enough RT model is available.

- (4) *Use the synergy of multiple observations, in particular the passive and active microwave observations* - The use of multiple satellite data to perform a retrieval was not common practice in the past, and it is still not widespread for several reasons. For instance, assimilation centres prefer to use the observations of satellites in an independent way, for commodity and practical issues. Project management in space agencies is also a limitation for the use of several instruments. The European Space Agency (ESA) has recently implemented the approach where two independent datasets of SM are merged *a posteriori* for the ESA CCI soil moisture [22] from passive and active microwave observations. Using the multi-instrument/multi-wavelength synergy by combining *a priori* the observations in the retrieval should, however, be more optimal than the *a posteriori* merging of retrieved products [1]. In fact, NN are very good tools to exploit the synergy of several instruments, in particular for SM [42]. Combining passive and active microwave observations is now in evidence (e.g. [43]).

These four propositions in [2] are still valid because the difficulties listed then (i.e., lack of a reliable RT model or large-scale empirical SM database) have not yet been solved. The NN approach is now being accepted and developed in agencies such as CNES (Centre National d'Etudes Spatiales), ESA (European Space Agency), NASA (National Aeronautics and Space Administration) or ECMWF (European Centre for Medium-Range Weather Forecasts).

Note that neural network and physical inversion scheme should not be opposed and could be combined. For instance, a retrieval scheme needs to be available soon enough after the launch of a new satellite. So a physical iterative inversion scheme might be necessary, or a radiative transfer model should be used to build a synthetic database to train a neural network inversion. After the acquisition of about one year of data, real observations should be enough to train a NN model.

Chapter 3

Datasets

In this paper we use ASCAT level 2 data and the latest ECMWF reanalysis ERA5 data of soil temperature, leaf area index, SM, and two-meter air temperature diurnal cycle amplitude [35]. The study period is spanning several years, from 1 January 2016 to 31 December 2019. Both ERA5 and ASCAT observations are projected into a common 0.25° -resolution grid.

3.1 ERA5 database

3.1.1 Soil Temperature (ST)

ST is provided by ERA5 in four vertical levels: 0-7cm; 7-28cm; 28-100cm; and 100-289cm. We will use the value of the upper layer to be more closely linked to the SM remotely-sensed by the ASCAT measurements. Fig. 3.1 shows an example of a ERA5 ST layer 1 map for June 11th 2016. Only some ascending and descending orbits are represented due to the coincidence with all the data sources. It can be seen that ST can facilitate the identification of surface types (vegetated areas, arid regions, latitudinal location, *etc.*) so it can potentially provide information on the presence of water in the soil.

3.1.2 Leaf Area Index (LAI)

Vegetation information has an impact of the backscatter coefficient of the ASCAT measurements. So it is important to consider it especially at the global scale, but also at the local level when vegetation has a seasonality [51]. Several variables describing vegetation could be considered such as the LAI or the Normalised Difference Vegetation Index (NDVI). The LAI is privileged here because it is more easily expressed in the model. Satellite datasets could be used here to document the vegetation (e.g. [29, 50, 34]),

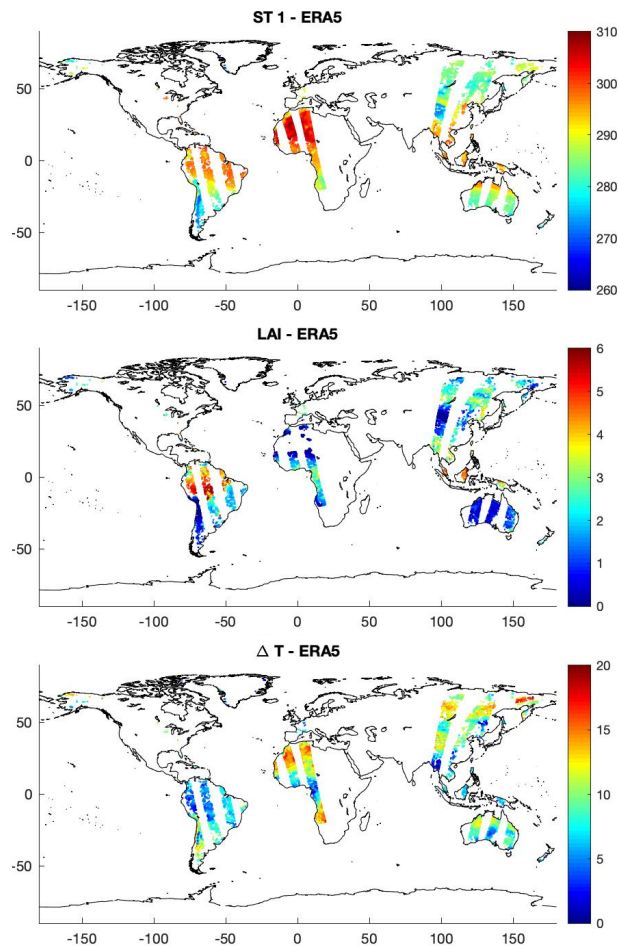


Figure 3.1: June 11th 2016, maps of auxiliary information from ERA5. From top to bottom: Soil Temperature at layer 1 (ST 1), Leaf Area Index (LAI), and amplitude of diurnal cycle of Temperature (ΔT).

as ancillary information. Since the objective is in this study to exploit the ASCAT SM information in the ECMWF LSM assimilation scheme, the ERA5 LAI is chosen. An example of ERA5 LAI is also presented in Fig. 3.1.

3.1.3 Amplitude of the diurnal cycle of surface temperature ΔT

As shown in [51], the amplitude of the diurnal cycle of temperature is strongly related to the SM and hence, can be used to quantify the presence of water in the soil. The higher the SM is, the higher the thermal inertia becomes, implying a lower ΔT . See Fig. 3.1 for an example of ERA5 ΔT on June 11th 2016.

3.1.4 Soil Moisture

SM is provided in ERA5 for several layers: levels 1 (0-7 cm); 2 (7-28 cm); 3 (28-100 cm); and 4 (100-289 cm). In the following, we will focus in the layer 1 since this is the level that ASCAT observations are most sensitive to. SM is provided as volumetric soil moisture in m^3/m^3 . Since SM is the objective in this study, this will be a strategic variable for the calibration of the retrieval schemes and for the evaluation of the SM information content. Fig. 3.2(top/left) represents an example of SM provided by ERA5, expressed in m^3/m^3 , for June 11th 2016.

3.2 ASCAT information

ASCAT is a C-band (5.255 GHz) scatterometer using vertically polarised antennas [26]. Ground echoes are received by the instrument and, after de-chirping, the backscattered signal is spectrally analysed and detected. On board ASCAT, two sets of three antennas measure the resultant electromagnetic backscatter, in two 500 km wide swaths, on each side of the satellite ground track. ASCAT instruments are onboard the Metop-A (launched on October 19th 2006), Metop-B (launched on 17th September 2012), and Metop-C (launched on November 7th 2018) satellites. The Metop series is constituted of polar orbiting meteorological satellites that form the space segment component of the EUMETSAT Polar System. Metop satellites are in a lower polar orbit, at an altitude of 817 km, to provide observations of the global atmosphere, oceans and continents.

ASCAT has been designed primarily to measure wind speed and direction over the oceans. However, ASCAT has also proved to be very useful for the

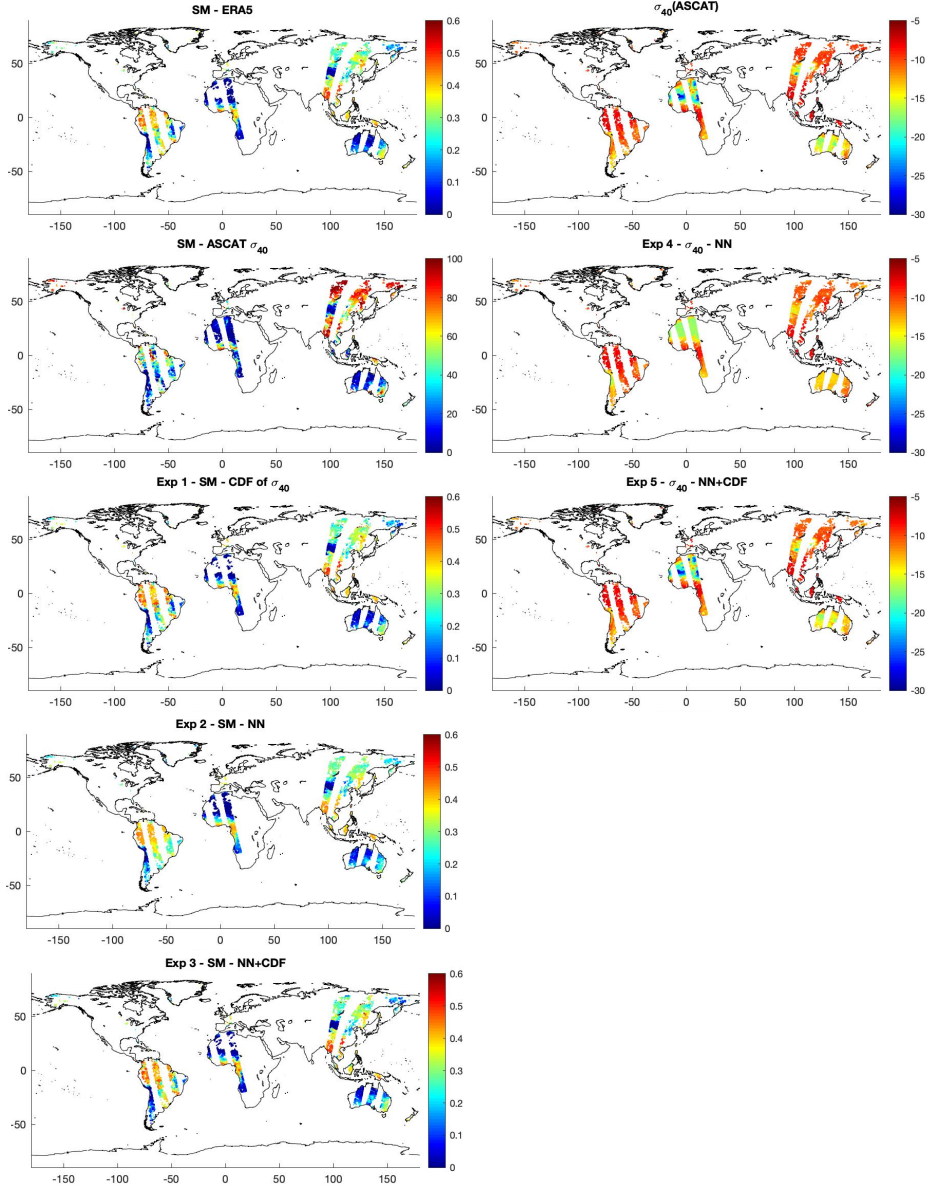


Figure 3.2: June 11th 2016. Left: from top to bottom, SM from ERA5, retrieval from ASCAT retrieval, CDF-matching of ASCAT σ_{40} (Experiment 1), NN retrieval (Experiment 2), and NN+CDF retrieval (Experiment 3). Right: from top to bottom, σ_{40} from ASCAT, from NN direct model (Experiment 4) σ_{40} , and NN+CDF direct model (Experiment 5)

retrieval of several other geophysical variables. In fact, the backscattering coefficient σ , measured with scatterometers, is dependent on the dielectric properties of the soil surface layer, surface roughness, and vegetation. In particular, increasing the fraction of water contained in soil increases the dielectric properties thus ASCAT measurements can be used to retrieve SM. With the rapid global coverage, day or night and all-weather operation, ASCAT offers a unique tool for long-term climate studies.

One huge advantage of active measurements (compared to passive microwave observations) is their spatial resolution, suitable for agricultural or small-scale hydrology applications e.g., [53]. The major issue when using this type of data is the contaminating effect of other parameters such as vegetation or roughness. For instance, Woodhouse and Hoekman [68, 69] estimated SM with an inversion method that incorporates both a model and *a priori* quantitative information, at the global scale. However, modelling such complex signals are very difficult so another strategy has been introduced to derive a SM index at the pixel level to produce a global SM database [66, 67]. This localisation of the retrieval allows for the suppression of the spatial variabilities in the contaminating effect.

The local incidence angle of the ASCAT measurements varies. In order to obtain a dataset in which the samples are comparable to each other, a linear regression has been performed to express the measured backscatter as a function of the incidence angle for each pixel. This regression function was then used to interpolate the measurements to a common incidence angle of 40° in a manner similar to the one proposed by [66], reducing the influence of the vegetation cover on the signal. This harmonisation of the ASCAT measurements are done in a very simple way (i.e. linear regression) and when using these transformed backscatters, the underlying hypothesis is that vegetation has no impact or remains constant along the time, which is not always satisfied. The transformed ASCAT data will be denoted by σ_{40} [66] in the following sections of this paper. Fig. 3.2(top/right) represents an example of σ_{40} for June 11th 2016.

This σ_{40} has been used to retrieve soil moisture. [67] define a Soil Wetness Index (SWI) between 0 and 100 by normalising observations of σ_{40} using their minimal and maximal values (estimated over a long time series): a linear transformation is used to set the historical extremes of $f(p)$ to 0 and 100. This type of approach is performed on a pixel-by-pixel basis since a dedicated retrieval is defined for each location. The SWI needs to be transformed into a SM using a reference, for instance from the ECMWF model: The $SWI = 0$ is associated to SM_{min} and $SWI=100$ is associated to SM_{max} , both SM values provided by the model. So in fact, the retrieval is calibrated on the model, at a pixel scale. The quality of the retrieval is based on several aspects: for

instance, a long time series of observations is needed to capture the extremes and a good estimation of SM should be provided by the reference model.

3.3 Preliminary analysis

A preliminary analysis of the ERA5 and ASCAT data is conducted at global scale, for the period 2016-2019, to assess their large-scale spatial and temporal variabilities. All the collocated samples of the database are gathered, whatever the location and the time, and then statistics can be performed to analyse the links among the variables. Fig. 3.3 represents the correlation

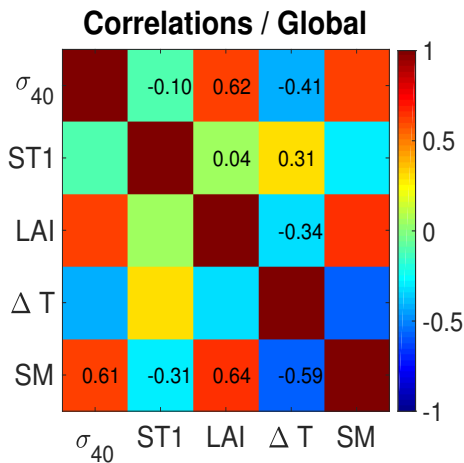


Figure 3.3: Correlation matrix between the variables present in the database, at the global scale, and for multiple years (January 2016 to December 2019).

matrix between the variables gathered in the database, at the global scale and for several years. The bottom row represents the correlations of the other variables with the ERA SM, so it provides an indication of the SM information that can be obtained from the ASCAT and the auxiliary variables. The σ_{40} backscatter from ASCAT is correlated at the 0.61 level with SM. ERA5 ST 1 is only negatively correlated at -0.31. The LAI vegetation variable is highly correlated with SM (0.64). The amplitude of the diurnal cycle of temperature, at the -0.59 level, is also a very good information for the SM. These numbers are coherent with the results in [51] and [42].

Since these correlations integrate all the locations, for all the seasons, this is only a very general measure. Correlations could be obtained at the pixel scale, sometimes with higher correlations. It would be misleading to directly compare the correlations in Fig. 3.3 and temporal correlations at

the pixel level because these two correlations measure different relationships. The correlations of Fig. 3.3 integrate the changes in σ_{40} that are related to other factors than the SM such as for instance the vegetation or the roughness of local conditions. The ASCAT σ_{40} is as sensitive to vegetation than to SM, meaning that if used at the global scale, both SM and vegetation information need to be delineated. These contaminating factors reduce the σ_{40} -to-SM link at the global scale. This is why when considering global relationships, it is necessary to use auxiliary (i.e. from ERA5) and even ancillary (i.e. independent from ERA5) information. Therefore, a choice needs to be made between global and local relationships, or a way to combine both approaches is necessary if we want to improve both aspects.

Chapter 4

Methodology for the exploitation of soil moisture information

CDF-matching is a pixel-based and simple model used to rescale satellite data towards the model, prior to assimilation. Several CDF-matching approaches will be presented here, including new ones. CDF-matching can also be used as a retrieval scheme, for simple cases like the retrieval of SM using scatterometer data. On the other side, NNs aim at a more complex, multi-parameter global relationship between the satellite observations and the SM. These two approaches have not the same objective but they share several features. Their similarities/differences and advantages/drawbacks will be discussed and compared. The possibility of combining them will also be presented.

