DOTTORATO DI RICERCA IN

*International Doctorate in Civil and Environmental Engineering*

CICLO XXXIV

COORDINATORE Prof. Luca Solari

# HIGH-RESOLUTION REMOTE SENSING FOR RAINFALL AND RIVER DISCHARGE ESTIMATION

Settore Scientifico Disciplinare: ICAR/02


**Dottorando**
Paolo Filippucci

**Tutore**
Prof. Dr. Carla Saltalippi

**Tutore**
Prof. Dr. Wolfgang Wagner

**Tutore**
Dr. Angelica Tarpanelli

**Coordinatore**
Prof. Dr. Luca Solari


Anni 2018/2021



Dissertation

# HIGH-RESOLUTION REMOTE SENSING FOR RAINFALL AND RIVER DISCHARGE ESTIMATION

*Ausgeführt zum Zwecke der Erlangung des akademischen Grades eines*
*Doktor der technischen Wissenschaften (Dr.techn.)*

*Unter der Leitung von*
*Univ.-Prof. Dipl.-Ing. Dr.techn. Wolfgang Wagner*
*Univ.-Prof. Dipl.-Ing. Phd. Carla Saltalippi*
*Dipl.-Ing. Phd. Angelica Tarpanelli*

*E120-01-1*
*Research Group Microwave Remote Sensing*

*Eingereicht an der Technischen Universität Wien TU*
*Fakultät für Mathematik und Geoinformation*

*von*

Paolo Filippucci

Matrikelnummer: e11934924

Florence, 20/06/2022

PhD committee: Dr. Carla Saltalippi

Prof. Dr. Wolfgang Wagner

Prof. Dr. Pere Quintana-Segui

Prof. Dr. Thierry Pellarin

Prof. Dr. Juraj Parajka

Prof. Dr. Josef Eitzinger

Summary

The European Union's Earth Observation program, Copernicus, aims to create vast amounts of global data from satellites and ground-based, airborne and seaborne measurement systems with the goal of providing information to help service providers, public authorities, and other international organizations improve European citizens' quality of life. With the aim of reaching this goal, a new family of missions called Sentinels has been developed by the European Spatial Agency, ESA, specifically for the operational needs of the program. These missions carry a range of cutting-edge technologies, such as radars and multi-spectral imaging instruments, for land, ocean and atmospheric monitoring. Multiple kinds of high-resolution data are now available to the scientific community, which is working to adapt and develop the existing models and algorithms to the new information.

With this objective, two of the most important variables that contribute to the water cycle, rainfall and river discharge, were selected in this thesis to be estimated by the use of the information obtained from Sentinel-1 and Sentinel-2 sensors. The monitoring of these two variables is fundamental in many hydrological applications, like flood and landslide forecasting and water resources management, and their impact is clearly visible from space. In-situ measurements are the traditional data source of them, but the worldwide declining number of stations, their low spatial density and the data access problem limit their use. Satellite sensors have been therefore adopted to support and, in some cases, substitute the existing gauge network in estimating river discharge and rainfall, thanks to the strong growth in technologies and applications. Two valuable examples of this are SM2RAIN algorithm, which allows to estimate rainfall from Soil Moisture (SM) observations by exploiting the inversion of the soil water balance equation, and the CM approach, a non-linear regression model capable of linking the ground measurements of river discharge to the near-infrared (NIR) reflectance ratio between a dry and a wet pixel chosen around the border of a river. Notwithstanding their usefulness, until now several limitations affected these two methodologies. The main issue with satellite derived rainfall data was their low spatial resolution which could not overcome 10 km, a quantity insufficient to obtain accurate information over many areas and posing important constraint on their use for many applications and fields, which require more and more detailed information. Similarly, the resolution of the available NIR data was not suitable to provide information for narrow rivers (< 250 m wide), nor to study river features and patterns that were here averaged within a single pixel. The recently available high-resolution data from Sentinel Missions of the Copernicus program offer an opportunity to overcome these issues.

The data from Sentinel-1 mission can be used to obtain a high spatial resolution SM product, named S1-RT1, which is adopted in this thesis to derive 1 km spatial resolution (500 m spacing) rainfall data over the Po River basin from it, through the algorithm SM2RAIN. The rainfall derived from the 25 km ASCAT SM product (12.5 km spacing), resampled to the same grid of S1-RT1, is compared to the latter to evaluate the potential benefits of such product. SM2RAIN algorithm needs to be calibrated against a benchmark, which poses important limitations on the applicability of the analysis in data scarce regions. In order to overcome this issue, a parameterized version of SM2RAIN algorithm is previously developed relying on globally distributed data, to be used along with the standard approach in the high-resolution rainfall estimation. The performances of each obtained product are then compared, to assess both the parameterized SM2RAIN capabilities in estimating rainfall and the benefits deriving from Sentinel-1 high spatial resolution.

For the river discharge estimation, the use of Sentinel-2 NIR reflectances within the CM approach is investigated to support the hypothesis that a higher satellite product's spatial resolution, i.e., 10 m (vs. a medium-resolution, i.e., 250 m), is able to better identify the periodically flooded pixels, more related to the river dynamics, with obvious advantages for river discharge estimation. Moreover, the improved resolution allows both a finer distinction between vegetation, soil and water and the characterization of water turbidity in the river area, which is important to correctly estimate the river discharge using this approach. A new formulation enriched by the sediment component is proposed along with a procedure to localize the periodically flooded pixels without the intake of calibration data, which is a first step towards a completely uncalibrated procedure for the river discharge estimation, fundamental for ungauged rivers.

The obtained results show that the high-resolution information from Copernicus actually increase the accuracy of the satellite derived products. Good estimates of rainfall are obtainable from Sentinel-1 when considering aggregation time steps greater than 1 day, since to the low temporal resolution of this sensor (from 1.5 to 4 days over Europe) prevents its application to infer daily rainfall. In particular, the rainfall estimates obtained from Sentinel-1 sensors outperform those from ASCAT in specific areas, like in valleys inside mountain regions and most of the plains, confirming the added value of the high spatial resolution information in obtaining spatially detailed rainfall. The use of a parameterized version of SM2RAIN produces performances similar to those obtained with SM2RAIN calibration, attesting the reliability of the parameterized algorithm for rainfall estimation in this area and fostering the possibility to apply SM2RAIN worldwide even without the availability of a rainfall benchmark product. Similarly, the river discharge estimation from Sentinel-2 reflectances from selected stations along two Italian rivers, the Po and the Tiber, confirms that reliable

performance can be obtained from high-resolution imagery. Specifically, over both the stations the new formulation improves the river discharge accuracy and over the Po River the best performances are obtained by the uncalibrated procedure. Google Earth Engine (GEE) platform has been employed for the data analysis, allowing to avoid the download of big amounts of data, fostering the reproducibility of the analysis in different locations.

*"Man must rise above Earth to the top of the atmosphere and beyond,*
*for only then will he fully understand the world in which he lives"*

*Socrates (469-399 BC)*

*To ?*

# ACKNOWLEDGEMENTS

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

## ACRONYMS AND ABBREVIATIONS

*ADCP* – Acoustic Doppler Current Profiler

*AMW* – Active MicroWave

*ASCAT* – Advanced SCATterometer

*ASTER* – Advanced Spaceborne Thermal Emission and Reflection radiometer

*ASTGTM* – ASTER Global digiTal elevation Model

*AWAP* – Australian Water Availability Project

*BIASr* – relative BIAS error

*CMORPH* – Climate prediction center MORPHing method

*DEM* – Digital Elevation Model

*DOY* – Day Of Year

*DPC* – Italian Civil Protection Department

*DPR* – Dual-frequency Precipitation Radar

*ECMWF* – European Centre for Medium-Range Weather Forecasts

*ECV* – Essential Climate Variables

*EO* – Earth Observation

*ER* – Early Run (IMERG)

*ERA5* – fifth generation ECMWF Atmospheric Reanalysis of the global climate

*ESA* – European Space Agency

*ETOPO5* – Earth TOPOgraphy 5 arc minute

*EUMETSAT* – EUropean organisation for the exploitation of METeorological SATellites

*FAR* – False Alarm Ratio

*FR* – Final Run (IMERG)

*GCOS* – Global Observing Systems information Center

*GEE* – Google Earth Engine

*GMES* –Global Monitoring for Environment and Security

*GNSS* – Global Navigation Satellite Systems

*GPCC* – Global Precipitation Climatology Centre

*GPM* – Global Precipitation Mission

*GPM-CO* – GPM mission Core Observatory

*GRDC* – Global Runoff Data Centre

*H SAF* – Satellite Application Facility on Support to Operational Hydrology and Water Management

*IPCC* – International Panel on Climate Change

*MCM* – Modified Conditional Merging

*MEaSUREs* – Making Earth System data records for Use in Research Environments

*MSI* – MultiSpectral Instrument

*IMD* – Indian Meteorological Department

*IMERG* – Integrated Multi-satellitE Retrievals for GPM

*JRC* – European Commission's Joint Research Centre

*LAI* – Leaf Area Index

*LR* – Late Run (IMERG)

*MCM* – Modified Conditional Merging algorithm

*MetOp* – Meteorological Operational Satellites

*NaN* – Not a Number (invalid data)

*NASA* – National Aeronautics and Space Administration

*NIR* – Near InfraRed

*NOAA CPC* – National Oceanic and Atmospheric Administration Climate Prediction Center

*PERSIANN* – Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks

*PMW* – Passive MicroWave

*POD* – Probability Of Detection

*OI* – Optimal Interpolation

*RGB* – Red-Green-Blue

*RMSE* – Root Mean Square Error

*RMSEr* – relative Root Mean Square Error

*RT1* – first-order Radiative Transfer model

*SAR* – Synthetic Aperture Radar

*SM* – Soil Moisture

*SMAP* – Soil Moisture Active Passive

*SM2RAIN* – Soil Moisture to RAINfall

*S1* – Sentinel 1

*S2* – Sentinel 2

*TC* – Triple Collocation

*TRMM* – Tropical Rainfall Measuring Mission

*TS* – Threat Score

*Tsoil* – Soil temperature

*US* – United States

VIS/IR – VISible/InfraRed

# 1. INTRODUCTION

*In this chapter, a brief introduction related to the current state of the water cycle monitoring strategies, and their relevance for hydrological applications, is reported. Particular attention is devoted to exploit remote sensing data as support tools to the ground monitoring that has still the main role in the observation of the principal variables of the water cycles. On this respect, the Copernicus program of the European Union offers a plethora of products useful for the Earth Observation and specifically in the evaluation of the water cycle variables at high-resolution. In this thesis, two important quantities are selected due to their importance in many aspects of the water cycle, the human lives and the climate change: rainfall and river discharge. By exploiting the satellite data extracted by the Copernicus program new approaches are here developed and adapted to the new data source with the final purpose to generate satellite-based products for rainfall and river discharge.*

### 1.1 State of the art

Floods, drought and landslides are the water related natural hazards that cause the most serious damage to the environment, people and properties. The occurrence of these events is related to climate: wet soil moisture (SM) conditions and intense rainfall are often the drivers of flood and landslide events (CIABATTA ET AL., 2016), whereas precipitation absence and low soil moisture are clear indicators of an ongoing drought. According to International Panel on Climate Change (IPCC) 5th report, climate change is expected to aggravate the occurrence of these phenomena since extreme weather and climate events will step up (IPCC, 2013), increasing both the risks of floods and landslides and the period of drought in many regions worldwide. The current climatic crisis is also expected to increase the water scarcity: water consumption has already been steadily increased in the last century (KUMMU ET AL., 2016), and the climate change could cause a reduction of water availability together with an increasing of irrigation demand, further exacerbating the situation. Many areas will experience water scarcity due to this phenomenon, as it is already happening (ROCKSTRÖM ET AL., 2012).

In this framework, the presence of an adequate monitoring network capable of providing accurate and spatially detailed information related to the water cycle is fundamental for multiple fields and users. High-resolution data are particular important for:

- stakeholders and companies involved in water management, to increase their intervention capacities and reduce wastage;
- researchers, to improve hydrological models performances and capabilities, whose need of high quality input data with sufficient resolution characteristic is increasing along with the models complexity (SILBERSTEIN, 2006; RAGETTLI ET AL., 2013). High spatial resolution data have also already proven useful to increase the hydroclimatic hazard prediction models capabilities (CENCI ET AL., 2017), since their forecasting ability is dependant from the

knowledge of their triggering conditions (HANNAH ET AL., 2011; PONZIANI ET AL., 2012), and the most important effects of floods, landslides and droughts occur at small scale;

- insurance companies, which are gradually demanding high spatial resolution data, even at monthly temporal scale, with the purpose of developing index-based insurance for small-scale agricultures (ENENKEL ET AL., 2019; BLACK ET AL., 2016).

Space-based Earth Observation (EO) data are a highly strategic asset that can be exploited to obtain the desired information. EO technology can provide a global understanding of many of the essential variables governing the water cycle and monitor their evolution from global to basin scales. In the past few decades, a significant development of EO system was observed, with the launch of many EO missions that provide an unprecedented capacity to quantify several variables on a routine basis. Also the European Union has recognized the crucial role Space can play in tackling the climate change and it has developed its own program for monitoring the Earth: Copernicus. Copernicus, previously known as GMES (Global Monitoring for Environment and Security), is coordinated and managed by the European Commission. It is implemented in partnership with the Member States, the European Space Agency (ESA), the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), the European Centre for Medium-Range Weather Forecasts (ECMWF), EU Agencies and Mercator Océan. Copernicus scope is to collect and process data from multiple sources, such as earth observation satellites and in-situ sensors, and to provide users with reliable and up-to-date information through a set of services related to environmental and security issues. The services address six thematic areas: land, marine, atmosphere, climate change, emergency management and security.