4.1 CDF-matching

CDF-matching is a general and simple idea where the CDF of one variable p is transformed using a function f to mimic that of another variable t . However, this can be done in various ways, and for multiple objectives.

4.1.1 Methodology

Several versions of the transformation f can be used for the CDF-matching:

Gaussian matching - The first approach is the classical CDF-matching technique used in NWP centres such as at ECMWF. The objective is to

transform p (satellite observation) so that the CDF of $f(p)$ is similar to the CDF of t , under the Gaussian hypothesis. In this case, prediction p and target t are of similar nature (e.g. SM information). Let μ_p and σ_p be the mean and standard deviation of p ; and μ_t and σ_t be the mean and standard deviation of the target t . If the p and t distributions are Gaussian, the objective is to get the mean and standard deviation of $f(p)$ closer to those of t . The linear transformation f is then chosen following:

$$f(p) = A + B \cdot p, \text{ with } \begin{cases} B = \frac{\sigma_t}{\sigma_p} \\ A = \mu_t - B \cdot \mu_p \end{cases} \quad (4.1)$$

This transformation f offers a perfect bias-correction and normalisation of p towards t , but the CDFs will agree only if the PDFs of p and t are Gaussian. This means that the CDF-matching will not be optimal if, for instance, there are more extremes than with a Gaussian distribution (i.e. positive kurtosis), or if the distributions are not symmetric (positive or negative skewness) which is often the case with variables ranging between fixed thresholds such as with $0 \leq \text{SM} \leq 1$. Linear functions do not improve correlation, so $\text{COR}(p,t) = \text{COR}(f(p),t)$. Actually, the relationship in Eq. (4.1) is performed at ECMWF at the monthly scale (with a 3-months moving window): This can then improve the correlation at the seasonal scale: $\text{COR}(p,t) \leq \text{COR}(f(p),t)$.

Binning approach - Another method focuses on an empirical representation of the CDF rather than on the Gaussian assumption used in the first approach. A binning is first performed (e.g. regular or by quantile) on both t and p variables. The transformation f is then represented by the association of each ordered bin: e.g. each data in the first quantile of p is associated to the average of the first quantile of t . This approach should preserve better the extreme values of t and this technique can handle non-symmetric PDFs because no assumption is made on the distributions of p or t . The obtained transformation $f(p)$ provides discretised values, but the discretisation resolution depends only on the chosen binning; and continuous functions (splines, polynomial) can be used to link the bins if continuity is an issue.

Function approximation - A function approximation can be used to obtain a relationship f from p to t . What is new in this approach is that the observed sample-to-sample link between p and t in the reference dataset is actually used, contrarily to the first two CDF-matching approaches that work independently on the p and t spaces to estimate the CDFs. This link can be interesting to exploit: It will be seen in the following that this is what the NN methodology does, as any machine learning technique. Another advantage is that a continuous transformation is obtained, and it is possible to control the

complexity of f based on physical constraints (monotonic function or not, smoothness of the physical link, *etc.*). As for any function approximation, it is not entirely sure that extreme events can be well preserved/represented, this needs to be checked for each particular case. Non-symmetric and non-Gaussian distributions can be processed if f is chosen to be non-linear. Since this method is richer (i.e. more freedom in the choice and complexity of the transformation f), this method is more complex to control and subject to instabilities, but constraints can be used to stabilise it.

Approach	Method	Matching object	Database	Auxil. info.	Link $x/y?$	Extremes	Linearity	Improv. corr.?
CDF-matching	Gaussian match	CDF mean & std	Pixel	No	No	Not really	Yes ⁽¹⁾	No ⁽¹⁾
	Binning	Empirical CDF			No	Yes	No	Yes
	Fct. approx.	Linear relationship			Yes	Not particularly	Yes	No
		Nonlinear relation.			Yes	Depend	No	Yes
Neural Network	MLP	Multivar. relation.	Global	Yes	Yes	Not particularly	No	Yes

Table 4.1: Feature comparison of the CDF-matching and NN approaches for function $x \rightarrow y$. The sign ⁽¹⁾ means “except if seasonal”.

Table 4.1 compares several features of these three CDF-matching approaches. One way to chose the right technique is to monitor the scatterplot between p and t . The symmetry of the PDFs, proportion of extreme cases, or the non-linearity of the relationship between the two variables should determine which method to use. Practical tests should also be performed to analyse the spatial pattern of the diagnostics. The transformation f can be linear or non-linear. A linear transformation can be good enough if the p -to- t link is simple enough at the pixel scale. In the case of a linear f , the correlation is not changed ($\text{Cor}(p,t) = \text{Cor}(f(p), t)$). If the scatter plots between p and t are complex, a nonlinear f can be chosen, and this can potentially improve the correlation: $\text{Cor}(p, t) \leq \text{Cor}(f(p), t)$.

The CDF-matching works at the pixel level to exploit mainly the temporal information of the satellite observation. This means that spatial patterns in the satellite observations can be transformed toward the LSM. The objective here is to use the dynamics of the satellite observation, and less its spatial pattern.

4.1.2 Rescaling tool

Let p be a prediction (e.g. a satellite SM retrieval) and t be the corresponding target retrieval (e.g. the reference SM). A “rescaling” technique intends to find a transformation $f(\cdot)$ of the satellite estimates so that $f(p)$ is “closer” to the model variable t . For instance, the goal is to perform a bias-correction and/or transform the range of variability. The most widely used rescaling technique in the context of assimilation is the CDF-matching. This method

is often used *prior* to the assimilation of the satellite information into LSMs. The general objective is to find a transformation f so that the $\text{CDF}(f(p))$ is similar to $\text{CDF}(t)$. This transformation is done at the pixel level. The fact that averages and ranges of values are put in agreement facilitates the assimilation of $f(p)$ by avoiding jumps in the model variables. In general, the CDF-matching does not change the ordering of the satellite data at the pixel scale, but only their dynamics. This is true if the CDF-matching transformation is strictly increasing which is the case in general. At ECMWF, the CDF matching performed using a different transformation for each month so the annual ordering could potentially be changed in this case. But the idea behind the CDF matching is really to not change the ordering of states locally at the pixel level. Temporal correlations can actually be changed after the CDF-matching, but mostly if we use a nonlinear transformation. However, it is important to note that the spatial pattern of the satellite information can be transformed, towards the spatial pattern of the model, because spatial ordering of the satellite data is not preserved. This can truly distort the spatial information from the satellite towards the patterns of the model.

Fig. 4.1 illustrates the CDF-matching for two SM examples and two σ_{40} examples. The first column represents the scatter plots between the estimate p (x-axis) and the target data t (y-axis). The agreement between the σ_{40} is not exceptional for these two locations (correlations of 0.53 and 0.62) probably because other surface parameters such as vegetation impact the backscatter. Correlations for the SM is higher (0.73 and 0.71). The goal of a CDF matching should be to transform the x-axis so that the scatter plot becomes as close as possible to the diagonal line. This figure illustrates the CDF-matching by function approximation: in the second column, linear and 3rd-order polynomial fits are used to find the optimal link between p and t . When applied to the estimate p , these fits f transform it towards the target t : in the third column, the original scatter plot in blue is translated toward the diagonal line for both CDF-matchings. As a consequence, the PDFs get closer (from blue to green) to the reference PDF (in red), in the fourth column, although the matching is not perfect as the quality of the original scatter plots determines the quality of the fitting that can be obtained by the CDF-matching procedure.

The last column in Fig. 4.1 represents the time series of the target, original and CDF-matched variables. The blue time series (prediction) is well translated in green to the black one (target). The linear fitting does not impact the dynamics of the variable, the non-linear fitting can potentially change part of this dynamic. This can be seen on the correlations before and after the CDF-matching, given for the four samples of Fig. 4.1 in Table 4.2. As expected, the non-seasonal linear CDF-matching has no impact on the

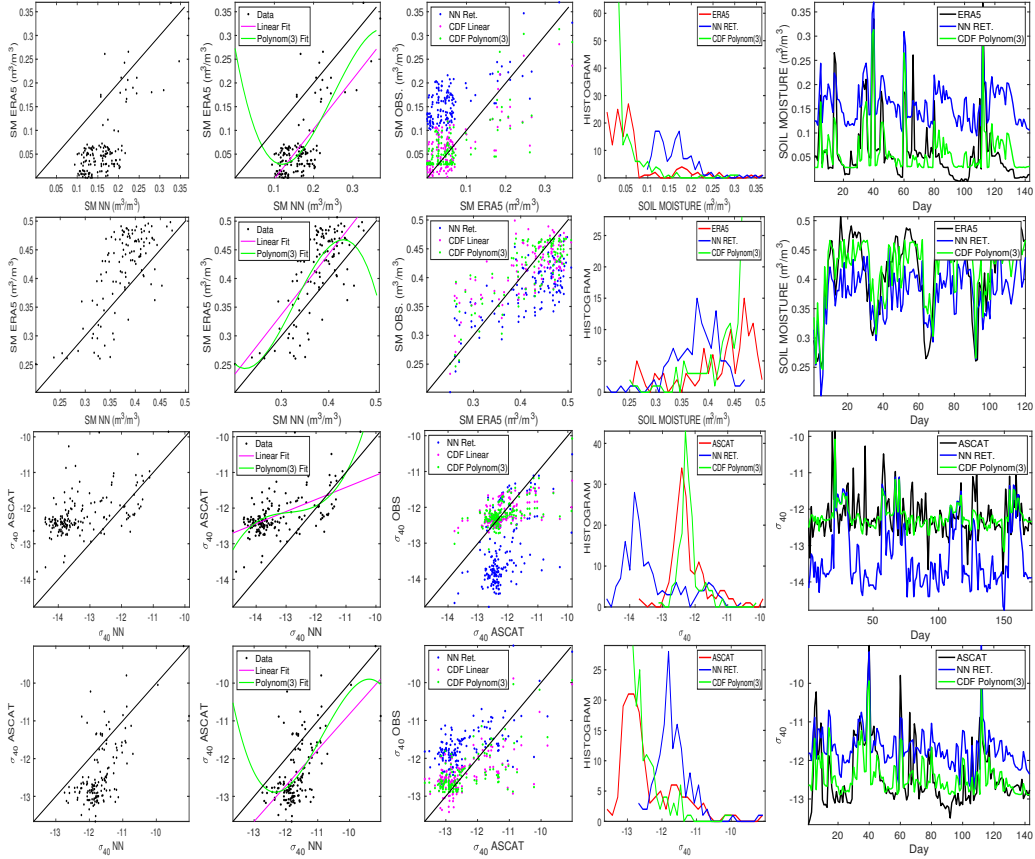


Figure 4.1: CDF-matching examples over two pixels for SM (two top lines), and σ_{40} (two bottom lines). From left to right: (1) The scatter plot of the NN estimate versus the target data; (2) Same but with a linear regression and 3rd-order polynomial fits; (3) Scatterplots of the targets with respect to the satellite observation (NN estimate, and CDF-matched with linear and polynomial fit); (4) The PDFs of the NN estimate and the corresponding CDF-matching; (5) The time series of the target, NN estimate, and CDF-matching (3rd-order polynomial fit).

correlation, it is just a rescaling technique, like the Gaussian CDF-matching approach. The non-linear CDF-matching on the contrary can increase the correlation, as expected.

Variable	COR(p,t)	COR(Lin(p), t)	COR($(P_3(p),t)$)
SM	0.729	0.729	0.763
SM	0.709	0.709	0.752
σ_{40}	0.533	0.533	0.580
σ_{40}	0.619	0.619	0.648

Table 4.2: The correlations before and after the CDF-matching (linear and 3rd-order polynomial fits) for the four samples of Fig. 4.1.

The CDF-matching by function approximation could be criticised because it can significantly change the satellite observations towards the target. It was commented earlier that this was a criticism for NNs trained on model data. However, the CDF-matching is a simple, mono-variate function, often monotonic (i.e. mostly an increasing function for CDF-matching). The CDF-matching does not change the nature of the satellite observations, except in the case of a negative correlation between p and t : with the Gaussian and binning techniques, the sign of the correlation cannot be changed, but the function approximation approach can invert the correlation sign by using a decreasing function f . However, if the two variables p and t (supposed to represent the same variable for a CDF-matching) are negatively correlated in some locations, it means that the satellite retrieval p is of very poor quality and should not be CDF-matched for assimilation.

CDF-matching is generally done at the pixel level to exploit the temporal information of the satellite data. Actually, the CDF-matching could also be done by gathering all the data for all the pixels, so that the CDF-matching is performed on both the temporal and the spatial variations. This would mean that the overall CDFs of the two datasets would be put in agreement (e.g. same global mean, same extremes) but this would not modify the spatial pattern of the satellite observations towards the model. In this way, the CDF-matching retrieval scheme could be more directly compared to the NN retrieval that will be presented in the following section.

4.1.3 Retrieval scheme

[67] define a Soil Wetness Index (SWI) between 0 and 100 by normalising observations of σ_{40} using their minimal and maximal values (estimated over a long time series): a linear transformation is used to set the historical extremes

of $f(p)$ to 0 and 100. The rescaling by extreme values is not based on the full CDF information so the $\text{CDF}(f(p))$ can differ from $\text{CDF}(t)$. In order to be exploitable for assimilation, this index is then generally CDF-matched to the model SM so that the SWI are not between 0 and 100 but rather between the minimal and maximal value of t (in each pixel). A direct and unique rescaling by extreme can be done directly on the σ_{40} to the minimal and maximal values of m . These two approaches are mathematically equivalent and the simpler approach should be preferred.

When the σ_{40}/SM relationship is simple enough, the CDF-matching can even be used as a simple retrieval scheme. When using such an approach, the underlying hypothesis is that the ASCAT backscatter σ_{40} might be related to the SM and other parameters; but when considered locally in one pixel, the σ_{40} is only dependent on the SM (the other variables do not vary). In this way, the link between the satellite observation σ_{40} and the variable SM becomes simple enough to be represented by a mono-variate relationship. This might be true for variables constant in time such as soil texture, but it is not valid everywhere for variables such as vegetation¹. So this approach can appear simplistic, but it might be adequate in some locations, and it is at the basis of the retrieval schemes developed for ASCAT retrieval of SM.

The retrieval here is based on a target t variable from the model. If a global dataset of real observations is available, the retrieval could be calibrated on it. But the training using a model reference is an interesting approach potentially helping the assimilation, as long as the temporal variability of the satellite observation is preserved in the $f(p)$. The retrieval is performed at the same time as the CDF-matching is done.

As for the CDF-matching, the function approximation can use a decreasing or increasing function f to change the sign of the correlation. Again, if the correlation between p and t has not always the same sign at the global scale, a problem in p is probable. For instance, it can mean that an auxiliary information is truly necessary to describe a more complex link between the two variables, and some additional work would be required so that the correlations have the same sign in all locations.

¹Remember that the ASCAT backscatter is estimated at 40° incidence angle because it is less sensitive to vegetation in this case, but this is not a perfect approximation of course

4.2 Neural Networks

4.2.1 Model

Among the NN techniques that can be used to perform satellite remote sensing, the MLP (Multi-Layered Perceptron) [60] is certainly the most widely used. We will use this model but other neural architectures could be chosen instead, such as deep learning networks. A MLP is a non-linear mapping model: Given an input X , it provides an output Y in a non-linear way. The neurons are organised in successive layers. Each layer in the NN is composed of individual neurons. A neuron performs first a weighted average of its inputs from the previous layer. The so-called synaptic weights are associated to each connection between two neurons. Every neuron has a connection with all the neurons of the previous layer for a fully connected network. Once this weighted average is done, a non-linear sigmoid function g is applied, the Tanh function is used here. The final output of a neuron i is then given by:

$$y_i = g \left(\sum_{j=1}^N w_{ji} x_j \right) \quad (4.2)$$

where $(x_j; j = 1, \dots, N)$ are the N inputs of neuron i , and w_{ji} is the synaptic weight between neuron j and i [13]. The MLP model is defined by the number of input neurons (i.e., the size of the inputs), the number of outputs, and the number of neurons in the hidden layers that control the complexity of the model. A study has to be conducted to define the optimal number of hidden layer(s) and the number of neurons in the hidden layer(s). A balance needs to be found: Too many free parameters in the model can result in over-learning (over-parameterisation) leading to degraded generalisation properties. On the contrary, too few free parameters will yield under-parameterisation and bias errors in the model. Provided that enough samples are available to train the NN, any continuous relationship, as complex as it is, can be represented by a MLP [37]. Furthermore, in [17], a theorem shows that a two hidden layer NN is able to represent any discontinuous function which can be an important feature for inverse problems.

The estimation of the optimal parameters of the NN is performed during the training stage. For that purpose, a loss function is first defined (the least-squares is used here as done usually). Then, a gradient descent algorithm allows to find optimal parameters so that the NN estimates are as close as possible to the actual outputs provided in the learning database.