In order to fulfil Copernicus scopes, ESA is building the largest earth observation satellite system in the world — six families of Sentinels have been defined, developed and put into orbit. Each Sentinel is composed using some of the most advanced technologies currently available. The vast majority of data and information delivered by the Copernicus Space infrastructure and the Copernicus services

are available and accessible to any citizen and any organization around the world on a free, full and open access basis. Data availability at this scale, frequency, and quality constitute a fundamental paradigm change in earth observation, particularly considering the open access strategy decided by the European Union. At the current state, three complete two-satellite constellations are in orbit (Sentinel-1, Sentinel-2 and Sentinel-3) plus an additional single satellite, Sentinel-5P, and the first of the two satellites composing Sentinel-6 mission. These sensors are producing around 15 Terabytes of data every day, which are constantly downloaded and processed by scientists and users all over the world.

However, this growing observational capacity is also increasing the need for dedicated research efforts aimed at exploring the potential offered by the synergies among different and complementary EO data records. With this purpose, this thesis proposes an analysis of the improvements that the use of the high quality data provided by Sentinel Missions can provide for the estimation of information related to the water cycle. Specifically, two of the most important variables of the water cycle are here selected to be estimated: precipitation, and more specifically its liquid fraction, the rainfall, and river discharge. These variables are indicated by the Global Observing Systems Information Center (GCOS) among the Essential Climate Variables (ECV), i.e., variables whose knowledge is necessary to understand the evolution of the climate, to assess the related risks and to develop mitigation and adaptation strategies. They have been selected for this study due to their crucial role in influencing human lives and their large impact on the territory, which allow to study these variables by using satellite data. Indeed, rainfall and river discharge are strictly linked to each other and to human activities, and their changes could have effects not only on local communities but also at global scale, e.g., causing variation on ocean salinity and thermohaline circulation (**Trenberth K. E., 2011; Nohara et al., 2006; Piecuch et al., 2018; Ahmed et al., 2020**).

Currently, in situ measurements are the most cost-effective and reliable options for both river discharge and rainfall monitoring at local scale (**Fekete et al., 2015; La Barbera et al., 2002**).

Accurate measurements of rainfall are generally obtained through devices such as rain gauges, disdrometers and weather radars. There exist different types of rain gauges, each with its own strength and weakness: accumulation gauges, tipping-bucket gauges, weighing gauges and optical gauges (STRANGEWAYS, 2006; TAPIADOR ET AL., 2012). According to the sensor type, it is possible to infer reliable precipitation or rainfall point measurements from each of them. Notwithstanding this, they all suffer from various source or errors, such as wind, evaporation, instrument error, spatio-temporal variation of the drop-size, percentage of frozen/liquid precipitation or splashing (MICHELSON, 2004; PETERSON ET AL., 1998). The more technologically sophisticated disdrometers and weather radars are instead able to detect the size of individual rain drops, increasing our understanding of the precipitation processes and our capacity to estimate and model it (SUN ET AL., 2017). Specifically, weather radars are able to survey large areas and to capture the spatial variability of rainfall fields, even if they are less precise than the punctual measurements obtained by rain gauge and disdrometers, due to errors sources as the variability of the drop size distribution, the reflections from non-weather targets and the beam lifting range effect, which can cause misdetection of precipitation (KIDD AND HUFFMAN, 2011). Moreover, their high cost is an obstacle for the development of a global radar network (HABIB ET AL., 2012; SAUVAGEOT, 1994). However, the rainfall spatial variability makes the current rain gauge network (Figure 1-1) inappropriate to describe in detail its distribution over the full globe, since the number of gauges is too scarce with respect to Earth surface and they are unequally distributed, being the majority of them located in the most developed countries (KIDD ET AL., 2017; VÖRÖSMARTY ET AL., 2001; DINKU, 2019).

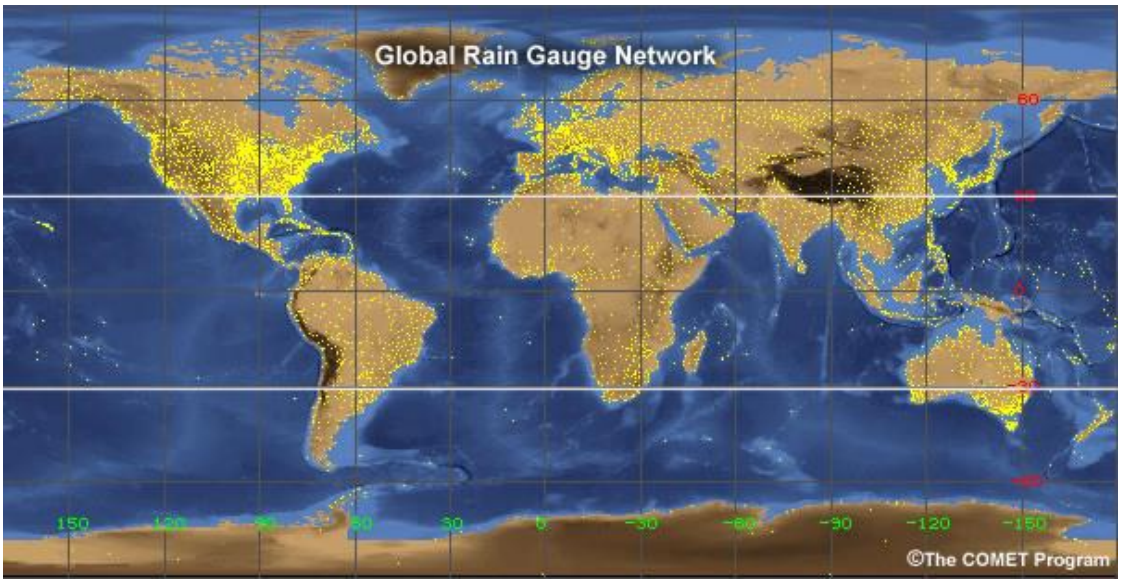


*Figure 1-1: Global Rain Gauge Network distribution (yellow dots). Produced by the Comet ® program.*

Similar discrepancies can be found also for the river discharge gauge network (Figure 1-2). River discharge is not directly measurable, being determined by the relationship between stream velocity and channel area. The most common way to measure it is the velocity-area method, which requires knowledge of water depth, channel width and flow velocity at each section of interest (**Herschy, 1995**). These data can be derived through a measurement campaign, by obtaining the flow velocity at single/multiple depths over multiple vertical cross-sections (**Pelletier, 1988**). River discharge is then calculated through the product between the stream velocity distribution and the channel cross-sectional area (**Gravelle, 2015**). The river discharge measured through these campaigns can be then associated to measurement of flow velocity (by current-meter or Acoustic Doppler Current Profiler, ADCP) or river stage through the use of rating curves. These mathematical formulas or calibrated relationships are derived by considering simultaneous measurements of river discharge and water velocity or level. Once the rating curve is calibrated, it is possible to monitor river discharge knowing only the water level or velocity. The effectiveness of the rating curves method is reduced for extreme river discharge and for channels in which the sediment transportation or deposition can alter the section area. Indeed, discharge is often not measured during flood events, thus limiting the effectiveness of the obtained relationships for estimate the relative flow and any change of the cross-

section area would introduce an error in the estimated river discharge. Notwithstanding this, the rating curves are used worldwide to estimate river discharge on gauging structures.

Nevertheless, number of working stations is declining worldwide (Figure 1-2) (CROCHEMORE ET AL., 2020; VÖRÖSMARTY ET AL., 2001). This is worsened by the fact that discharge estimates either do not exist or are not promptly available over many rivers, sometimes due to political reasons or economic evaluations (FEKETE AND VÖRÖSMARTY, 2007).

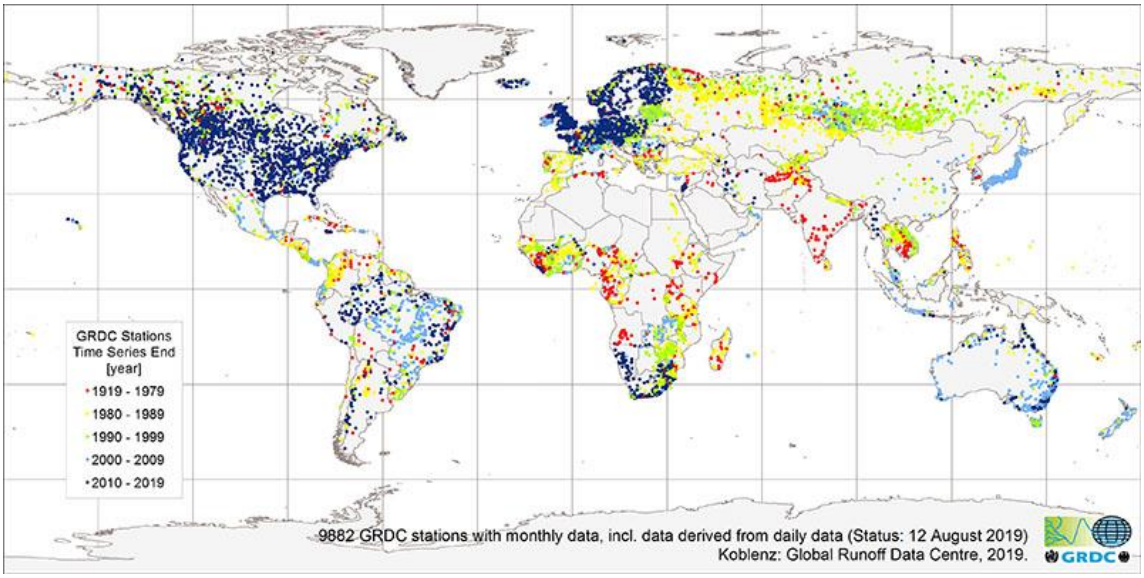


*Figure 1-2: Global Runoff Data Centre (GRDC) stations distribution over the last century.*

While in-situ networks are worldwide declining, technological progress in remote sensing sensors enables us to obtain data on the global scale, remarkably expanding humanity's understanding of its own living environment from spatial and temporal perspectives. The information of the ground-based networks can therefore be supported, integrated and in some cases substituted by EO data, which have proven to be a valuable source of information variables such as rainfall and river discharge (BARRET AND BEAUMONT, 1994; KIDD AND LEVIZZANI, 2011).

In the following sections, a brief review of the current state of the art of satellite rainfall (section 1.2) and river discharge (section 1.3) estimation methodologies will be given, together with the major limitations of the described approaches.

### 1.2 Remote sensing of rainfall

The classical remote sensing-based technique to estimate rainfall is the "top-down" approach (Brocca et al., 2014a), where the upwelling radiation or backscatter from clouds measured by satellite sensors are used to estimate the surface instantaneous rainfall rate. To date, many satellite sensors for observing the atmosphere are available, which can be classified into three categories according to the spectral sensitivity: VISible/InfraRed (VIS/IR), Passive MicroWave (PMW) and Active MicroWave (AMW) sensors. Different methods to derive precipitation have been developed for each type of sensor and also by considering their merging. VIS/IR based techniques are established on the principle that cold cloud tops are related to greater rainfall events, hence allowing to link IR cloud top temperature to rainfall rate through different algorithms, such as the Griffith-Woodley algorithm (Griffith et al., 1978), the Convective/Stratiform technique (Adler and Negri, 1988), and also through neural networks, as for the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) product (Nguyen et al., 2019). Still, the relation between precipitation and IR cloud temperature is indirect and the quality of the obtained precipitation data is erratic. Conversely, PMW radiometers proved to be a more direct method to infer precipitation from space, since this radiation is able to sense the size of the precipitation through clouds. The first PMW-based techniques to retrieve rainfall measurements were based on simple regression function applied between observed rainfall rates and sensed brightness temperature. In recent years, many different methods exploiting probabilistic physical and iterative algorithms have been developed to improve the quality of the estimated rainfall (Bauer et al., 2001; Casella et al., 2015; Wentz and Spencer, 1998). Notwithstanding this, the temporal frequency of PMW sensors is usually low (>3 hours), limiting the reliability of the total rainfall measurements. Conversely, VIS/IR observations are far more frequent (30 minutes or less), being most of the VIS/IR sensors onboard of geostationary satellites. Therefore, VIS/IR and PMW rainfall estimates are often merged together to increase accuracy, coverage and spatial and temporal resolution of the estimated precipitation. A well-known example of such methodologies is the Climate

Prediction Center morphing method (CMORPH), which derive motion vectors from IR data and use them to improve the precipitation estimates generated by PMW retrievals (JOYCE ET AL., 2004). Finally, the launch of the Tropical Rainfall Measuring Mission (TRMM) mission in 1997 introduced the first AMW observations from satellites: a single-frequency radar operating at 13.8 GHz capable of measuring the 3D rainfall distribution over both land and water. This led to the generation of the first precipitation product obtained from VIS/IR, PMW and AMW information (KUMMEROW ET AL., 2000). TRMM was followed by the Global Precipitation Measurement (GPM) mission, based on a constellation of satellites mounting different type of sensors. Their data were combined to develop the Integrated Multi-satellitE Retrievals for GPM (IMERG) precipitation product. The combined use of data from multiple polar and geostationary satellites, along with a new Dual-frequency Precipitation Radar (DPR) and an accurate radiometer (GMI) (HOU ET AL., 2014), allows to provide precipitation estimates characterized by high spatial and temporal resolutions (0.1 degree and 30 minutes, respectively) and global coverage. Despite the good level of accuracy achieved, the difficulties in obtaining and inter-calibrating near-real time observations from multiple agencies and the overall high cost for operation and maintenance of the whole satellite constellation are obstacles still to be overcome in order to guarantee the data continuity. Moreover, additional rainfall datasets are still needed to understand residual uncertainties and errors (MASSARI ET AL, 2017; CHEN ET AL., 2020) like, e.g., seasonal and local bias (MAGGIONI AND MASSARI, 2018). The integration of IMERG with alternative rainfall products can be also carried out to reduce uncertainties, as in MASSARI ET AL. (2020).