4.2.2 NN database

Once the NN architecture is defined, a training dataset including input/output couples is required to determine the weights W in the model. A quality criterion is defined to specify how well the NN reproduces this training dataset (e.g. least squares). An optimisation algorithm improves this quality criterion (by reducing the errors) in order to find the optimal parameters W . The optimisation algorithm is the back-propagation algorithm from [60]. It is a stochastic gradient descent that has been optimised to the hierarchical structure of the NN to speed up the learning stage.

Learning databases to train NN are generally built in two ways. First, a collection of states (e.g. SM plus other variables) are used as inputs for a RT model to obtain associated simulated satellite observations (e.g. σ_{40}). We call this approach “physical” as it uses a physical model to construct the database. As mentioned earlier, such forward RT models for surface properties are not yet fully satisfactory. Second, a collection of true satellite observations is put in coincidence with a large and representative collection of true SM *in situ* measurements. Again, this approach is not possible in our case because such collection of SM measurements is not available.

It was proposed in [2] to use real satellite observations (σ_{40}) with SM from a reanalysis (such as SM from ERA5). Over the whole database constructed to define the NN, representing more than 1 million points, 60% are kept for the learning, 20% for the testing, and 20% for the validation. These three datasets have been selected randomly. Since the number of samples in each one is very large, their statistical characteristics are similar. (1) The training of the NN is performed on the learning database. (2) The testing database is used in parallel to the training process: at each learning step, the learning database is used to improve the parameters of the NN model. The testing database is not directly used to change the NN parameters, but it is used in each step of the learning to monitor the results of the NN in an independent database. This allows measuring the generalisation capacity of the NN, i.e., its ability to perform retrievals on new data. During the learning, the generalisation errors are monitored and the learning is stopped when, after a decrease, they start to increase. This procedure avoids the over-training of the NN, i.e., the problem in which an algorithm performs very well on its learning data set but poorly on new data. (3) The testing database is used several times, first to measure the generalisation capacities of the NN on each step of the learning process, second for all the tested NN configurations. As a consequence, the whole NN selection process could “learn” the testing database (i.e., be biased toward it) and the evaluation of the generalisation errors could become misleading. To avoid this problem, another independent

database is used: The validation database. It is only used to estimate the retrieval errors on an independent data, once the learning is complete, and once the model is chosen.

There are several advantages of the NN training of SM-retrieval based on reanalysis data. For instance, the obtained SM is “compatible” with the SM of the reanalysis model: the overall histogram of SM at the global scale should be relatively close to that of the model. This means that definition of SM and its units are compatible. Contrarily to the CDF-matching done at the pixel-level, this is only true at the global scale. It means that systematic biases can be present regionally: (1) The satellite observations might not have enough information to change their estimate at the regional scale and the model has more information. In this case, some auxiliary information could be added to the NN retrieval to compensate for this. (2) The model can also be wrong, and the retrieval which is based on the observation does not agree with it. What is important is that the NN retrieval of SM should be based on the satellite information, because this is what we want to exploit in the assimilation.

Another advantage of learning on reanalysis data concerns the use of multiple instruments. Instruments can be different at the same time (like passive and active microwave observations), or they can be similar ones but consecutive in a long time-record. Having a same reference (i.e. the reanalysis) to train each retrieval is a true advantage for their inter-calibration.

4.3 Quality criteria

We want to compare several technical configurations using CDF-matching or NN to extract SM information from the ASCAT instrument. Some diagnostics are therefore needed to measure the quality of the proposed solutions. It is not easy to compare agreements in the satellite observation space (p) and on the model variable space (t). The following quality criteria measure some type of coherency between the satellite observations and the model. This is required to assimilate the satellite observations into the model. However, the objective cannot be to get too close to the model, otherwise there would be no additional information in the satellite observations. We will comment, in the discussion, on the spatial and temporal information included in the satellite observations and how they are preserved by the different approaches.

Three types of metrics will be considered in the following:

- The temporal correlation between the satellite observation and the model SM, on each pixel. This measure indicates if the dynamics in the satellite observation is coherent enough with the ERA5 data. As

mentioned in the introduction, the dynamics of satellite observations are the most important feature for assimilation so some calibration techniques are often used in NWP centres to suppress differences in absolute values but preserving the dynamical behaviour.

- The range of variability of the satellite observations compared to the model will also be assessed. The minimal, maximal and range of values will be compared between the model and the satellite observations, to make sure that variables are actually similar (several units can be considered for measuring SM), and that they comparable extreme distributions. Regression model predictions are always inside the range of variability of the dataset used to train them.
- Bias and standard deviation errors will also be measured between the model and the satellite observations. These diagnostics indicate differences between the two estimates not in relative terms (like the correlation) but in terms of absolute values.

Each time, the same sample dataset will be used for the satellite observations and the model so that if one data is missing for one, it will not be taken for the other. In this way, statistics can reliably be compared to each other.

NN are un-biased estimator model. This means that the average of the estimations is equal to the average of the targets present in the learning database, so there should be no biases between the target (model) and the estimated (NN) estimates. But this is true only at the global scale, when mixing all pixels and all time steps. Local biases can be encountered regionally due to very particular conditions and/or missing auxiliary information. CDF-matching, by definition, do not have biases at the pixel level because the calibration has been done pixel by pixel for that purpose.

As any other statistical models, NN will always tend to dampen the range of variability of the target dataset (model). This is due to the fact, again, that some information on extreme cases is missing so the statistical model provides “averaged” values for these extremes. The CDF-matching technique, in opposition, is by definition following the range of variability of the target dataset (model), minimal and maximal values will be respected, as long as the training dataset for the CDF-matching is large enough to include these extremes. However, it does not mean that extremes will be well retrieved, as extremes on both datasets (objective and estimation) can occur in totally different times of the dataset record.

Again, if the target (model) is wrong, it does not mean that the NN will reproduce these errors. In fact, the NN model tries to obtain global physical relationships that link the inputs and the output, at the pixel scale.

This means that no spatial or temporal error structures of the model would be reproduced by the NN because it actually has no information about the localisation of each pixel. The NN can be trained using the ERA5 model but if ERA5 is wrong, it will correct it because after the training stage, the spatial and temporal variability in the NN outputs are only driven by the observations, not the model. If a regional incoherency is present in the model, the NN would try to correct this incoherency between the regional and the global scales.

It is clear that each diagnostic provides different information on the similarities of the datasets (range of values), on the statistics (bias and STD errors) and on the agreement (correlation) between the datasets. Table 4.1, indicates the methods that can improve the correlation or not, or those that represent better the extremes.

Chapter 5

Experiments and results

Mode	Experiment number	Inputs	Model	Learning outputs	Inputs to assimilation
Inverse	1	ASCAT σ_{40}	CDF-Matching	ERA5 SM	SM
	2	ASCAT σ_{40} +ST1+LAI+ ΔT	NN		
	3	ASCAT σ_{40} +ST1+LAI+ ΔT	NN+CDF		
Forward	4	ERA5 SM +ST1+LAI+ ΔT	NN	ASCAT σ_{40}	σ_{40}
	5	ERA5 SM +ST1+LAI+ ΔT	NN+CDF		

Table 5.1: Description of Experiments 1, 2, 3, 4, and 5. SM is soil moisture.

In this paper, five experiments will be conducted. Three of them (Exp. 1, 2, 3) are designed to perform a retrieval, i.e. estimate SM based on the satellite information; and two (Exp. 4, 5) aim at representing a forward model, from the SM to the satellite observation (see Table 5.1). These inverse and forward models are investigated to help making the decision between the assimilation of retrieved SMs or direct observations σ_{40} . Experiments use the CDF-matching approach (Exp. 1), the NN method (Exp. 2, 4), or a combination of them (Exp. 3, 5). The scheme in Fig. 5.1 illustrates how these five experiments try to relate the two spaces: satellite observations and SM.

In the following, the computational coast will not be mentioned as both the CDF-matching and NN are very fast and negligible for any NWP center. The NN method is slightly faster than the CDF-matching, but the CDF-

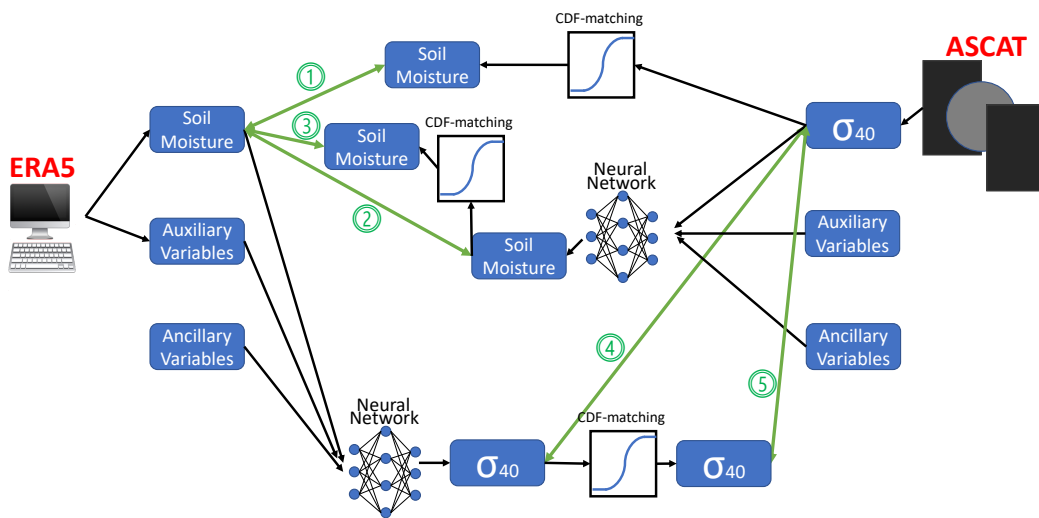


Figure 5.1: Scheme representing the five strategies (in green) to exploit the ASCAT information for SM. The green arrows represent the experiments with the input and output variables that are used in each model. Experiments 1, 2, and 3 use ASCAT data to estimate a SM compatible with ERA5, and experiments 4 and 5 intend to build a forward model using input ERA5 data to estimate a backscatter σ_{40} compatible with ASCAT.

matching speed depends on the used approach (Gaussian matching, binning approach, function approximation), see Section 4.1.1.

5.1 Retrieval schemes

Experiment 1 uses the retrieval scheme based on the CDF-matching technique (section 4.1.3). The CDF-matching here is not only a rescaling but an inversion process. This approach intends to find a direct and simple relationship (i.e. third-order polynomial function) between the ASCAT σ_{40} and the ERA5 SM. Only the ASCAT backscatter is used as an input here, but the inversion is performed at the pixel level to compensate for this lower information content compared to other retrieval schemes. Fig. 3.2(left column, third row) represents the SM retrieved using this scheme: the general spatial pattern is quite well represented by the CDF-matching retrieval when compared to the ERA5 SM.

Experiment 2 uses the NN approach presented in section 4.2. This is the NN solution introduced in [2], used later on in [42, 43] and [58]. The inputs for the retrieval are the ASCAT real observations σ_{40} plus the ERA auxiliary information (ST1, LAI and ΔT). The NN uses four inputs, ten neurons in the hidden layer, and one output coding the SM. Fig. 3.2 (left column, fourth row) represents a daily retrieval. The spatial pattern is very close to that of the ERA5 SM (top row). It can be noted that compared to the CDF-matching retrieval of Exp. 1, the extreme positive values (larger SM) are less well retrieved by the NN. This is to be expected because the NN uses a compromise between all the pixels at the global scale, and this means that extremes are reduced. Additional information would be required in order to better characterise these extremes everywhere.

This is exactly the objective of *Experiment 3* that combines the NN of Section 4.2 and the CDF-matching of section 4.1.2. As in Exp. 1, the extreme values are a better fit to ERA5 SM.

Fig. 5.2 represents on the left the correlations between the target (ERA5 SM for Exp. 1, 2, 3) and the estimation. For SM, the Exp. 1 CDF-matching of the σ_{40} appears to have overall the best correlations except in some regions (Europe, Northern part of South America, Eastern USA, mountains (Himalaya and the Andes) and part of the Sahara desert) where Exp. 3 can have higher correlations. Bias errors are equal to zero for the Exp. 1 and 3 that use CDF-matching. This is logical and we see that the purely NN retrieval of Exp. 2 that intends to find a global relationship has some limited regional biases, as mentioned in Section 4.2. The third column in Fig. 5.2 represents the Standard Deviation (STD) of the errors. It is noticeable that

Exp. 2 and 3 have almost the same STD errors (only the regional bias errors are corrected by the CDF-matching). The regions where Exp. 3 seems better than Exp. 1 are the same as for the correlation. The fact that the global NN performs better in these regions means that the auxiliary information (ST1, LAI and ΔT) is important for the SM retrieval and that σ_{40} alone is missing this information.

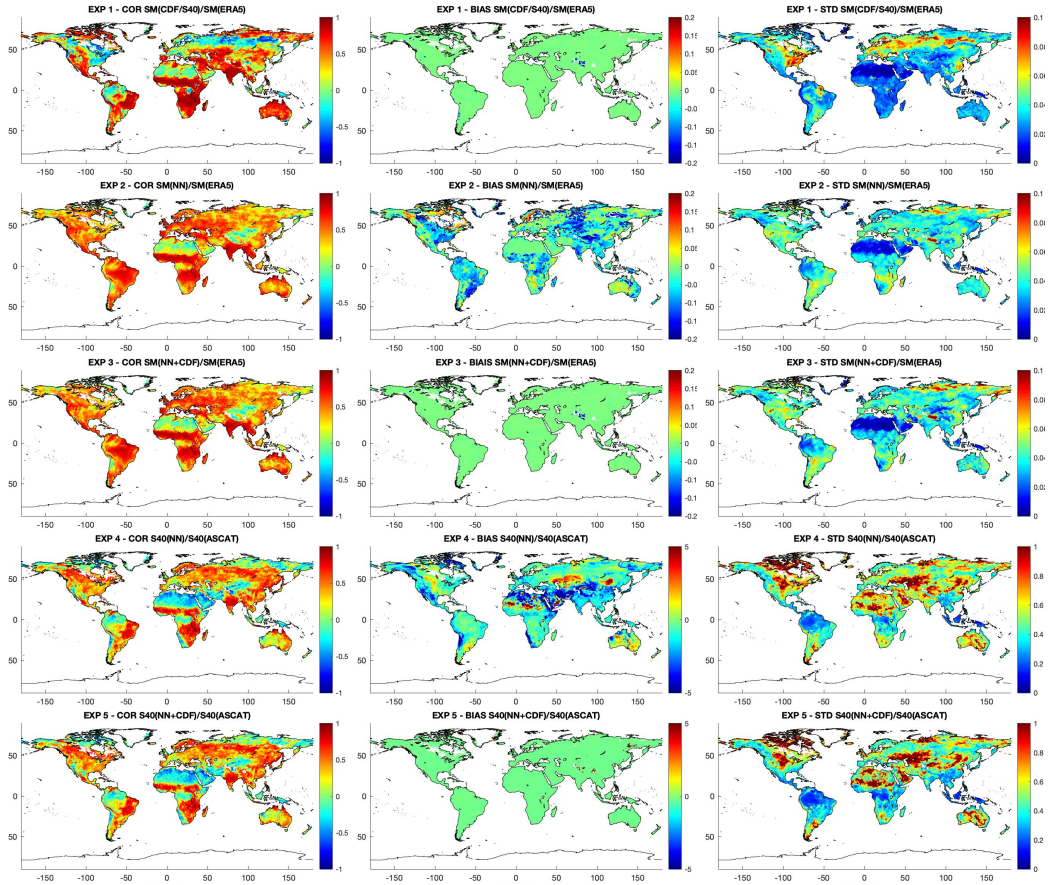


Figure 5.2: Correlations (left), bias (center) and STD (right) errors at the global scale, from top to bottom for Exp. 1, 2, 3, 4 and 5.

Fig. 5.3(third row) shows that the NN approach is able to extract information from the ASCAT observations and the auxiliary information to fairly well reproduce the extremes of the ERA5 SM. The fourth row shows that combining the NN and CDF-matching approaches enables the extremes of the ERA5 SM to be better captured. Using one or the other really depends on what we want to achieve with the retrievals. In the context of

weather forecasting in NWP, the CDF-matching is desirable to facilitate the assimilation.