The recently introduced "bottom-up" approach points toward addressing these problems by inferring or correcting rainfall estimation over land using SM observations from satellite or gauges. This method provides accumulated rainfall estimates (CROW ET AL., 2009; BROCCA ET AL., 2013; PELLARIN ET AL., 2013) instead than the instantaneous rate, as for the "top-down" products. Many methods based on this approach share the same limitations, linked to the limits of measuring SM from space: rainfall estimated only over land, low accuracy in presence of dense vegetation or complex

topography and difficulties in estimating rainfall in case of soil saturation. Among the algorithms based on this approach, SM2RAIN (Brocca et al., 2014a) has distinguished itself for its versatility and simplicity. By inverting the soil water balance equation, this algorithm allows the estimation of the rainfall occurring between two SM measurements. It has already been applied worldwide to both regional (Ciabatta et al., 2015; Tarpanelli et al., 2017) and global (Ciabatta et al., 2018; Brocca et al., 2019) satellite SM products, obtaining satisfactory results, in particular over regions characterized by scarce rain gauge density (Massari et al., 2020). Different studies have shown the usefulness of SM2RAIN derived products for hydrological application such as flood and landslide prediction (Brunetti et al., 2018; Camici et al., 2018; Brocca et al., 2020).

Nevertheless, the major limitation of satellite observations, regardless of the adopted approach, is the inherent "technological" compromise between temporal and spatial coverage: SM and rainfall satellite products are usually characterized by a frequent revisit time (<1 day) and a coarse spatial resolution (~10-50 km). The multiple geographical features that appear inside each pixel's area (Figure 1-3a) affect the information obtained from the remote sensing sensor. Taking as example the SM estimation, when these features are characterized by different SM values from their surrounding (e.g., irrigated fields), low quality of SM estimation (e.g., mountainous area or cities) or even no SM value at all (e.g., water bodies), the overall quality of the obtained SM estimates decreases. By moving at higher spatial resolution (Figure 1-3b), the quality of the remote sensing SM product should increase, due to the possibility to better distinguish each geographical feature.

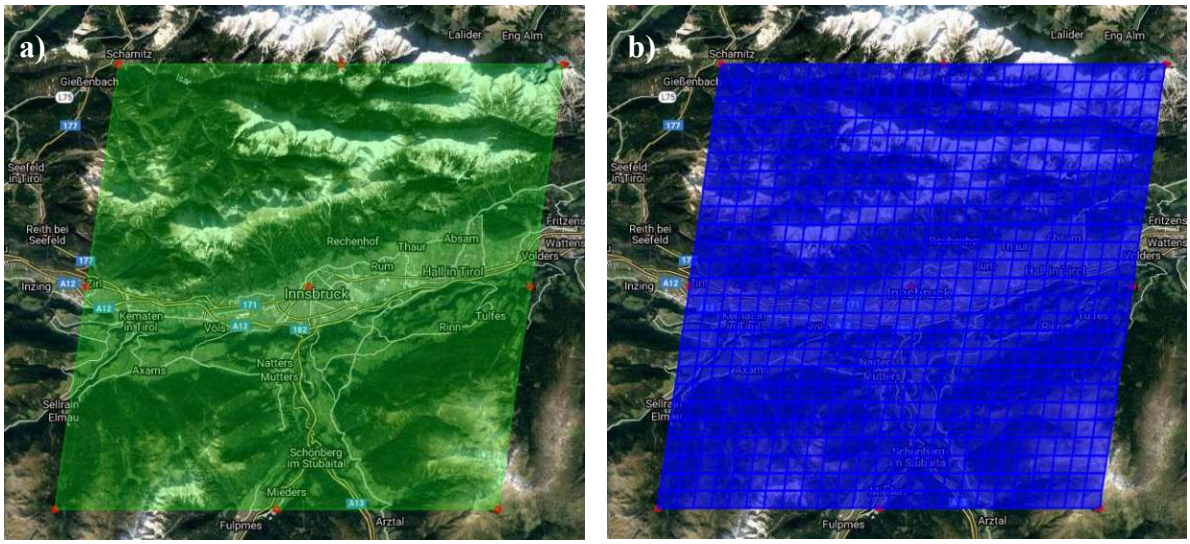


*Figure 1-3: Example of one pixel at coarse resolution (25 km spatial resolution) covering the city of Innsbruck, Austria, (panel a) and the correspondence 1km spatial resolution pixels (panel b)*

Therefore, it is of primary importance to obtain data with high temporal and spatial resolution, in order to enhance the prediction capability of hydrological models requiring high-resolution input data (MERLIN ET AL., 2008) and the spatial accuracy of the information related to water resources. The first attempt to accomplish the objective of high spatial resolution was the use of downscaling procedures. Many different approaches, from geostatistical analysis to data fusion, have been developed in the last years in order to obtain sub-pixel information from coarse resolution products (PENG ET AL., 2017) to be used in different applications (e.g., DARI ET AL., 2020). However, their results were often unsatisfactory, because of the limitations of the auxiliary data (e.g., cloud cover for optical data and model errors when using model data) and the uncertainties of downscaling algorithms (PENG ET AL., 2017).

Recently, the new launched Sentinel Missions of the Copernicus program, has opened new possibilities to overcome these issues. Specifically, Sentinel-1 (S1) mission is composed of two satellites that share the same orbit 180° apart and follow a strict acquisition scenario with a 12-days repeat cycle (6-days by considering both satellites), each carrying an identical C-band Synthetic-Aperture Radar (SAR) sensor capable of sensing high-resolution microwave backscatter (down to 5 meters). This setup leads to a revisit frequency of 1.5-4 days over Europe, thanks to the overlap of

the orbits. The condition of high spatial and medium temporal resolution is, for the first time, met by the two S1 satellites currently in orbit. SM measurements with 1 km spatial resolution can be obtained from this mission (Bauer-Marschallinger et al., 2018).

The application of SM2RAIN to such data could therefore allow to obtain high-resolution (1 km) rainfall estimates over land. This approach, however, is limited by the need to calibrate SM2RAIN algorithm against a rainfall product with spatio-temporal characteristics similar to those of the input SM. Datasets with such spatio-temporal characteristics are rarely available, due to the already mentioned scarce density of ground-based networks, thus limiting the calibration and validation of high-resolution rainfall from SM2RAIN only over a few selected areas.

### 1.3 Remote sensing of river discharge

Many EO satellite-based techniques for estimating river flow have been developed in the last years. These techniques can be divided in those who need to be calibrated and those which instead does not need calibration and can be applied also in ungauged basins (Gleason et al., 2020). Most of the latter are based on a simple evaluation of the excess of the water balance equation from hydrologic models (Parr et al., 2015; Lopez et al., 2018) or on the applying of hydraulic models without the use of any in-situ data (Schumann et al., 2007; Di Baldassarre et al., 2009; De Frasson et al., 2019). Sometimes these models are coupled with other satellite-based variables such as river width, water surface, river stage or slope measurements (Sun et al., 2015; Andreadis et al., 2007; Durand et al., 2008). Even though these methods are characterized by low accuracy, they are able to improve the poor information available on the ungauged rivers. The calibrated models are instead more reliable, being adjusted using in-situ data or a calibrated hydrologic model. They are useful to extend the river discharge knowledge in space and time. The majority of the EO satellite-based techniques to retrieve river discharge belong to this category. Hydrologic models can be used to improve the information of river discharge, also by using remote sensing information as input of the water balance equation, as, e.g., in Dziubanski et al. (2016), where a hydrologic model with

a remote sensing-based snow routine is used to improve the estimated river discharge of seven watersheds in the Upper Mississippi River basin. This is just one example of a rich literature in which one or more satellite-based signals are used in calibration/assimilation to retrieve river discharge information at regional (JODAR ET AL., 2018; SYED ET AL., 2005; WANG ET AL., 2015) but also at global scale (CHANDANPURKER ET AL., 2017; LIN ET AL., 2019; ZHANG ET AL., 2016). Another group of techniques relies on the use of remotely sensed hydraulic variables and in-situ data to derive river discharge information using hydraulic models to infer river discharge from functions of river width, depth and/or slope (e.g., BJERKLIE ET AL., 2003; 2005; 2007), sometimes coupled with hydrological models (e.g., NEAL ET AL., 2009). Finally, many methodologies are based on the concept of the space rating curves, where satellite-based measurements of river level, river width or other empirical fluvial geomorphic phenomena are coupled to the observed river discharge to obtain a relationship that can be used to estimate the latter. In this framework, river stage measurements from altimetric data are widely used because of the direct relation between them and river discharge (ABDALLA ET AL., 2021; BELLONI ET AL., 2021; PARIS ET AL., 2016; TOURIAN ET AL., 2013; ZAKHAROVA ET AL., 2020), but the spatial-temporal sampling of the altimetry missions is currently a limitation, leading to the need of applying different strategies to densify the data (TOURIAN ET AL., 2016; 2017; BOERGENS ET AL., 2017; SCHWATKE ET AL., 2015). The combination of altimetry and other satellite information can also be exploited to improve the river flow estimations quality (see TARPANELLI ET AL., 2021 for a review). In this respect, satellite optical sensors play an important role along with radar altimetry for improving the river monitoring. A widespread methodology to infer river discharge from optical data is the extrapolation of river width measurements from satellite imagery. This is possible thanks to the different behavior of water and soil in the optical bands, which make these features quite clearly distinguishable (YANG ET AL., 2020). River discharge measurements can be then obtained from the estimated river width by using empiric formulas, rating curves or models (ELMI ET AL., 2021; FENG ET AL., 2019; PAVELSKY ET AL., 2014), also in combination with altimetric data (SICHANGI ET AL., 2016). Notwithstanding this, river width

measurement is possible only in free clouds days (reducing the temporal availability of the obtained datum) and there are often misclassifications of water pixels due to cloud shadows or in areas characterized by steep terrain or dense urban environments (MUELLER ET AL., 2016). Focusing on daily measurements of river discharge, TARPANELLI ET AL. (2013; 2017; 2020) used passive remote sensing data from MODIS, an optical sensor characterized by spatial resolution up to 250 m, to develop a linear regression approach first introduced by BRACKENRIDGE ET AL. (2007), which is effectively used to monitor high and low flow in medium sized basins (<100'000 $km^2$). The approach (hereafter named $CM$) is based on the different behavior in the Near-InfraRed (NIR) region between a Calibration ($C$) and a Measurement ($M$) pixels chosen over land and water, respectively. Their ratio can be correlated with the in-situ measurements of river discharge. Indeed, in a bare soil or urban area the $C$ pixel has a reflectance behavior quite stable in time, sensitive mainly to atmospheric noises and factors like the bidirectional reflectance effects. Differently, a pixel around the bank of a river ($M$ pixel), often inundated, has generally lower reflectance values due to the presence of water. The ratio between the two pixels can be therefore related to river discharge (TARPANELLI ET AL., 2013; 2017).

Both pixels are affected by noises due to different sources, such as the atmosphere or the acquiring process, but the calculation of their ratio allows to account for them and therefore it is well correlated with the in-situ measurements of river discharge. Notwithstanding this, the presence of sediments in the river induces an increase in the water reflectance (MALINOWSKI ET AL., 2015), which corresponds to an underestimation of the river discharge assessed through $CM$ approach. The analysis of this phenomenon for medium-narrow rivers (100 - 300 m) can be facilitated by the use of high-resolution multispectral sensors to better represent the dynamic of such rivers. With respect to the medium resolution images used in the scientific literature(i.e. MODIS, MERIS and OLCI), data with higher spatial resolution coming from the new launched Sentinel-2 (S2) Mission of the Copernicus program (10, 20 and 60 m spatial resolution), have opened new possibilities to investigate small targets, as can be easily deduced by comparing two images in Red-Green-Blue (RGB) obtained by

MODIS (Figure 1-4a) and S2 (Figure 1-4b) at the Pontelagoscuro station, over the Po river in Italy during January 2019.