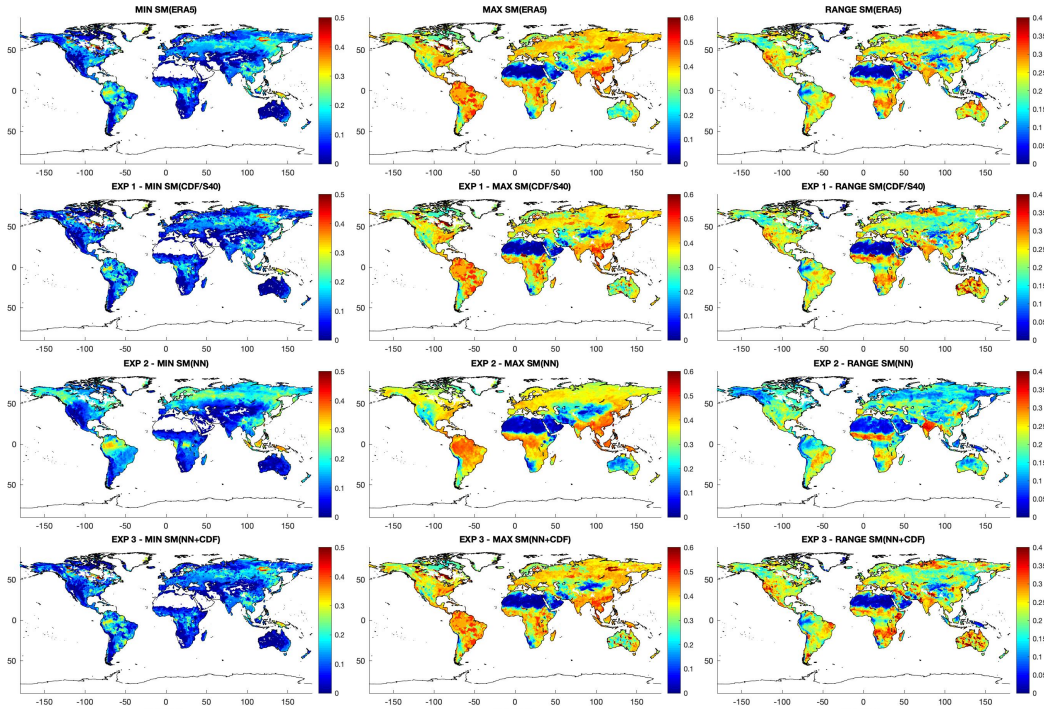


Figure 5.3: Minimum (left), maximum (center), and range (right) for SM. From to bottom for ERA5, Exp. 1, 2 and 3.

Annex 8 provides some statistics about minimal, maximal and range of variability for these three experiments.

5.2 Forward modeling

Physical [27] and empirical [10] RT forward models are available, but their accuracy is not entirely satisfactory yet and simulations still have significant errors compared to real observations, challenging the strategy of assimilating raw observations such as the ASCAT σ_{40} . One reason is that input parameters are not available at the global scale. Furthermore, processes are so complex that it is difficult to characterise them at the global scale. This explains why assimilation of retrieved SM has gained interest [57]. Is the NN ability to capture simple but general relationships at the global scale good enough to simulate with enough accuracy the ASCAT backscatter σ_{40} ?

Forward RT modelling using NN has been tried (e.g., [15]) but this approach has been rarely used in the past compared to the inversion. One explanation is that physical models are available in many applications. This is true for atmospheric variables for which RT models are more reliable, but less true for surface variables. In addition, it is generally admitted that numerous input parameters are necessary to simulate the satellite observations and some are very difficult to obtain at the global scale. Since the SM/ σ_{40} relationship is simple in the retrieval model, especially when localised at the pixel scale, it is worth investigating if a simple NN forward model could be constructed.

The two NN forward modelling schemes (experiments 4 and 5) are presented in Fig. 5.1 and Table 5.1). In order to simulate correctly ASCAT σ_{40} , several input parameters are required. Surface SM is of course the variable that we are interested in, but vegetation (LAI) is a variable highly impacting the ASCAT backscatter so it needs to be included. Soil Temperature (ST) and the amplitude of the diurnal cycle of temperature are also added. In addition to the NN model, experiment 5 also includes a CDF-matching.

Fig. 4.1 represents on the left the correlations between the target ASCAT σ_{40} and the estimation (Exp. 4 and 5). The overall structure is similar to the correlations in the SM space (Exps. 1, 2 and 3). However, the magnitude of the correlations are however lower. Some spatial features are noticeable such as a strong dipole in South America or negative correlations over the Sahara. As expected, the NN retrieval can be regionally biased (in the Sahara and middle East) but this problem is solved when adding the CDF-matching. The standard deviation is, however, not impacted by the CDF-matching.

Finally, when comparing to the error structures and correlation levels, the retrieval of the SM seems more reliable than the simulation of the forward model. To obtain better results, additional input information would probably be required, potentially all the geophysical parameters impacting the backscatter. However some of these parameters (structure of the vegetation, soil texture, etc.) can be very difficult to obtain at the global scale with enough accuracy. Note that the assimilation of the outputs of such forward models would require a transformation of these σ_{40} into the model state variable (i.e., SM), through the Jacobians of the RT, and this would add up even more errors. In addition, in the current methodology the observations are interpolated to model grid point locations which may favour the inverse methods. For conventional forward models the model values are usually interpolated to the observation locations before running the forward model. It would be interesting to investigate whether the interpolation method makes a difference to the performance of the forward NN model. This will be the subject of a future study.

5.3 Comparison of the CDF-matching techniques

In Section 4.1.1, several approaches were presented to perform CDF-matching. Fig. 5.4 represents the distributions of the correlations between the CDF-matching of ASCAT σ_{40} and ERA5 SM (TOP), the CDF-matching of the NN SM-retrieval and ERA5 SM (MIDDLE), and the CDF-matching of the NN-estimation backscatter and the ASCAT σ_{40} (BOTTOM). Three CDF-matching methods are compared, the Gaussian, binning and function approximation (i.e. third-order polynomial) approaches and the averaged correlations are provided. The Gaussian CDF-matching does not improve the correlation of the original data because it is a linear method. The binning and function approximation use a non-linear function so they can increase the correlation. The highest improvement is given when using the function approximation (averaged correlations from 0.53 to 0.65, from 0.57 to 0.61 and from 0.43 to 0.54). This improvement comes from the fact that negative correlations (present in some locations, see Fig. 5.2) can be transformed into positive ones using a negative function. As a consequence, no negative correlations can be seen for the function approximation.

5.4 Combining algorithms

A priori screening of the assimilated satellite data is already in place at ECMWF. Therefore, a dedicated screening could be developed specifically for retrieval SM. It is proposed here to combine Exps. 1 and 3 by defining in which pixels one is better than the other one. To illustrate the potential of this approach for the assimilation, Fig. 5.5 represents in the first maps the correlations of pixels where Exp. 1 is better, the second map when the Exp. 3 is better, and the second map when both are combined. It can be seen that this combination could be highly beneficial, in particular in half of South Africa, most of Europe, large regions over North America and South-East Asia. Implementation of such strategy should not be difficult, less for instance than the inclusion of a RT model in the assimilation system.

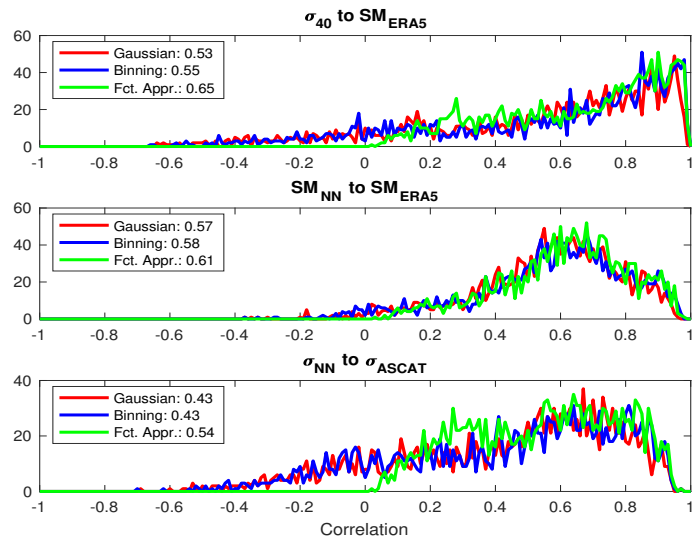


Figure 5.4: Distributions of the correlations between: the CDF-matching of ASCAT σ_{40} and ERA5 SM (TOP), the CDF-matching of the NN SM-retrieval and ERA5 SM (MIDDLE), and the CDF-matching of the NN-estimation backscatter and the ASCAT σ_{40} (BOTTOM). The CDF-matching is performed using the three approaches: Gaussian transformation, binning, and function approximation. The mean correlation is provided in the legend for each experiment.

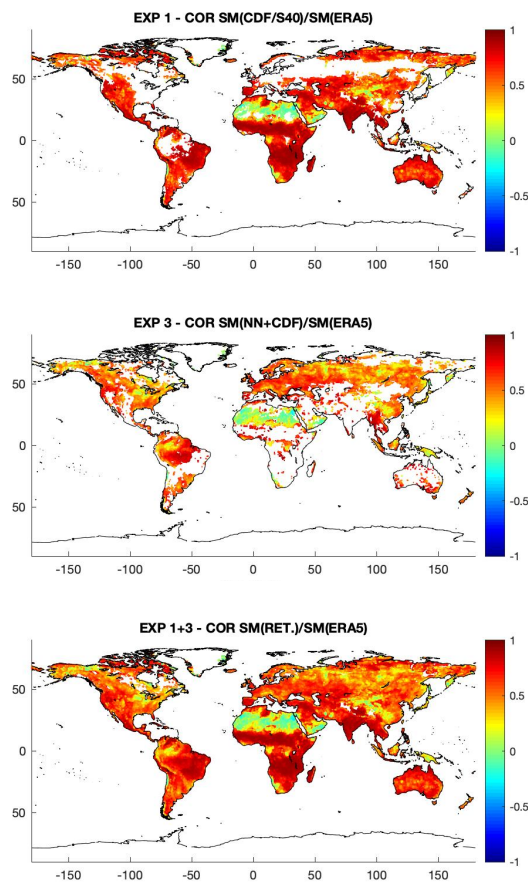


Figure 5.5: Illustration of the combination of SM retrievals. Top: correlation between retrieved and ERA4 SMs on pixels where Exp. 1 is better; Middle: Same for Exp. 3; Bottom: combination of these pixels.

Chapter 6

Evaluation

We will focus here on the USA because it is the region with most *in situ* observations for the validation of the soil moisture. It is also a contrasted region with better results over the west and results of lower quality in the east.

We have seen previously that the NN retrieval scheme could be combined with a CDF-matching to obtain lower local biases. Another “localization” approach will be introduced here.

6.1 Localization of the NN model

NN are statistical models that are trained over a large database of samples \mathcal{B} . The training consist in finding the parameters of the NN that allow to reach the best compromise, where the NN has lowest errors overall over \mathcal{B} . The NN is an un-biased estimator, but this is only true on the whole training database. It does not mean that local biases cannot be present in some part of the \mathcal{B} domain. For instance, the retrieval of soil moisture might systematically be under- or over-estimated in some complex areas. Often, these local biases are interpreted as a lack of examples in these regions but this is rarely the case and uniformizing the distribution of samples in \mathcal{B} to better represent these complex areas would not solve the problem.

The true cause of such local biases is the lack of information required to better describe the local conditions in these complex regions. Two situations can originate such difficulties. First, necessary missing data is lacking but this information might be difficult to obtain, in particular at a global scale. For instance, the vegetation state or rugosity at the surface has an impact on a scatterometer measurement so this information can be necessary. Second, the observations might just be inadequate for some conditions. For instance,

scatterometer measurements are inadequate for the retrieval of surface parameters over highly vegetated areas. Such lack of information can translate into local biases, but also random errors, or a false stable dynamics of the retrieved product.

We will refer in the following to the term “localization” when information is added to the NN model (this can be done in several ways) so that it changes its behaviour in different locations of the spatial domain. Two general forms of information can be exploited for the localization of the NN:

- When the localization information is a physical variable that better constrains the radiative transfert, then we expect to obtain a more general physical NN that has the potential to be used globally, which is very convenient in an assimilation context, but this is a more ambitious task. Localization means here that additional information is added to the NN model so that he better represent local specific conditions.
- When the localization information is an information to better isolate smaller regions from the global domain, with the intend to simplify the physics, we obtain a model that is more statistical in nature. The actual relationship between the observations and the geophysical variables is simplified because unknown impacting factors are not varying anymore.

The choice between these two approaches is dictated by the application. Using a purely statistical approach in a context of assimilation is totally acceptable. A parallel can be done with the CDF-matching that transforms in a purely statistical way the satellite observations at the pixel level (the highest localization resolution), before they are assimilated in the model.

Note that even if we try to introduce more physics in the NN by using pertinent RT variables, they might be used by the NN in a purely statistical way. This can be observed by inspecting the Jacobians of the obtained NN, these sensitivities inferred by the NN are often disappointing in terms of physics (but they can actually be improved using physical constrains [9]). The NN is after all a statistical model: the introduction of additional physical inputs might only be a way to reduce the complexity and obtain a more straightforward solution, locally, without trying to describe the complex dependency on auxiliary information. As a machine to find best compromises, the NN is often a good tool to obtain the first order relationship, but obtaining more subtle dependencies might require a dose of “regularisation” (by introducing physical constraints) [9]. The extreme physically based approach would be the construction of hybrid models where physics and neural networks are combined: The physical side provides some structural constraints (pertinent inputs, hierarchical dependencies, constraints on some parameters, etc.), and

the NN side is the use of the machine learning that infer the parameters of the physical relationships from a set of samples.

Some localisation strategy needs then to be used to reduce regional biases of the NN retrieval, in order to facilitate the assimilation. It would also be seen that localisation can actually improve also the dynamics of the retrieved product, and this is a very important feature for the assimilation because most of the impact of satellite observations in assimilation mode comes from this dynamic./

6.1.1 Dataset

Three years of data have been gathered for this experiment. The list of variables that have been selected is described in Table 6.1. All these variables have been obtained and put in coincidence at ECMWF.

Lat	°
Lon	°
Time	
σ_{b1}	
σ_{b2}	
σ_{b3}	
Inc_{ang1}	°
Inc_{ang2}	°
Inc_{ang3}	°
σ_{40}	
ERA5 SM Level 1 index	0-1
ERA5 SM Level 1	m3/m3
ERA5 SM Level 2	m3/m3
ERA5 SM Level 3	m3/m3
Land-sea mask	0-1
Soil Temp 1	K
Soil Temp 2	K
Soil Temp 3	K
LAI	0-6
Proxy NDVI	0-1
dT	K

Table 6.1: Variables of the database

Tab. 6.2 represents the averaged (among the pixels) correlations between the variables (and their anomaly) considered in this study. These values can indicate us what are the variables that can be used to retrieve SM for

	s40	SM	ST	NDVI	dT
s40	1.00	0.47	0.60	0.62	-0.36
	1.00	0.38	0.23	0.27	-0.22
SM	0.47	1.00	0.43	0.51	-0.47
	0.38	1.00	-0.42	-0.17	-0.38
ST	0.60	0.43	1.00	0.97	-0.22
	0.23	-0.42	1.00	0.61	0.16
NDVI	0.62	0.51	0.97	1.00	-0.24
	0.27	-0.17	0.61	1.00	0.07
dT	-0.36	-0.47	-0.22	-0.24	1.00
	-0.22	-0.38	0.16	0.07	1.00

Table 6.2: Averaged (among the pixels) correlations between the variables (and their anomaly) considered in this study

instance: the backscatter, vegetation and amplitude of the diurnal cycle are good sources of information for the raw SM time series. It means that they share similar seasonality and this can be exploited for the retrieval of SM. However, for obtaining more information than the seasonality, correlation among the time series anomalies need to be analysed too, and of course, their absolute values is decreased compared to the raw correlations, showing that this signal is more difficult to obtain.

Fig. 6.1 shows the PDF of the crossed correlations between the anomalies of the variables listed in Table 6.1. Some are negative, some are positive, with a reasonable behaviour (e.g. the higher the diurnal cycle of temperature, the less soil moisture is, resulting in a negative correlation). However, it can be seen that some of the correlations can be positive in some pixels and negative in other ones, and that the correlation values are spread. This is true for instance for the link between SM and NDVI. Clearly, the retrieval of SM using the other variables and observations is not a straightforward task because this link is dependent on the situation and can vary a lot, from one pixel to the next. The retrieval technique needs to be state-dependent and should try to resolve these complex links. This is a characteristics of NN that are non-linear by nature but still, it can be hard to catch this state-dependency if not enough information is provided. It shows the difference between wanting to build an overall general model (with a large variety of correlations to infer) and building a model at the pixel scale (where only one set of correlations might suffice). This is done by using some form of localization.

Two configurations will be tested here: (1) a global NN that will try to

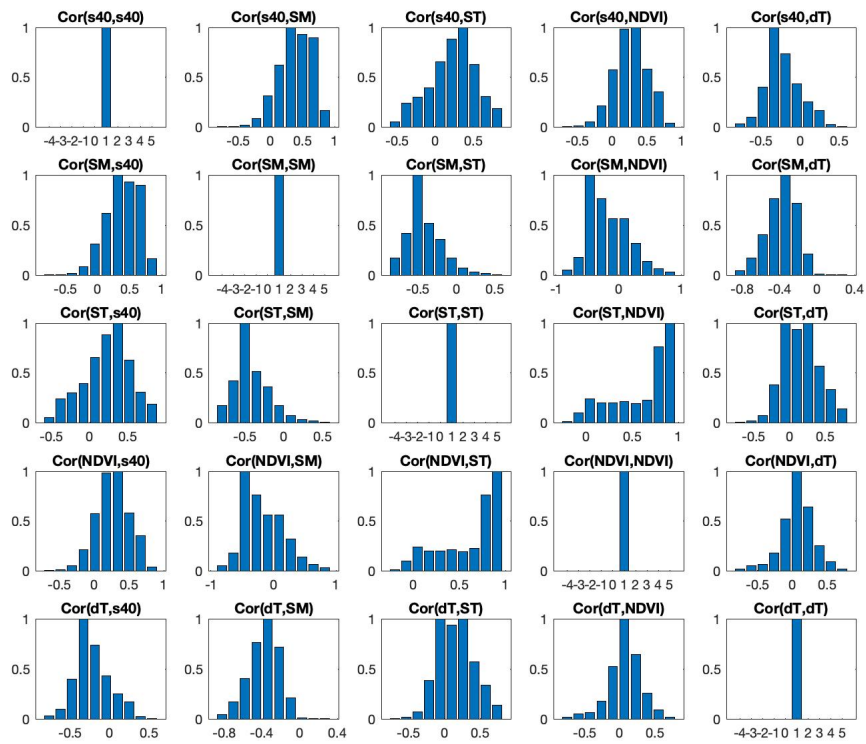


Figure 6.1: PDF of the correlations of the anomalies of the variables considered in this study

infer the input-output relationship in general terms, on the ensemble of the spatial domain. (2) A specific NN for each pixel. This is an extreme form of localization but there are enough samples to train such simple NNs and we want to investigate the potential of a localization strategy.

6.1.2 SM retrieval results over USA

Fig. 6.2 represents some statistics of the two configurations (global or localized) of the retrieval. Bias, STD and RMS error maps are represented for the global retrieval (left) and the localised one (right). The localized retrieval clearly outperforms the global retrieval, for the three diagnostics. By definition, there is no bias errors for the localized retrieval, and the STD and RMS errors are very significantly reduced compared to the global retrieval. This is clearly the best

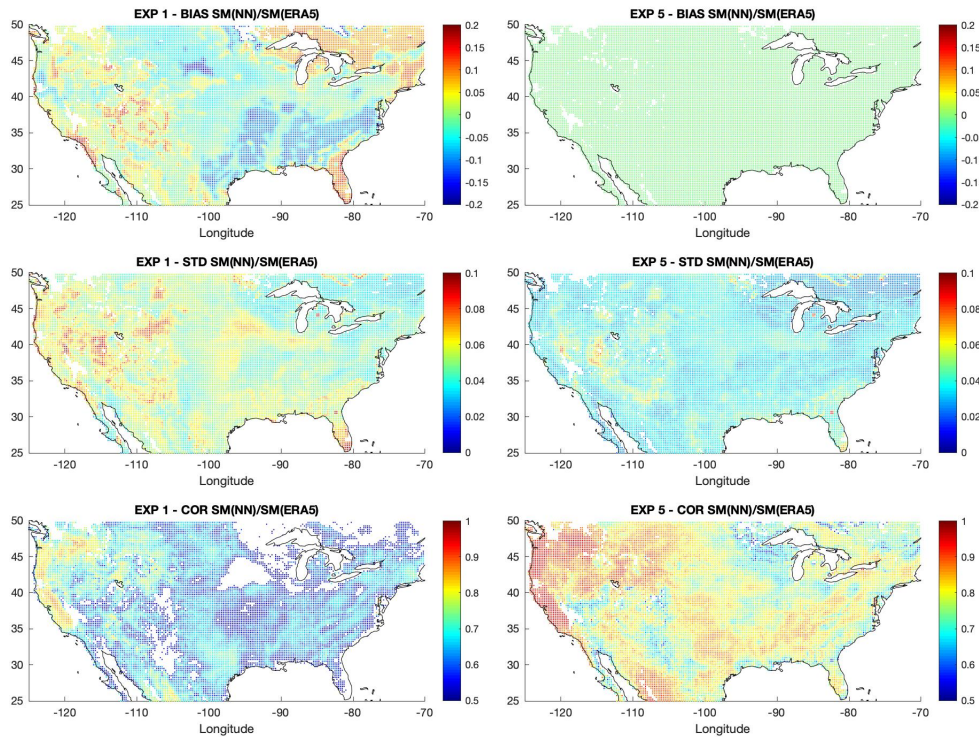


Figure 6.2: Bias error (top), STD error (middle) and CORrelation (bottom) between ERA and retrieved SM from the “global-area” retrieval (left) and the pixel-wise retrieval (right).

6.2 NN uncertainty estimation

Uncertainty on NN products is important to assess the quality of the NN simulations (in the forward or inverse mode) and to assimilate them in Numerical Weather Prediction (NWP) forecasting frameworks. In the forward mode, they will represent the forward uncertainty, or the observations uncertainty in the inverse mode. No easy solution has been proposed so far to estimate the NN uncertainties. A simple, pragmatic and easy implementation scheme has been proposed in [7] to achieve this task, based on two ideas: the use of the learning database \mathcal{B} to estimate the uncertainties and estimate these uncertainties on clusters or bins of the NN input space. We present here an alternative following the similar ideas.

6.2.1 Using the learning database \mathcal{B} as a reference

The first important point is the use of the training database to estimate these NN uncertainties. A reference needs to be used to estimate the errors and which better database for this purpose than the database that we chose to train the NN?

Several sources of uncertainty can cause the NN errors:

- *Observation errors*: NN inputs can be corrupted by instrumental noise, they are included in the learning database if it is constituted of real observations, or they can be introduced artificially using some specifications if it is synthetic data.
- *Structural errors*: The architecture of the NN might not be adequate (e.g., not enough or too complex mapping function); the learning scheme used to train the NN might not be perfect (e.g., bad initial conditions, insufficient optimisation algorithm); the learning dataset might be deficient (i.e., insufficient number of samples or lack of diverse data). All these reasons cause the NN to be un-optimal for the retrieval. Again, these errors are included in \mathcal{B} .
- *Insufficient input information*: Additional inputs might be necessary to obtain better outputs.
- *Representation errors*: They are linked, for instance, to the inherent smoothing of a discrete description of profiles (i.e., atmospheric layers instead of an infinity of levels) [?].
- *Output errors*: Errors can also be present in the outputs in \mathcal{B} .

- *Collocation errors*: The database might be constructed empirically gathering data at different space/time resolution. These temporal delays or spatial distances introduce errors too in the database \mathcal{B} .

This non-exhaustive list describes real physical sources of uncertainty that can cause the NN errors. It is important to take into account all these sources of information and this is why using \mathcal{B} , that includes most of these error sources, is a good strategy. If all sources of uncertainties are represented in the learning database, then the goal in the following will be (1) to compute this uncertainty on \mathcal{B} and (2) to build a model able to estimate it.

6.2.2 Estimation of the “empirical” uncertainty

To obtain uncertainty statistics such as the standard deviation of the error distribution, samples need first to be gathered. In [?], this is done by grouping samples using a clustering on the input space. Such an approach was tested here but results were not convincing. Therefore, another approach was attempted: the grouping is made in space and time. In space because the retrieval errors are estimated individually on each pixels. In time because we use a sliding temporal window, all the samples on that window are gathered to estimate the retrieval STD. The sliding window is chosen with size 50 days. It provides a slowly varying STD estimate which is compatible with the fact that uncertainties are related to weather regimes.

6.2.3 Uncertainty NN model

An uncertainty estimate was associated to all the samples in the learning database \mathcal{B} , using the sliding temporal window. The goal now is to develop a model that is able to estimate these uncertainties from the information available at the time of the retrieval, namely the input parameters.

Any source of information could be used as inputs to the uncertainty model, but the most important is the information that is used to perform the retrieval itself. In addition, any localisation (spatial or temporal) information can be useful too. This is why the model chosen here is a regular NN, with the same inputs as the retrieval, on the pixel-wise configuration (i.e. full localization) is the perfect strategy chosen here.

6.2.4 Uncertainty estimation results

In Fig. 6.3, the time series of STD of retrieval uncertainties is represented for two locations. The blue curve is measured on the learning database, it would

be the objective to estimate these values. The red curve is the estimation obtained with the methodology described in the previous sections. It can be noted on the left that the evolution of the uncertainty is quite well reproduced. The pixel on the right is less satisfactory but part of the dynamics is reproduced by the estimate, even if the amplitude of the dynamics is not as large.

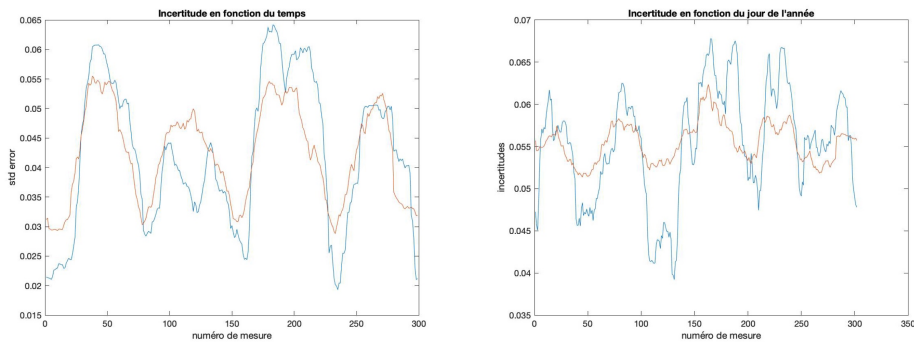


Figure 6.3: Time series of the STD of retrieval uncertainties, the true STD is blue and the estimated

Another way to investigate the quality of the retrieval uncertainties is to look at their spatial pattern. Fig. 6.4 represents the true (left) and estimated (right) STD of retrieval error for winter (top) and summer (bottom) over Europe. The spatial patterns are well retrieved by the STD estimation. Again, the amplitude is not perfect, especially for winter but the state-dependency of the estimation is very interesting and should be exploitable for two purpose: flag situations with too large errors, and quantify uncertainties for facilitating the assimilation of the retrieved quantities. More work needs to be done in order to improve these state-dependent uncertainty estimations, and many more tests have been conducted in this study (not shown). For instance the binning of the input space was tested, following [7] but the sliding temporal window appeared more efficient in this context. This work will be investigated more in the future, it is a true research topic in itself, but we proposed here several ways to start this investigation. The uncertainties are a second-order information that we can obtain from the satellite observations (the first order information is the SM estimation itself). Errors are actually the SM information what the NN was not able to estimate. It is therefore natural to have more difficulties to estimate them. Note however that we do not intend here to estimate the errors themselves, but a statistics on them (STD).

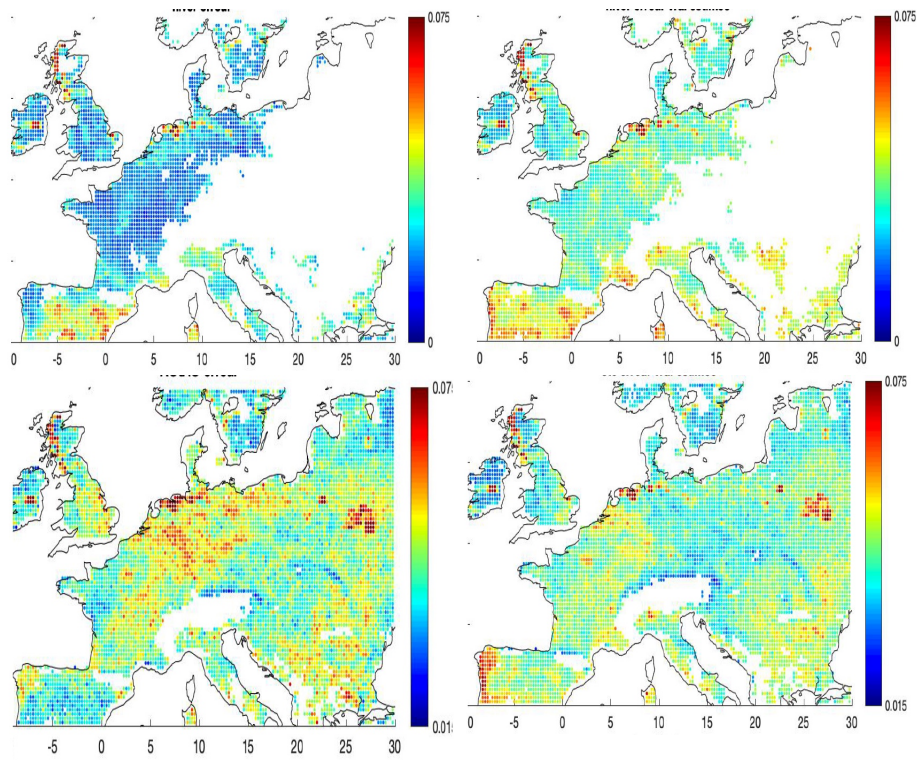


Figure 6.4: True (left) and estimated (right) STD of retrieval error for winter (top) and summer (bottom) over Europe.

6.3 Validation using in situ Measurements

6.3.1 In situ soil moisture database

The database of in situ measurements of soil moisture over USA is composed of three network of observations: SCAN, SNOTEL, USCRN. For our experiment, four years of data have been gathered, from 2016 to 2019.

The location of the stations of these three networks is represented in Fig. 6.5. On the left the number of observations gathered for each station, and the averaged soil moisture is represented on the right.

6.3.2 Evaluation

Fig. ?? represents over the in situ pixel measurements, the correlation (top) and the standard deviation of errors (bottom) for the global retrieval (left) and the pixel-wise retrieval (right). There is a clear increase of correlations when using the localisation of the pixel-wise approach. This will be confirmed in the following figure. The standard deviation is also decreased in the localized retrieval.

Fig. 6.7 shows the histograms of the pixel-wise correlations between the in situ and the retrieved soil moistures of Fig. 6.6. The top plot is for the global-scale retrieval and the bottom one for the pixel-wise retrieval. The averaged correlations is also indicated, and it is clear that the localization strategy has improved very significantly the statistics that go from 0.37 to 0.51. Two comments need to be made. First, these correlations do not seem very high, but we know that it is always extremely difficult to compare a satellite pixel measurement and a point-measurement in one particular location. The best agreements with in situ stations is always obtained with the reanalysis itself and scores typically reach about a 0.6 correlation level. The second aspect is that the localization (here in the form of using a NN for each pixel) not only changes the biases errors that a statistical global method can introduce regionally, it also changes the dynamics of the SM retrieved product, and it changes it in the right direction, closer to the actual in situ data. This means that the location of the NN is actually even more interesting than the CDF matching that do not change the correlations, except if the CDF matching model is non-linear. Even if non-linear, the CDF-matching is mono-variate, and not been multivariate is difficult to model complex local relationships.