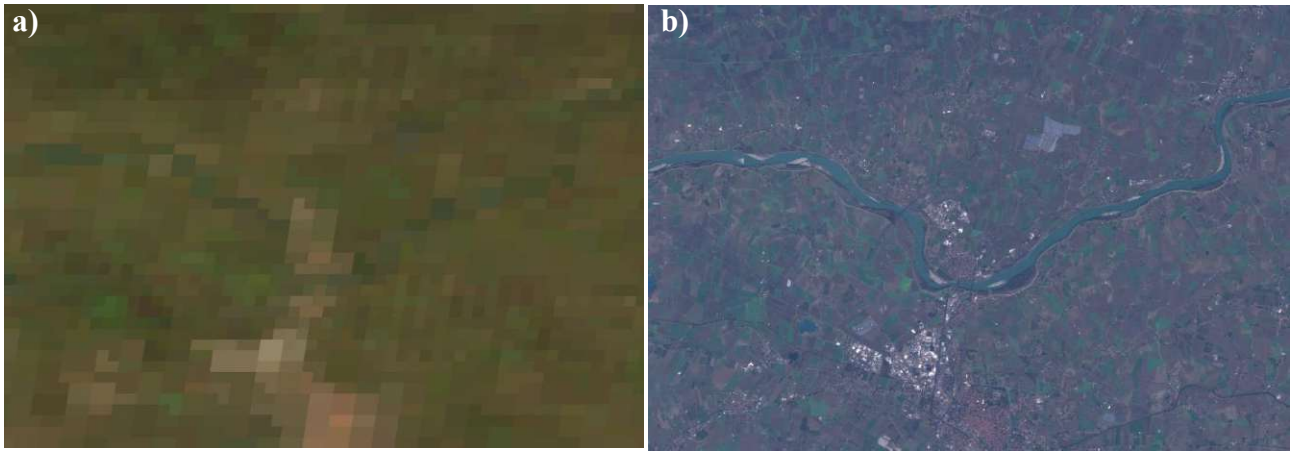


*Figure 1-4: RGB image of Pontelagoscuro, Po River, Italy, from MODIS (panel a) and S2 (panel b) sensors captured the 29-01-2019*

### 1.4 Purposes of the thesis

The aim of this thesis is to take advantage of the cutting-edge information available from the Sentinel missions of the European Union's EO program, Copernicus, to improve the capacities of the existing models and algorithms in observing the Earth, and, more specifically, the water cycle, with the final goal to support service providers, public authorities, and other international organizations involved in natural hazard mitigation and water resources management. With this objective, the data of S1 and S2 were selected to be used for the estimation of rainfall and river discharge, two variables fundamentals to understand and predict the evolution of climate, to develop mitigation strategies and to study adaptation measures. The analysis can be therefore conceptually divided in two branches: the use of SM2RAIN algorithm for rainfall estimation from a high spatial resolution SM product derived from S1 and the use of $CM$ approach for river discharge estimation from S2 data.

As mentioned above (Section 1.2), one limitation of SM2RAIN is the need of a benchmark dataset with spatio-temporal characteristics similar to those of the selected SM product to calibrate the algorithm, which allows to estimate rainfall only over areas where a rainfall product is already available. The first target of this work was therefore to develop a methodology to estimate the

SM2RAIN parameters values independently from a reference, i.e., a self-calibrated SM2RAIN product. Several datasets globally available of climatic and land descriptors were then collected to be used as potential predictors for the development of parametric relationships capable of retrieving SM2RAIN parameters values, to be applied to data from the Advanced SCATterometer (ASCAT), a medium resolution SM product (25 km spatial resolution). The descriptors were then analyzed on a subset of points selected over areas where high quality rainfall observations were available, and, once the relationships were developed, they were validated worldwide through comparison with different global rainfall products: GPM IMERG Final Run, GPM IMERG Late Run and the Global Precipitation Climatology Centre, (GPCC). The performances of the parameterized SM2RAIN product were hence compared with those resulting from the calibration of SM2RAIN in the same period, with the aim of assessing whether the two products performances were in line with each other.

Once the reliability of the parametric relationships was assessed, both the parameterized and calibrated versions of SM2RAIN algorithm were applied to a high-resolution (1 km) SM product derived from S1 over the Po River basin in northern Italy, with the final scope of evaluating their capabilities in reproducing rainfall at such spatial resolution. This SM product, from here on named S1-RT1, was obtained by using a retrieval algorithm based on the first order solution of Radiative Transfer equation (RT1, **Quast et al., 2019**) on S1 data, and is characterized by 500 m spatial sampling and 1.5-4 days temporal resolution. Po River basin area was selected as study area because of the availability of a ground-based rainfall dataset with very high spatial and temporal resolution and because it comprehends many geographical features, such as mountains, hills, lakes, rivers and plains, which make it a good test area for this analysis. Both SM2RAIN versions were applied also to ASCAT-derived SM, after it was regridded to S1-RT1 coordinates, in order to assess the benefits derived from the use of high-resolution SM by comparing the resulting rainfall products performances.

The second target of this thesis is to understand the benefits of high-resolution imagery for river discharge estimation, by extending the procedure to small targets and by testing its limits. In particular, three main targets were identified:

- estimate river discharge over medium/narrow rivers, specifically at Pontelagoscuro gauged station over the Po River in Northern Italy, characterized by a dense record of hydraulic variables, and at Montemolino gauged station over the Tiber River in Central Italy, selected for its size of about 50-60 m width;
- modify the $CM$ approach to take into account the impact of the solid transport, exploiting the high-resolution of S2, so as to improve the estimation of river discharge;
- investigate the characteristics of those pixels that are more sensitive to river discharge variations, making a first step toward an uncalibrated procedure to select $M$.

In order to foster the analysis to larger time scales and in other areas, the code was implemented in the Google Earth Engine, GEE, platform (GORELICK ET AL., 2017), reducing by far the computational time and capacity required to apply the methodology.

The present work is a collection of papers published in or submitted to peer-reviewed journal, showing the results and the insights found during the Doctorate research period. The thesis is structured as follows.

Chapter 2 contains the methods explanation: the original SM2RAIN algorithm and the methodology for the parametric relationships' development are descripted in section 2.1, section 2.2 illustrates the original $CM$ approach for river discharge estimation and the proposed modifications, and the performance metrics and analysis used for the validation of the products of rainfall and river discharge are presented in section 2.3.

The SM2RAIN relationships development analysis is reported in chapter 3. First, all the data employed or tested in the procedure are descripted in section 3.1, then the descriptors' selection

procedure is explained (section 3.2) and the obtained parametric relationships are shown and commented in sections from 3.3 to 3.5. Finally, the results of the parameterized SM2RAIN procedure and the comparison between the parameterized and calibrated SM2RAIN products are illustrated for the local (section 3.6) and global (3.7) analysis.

Chapter 4 illustrates the rainfall estimation from high-resolution SM from S1 through the parameterized SM2RAIN. Specifically, section 4.1 describes the Po River basin characteristic, section 4.2 presents the datasets employed in the analysis, section 4.3 shows the results of the rainfall estimation and, finally, the evaluation of the benefits from the high-resolution information is descripted in section 4.4.

The river discharge analysis is explained in chapter 5. The two study areas and the satellite and ground datasets are introduced in section 5.1 and section 5.2 respectively, followed by the cloud platform GEE description in section 5.3. The results of the proposed model application for Pontelagoscuro station are shown in section 5.4, while the ones obtained at Montemolino station are presented in section 5.5.

Chapter 6 presents the main conclusions of the work and an outline for future research and applications.

Finally, chapter 7 contains the Appendix to chapter 3 and 4, briefly deepening the TC analysis of the parameterized SM2RAIN algorithm applied to ASCAT (section 7.1), the causes of low rainfall accuracy over the Po River basin (section 7.2) and the time accumulation strategy for the analysis of the different high-resolution rainfall products (section 7.3).

# 2. METHODS FOR RAINFALL AND RIVER DISCHARGE ESTIMATION

*In this chapter, the methods adopted for the estimation of the rainfall and the river discharge are descripted, together with the indexes used to estimate their performances. More in details, section 2.1 describes SM2RAIN algorithm for rainfall estimation, the procedure derived to obtain the parametric relationships and the experiment designed for their validation; section 2.2 presents the original CM approach for river discharge estimation, the applied modifications and its new formulation; and section 2.3 illustrates both the classical scores adopted to estimate the performances of the estimated products and the triple collocation methodology.*

## 2.1 Rainfall estimation from Soil Moisture: SM2RAIN algorithm

---

This section will present the algorithm adopted for rainfall estimation, SM2RAIN. After describing the concept behind the technique, the procedure to derive the parametric relationships over selected pixels will be explained, followed by the experiment designed for their validation worldwide. The section is based on the following publication, published by Journal of Hydrology:

*Filippucci, P., Brocca, L., Massari, C., Saltalippi, C., Wagner, W., and Tarpanelli, A.: Toward a self-calibrated and independent SM2RAIN rainfall product, Journal of Hydrology, 603(A), 126837, doi: 10.1016/j.jhydrol.2021.126837, 2021.*

---

*Original SM2RAIN algorithm*

SM2RAIN is an algorithm developed by BROCCA ET AL. (2013; 2014A) to estimate the accumulated rainfall precipitated between two SM measurements. This result can be achieved by inverting the soil water balance equation. It was successfully applied to different satellite and in situ SM dataset (CIABATTA ET AL., 2018; BROCCA ET AL., 2019; FILIPPUCCI ET AL., 2020) offering good results, especially in poorly gauged regions (MASSARI ET AL, 2020).

Considering a layer characterized by a soil depth $Z$ [mm] and a soil porosity $n$ [$m^3/m^3$], the soil water balance equation can be written as:

$$Zn\, {dSM(t)}/{dt} = p(t) - r(t) - e(t) - g(t) \quad (1)$$

where $SM(t)$ is the relative SM [-], i.e., the soil moisture saturation fraction, $p(t)$ is the rainfall rate [mm/d], $r(t)$ is the surface runoff rate [mm/d], $e(t)$ the evaporation rate [mm/d] and $g(t)$ the drainage rate [mm/d]. During rainfall events and unsaturated conditions, evaporation and surface runoff rates can be considered negligible (BROCCA ET AL., 2015). Equation (1) can therefore be rewritten, by using FAMIGLIETTI AND WOOD (1994) relationship to express the drainage rate, as:

$$p(t) = Z^* \, {dSM(t)}/{dt} + a\, SM(t)^b \quad (2)$$

with $Z^* = Zn$, $a$ [mm/d] is the saturated hydraulic conductivity and $b$ [-] is the exponent of the Famiglietti and Wood equation. Remotely sensed SM tends to be noisy and it is sensitive to a thin topsoil layer (few centimetres). Therefore, the exponential filter approach (WAGNER ET AL., 1999; ALBERGEL ET AL., 2008) is applied to satellite SM observations before their use in equation (2). The estimation of rainfall is therefore obtained by the knowledge of two consecutive SM measurements together with 4 parameters: $Z^*$, $a$, $b$ and $T$, the time constant of the exponential filter. In its standard application, the parameter values are estimated by calibrating SM2RAIN against reference rainfall data with similar spatial and temporal resolution, with the objective of minimizing the Root Mean Square Error (RMSE). BROCCA ET AL. (2019) introduced also a modified version of the exponential

filter, in which the characteristic time length, $T$ is decreasing with increasing SM according to a 2-parameter power law (Brocca et al., 2019). This version allows to slightly improve the obtained rainfall performances but increase the number of required parameters. It was therefore discarded for the estimation of the parametric relationships, while it was adopted for comparison in chapter 5, in order to assess the parameterized SM2RAIN results against the best version of the calibrated SM2RAIN.

*Procedure for the parametric relationships*

The methodology used to obtain the four parametric relationships is described here. As a first step, 1009 points uniformly distributed over the areas covered by rain gauges in Australia, India, Italy and US were selected (see Figure 2-1 and **BROCCA ET AL., 2019**). Each point is representative of an area of 25 x 25 km$^2$, and the spacing between the points is around 1 degree.