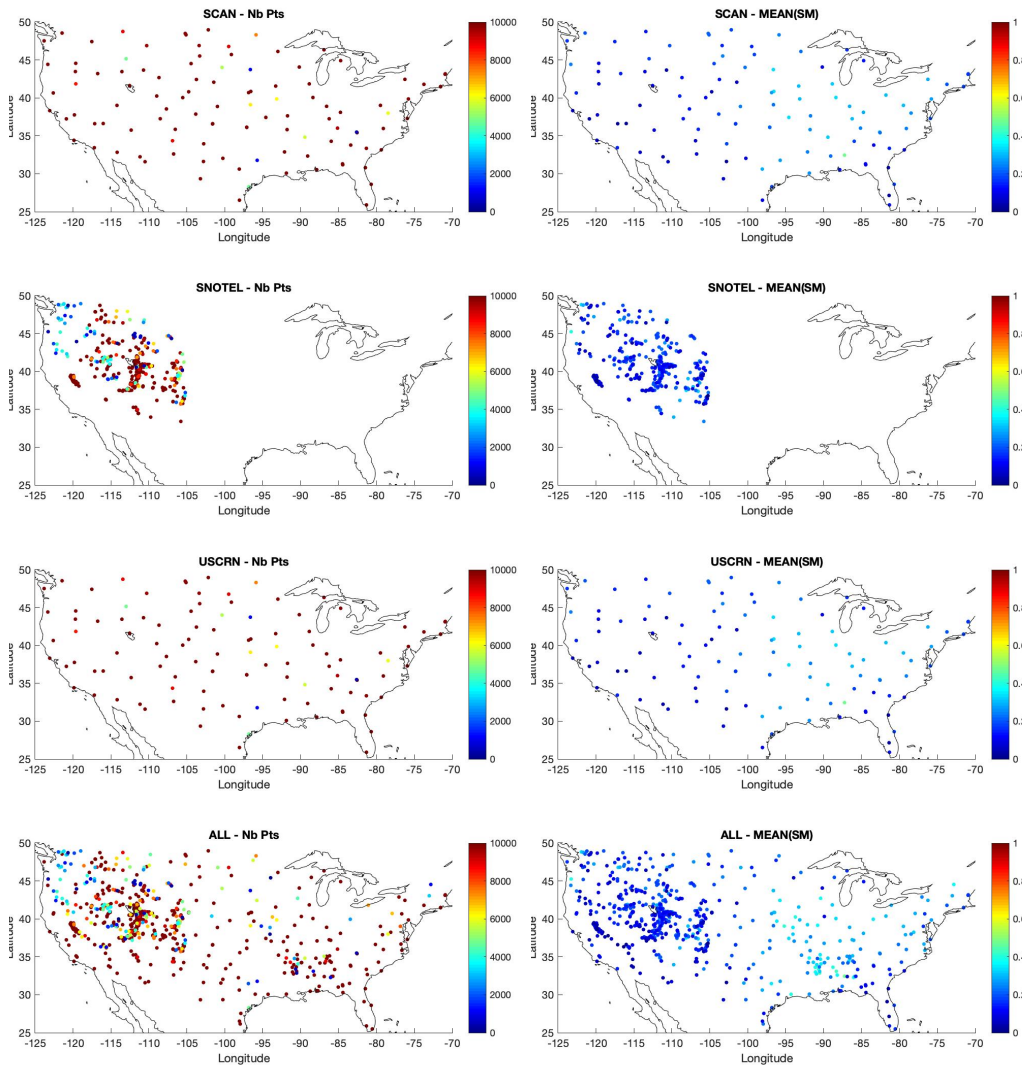


Figure 6.5: Number of samples (left) and averaged soil moisture (right) for all the in situ stations, from the top to bottom, for the SCAN, SNOTEL, USCRN and ALL networks.

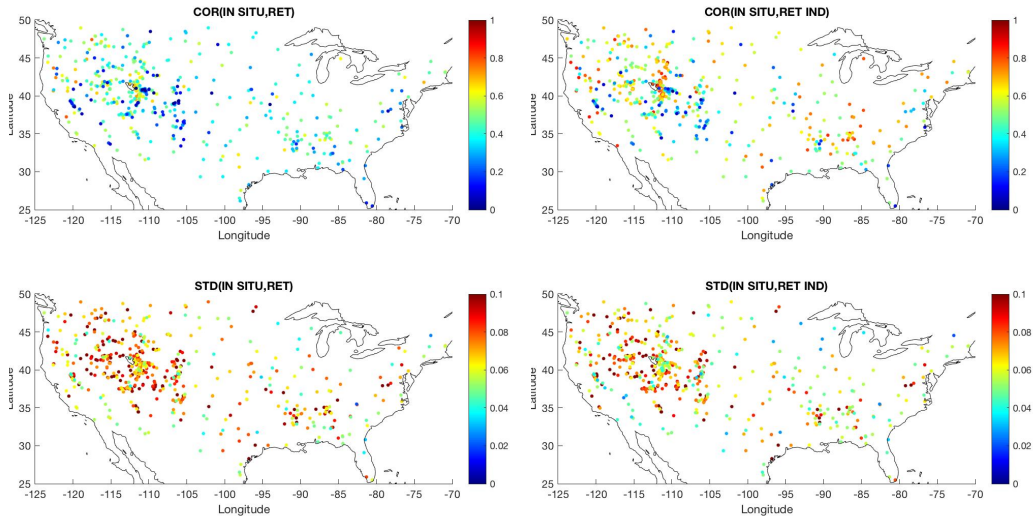


Figure 6.6: Correlation (top) and the standard deviation of errors (bottom) for the global retrieval (left) and the pixel-wise retrieval (right) over the in situ soil moisture stations.

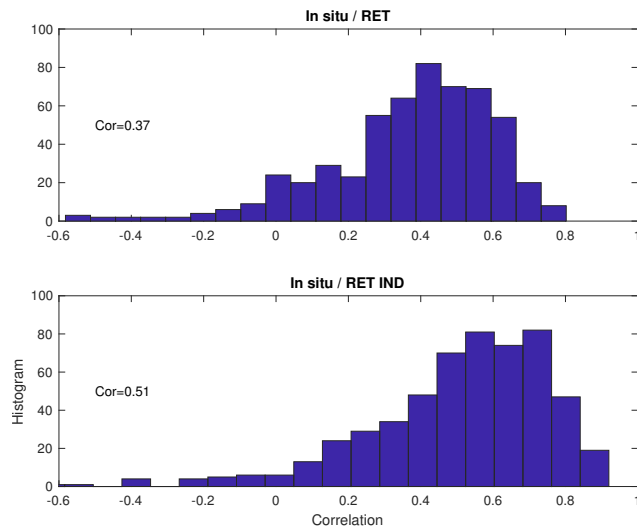


Figure 6.7: Histogram of the correlations between the in situ and the retrieved soil moistures over the USA. Left is for the global-scale retrieval and the right for the pixel-wise retrieval.

Chapter 7

Conclusion and perspectives

This study is intended to inform long-term strategies for the assimilation of satellite-derived observations of surface variables into LSMs using the ASCAT derived soil moisture estimates and ERA5 data. Some general approaches have been proposed that could be used to other satellite observations, other NWP frameworks, and other geo-physical variables.

The first aspect discussed in this paper concerns the CDF-matching techniques. In addition to the classical Gaussian approach commonly used in NWP centres, we proposed two new methods: the binning and the function approximation. These three techniques share some features and each has its own advantages and drawbacks, with different levels of agreement with the target model reference. Among the characteristics that are interesting to consider is the ability of the CDF-matching to increase correlation with the target provided by non-linear approaches. The choice of the CDF-matching should really be dependent on the considered application. However, it would be interesting to test new CDF-matching techniques in an assimilation framework to fully understand their potential.

The limitations of RT models have inhibited the direct assimilation of satellite observations sensitive to surface variables such as SM. Global statistical relationships have been proposed in the past to overcome this difficulty [2]. In this paper, the NN and CDF-matching methods have been reviewed and conceptually compared. We have shown that CDF-matching can be used as a local retrieval scheme. Assimilation performs a weighting of the information from the satellite and the state of the LSM, based on their respective accuracies. The CDF-matching approach is a mono-variate model highly calibrated towards the LSM, in contrast to the NN which is a multi-variate global model. Both methods have pros and cons: The CDF-matching uses the dynamical information of the satellite observations, but any spatial information is lost. The NN intends to obtain physical relationships at the

global scale, which is a more ambitious task. The NN is less calibrated towards the model, as it can correct the LSM dynamical behaviour and spatial patterns. The optimal choice between these two methods is truly dependent on some aspects of the application: quality of the satellite observation, simplicity of the satellite/variable physical relationship, availability of auxiliary information, or desire to exploit the synergy of multiple instruments.

For the retrieval of SM from the ASCAT backscatter, the CDF-matching appears to perform well. This comes from the fact that the ASCAT backscatter σ_{40} dynamics are mainly dependent on the local SM conditions and less on the other factors (vegetation, surface roughness, etc.). The relationship between σ_{40} and SM is simple enough that CDF-matching is a good strategy. This would probably not be the case for passive microwave observations for instance because the relationship would be more complex and dependent on other factors [51].

NN and CDF-matching are not dependent on each other: NN do not necessitate CDF-matching because their SM agrees better with the model, and CDF-matching does not require the NN modelling. However, we have shown that NN and CDF-matching can actually be combined, which is an advantage if multi-instrument data needs to be combined.

Using a NN for the forward RT model does not perform as well as for the inverse model in this study. This may seem surprising because in an inverse problem such as remote sensing, forward models are mathematically well-posed but the inverse model can become ill-posed (in terms of existence, uniqueness and stability of the solution). However, the inverse problem here (this might be different for another application) is simple as only local and statistical relationships are sufficient. The forward model requires complex deterministic relationships, using several auxiliary variables, which are not easy to obtain at the global scale. Furthermore, in the forward modelling for LSM assimilation strategy, knowledge of the observation errors is required (instrument plus NN forward modelling errors), and the inversion still needs to be performed in the assimilation (related to the inversion of the Jacobian matrix in the assimilation formula, see [55]) so that observation variables are transformed into model variables such as SM. This “inversion” also comes with uncertainties. We hypothesise that adding instrument, forward and inversion errors might be less optimal than inverting directly the satellite data into SM and assimilating it. This would be different if the forward model was very reliable, like in atmospheric applications. This needs, however, to be explored with assimilation experiments.

It was shown that localization strategy is a good way to adapt the NN model to local conditions. Several localization schemes could be implemented, for instance more parameters could be introduced so that the NN

can change its behaviour in different regions of the world. This is a more ambitious task, that intend to obtain a more physically-oriented model, but sometimes this additional information is simply not available. We tested here a different approach where a NN is trained for each pixel. This simplified the physics that is involved, and unknown parameters can become useless if they do not vary in this particular location. We have shown that this localization strategy decreases regional biases of the NN, and also improves the dynamics of the retrieval. Both of these results are an excellent news for the assimilation. Evaluation over the USA where many in situ measurements are available shows that the localization truly improves the NN behaviour and soil moisture retrieval.

Finally, a new scheme has been proposed to estimate the NN uncertainties. Obtaining such uncertainties of the NN outputs is mandatory if we want to assimilate them in a NWP centre. We proposed here a solution where empirical uncertainties are first estimated using a sliding time window, and then a NN is trained to estimate the resulting NN error STDs. Also not perfect, these experiments show that some state-dependency could be exploited, either as a flag to reject the assimilation of problematic data, or as a way to weight data based on their uncertainties.

Perspectives for this work are numerous. The next step will be to integrate these developments and to conduct data assimilation experiments to evaluate the impact on NWP forecasts. The results from an assimilation experiment do not always determine whether one retrieval scheme is better than another, as assimilation systems might need further developments to fully take advantage of some additional information. Indeed, it is often difficult to obtain a positive forecasting impact when assimilating very new satellite observations as numerous adjustments are required to optimally exploit the new information. However, the inverse and forward approaches could be tested in exactly the same conditions (screening, amount of information) so the comparison should really indicate which method seems more promising. Furthermore, it is envisaged that the NN approach will be employed by the EUMETSAT Satellite Applications Facility for Hydrology (H SAF). It would complement the change detection approach that is currently used in the ASCAT surface SM derivation [30, 32]. Furthermore, high quality global H-SAF root-zone SM products are generated by assimilating the ASCAT surface SM into an LSM [31, 33]. Complementary root-zone SM products are planned at 10 km resolution, which will benefit from assimilating NN based ASCAT surface SM. As already mentioned, the ASCAT data have been pre-processed using a linear regression to estimate the backscatter at a 40° incidence angle. This was done because it was shown that impact of vegetation is less important at this angle but this is only partially true. If

the vegetation was also retrieved at the same time, it might be possible to avoid such a pre-processing step and invert more or less directly the ASCAT backscatter at its original incidence angle.

The NN forward model would greatly benefit from the collaboration with RT experts in several ways. For instance, all the information used by such physical models could be used as inputs to the NN, however NN expertise is always required to really exploit, in an optimal way, this information. As mentioned previously, using an approximate forward model and then assimilating the satellite observations directly (requiring some kind of “inversion” in the assimilation formalism through the inversion of the RT Jacobians matrices), adds up two uncertainties; where the statistical inversion produces output data that is already in model space. These aspects need additional analysis.

NN are perfect tools to exploit the synergy of multiple instruments with multi-wavelengths [6, 4, 1, 42]. Combining passive and active microwave observations for SM retrieval is now in evidence, see e.g. [2] or [43]. The fact that the retrievals can be done using the same methodology is a tremendous advantage, and it was shown that the *a priori* combination of information for the retrieval is more advantageous than the *a posteriori* combination of individual retrievals. Attempts should then be made to combine ASCAT and SMOS data in the framework presented here. SMOS NN retrievals are already quite mature [57] so this could be an interesting extension to the work presented here.

Finally, these assimilation ideas could be exploited for other surface variables than soil moisture. For instance, statistical retrievals of land surface temperatures and surface emissivities have been proposed in [3] in the microwaves and in [48] in the infrared, and a similar approach was proposed too for the surface fluxes [40]. Since vegetation also impacts the satellite observations [50], it would be beneficial to retrieve this quantity too, simultaneously.

Chapter 8

Annexe

Similarly to Fig. 5.3, the minimum (left), maximum (center), and range (right) for σ_{40} is provided in Fig. 8.1, from top to bottom for ASCAT, Exp. 4 and 5. Range and correlations statistics are also provided for SM and σ_{40} for the 5 experiments over: Europe (Fig. 8.2), USA (Fig. 8.3), Australia (Fig. 8.4), South America (Fig. 8.5), Africa (Fig. 8.6) and Asia (Fig. 8.7).

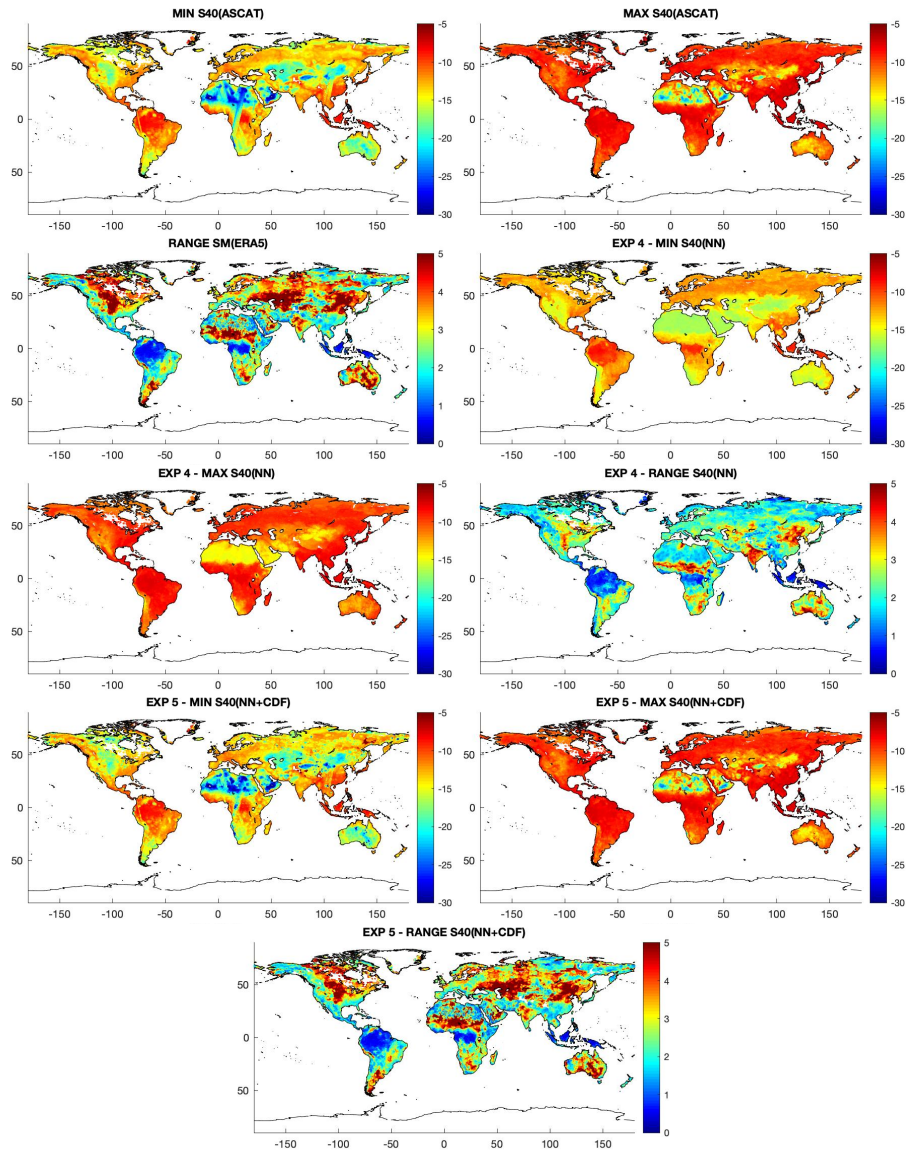


Figure 8.1: Minimum (left), maximum (center), and range (right) for σ_{40} . From top to bottom for ASCAT, Exp. 4 and 5.

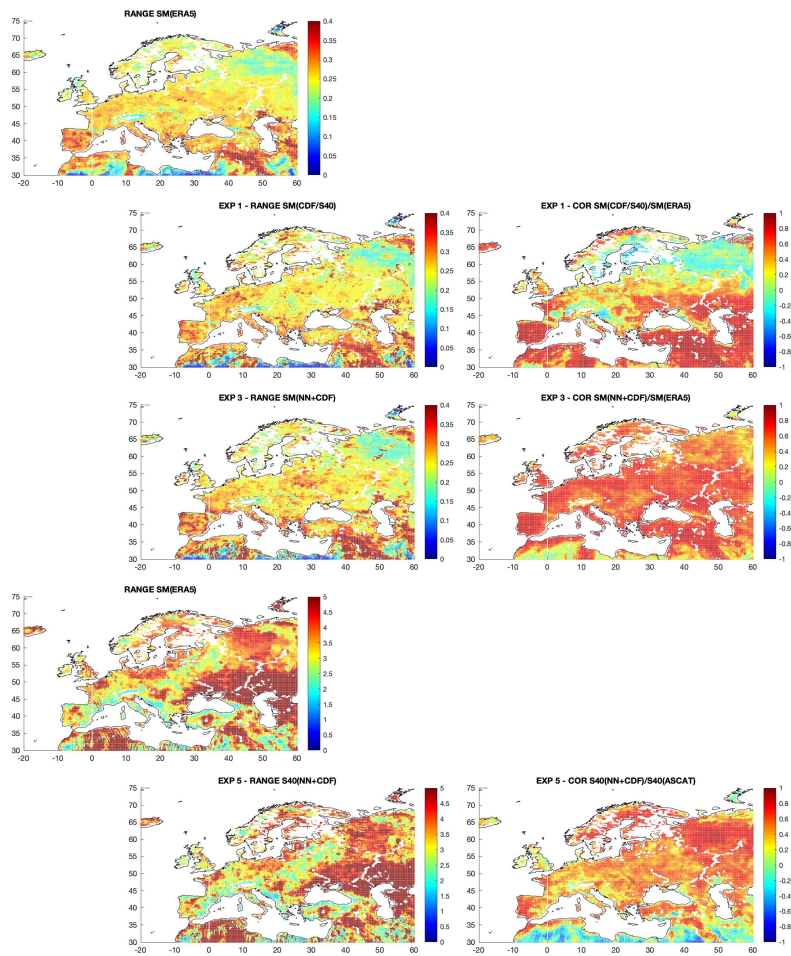


Figure 8.2: Statistics Europe

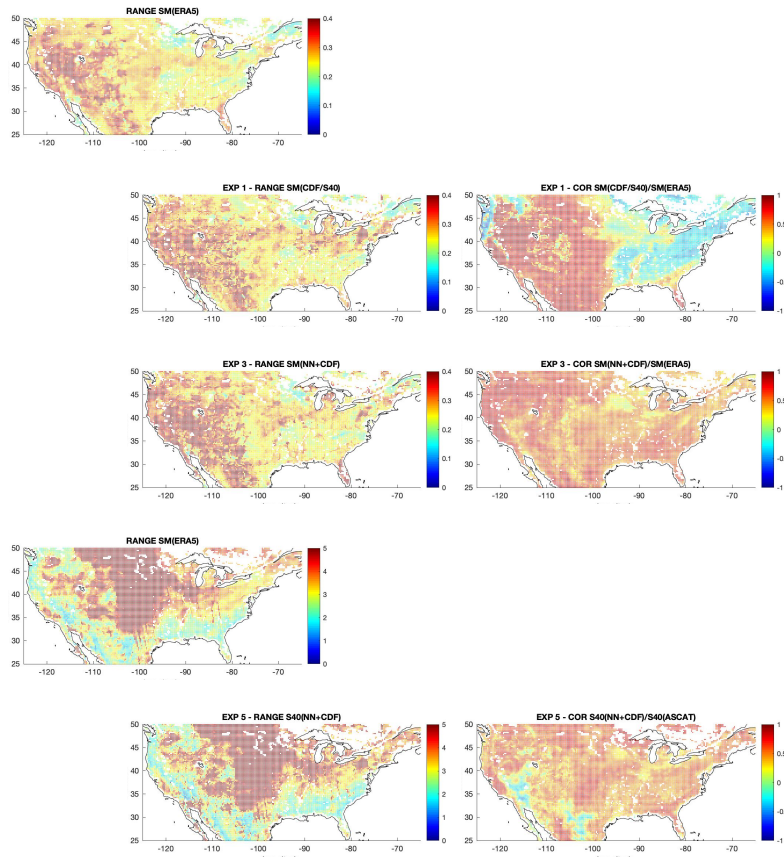


Figure 8.3: Statistics USA

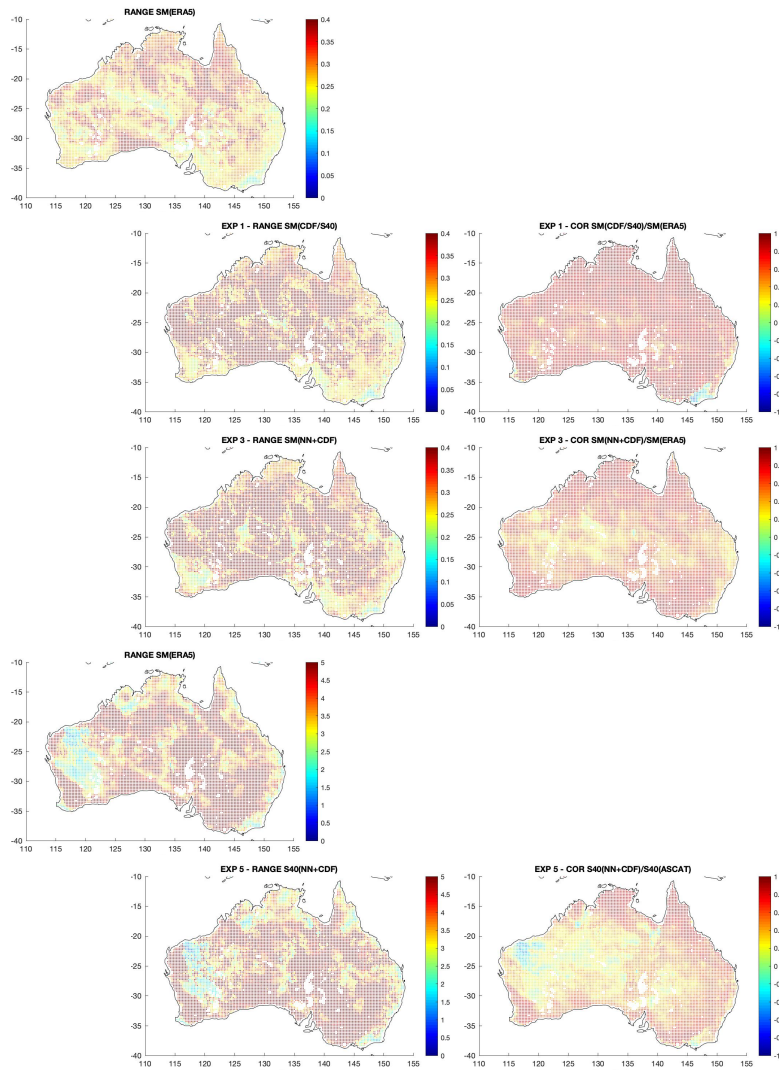


Figure 8.4: Statistics Australia

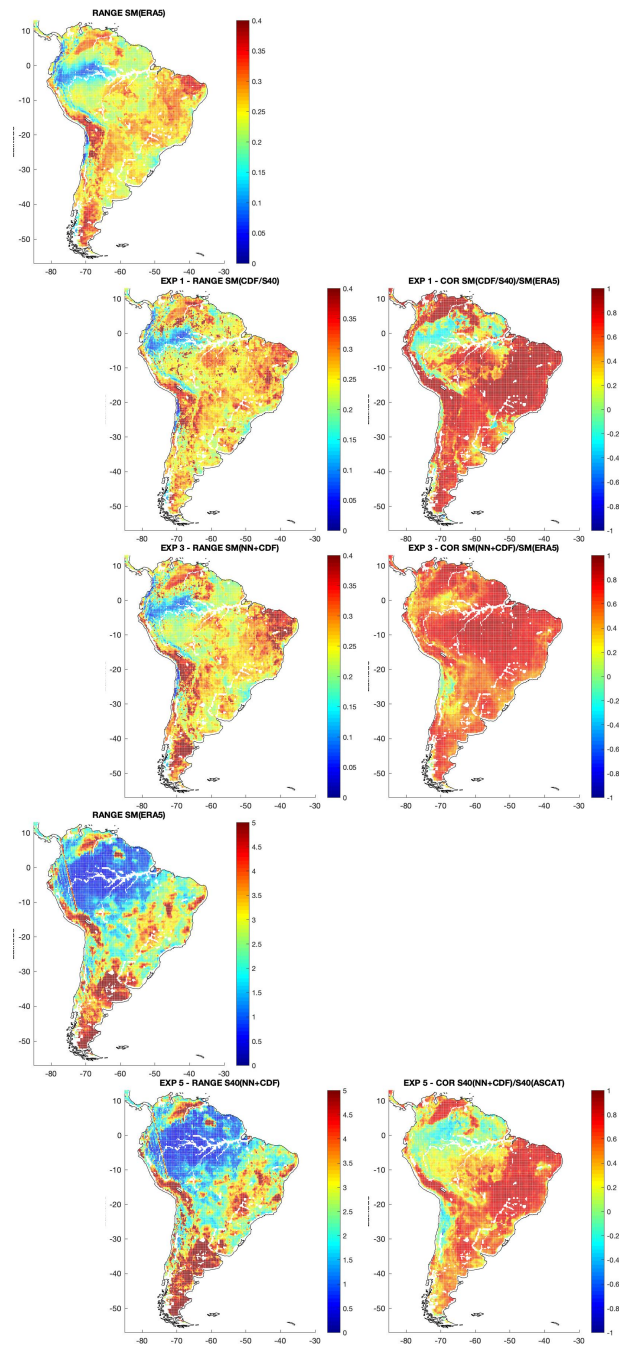


Figure 8.5: Statistics South America

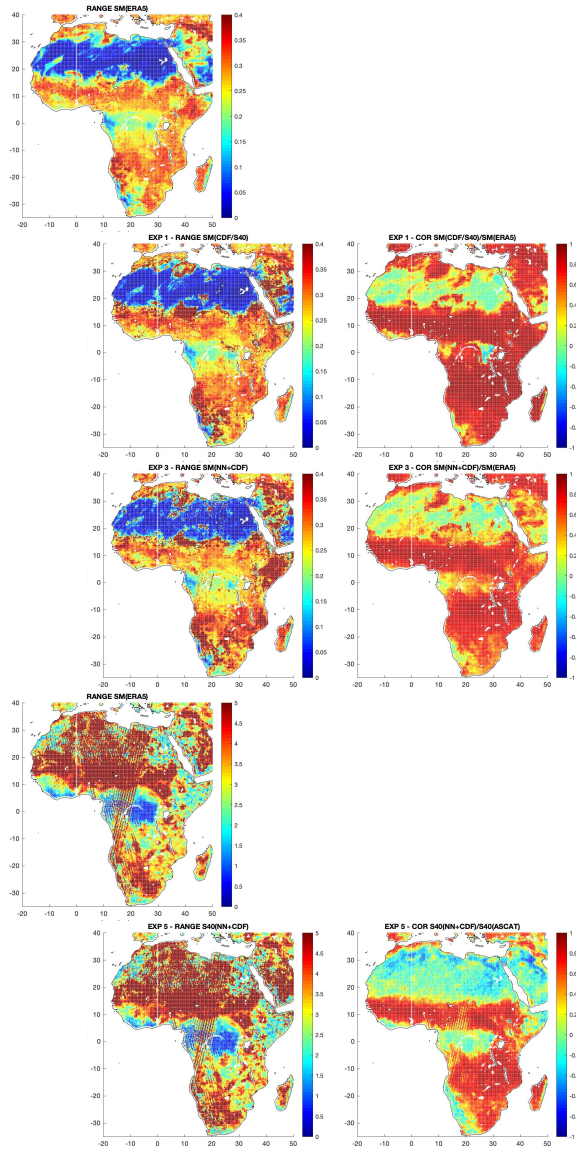


Figure 8.6: Statistics Africa

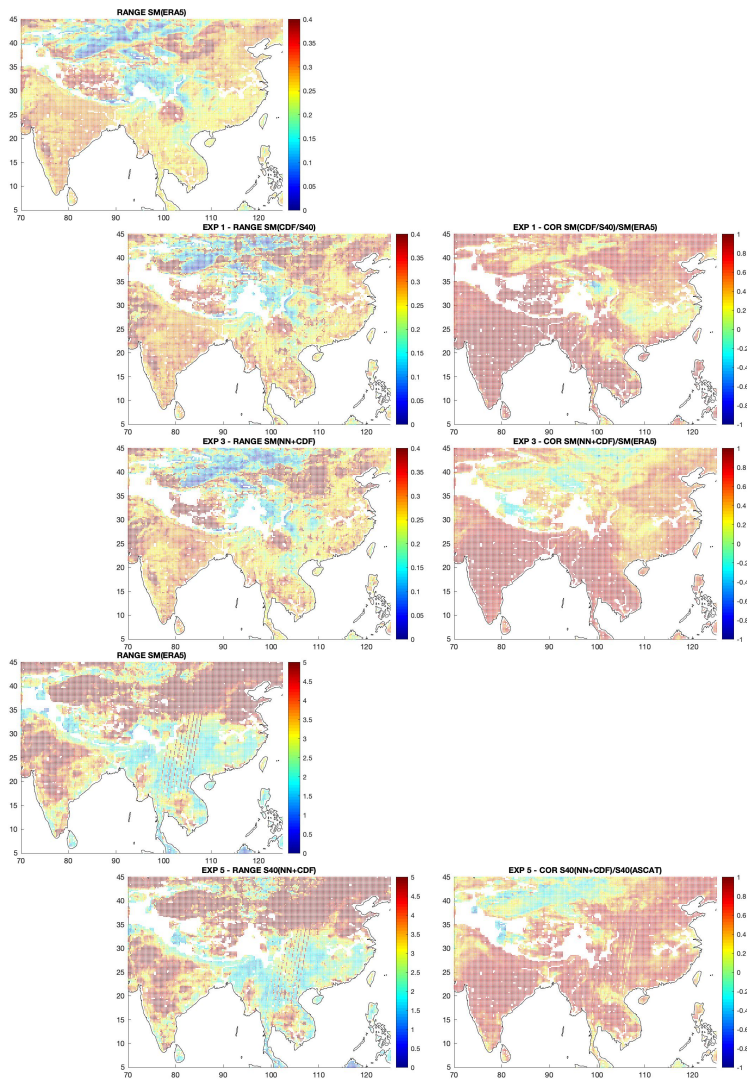


Figure 8.7: Statistics Asia

Bibliography

- [1] F Aires, O AZNAY, C Prigent, M Paul, and F Bernardo. Synergistic multi-wavelength remote sensing versus a posteriori combination of retrieved products: Application for the retrieval of atmospheric profiles using MetOp-A. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 117(D18), September 2012.
- [2] F Aires, C Prigent, and W B Rossow. Sensitivity of satellite microwave and infrared observations to soil moisture at a global scale: 2. Global statistical relationships. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 110(D11), June 2005.
- [3] F Aires, C Prigent, W B Rossow, and M Rothstein. A new neural network approach including first guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature, and emissivities over land from satellite microwave observations. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 106(D14):14887–14907, July 2001.
- [4] Filipe Aires. Measure and exploitation of multisensor and multiwavelength synergy for remote sensing: 1. Theoretical considerations. *Journal of Geophysical Research: Atmospheres*, 116(D):2301, January 2011.
- [5] Filipe Aires, Alain ChéDin, Noëlle A Scott, and William B Rossow. A Regularized Neural Net Approach for Retrieval of Atmospheric and Surface Temperatures with the IASI Instrument. *Journal of Applied Meteorology*, 41(2):144–159, February 2002.
- [6] Filipe Aires, Maxime Paul, Catherine Prigent, Bjorn Rommen, and Marc Bouvet. Measure and exploitation of multisensor and multiwavelength synergy for remote sensing: 2. Application to the retrieval of atmospheric temperature and water vapor from MetOp. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 116(D2):D02302, January 2011.