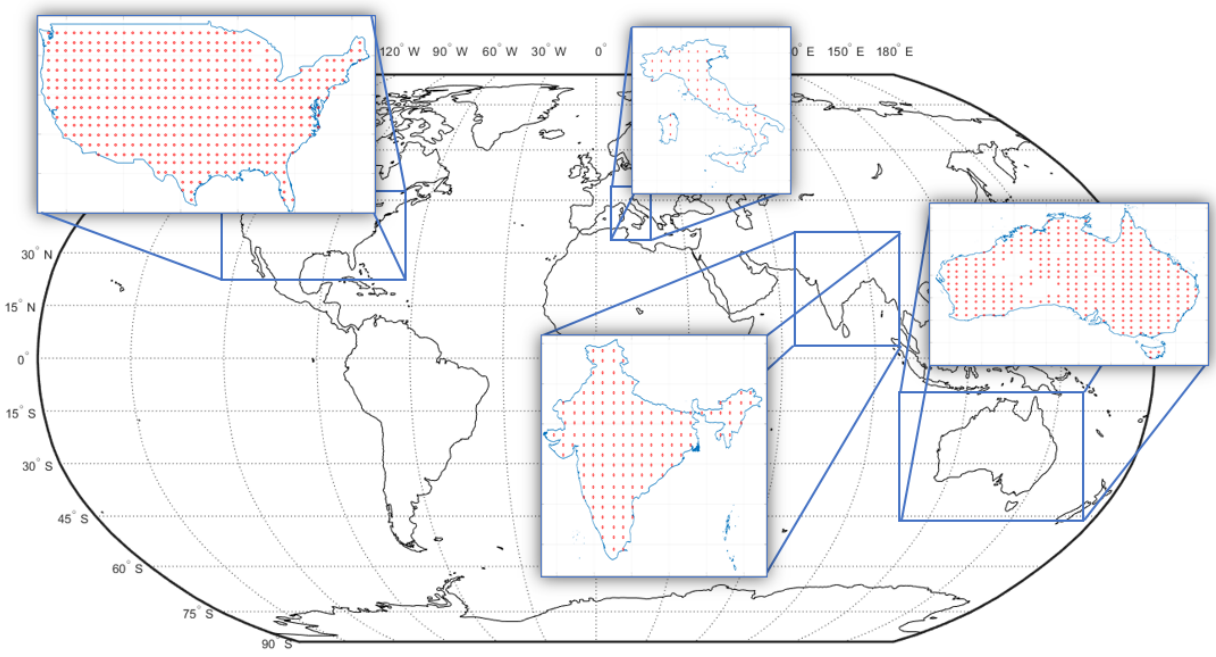


*Figure 2-1: 1009 points grid for the local analysis, uniformly distributed over the areas covered by rain gauges in Australia, India, Italy and USA.*

Secondly, the climatic and land descriptors selected for this study, describing climatic (rainfall and evapotranspiration) and land (soil texture and soil type, SM, soil temperature, topography and vegetation cover) characteristics (see chapter 3), were spatially interpolated to the 1009 points. For

the time-varying quantities (e.g., rainfall and soil moisture), different statistics were computed for each point to obtain the descriptors. Specifically, the daily annual average and the average number of rainy days per year was calculated for rainfall, therefore obtaining information about the climatology of each point. To obtain them, first the percentage of rainy days (rainfall > 0) and the average rainfall were calculated for each day of the year (DOY) using the available years, then the average values were calculated to obtain the annual average values. For SM, SM noise, soil temperature and evapotranspiration, the mean, median, maximum, minimum, standard deviation and coefficient of variation were computed in each point of the selected area. The temporal difference of consecutive SM and SM noise measurements was also considered, since the variation of SM is exploited in SM2RAIN to obtain rainfall. The mean, median, maximum, minimum, standard deviation and coefficient of variation were therefore calculated also over these differences, considering both the actual and the absolute values, and just the positive and negative variations, since these should be related to different mechanisms of wetting and drying of the soil (e.g., the average value of the positive variation of SM, or the maximum of the absolute variation of SM noise). For the elevation, the mean and the standard deviation within each pixel was computed; particularly the standard deviation of elevation is an indicator of topographic complexity. The latter decreases the accuracy of soil moisture retrievals, due to shadowing effects and layover (a distortion that occurs in radar imaging when the signal reflected from the top of a tall feature is received by the emitter before the one of the base, **ULABY ET AL., 1981**).

Thirdly, the potential relationship between the descriptors and each SM2RAIN parameter values was analyzed through the Spearman's correlation index. Spearman's correlation index was calculated between each parameter and each descriptor. In order to obtain the parametric relationships, only the descriptors who showed a high absolute value of Spearman's correlation (greater than 0.6) with the related parameter were considered.

SM2RAIN parametric relationships were finally obtained through a stepwise non-linear backward approach: all the possible additive and multiplicative combinations between the selected descriptors were initially used in a multilinear regression algorithm to obtain a first estimation of the relationship. An exemplary formula for the combination of $n$ descriptors is here reported:

$$par = \alpha_0 + \sum_{i=1}^{n} \beta_i d_i + \sum_{i=1}^{n} \sum_{j=1}^{n-\{i\}} \gamma_i d_i d_j + \cdots + \omega \prod_{i=1}^{n} d_i \quad (3)$$

where $\alpha_i$, $\beta_i$, $\gamma_{ii}$, and $\omega$ are the coefficients to be estimated and $d_i$ is a descriptor. The procedure was repeated iteratively by eliminating, at each step, the less significant factor, until an optimal combination of limited number of coefficients (minor or equal to 3) and good performance (drops in Spearman's correlation in comparison with the previous step < 0.015) was reached. It was also verified that the Spearman's correlations between each factor used in the relationship were below 0.2, in order to avoid the cross-correlation between the factors used in the parametric relationships.

*Procedure for the parametric relationships validation*

Once obtained the parametric relationships, those were validated by applying them worldwide using ASCAT SM data, a satellite derived SM product available from 2007 onward (see chapter 3 for the data description). In order to assess the goodness of the obtained results, the parameterized SM2RAIN rainfall product was compared with the SM2RAIN-ASCAT rainfall product obtained by calibrating SM2RAIN with the standard approach using ERA5 rainfall as reference (see chapter 3 for the data description). SM2RAIN-ASCAT was calibrated in the full available period 2013-2019, in order to compare the parameterized product with the best possible SM2RAIN version. It has to be noted that the standard calibration results presented in this chapter are different from those obtained in **BROCCA ET AL. (2019)** who applied a different filtering approach and the climatology correction based on ERA5 (not considered here). Both the SM2RAIN derived rainfall products were compared with a benchmark dataset and by using TC (section 2.3). All the products involved in the validation were re-gridded to ASCAT grid (12.5km spacing), using the weighted average of the overlapping areas for datasets with a spatial resolution finer than 25 km and the weighted average of the four nearest pixels

for the remaining datasets. SM2RAIN method is applicable everywhere, but the reliability of the estimated rainfall depends on the reliability of the estimated SM. This excludes all the areas with high vegetation regime, where C-band microwave measurements cannot reach the soil, coastal areas, wetlands, topographically complex areas, region characterized by subsurface scattering (**MORRISON ET AL., 2019**), and frozen or snow cover terrains (**HAHN ET AL., 2018**). It was therefore defined a committed area with high confidence in the successful retrieval of surface soil moisture from ASCAT by excluding the aforementioned categories. The committed area is obtained from the EUMETSAT H SAF product validation report (**HAHN ET AL., 2018**).

## 2.2 River discharge estimation from Near-InfraRed data: CM approach

---

This section describes the method to obtain river discharge from NIR data, the CM approach. Following the original method explanation, the concepts behind the approach development will be illustrated, and finally the new formulation will be given. This section is based on a preprint of the following publication, submitted to Remote Sensing of Environment:

*Filippucci, P., Brocca, L., Bonafoni, S., Saltalippi, C., Wagner, W., and Tarpanelli, A.: Sentinel-2 high-resolution data for river discharge monitoring, submitted to Remote Sensing of Environment, 2021.*

---

*Original CM approach adapter to S2 data*

TARPANELLI ET AL. (2013) developed an approach applied to estimate river discharge from reflectance data based on the different response of water and soil to radiation in the NIR band. In this region of the electromagnetic spectrum, the reflectance of soil and vegetation is much greater than that of pure water. Therefore, the reflectance of a "wet" pixel, located near the boundaries of the river, changes with the river discharge variation and, specifically, it decreases when river discharge increases and the pixel is flooded. Changes in reflectances are not caused only by changes in river discharge, but could be also related to other sources, such as vegetation, atmosphere composition, soil moisture variations and bidirectional reflectance effects, not negligible over water bodies (KREMEZI AND KARATHANASSI, 2019). A "dry" pixel not affected by river discharge variation is then selected to correct the measure, assuming to be affected by the same sources of noise, as mentioned above. By defining the "dry" pixel as Calibration ($C$) pixel and the "wet" pixel as Measurement ($M$) pixel, it is possible to obtain a proxy of river discharge by calculating the NIR reflectance ratio $C/M$. The ratio is then filtered with an exponential filter (WAGNER ET AL., 1999) to reduce the effect of residual noises. The locations of $C$ and $M$ are generally calibrated by calculating all the possible combinations of the reflectance ratio $C/M$ on a certain area in the vicinity of the gauged stations and comparing each single combination with the in-situ measurements of river discharge. The reflectance ratio providing the highest correlation defines the best location for $M$ and $C$ (see TARPANELLI ET AL., 2013 for details).

The passage from medium to high spatial resolution images requires several adjustments. The first is related to the sensitivity to cloud of the NIR band. The solution adopted by TARPANELLI ET AL. (2013) for the medium resolution product consisted in eliminating all the images affected by clouds. In this work, instead, the S2 cloud probability product is used to mask each image's pixels with a probability greater than 50% of being cloudy in each S2 image. The threshold of 50% was chosen by noting that in the selected areas the cloud detection algorithm may occasionally confuse river pixels with clouds and assign them a high cloud probability value. A visual inspection of each NIR images

was carried out to verify the chosen threshold and to discard those in which the river is not clearly visible, i.e., in case of failure of the cloud detection algorithm or in case of too few valid pixels in the image.

The cloud masking has the side effect to replace clouds with "holes" of no-data in each image. The selection of a single pixel for $C$, as in the original formulation, should be therefore avoided in order to prevent the loss of data in case a cloud appears over the selected "dry" pixel. Following TARPANELLI ET AL. (2017A), the $C$ timeseries is defined as the average reflectance value of the pixels that have a low temporal coefficient of variation (specifically lower than 5th percentile) with respect to all the pixels of the selected area.

Regarding the $M$ pixel, as previously noticed, it should be selected over the river boundaries, since this area is the interface between water and soil and it is particularly sensitive to river discharge variations. The selection is constrained to an extended water mask that includes the river and a buffer around it. The European Commission's Joint Research Centre (JRC) Global Surface Water map (PEKEL ET AL., 2016) was used to identify the extended area and a buffer of 3 pixels is used to enlarge the area. To assess the best location for $M$, each pixel within this extended water mask was selected to calculate the ratio $C/M$ (except the pixel already selected for the calculation of $C$ if any are included in the water mask). Then, the exponential filter was applied, using a value of the filter $T$ parameter equal to 5, which is the median number of days between two consecutive acquisitions of S2 images. The Spearman's correlation between the observed river discharge and the obtained filtered $C/M$ series was then calculated, as shown in Figure 2-2b for Pontelagoscuro station. By comparing Figure 2-2a with Figure 2-2b, it is evident that the highest correlations for the $M$ pixel coincide with the sediment deposit not vegetated near the river bank.

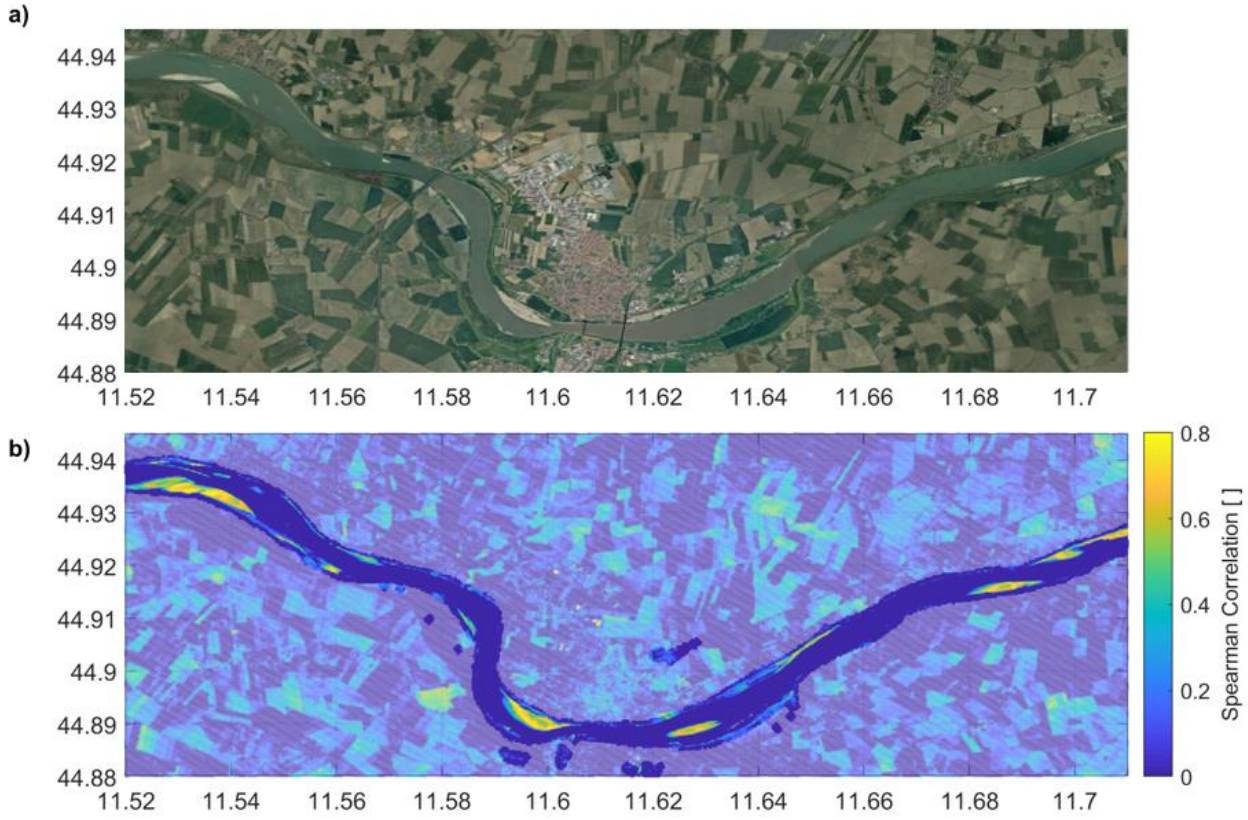


*Figure 2-2: Po River at Pontelagoscuro station. Panel a) represents the location of the area from Google Earth Copyright ©2021, CNES / Airbus, European Space Imaging, Landsat / Copernicus, Maxar Technologies). Panel b) shows Spearman's correlation between filtered C/M and the ground observed river discharge, obtained varying the location of pixel M and fixing C as the average of the pixels with coefficient of variation lower than 5th percentile calculated between January 2018 - September 2019.*

*Modifications of CM approach*

The expected benefits of the use of satellite imagery at high spatial resolution consist in the possibility to observe and account for aspects of river discharge estimation that are not clearly observable from lower resolution products. In this analysis, two main factors were considered to improve the $CM$ approach: the presence of sediments in the water and the selection of multiple pixels for $M$. In order to better observe these phenomena, the analysis was carried out in a smaller area of the Po River (Figure 2-3a), within the coordinates [11.588 °E, 11.605 °E, 44.887 °N, 44.9 °N], in the period January 2018-September 2019.

The reflectance of clear water in the NIR band is very close to zero, but the reflectance of the river generally appears to be different from zero, although it remains low compared to soil or vegetation.

During high flows, rainfall events mobilize large amounts of fine sediment into river systems (KEESSTRA ET AL., 2019), which contribute to the suspended sediment loads in the river and thus to increase the reflectance (AHN AND PARK., 2020). In these cases, the expected decrease in $M$ pixel reflectance due to inundation is hindered by the reflectance increase caused by the change in suspended sediments in the river. As a consequence, the ratio $C/M$ fails in high flow estimation, when the measure is most useful. This phenomenon is clearly visible in Figure 2-4a, where the hydrograph of the river discharge at Pontelagoscuro during the flood event of October-November 2018 is shown along with the ratio $C/M$ calculated by selecting $M$ as a single pixel, for which the Spearman's correlation with the observed discharge is the highest (see the location of $M$ in Figure 2-3b). By the observation of two images acquired before and in concomitance with the flood event (Figures 2-4b, 2-4c), the river becomes larger and strongly turbid, increasing its reflectance up to values very close to the surrounding soils (see Figure 2-4c). During and after the flood event, the suspended sediments decrease due to the catchment sediment dynamic (KEESSTRA ET AL., 2019; GENTILE ET AL., 2010) that is different from the river discharge dynamic. Consequently, $M$ decreases and the ratio $C/M$ increases (see Figure 2-4a), leading therefore to the late detecting of the inundation. In order to correct the problem, particularly important during the hydrograph rising limb, a correction should be introduced in the formulation taking into account the role of the suspended sediments.

Another important aspect is strongly related to the pixel dimension of S2. Due to the high frequency of the NIR electromagnetic waves and to the water absorption properties in this band, the penetration depth of the NIR radiation into water is in the order of a few centimetres. A completely flooded pixel would therefore lose its sensitivity to the water level increasing and thus, to the river discharge variation. A single pixel of S2 is too small to account for the variation of the river discharge over the study area because the width of the river during a flood event is much larger than the dimension of a single S2 pixel (compare Figure 2-4b and 2-4c for the Po River width variation). Indeed, by looking at Figure 2-4a, the $C/M$ ratio obtained from the maximum correlated pixel shows a limited variability

during the low flow from July to October 2018, being the chosen *M* pixel located over completely dry soil. In the high-resolution product analysis, the single pixel can be affected by both long period of dry condition (with reflectance very similar to the soil) alternated with period of saturated conditions (with reflectance very similar to the water inside the river). The flow variation signal, in such case, appears only when the chosen pixel rapidly changes its condition from dry to wet and cannot be used to properly represent the comprehensive river discharge variation.

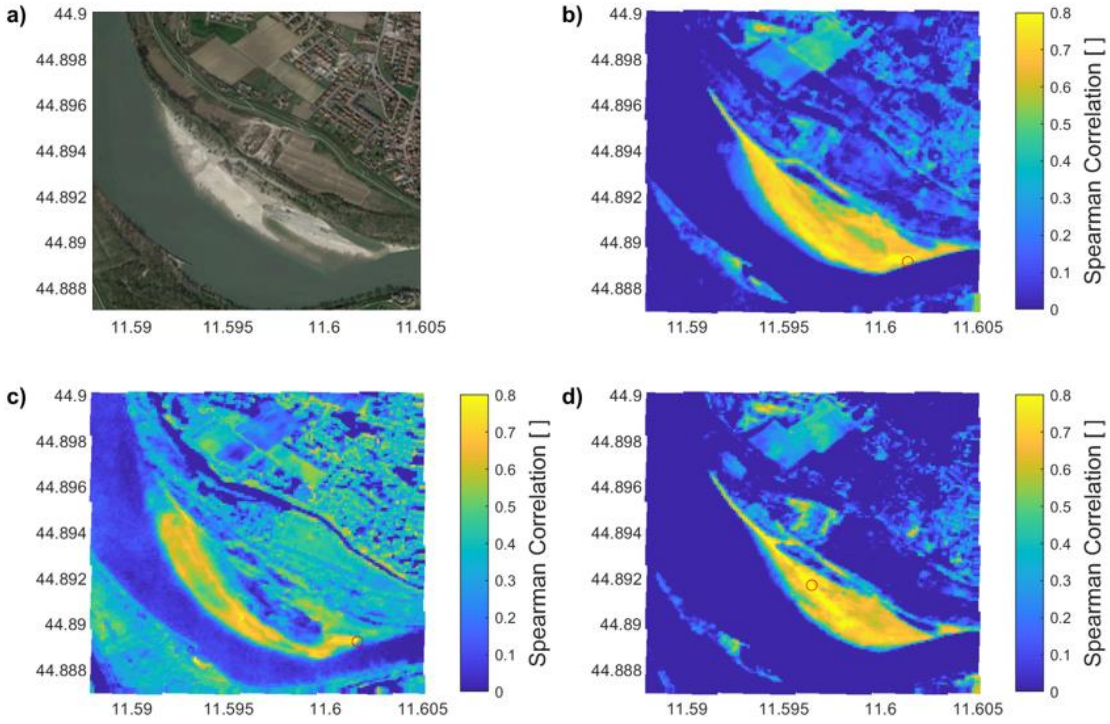


*Figure 2-3: Po River at Pontelagoscuro station. Panel a) represents the area from Google Earth (Copyright ©2021 Immagini ©2021, CNES / Airbus, Maxar Technologies). Panels b), c) and d) show Spearman's correlation between filtered C/M and the ground observed total discharge (b), discharge below the 50$^{th}$ percentile (c), high discharge above the 50$^{th}$ percentile (d). The red circles in panels b, c, d represent the position of the M pixel where the Spearman's correlation with the ground-based observations is highest.*

This is confirmed by the location of the pixel that is maximum correlated with the ground observed river discharge: if we consider the ground observed river discharges below the threshold of 50th percentile (around 1000 m$^3$/s in the selected period), the highest Spearman's correlations are obtained more internally into the river, as shown in Figure 2-3c. If we consider instead high flows, above the threshold of 50th percentile, the Spearman's correlations are higher in the external area (Figure 2-3d). Based on these considerations, the choice of a single pixel for all the river conditions is not

advisable: $M$ should be based on multiple pixels, chosen so as to take into account nearly all the river regimes.

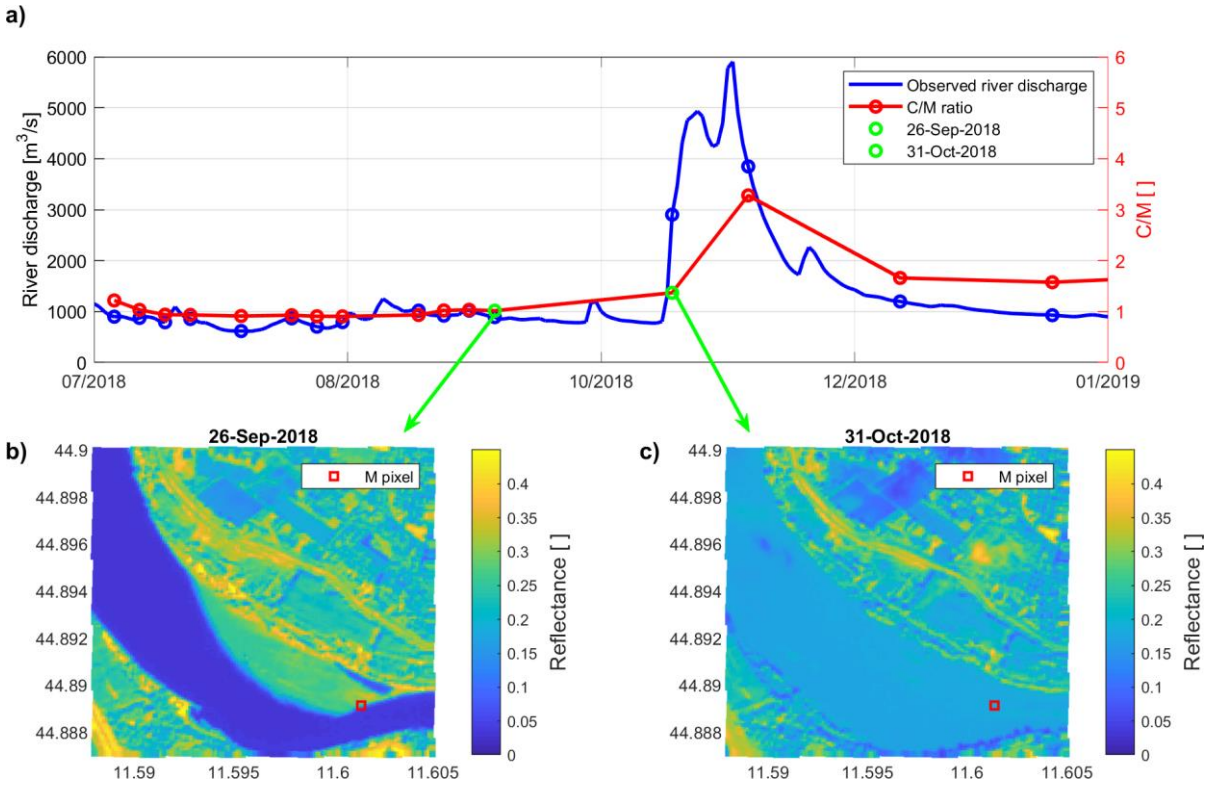


*Figure 2-4: Po River at Pontelagoscuro station. Panel a) shows the hydrograph of the river discharge and the C/M ratio calculated for the analyzed area. The circles on the temporal series represent the days when the satellite overpasses. Panel b) and c) show the S2 reflectance images before and during the October-November 2018 flood event, respectively. The red squares in the panels b and c represent the position of the M pixel, where the Spearman's correlation with the ground-based observation is the highest.*

*A new formulation for the CM approach*

Following the considerations described in the previous paragraph, two main modifications of the $CM$ approach are required: the management of suspended sediments effect and the use of multiple pixels for $M$ selection.

The role of suspended sediments is important for a correct evaluation of the reflectance of the $M$ pixel. As already specified, the reflectance of $M$ during high flows should decrease, therefore its value has to be corrected by the presence of suspended sediments that make the water turbid and increase its reflectance. The reflectance of the turbid water can be obtained by averaging multiple water pixels, $W$, located over a region of the river constantly wet (inner part of the river, with water always present,

also during low flows). Assuming $W$ reflectance as a proxy of the presence of suspended sediments, $M$ should vary between $C$ for low flows (dry conditions) and $W$ for high flows (wet conditions).

The original $CM$ approach for the estimation of the river discharge, Q, assumed as:

$$Q \propto CM = \frac{C}{M} \tag{4}$$

can be then reformulated as:

$$Q \propto CMW = \frac{C}{M-W+z} \tag{5}$$

where $z$ is a numerical coefficient: in condition of high flow, $M$ and $W$ are in fact very close each other and any noise within their measurements can cause their difference being lower than 0. Furthermore, in case the selected $M$ pixel become completely flooded, the condition of $M = W$ would be verified and the $CMW$ ratio tends to infinity. For both these reasons, we introduced $z$ as

$$z = \max(W - M) + \min(M) \tag{6}$$

$W$ pixels have to be selected within the river. This condition can be ensured by considering only the pixels included within a water mask, and by calculating, for each pixel, the product between the temporal mean of the reflectance (the lower, the more probable presence of water) and its temporal standard deviation (the lower, the less variation occurs in the pixel), in order to identify the pixels where water is always present. The $W$ reflectance of Eq. (5) was then calculated by averaging the NIR reflectance of the pixels, for which the product above defined was below the 5th percentile.

The timeseries of $M$ should be obtained by averaging reflectances from multiple pixels located in an area within the water mask that is particularly sensitive to variation in river discharge. In this analysis, we investigated several configurations to analyze the best strategy: 1) single pixel; 2) 3x3 pixels; 3) 9x9 pixels; 4) multiple pixels. In particular, configuration 1) was investigated to be compared to the original formulation. Configurations 2) and 3) were investigated to identify possible improvements in the evaluation of multiple pixels. Finally, configuration 4) was included for the purpose of not

calibrating the location of the $M$ pixels with the ground observed discharge timeseries, but to approach an automated procedure to select an area or a set of pixels that is particularly suitable for describing the river discharge range variation. This area was identified following the concept that the NIR reflectance of pixels part of it should be uncorrelated with $W$ or/and $C$ reflectance timeseries. Specifically, pixels sensitive to low flow should be poorly correlated with $C$ (because they are inundated most of the time), while those sensitive to high flow should be poorly correlated with $W$ (because they are dry most of the time). Therefore, the condition that a pixel has to meet for being selected as one of the optimal $M$ pixels is a very low Spearman's correlation in reflectance with the pixels of $W$ or $C$ (lower than the 10th percentile). In addition, to avoid selecting permanently flooded pixels (because they are poorly correlated with $C$), the Spearman's correlation with $W$ must be less than 0.7.