- [7] Filipe Aires and Victor Pellet. Estimating retrieval errors from neural network schemes - Application to the retrieval of temperature profiles from IASI. *IEEE Transactions on Geoscience and Remote Sensing*, 2020.
- [8] Filipe Aires, Catherine Prigent, Stefan A Buehler, Patrick Eriksson, Mathias Milz, and Susanne Crewell. Towards more realistic hypotheses for the information content analysis of cloudy/precipitating situations – Application to a hyperspectral instrument in the microwave. *Quarterly Journal of the Royal Meteorological Society*, 145(718):1–14, January 2019.
- [9] Filipe Aires, Michel Schmitt, Alain ChéDin, and Noëlle Scott. The “Weight Smoothing” Regularization of MLP for Jacobian Stabilization. *IEEE transactions on Neural Networks*, 10(6):1502–1510, November 1999.
- [10] Nicolas Baghdadi, Mohammad El Hajj, Mehrez Zribi, and Safa Bousbih. Calibration of the Water Cloud Model at C-Band for Winter Crop Fields and Grasslands. *Remote Sensing 2013, Vol. 5, Pages 687-715*, 9(9):969–, September 2017.
- [11] Yifang Ban, Puzhao Zhang, Andrea Nascetti, Alexandre R Bevington, and Michael A Wulder. Near Real-Time Wildfire Progression Monitoring with Sentinel-1 SAR Time Series and Deep Learning. *Nature Publishing Group*, 10(1):1322, January 2020.
- [12] A L Barbu, J C Calvet, J-F Mahfouf, and S Lafont. Integrating ASCAT surface soil moisture and GEOV1 leaf area index into the SURFEX modelling platform: a land data assimilation application over France. *Hydrol. Earth Syst. Sci.*, 18(1):173–192, 2014.
- [13] C M Bishop. *Neural networks for pattern recognition*. Clarendon Press - Oxford, 1996.
- [14] N Bormann, A Collard, and Peter Bauer. Observation errors and their correlations for satellite radiances. ECMWF Newsletter, 2011.
- [15] F Chevallier, J-J Morcrette, F Cheruy, and N A Scott. Use of a neural-network-based long-wave radiative-transfer scheme in the ECMWF atmospheric model. *Quarterly Journal of the Royal Meteorological Society*, 126(563):761–776, January 2000.

- [16] P Courtier, E Andersson, W Heckley, J Pailleux, D Vasiljevic, M Hamrud, A Hollingsworth, F Rabier, and M Fisher. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quarterly Journal of the Royal Meteorological Society*, 124(550):1783–1807, July 1998.
- [17] G Cybenko. Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.*, 2:303–314, 1989.
- [18] P de Rosnay, C Albergel, G Balsamo, and J Muñoz Sabater. ASCAT data assimilation at ECMWF. *EGUGA*, page 5488, 2012.
- [19] Patricia de Rosnay, Joaquin Muñoz Sabater, Clément Albergel, Lars Isaksen, Stephen English, Matthias Drusch, and Jean Pierre Wigneron. SMOS brightness temperature forward modelling and long term monitoring at ECMWF. *Remote Sensing of Environment*, 237:111424, February 2020.
- [20] Dick Dee. Variational bias correction of radiance data in the ECMWF system. Technical report, ECMWF, 2004.
- [21] I Dharssi, K J Bovis, B Macpherson, and C P Jones. Operational assimilation of ASCAT surface soil wetness at the Met Office. *Hydrol. Earth Syst. Sci.*, 15(8):2729–2746, 2011.
- [22] W A Dorigo, A Gruber, R A M de Jeu, W Wagner, T Stacke, A Loew, C Albergel, L Brocca, D Chung, R M Parinussa, and R Kidd. Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sensing of Environment*, 162(C):380–395, June 2015.
- [23] C S Draper, R H Reichle, G J M De Lannoy, and Q Liu. Assimilation of passive and active microwave soil moisture retrievals. *Geophys. Res. Lett.*, 39(4):n/a–n/a, February 2012.
- [24] J Escobar, A Ch ’edin, F Cheruy, and N A Scott. Réseaux de neurones multicouches pour la restitution de variables thermodynamiques atmosphériques à l’aide de sondeurs verticaux satellitaires. *CRAS*, pages 911–918, 1993.
- [25] David Fairbairn, Patricia de Rosnay, and Philip A Browne. The New Stand-Alone Surface Analysis at ECMWF: Implications for Land–Atmosphere DA Coupling. *Journal of Hydrometeorology*, 20(10):2023–2042, October 2019.

- [26] J Figa-Saldaña, J J W Wilson, E Attema, R Gelsthorpe, M R Drinkwater, and A Stoffelen. The advanced scatterometer (ASCAT) on the meteorological operational (MetOp) platform: A follow on for European wind scatterometers. *Canadian Journal of Remote Sensing*, 28(3):404–412, June 2002.
- [27] Adrian K Fung, Zongqian Li, and K S Chen. Backscattering from a randomly rough dielectric surface. *IEEE Transactions on Geoscience and Remote Sensing (ISSN 0196-2892)*, 30(2):356–369, March 1992.
- [28] Alan J Geer, Katrin Lonitz, Peter Weston, Masahiro Kazumori, Kozo Okamoto, Yanqiu Zhu, Emily Huichun Liu, Andrew Collard, William Bell, Stefano Migliorini, Philippe Chambon, Nadia Fourrié, Min-Jeong Kim, Christina Köpken-Watts, and Christoph Schraff. All-sky satellite data assimilation at operational weather forecasting centres. *Quarterly Journal of the Royal Meteorological Society*, 144(713):1191–1217, April 2018.
- [29] G G Gutman. On the use of long-term global data of land reflectances and vegetation indices from the advanced very high resolution radiometer. *J. Geophys. Res.*, 104:6241–6255, 1999.
- [30] H-SAF. Metop ASCAT Soil Moisture at 25 km Swath Grid (H102, H103, H105), 2009.
- [31] H-SAF. ASCAT Root Zone Soil Moisture Profile Index 25 km resolution in NRT (H14), 2012.
- [32] H-SAF. Algorithm Theoretical Baseline Document (ATBD) Soil Moisture NRT, Metop ASCAT Soil Moisture Orbit, 2016.
- [33] H-SAF. Scatterometer Root Zone Soil Moisture (RZSM) Data Record 10km resolution - Multimission, 2020.
- [34] M C Hansen, R S DeFries, J R G Townshend, M Carroll, C Dimiceli, R A Sohlberg, M C Hansen, R S DeFries, J R G Townshend, M Carroll, C Dimiceli, and R A Sohlberg. Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm. *Earth Interactions*, 7(10):1–15, July 2009.
- [35] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquin Muñoz Sabater, Julien Nicolas, Carole Peubey, Raluca

- Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Saleh Abdalla, Xavier Abellan, Gianpaolo Balsamo, Peter Bechtold, Gionata Biavati, Jean Bidlot, Massimo Bonavita, Giovanna Chiara, Per Dahlgren, Dick Dee, Michail Diamantakis, Rossana Dragani, Johannes Flemming, Richard Forbes, Manuel Fuentes, Alan Geer, Leo Haimberger, Sean Healy, Robin J Hogan, Elías Hólm, Marta Janisková, Sarah Keeley, Patrick Laloyaux, Philippe Lopez, Cristina Lupu, Gabor Radnoti, Patricia Rosnay, Iryna Rozum, Freja Vamborg, Sebastien Villaume, and Jean-Noël Thépaut. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.*, 146(730):1999–2049, June 2020.
- [36] E V Holm and T Kral. Flow-dependent, geographically varying background error covariance for 1D-var applications in MTG-IRS L2 processing. Technical report, ECMWF, 2012.
- [37] K Hornik, M Stinchcombe, H White Neural networks, and 1989. Multilayer feedforward networks are universal approximators. *cs.cmu.edu*, 1989.
- [38] T Janjić, N Bormann, M Bocquet, J A Carton, S E Cohn, S L Dance, S N Losa, N K Nichols, R Potthast, J A Waller, and P Weston. On the representation error in data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 144(713):1257–1278, April 2018.
- [39] Carlos Jiménez, Douglas B Clark, Jana Kolassa, Filipe Aires, and Catherine Prigent. A joint analysis of modeled soil moisture fields and satellite observations. *J Geophys Res-Atmos*, 118(12):6771–6782, June 2013.
- [40] Carlos Jiménez, Catherine Prigent, and Filipe Aires. Toward an estimation of global land surface heat fluxes from multisatellite observations. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 114(D6), March 2009.
- [41] E Kalnay, M Kanamitsu, R Kistler, W Collins, and D Deaven. The NCEP/NCAR 40-Year Reanalysis Project, *B. Am. Meteorol. Soc.*, 77, 437–472, 1996.
- [42] J Kolassa, F Aires, J Polcher, C Prigent, C Jimenez, and J M Pereira. Soil moisture retrieval from multi-instrument observations: Information content analysis and retrieval methodology. *Journal of Geophysical Research: Atmospheres*, 118(1):4847–4859, May 2013.

- [43] J Kolassa, P Gentine, C Prigent, and F Aires. Soil moisture retrieval from AMSR-E and ASCAT microwave observation synergy. Part 1: Satellite data analysis. *Remote Sensing of Environment*, 173(C):1–14, February 2016.
- [44] J Kolassa, R H Reichle, Q Liu, S H Alemohammad, P Gentine, K Aida, J Asanuma, S Bircher, T Caldwell, A Colliander, M Cosh, C Holifield Collins, T J Jackson, J Martínez-Fernández, H McNairn, A Pacheco, M Thibeault, and J P Walker. Estimating surface soil moisture from SMAP observations using a Neural Network technique. *Remote Sensing of Environment*, pages 1–0, November 2017.
- [45] Y Y Liu, R M Parinussa, W A Dorigo, R A M de Jeu, W Wagner, A I J M van Dijk, M F McCabe, and J P Evans. Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. *Hydrol. Earth Syst. Sci.*, 15(2):425–436, 2011.
- [46] M Matricardi, F Chevallier, and S Tjemkes. An improved general fast radiative transfer model for the assimilation of radiance observations. Technical report, ECMWF, 2001.
- [47] J Muñoz Sabater, H Lawrence, C Albergel, P Rosnay, L Isaksen, S Mecklenburg, Y Kerr, and M Drusch. Assimilation of SMOS brightness temperatures in the ECMWF Integrated Forecasting System. *Q. J. R. Meteorol. Soc.*, 145(723):2524–2548, June 2019.
- [48] M Paul, F Aires, C Prigent, I F Trigo, and F Bernardo. An innovative physical scheme to retrieve simultaneously surface temperature and emissivities using high spectral infrared observations from IASI. *J. Geophys. Res.*, 117(D11):D11302, June 2012.
- [49] C Prigent, F Aires, F Bernardo, J C Orlhac, J M Goutoule, H Roquet, and C Donlon. Analysis of the potential and limitations of microwave radiometry for the retrieval of sea surface temperature: Definition of MICROWAT, a new mission concept. *J. Geophys. Res. Oceans*, 118(6):3074–3086, June 2013.
- [50] Catherine Prigent, Filipe Aires, William Rossow, and Elaine Matthews. Joint characterization of vegetation by satellite observations from visible to microwave wavelengths: A sensitivity analysis. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 106(D18):20665–20685, September 2001.

- [51] Catherine Prigent, Filipe Aires, William B Rossow, and Alan Robock. Sensitivity of satellite microwave and infrared observations to soil moisture at a global scale: Relationship of satellite observations to in situ soil moisture measurements. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 110(D7), April 2005.
- [52] Catherine Prigent, Ina Tegen, Filipe Aires, Béatrice Marticorena, and Merhez Zribi. Estimation of the aerodynamic roughness length in arid and semi-arid regions over the globe with the ERS scatterometer. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 110(D9), May 2005.
- [53] A Quesney, S Le Hégarat-Masclé, O Taconet, D Vidal-Madjar, J P Wigneron, C Loumagne, and M Normand. Estimation of watershed soil moisture index from ERS/SAR data. *Remote Sensing of Environment*, 72(3):290–303, 2000.
- [54] C D Rodgers. Retrieval of Atmospheric Temperature and Composition From Remote Measurements of Thermal Radiation. *Reviews of Geophysics and Space Physics*, 14(4):609, November 1976.
- [55] C D Rodgers. Characterization and error analysis of profiles retrieved from remote sounding measurements. *Journal of Geophysical Research*, 95(D5):5587–5595, 20.
- [56] Clive D Rodgers. Inverse Methods for Atmospheric Sounding: Theory and Practice. *Inverse Methods for Atmospheric Sounding: Theory and Practice. Edited by RODGERS CLIVE D. Published by World Scientific Publishing Co. Pte. Ltd.*, 2:9789812813718, 2000.
- [57] Nemesio Rodríguez-Fernández, Patricia de Rosnay, Clément Albergel, Philippe Richaume, Filipe Aires, Catherine Prigent, and Yann Kerr. SMOS Neural Network Soil Moisture Data Assimilation in a Land Surface Model and Atmospheric Impact. *Remote Sensing 2013, Vol. 5, Pages 687-715*, 11(11):1334–23, June 2019.
- [58] Nemesio J Rodriguez-Fernandez, Filipe Aires, Philippe Richaume, Yann H Kerr, Catherine Prigent, Jana Kolassa, Francois Cabot, Carlos Jiménez, Ali Mahmoodi, and Matthias Drusch. Soil Moisture Retrieval Using Neural Networks: Application to SMOS. *IEEE Transactions on Geoscience and Remote Sensing*, 53(11):5991–6007, August 2015.
- [59] Nemesio J Rodriguez-Fernandez, Joaquin Muñoz Sabater, Philippe Richaume, Patricia de Rosnay, Yann H Kerr, Clément Albergel,

- Matthias Drusch, and Susanne Mecklenburg. SMOS near-real-time soil moisture product: processor overview and first validation results. *Hydrol. Earth Syst. Sci.*, 21(10):5201–5216, 2017.
- [60] D E Rumelhart, G E Hinton, and R J Williams. *Learning Internal Representations by Error Propagation*. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. I, Foundations. MIT Press, Cambridge, 1986.
- [61] R W Saunders, M Matricardi, and P Brunel. An improved fast radiative transfer model for assimilation of satellite radiance observations. *Quart. J. Roy. Meteor. Soc.*, 125:1407–1425, 1999.
- [62] Roger Saunders, James Hocking, Emma Turner, Peter Rayer, David Rundle, Pascal Brunel, Jerome Vidot, Pascale Roquet, Marco Matricardi, Alan Geer, Niels Bormann, and Cristina Lupu. An update on the RTTOV fast radiative transfer model (currently at version 12). *Geoscientific Model Development*, 11(7):2717–2737, July 2018.
- [63] K Scipal, M Drusch, G Balsamo, and P de Rosnay. Assimilation of global ASCAT soil moisture observations in the ECMWF Numerical Weather Prediction Model. *EGUGA*, page 12881, 2019.
- [64] Jiancan Tan, Nusseiba NourEldeen, Kebiao Mao, Jiancheng Shi, Zhao-liang Li, Tongren Xu, and Zijin Yuan. Deep Learning Convolutional Neural Network for the Retrieval of Land Surface Temperature from AMSR2 Data in China. *Sensors*, 19(13):2987–20, July 2019.
- [65] S M Uppala, P W Kallberg, A J Simmons, U Andrae, V Da Costa Bechtold, M Fiorino, J K Gibson, J Haseler, A Hernandez, G A Kelly, X Li, K Onogi, S Saarinen, N Sokka, R P Allan, E Andersson, K Arpe, M A Balmaseda, A C M Beljaars, L Van De Berg, J Bidlot, N Bormann, S Caires, F Chevallier, A Dethof, M Dragosavac, M Fisher, M Fuentes, S Hagemann, E Holm, B J Hoskins, L Isaksen, P A E M Janssen, R Jenne, A P McNally, J-F Mahfouf, J-J Morcrette, N A Rayner, R W Saunders, P Simon, A Sterl, K E Trenberth, A Untch, D Vasiljevic, P Viterbo, and J Woollen. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612):2961–3012, October 2005.
- [66] W Wagner, J Noll, M Borgeaud, and H Rott. Monitoring soil moisture over the Canadian Prairies with the ERS scatterometer. *IEEE Transactions on Geoscience and Remote Sensing*, 37(1):206–216, 1999.

- [67] Wolfgang Wagner, Klaus Scipal, Carsten Pathe, Dieter Gerten, Wolfgang Lucht, and Bruno Rudolf. Evaluation of the agreement between the first global remotely sensed soil moisture data with model and precipitation data. *Journal of Geophysical Research*, 108(D):4611, October 2003.
- [68] I H Woodhouse and D H Hoekman. A model-based determination of soil moisture trends in Spain with the ERS-scatterometer. *IEEE Transactions on Geoscience and Remote Sensing*, 38(4):1783–1793, 2000.
- [69] I H Woodhouse and D H Hoekman. Determining land-surface parameters from the ERS wind scatterometer. *IEEE Transactions on Geoscience and Remote Sensing*, 38(1):126–140, 2000.