Summing up, the new procedure to estimate river discharge from high-resolution NIR reflectance data consists into the following steps (see also Figure 2-5):

1. The water mask is delineated through the "max extent" from JRC Global Surface Water map by considering a buffer of 3 pixels around each pixel of the water map;
2. each image where the field "percentage of cloudy pixels" of the S2 product indicates a percentage of cloudy pixels greater than 70% is directly discarded and each image's pixels in which the cloud probability is greater than 50% is masked out. Furthermore, the images with less than the 20% (empiric threshold) of valid pixels over the extended water mask are also discarded;
3. the $C$ timeseries is calculated by averaging the reflectance of those pixels showing a temporal coefficient of variation lower than the 5$^{th}$ percentile;
4. the $W$ timeseries is calculated by averaging the reflectance of those pixels, included in the water mask (step 3), whose product between temporal mean and temporal standard deviation of reflectance values is below the 5$^{th}$ percentile;

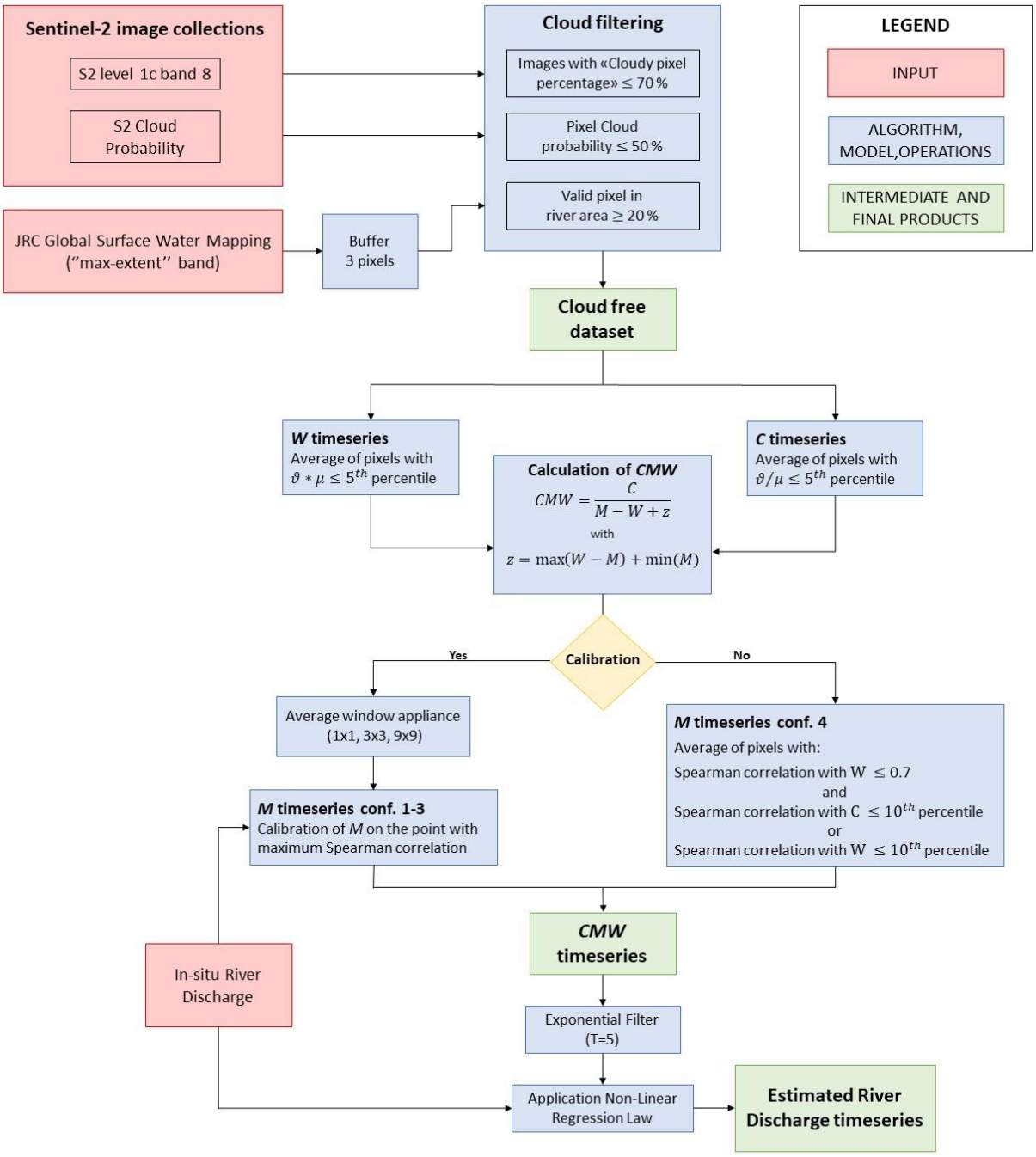


*Figure 2-5: Flowchart of the CMW approach*

5. the $M$ timeseries is obtained by considering the four configurations mentioned above. In the first three configurations, each pixel within the water mask (step 3) is used to calculate the ratio $CMW$ according to Eq. 5: in the case of configuration 1, the analyzed timeseries coincide with those of S2 data, whereas for configuration 2 and 3 each pixel value is resampled with a moving average spatial window of 3x3 pixels and 9x9 pixels, respectively. The pixel where the Spearman's correlation between $CMW$ and the ground

observed discharge is maximized, is selected as $M$ pixel. Alternatively, for configuration 4, multiple pixels in the water mask having Spearman's correlation with $W$ or $C$ below the 10$^{th}$ percentile and with $W$ below the value of 0.7 are selected and averaged.

6. $C$, $W$ and $M$ are used to calculate $CMW$ index, which is then filtered through the exponential smoothing function with $T$ equal to 5 days.

In order to compare the improvements of the new $CMW$ approach, the $CM$ approach was also tested, following the same steps except for the 4th, because the $W$ is not needed in the $CM$ approach, and replacing $CMW$ index by $C/M$ ratio in the step 5.

Finally, the estimation of the river discharge from $CMW$ or $CM$ was performed by a fitting function (cubic function) described for $CMW$ as:

$$Q_e = a + b \cdot CMW + c \cdot (CMW + d)^3 \quad (7)$$

where $Q_e$ is the estimated river discharge, a, b, c and d are parameters, with b and c being positive so as to ensure the monotonicity of the function.

## 2.3 Performance evaluation

---

This section illustrates the technique adopted to assess the performance of the estimated products of rainfall and river discharge. Specifically, the classical performance indices used to assess both rainfall and river discharge are initially discussed, following by the categorical indices and triple collocation method used to evaluate the rainfall performances and the fitting function and Nash-Sutcliffe index applied for the estimated river discharge assessment. The section is based on the following publication, published by Journal of Hydrology:

*Filippucci, P., Brocca, L., Massari, C., Saltalippi, C., Wagner, W., and Tarpanelli, A.: Toward a self-calibrated and independent SM2RAIN rainfall product, Journal of Hydrology, 603(A), 126837, doi: 10.1016/j.jhydrol.2021.126837, 2021.*

and on a preprint of the following publication, submitted to Remote Sensing of Environment:

*Filippucci, P., Brocca, L., Bonafoni, S., Saltalippi, C., Wagner, W., and Tarpanelli, A.: Sentinel-2 high-resolution data for river discharge monitoring, submitted to Remote Sensing of Environment, 2021.*

---

*Classical Performance scores*

Continuous metrics were applied to compare the estimated variables of rainfall and river discharge to the with the dataset taken as the 'standard' or to the observed data. In particular:

*Linear Pearson's correlation (R):* Pearson's correlation is the most common way to characterize statistical dependency between two datasets. It can be obtained from the ratio between the covariance of two dataset and the product of their standard deviation. It varies between -1 and +1, where -1 means negative linear relationship, +1 means positive linear relationship and 0 means no statistical dependency.

*Spearman's correlation (S):* Spearman's correlation is another common way to characterize statistical dependency between two datasets. While Pearson's correlation assesses linear relationships, Spearman's correlation assesses monotonic relationships. It can be obtained from the ratio between the rank covariance of two dataset and the product of their standard deviation. It also varies between -1 and +1, where -1 means negative monotonic relationship, +1 means positive monotonic relationship and 0 means no statistical dependency.

*Relative BIAS (BIASr):* Relative BIAS index can be calculated as the mean difference between two datasets, divided by the mean value of the reference dataset. It describes whether there is a systematic over or under-estimation with respect to the reference data. In this thesis the difference is performed between the estimated and the observed variable. Therefore, negative BIAS values mean that the product underestimates the analyzed quantity, while positive BIAS values indicate overestimation.

*Root Mean Square Error (RMSE) and Relative Root Mean Square Error (RMSEr):* Root mean square error (RMSE) can be calculated as the average deviation between single measurements of two datasets. It comprehends three sources of error: decorrelation, BIAS and random error. It should be noted that, since there is no "true" measure of a quantity, RMSE reliability strongly

depends from the reference dataset accuracy. Relative RMSE (RMSEr) is obtained dividing RMSE by the mean value of the reference dataset.

*Rainfall performances*

In order to estimate the performance of rainfall, two methods were applied: the categorical indices analysis and the TC method.

Categorical indices were computed to measure the performances in detecting rainfall for different rainfall classes. Five classes were selected, dividing the rainfall events in those greater than the 10$^{th}$, the 30$^{th}$, the 50$^{th}$, the 70$^{th}$ and the 90$^{th}$ percentile for each point of the grid. The categorical indices were calculated for each of those classes. Naming $H$ the number of successfully predicted events, $F$ the number of falsely detected events and $M$ the number of missed events, we can define:

*False Alarm Ratio (FAR)* refers to the fraction of erroneously detected events for each class. The optimum value is 0.

$$FAR = {F}/{(H + F)} \tag{8}$$

*Probability Of Detection (POD)* refers to the fraction of correctly predicted events for each class. The optimum value is 1.

$$POD = {H}/{(H + M)} \tag{9}$$

*Threat Score (TS)* is an integrated measure of the overall performances, giving the fraction of successfully detected events over the total missed and detected events for each class. The optimum value is 1.

$$TS = {H}/{(H + F + M)} \tag{10}$$

The classical methods described above permit to assess the similarities between the analyzed dataset and a reference one. Therefore, the performances reliability is dependent on the accuracy of the

reference, but since no dataset has zero-error measurement (VILLARINI ET AL., 2008), not even gauges (PETERSON ET AL., 1998; KIDD ET AL., 2017) the obtained performances are subjected to error. TC method, instead, permits the assessment of uncertainties of three different products against an unknown true reference. Here a brief explanation of the theory behind the method is presented. For further information, the reader is referred to MASSARI ET AL. (2017) and STOFFELEN (1998).

Each measure related to a quantity is characterized by both a random and a systematic error:

$$X = \alpha_X + \beta_X \theta + \varepsilon_X \tag{11}$$

where $X$ is the measure, $\theta$ is the unknown truth, $\varepsilon_X$ the random error and $\alpha_X$ and $\beta_X$ are respectively the additive and multiplicative component of the systematic error. Taking into consideration three different datasets whose errors are uncorrelated, the random error of each dataset can be considered Gaussian distributed with zero mean. The error variance of each dataset can therefore be written as (MCCOLL ET AL., 2014):

$$\sigma_\varepsilon = \begin{bmatrix} \sqrt{V_{11} - {V_{12}V_{13}}/{V_{23}}} \\ \sqrt{V_{22} - {V_{12}V_{23}}/{V_{13}}} \\ \sqrt{V_{33} - {V_{13}V_{23}}/{V_{12}}} \end{bmatrix} \tag{12}$$

where $V_{ij}$ is the covariance between the dataset $i$ and $j$. McColl underlined that, although gaussianity ensures that the RMSE is well descripted and assists in the interpretation, Gaussian data are not required for the TC, as it is often applied to non-Gaussian data such as SM. By using the definitions of correlation and covariance, it can be derived:

$$R_{TC} = \begin{bmatrix} \sqrt{{V_{12}V_{13}}/{V_{11}V_{23}}} \\ \sqrt{{V_{12}V_{23}}/{V_{22}V_{13}}} \\ \sqrt{{V_{13}V_{23}}/{V_{33}V_{12}}} \end{bmatrix} \tag{13}$$

$R_{TC}$ is the TC correlation against the unknown truth. This measure should not be taken as an absolute measure but as a relative measure between the three datasets.

*River discharge performances*

Regarding the river discharge performance scores, the coefficients of correlation of Spearman, S, and Pearson, R, were calculated between the $CMW$ or $CM$ timeseries and the ground observations of river discharge. Three performance scores were instead computed to quantify the error in the final estimation of the river discharge: i) the Root Mean Square Error (RMSE), ii) the relative RMSE (rRMSE) and iii) the Nash-Sutcliffe efficiency, NS.

*Nash Sutcliffe (NS):* the Nash–Sutcliffe model efficiency coefficient is an index designed specifically for hydrological model assessment (**NASH AND SUTCLIFFE, 1970**). It can be calculated by the following equation

$$NS = 1 - \frac{\sum_{t=1}^{T}(Q^t - Q_e^t)^2}{\sum_{t=1}^{T}(Q^t - \bar{Q})^2} \quad (14)$$

Where $Q^t$ is the observed river discharge at time *t*, $\bar{Q}$ is its average and $Q_e^t$ the estimated river discharge at time *t.* In case of perfect model without estimation errors, the NS is equal to 0.

# 3. DEVELOPMENT OF THE PARAMETERIZED SM2RAIN

*In order to use SM2RAIN algorithm for rainfall estimation, four parameters need to be calibrated. This chapter propose a methodology to avoid this step, by developing four parametric relationships that link each SM2RAIN parameter to some descriptors of the pixel. Specifically, several datasets of land and climatic were downloaded, resulting in considering more than 50 potential descriptors in this analysis. A two-phases methodology was adopted to identify the best relationships: first the descriptors better related to each SM2RAIN parameter were identified, by the use of the Spearman's correlation index. Then a stepwise non-linear backward approach was applied to obtain the desired relationships. The four parametric equations were then applied worldwide in order to validate them. The resulting rainfall product was then compared with the one obtained by calibrating SM2RAIN and with another satellite derived rainfall product GPM-FR, through the use of classical performance scores such as Pearson and Spearman's correlation, RMSE and the categorical indexes, FAR, POD and TS. Finally, Triple Collocation was applied to assess the capability in rainfall estimation of the parameterized SM2RAIN product against an unknown true reference.*

This chapter is based on the following publication, published by Journal of Hydrology:

*Filippucci, P., Brocca, L., Massari, C., Saltalippi, C., Wagner, W., and Tarpanelli, A.: Toward a self-calibrated and independent SM2RAIN rainfall product, Journal of Hydrology, 603(A), 126837, doi: 10.1016/j.jhydrol.2021.126837, 2021*

### 3.1 Data

Multiple descriptors are considered for the estimation of SM2RAIN algorithm parameters values through a regression-based regionalization approach (JAKEMAN ET AL., 1992; POST AND JAKEMAN, 1999; SEFTON AND HOWARTH, 1998; SEIBERT, 1999; WAGENER ET AL., 2004): several datasets are selected describing climatic (rainfall and evapotranspiration) and land (soil texture and soil type, SM, soil temperature, topography and vegetation cover) characteristics. The datasets were chosen for different reasons, including their relation with soil state and their availability worldwide. In the following, the datasets description is provided (see Table 3-1).

*Climatic data*

*Regional rainfall datasets*

Regional gauge-derived rainfall datasets were collected for the 1009 points uniformly distributed (0.25-degree resolution) over the areas covered by rain gauges in Australia, Italy, US and India (Figure 2-1). The regional rainfall datasets are used as reference for SM2RAIN calibration at the points for which the parametric relationships are developed. For each region of the study area, the data are collected for the period 2013-2017. In particular:

- The Australian Water Availability Project (AWAP) rainfall product was downloaded for the Australia region. This gridded dataset is obtained from the interpolation of daily measurements of the Australian Bureau of Meteorology rain gauge network, performed by using an optimized Barnes successive correction technique. Its spatial resolution is about 5 km (0.05-degree) with a daily temporal resolution.

- The rainfall dataset of the Italian Civil Protection Department (ITA - DPC) is an interpolation of more than 3000 rain gauges distributed over the Italian territory. The interpolation is carried out using the Random Generator of Space Interpolations from Uncertain Observations

*Table 3-1: description of the dataset downloaded and processed but not selected for the derivation of SM2RAIN parametric relationships*

| VARIABLE | SOURCE | TEMPORAL RESOLUTION | SPATIAL SAMPLING | ADDITIONAL INFORMATION |
|---|---|---|---|---|
| Soil Temperature (0-7 cm) | ERA5 Land | 1 h | 0.1° | https://www.ecmwf.int/en/era5-land |
| Evapotranspiration | ERA5 | 1 h | 0.25° | https://www.ecmwf.int/en/era5 |
| Rainfall | ERA5 | 1 h | 0.25° | https://www.ecmwf.int/en/era5 |
| Rainfall | IMERG Late Run | 0.5 h | 0.1° | https://gpm.nasa.gov/data/directory |
| Rainfall | IMERG Final Run | 0.5 h | 0.1° | https://gpm.nasa.gov/data/directory |
| Rainfall | GPCC | 1 d | 1° | Schamm et al. (2014) |
| Rainfall | AWAP | 1 d | 0.05° | http://www.bom.gov.au/jsp/awap/rain/index.jsp |
| Rainfall | IMD | 1 d | 0.25° | http://www.imd.gov.in/pages/service_hydromet.php |
| Rainfall | CPC | 1 d | 0.25° | https://psl.noaa.gov/data/gridded/data.unified.daily.conus.html |
| Rainfall | ITA-DPC | 1 d | 0.1° | Ciabatta et al. (2017) |
| Soil Composition Data | HWSD | / | ~0.008° | http://www.fao.org/land-water/databases-and-software/hwsd/en/ |
| Soil Moisture | ASCAT | ~12 h | ~0.125° | Wagner et al. (2013) |
| Soil Moisture (0-7 cm) | ERA5 Land | 1 h | 0.1° | https://www.ecmwf.int/en/era5-land |
| Topography | ETOPO5 | / | ~0.08° | https://www.ngdc.noaa.gov/mgg/global/etopo5.HTML |
| Vegetation Continuous Fields | VCF5KYR | / | 0.05° | https://lpdaac.usgs.gov/products/vcf5kyrv001/ |

(GRISO, PIGNONE ET AL., 2010) algorithm to spatially interpolate the measurements on a grid with about 10 km (0.1-degree) spatial resolution and aggregating the hourly data to the daily time step.

- For the US region, the National Oceanic and Atmospheric Administration Climate Prediction Center (NOAA CPC) Daily US UNIFIED Precipitation was downloaded. This rainfall product is characterized by an improved quality obtained by combining all information sources available at CPC and by taking advantage of the optimal interpolation (OI) objective analysis technique (XIE ET AL., 2007). Its spatial resolution is about 25 km (0.25-degree) with a daily temporal resolution.

- India region daily rainfall was obtained by downloading the India Meteorological Department (IMD) gridded dataset. This product combines daily rainfall data from 6955 gauges, using the Shepard method (**PAI ET AL., 2014**) to interpolate them, and it is characterized by a spatial resolution of about 25 km (0.25-degree).

The mentioned datasets were all temporally interpolated from their local time to 00:00 UTC, accepting the resulting uncertainty to obtain regular time spacing, in order to simplify the intercomparison with satellite-derived products (available at 00:00 UTC).

*Global Rainfall datasets*

Different global rainfall datasets were downloaded to obtain and validate the new SM2RAIN-ASCAT parameterized rainfall product, for the period 2013-2019:

- GPCC rainfall product (First Guess) is obtained from ~7000 quality-controlled stations all over the world (**SCHAMM ET AL., 2014**). Its spatial resolution is 1 degree, with a daily temporal resolution. Since it is based on ground observation, the accuracy of the dataset is greater over region with high gauge density, i.e., Europe and US.

- The IMERG algorithm estimates precipitation over the majority of Earth's surface by inter-calibrating the available Passive Microwave satellite precipitation estimates to the Combined Radar-Radiometer precipitation estimates from the GPM mission Core Observatory (GPM-CO) and then by merging and interpolating together these estimates with other precipitation estimates from infrared geostationary sensors (**HUFFMAN ET AL., 2020**). Morphing and Kalman filtering interpolation are used to provide the precipitation estimate if no valid microwave data are available. The resulting product spatial resolution is 0.1-degrees, and the temporal resolution is 30 minutes. Three Runs of IMERG are available to the users, based on increasing latency and accuracy: Early Run (IMERG-ER; latency of 4–6 h after observation), Late Run (IMERG-LR; latency 12–18 h) and Final Run (IMERG-FR; latency of about 3

months). Final Run V06 product, with a monthly adjustment based on GPCC, and Late Run V06 product, are used here. In this study, the 30 minutes rainfall data were accumulated to obtain daily precipitation estimates.

- ECMWF Reanalysis 5th Generation (ERA5) provide hourly data of various global atmosphere, land surface and sea-state variables, combining models with observations. It was developed within the Copernicus program Climate Change Service (C3S) and it replaces the previous ERA-interim reanalysis product. Its spatial resolution is around 36 km, resampled on a regular 0.25-degree grid, and the temporal resolution is 1-hour (HERSBACH ET AL., 2020). The hourly rainfall was calculated by subtracting the snowfall fraction to the total precipitation and accumulated to daily scale in this study.

*Evapotranspiration*

Hourly evapotranspiration data from ERA5 were obtained for each point of the study area. The hourly data were accumulated on windows of 12 hours centred at 00:00 UTC and 12:00 UTC, to obtain a temporal resolution aligned with the SM datasets (12 hours spacing).

*Land data*

*Satellite Soil Moisture and Soil Moisture Noise*

Advanced SCATterometer (ASCAT) is an active microwave sensor onboard of MetOp-A (launched 19/10/2006), MetOp-B (launched 17/09/2012) and MetOp-C (launched 07/11/2018) satellites. It uses two sets of three vertically polarized antennae, one on each side of the satellite ground track, and it senses backscatter radiation at 5.255 GHz (C-band). The sensor was originally developed to sense wind speed over oceans, but it turned out to be also sensitive to the amount of water in the soil, leading to the development of one of the longest satellite SM products available nowadays (from 2007 onward). ASCAT retrievals have a spatial resolution of 25 km, sampled at 12.5 km (~0.125°). Relative SM estimates and their related noise were downloaded from the EUMETSAT Satellite

Application Facility on Support to Operational Hydrology and Water Management (H SAF) H115 and H116 products for the period 2013-2019. In these years, the contemporary availability of the satellites MetOp-A and B permitted a sub-daily temporal resolution over most of the Earth (WAGNER ET AL., 2013). When the surface state was indicated as frozen, SM estimates were discarded. ASCAT measurements were linearly interpolated every 12 hours, to obtain regular time spacing. If no data were found within 5 days, each datum in the interval was set to Not a Number (NaN).

*Modelled Soil Moisture*

Hourly SM in the first soil layer (0 - 7 cm) of the ECMWF Integrated Forecasting System data from ERA5-Land were downloaded for the analysis period. ERA5-Land was produced by regridding the land component of the ECMWF ERA5 climate reanalysis with a finer spatial resolution (0.1-degree). SM was subsampled every 12 hours to obtain the same temporal resolution of ASCAT data.

*Topographic data*

Elevation data from Earth topography 5 arc minute (ETOPO5) were downloaded. Although this Digital Elevation Model (DEM) is available on a regular grid of 5 -minutes (~0.08°), the resolution of the source data base varies from 5-minute for the ocean floors, USA, Europe, Japan and Australia to 1 degree in data-deficient parts of Asia, South America, northern Canada and Africa.

*Soil Temperature*

Soil Temperature data in the first soil layer (0 - 7 cm) of the ECMWF Integrated Forecasting System data from ERA5-Land were downloaded for the analysis period. The hourly data were subsampled every 12 hours to match the temporal resolution of ASCAT data.

*Soil Composition Data*

The Harmonized World Soil Database v1.2 (WIEDER ET AL., 2014) contains worldwide soil composition information derived from regional and national data. Several soil parameters were downloaded for this analysis for the nominal year of 2000, including soil depth, sand-silt-clay

fraction, reference soil depth, carbon content and bulk density, at a spatial resolution of 5 arcminutes (~0.08°).

*Vegetation Continuous field*

Global fractional vegetation cover data VCF5KYRv001 was downloaded from NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs). The dataset relative to the nominal year of 2015 was downloaded, containing information of tree cover vegetation, bare ground and non-tree cover vegetation area percentage, with a spatial resolution of 0.05°.

All the data were resampled to ASCAT grid in order to allow their intercomparison. The chosen interpolation method consists in the nearest neighbour technique for evapotranspiration data, weighted average of the overlapping areas for datasets with a spatial resolution finer than 25 km and weighted average of the four nearest pixels for the remaining datasets.

### 3.2 Descriptors selection

The objective of this chapter is to find and validate four parametric relationships to estimate the SM2RAIN algorithm parameters values from climatic and land descriptors readily available worldwide. Through these relationships, SM2RAIN can be easily applied without the need of a reference rainfall dataset. By using the 1009 points, a local analysis was performed to find the parametric relationships formulation.

The number of potential descriptors obtained from soil data, vegetation continuous field, topography data and the statistic of time-varying quantities, exceed 50. Spearman's correlation values between each of them and SM2RAIN parameters were therefore calculated, in order to reduce the number of descriptors, by selecting for each parameter the quantities that are better related to it. An example of the procedure can be found in Figure 3-1, where three scatter density plots between the parameter $Z^*$ and three representative descriptors are shown. In the example, it can be seen how the soil water storage capacity values, obtained from the Harmonized World Soil Database, does not show

significant correlation with the $Z^*$ parameter (Figure 3-1c), contrary to the expectation. Greater absolute values of Spearman's correlation were obtained from the annual average daily rainfall (Figure 3-1a) and the standard deviation of the soil temperature (Figure 3-1b), with the latter showing an inverse relationship with the analyzed parameter.

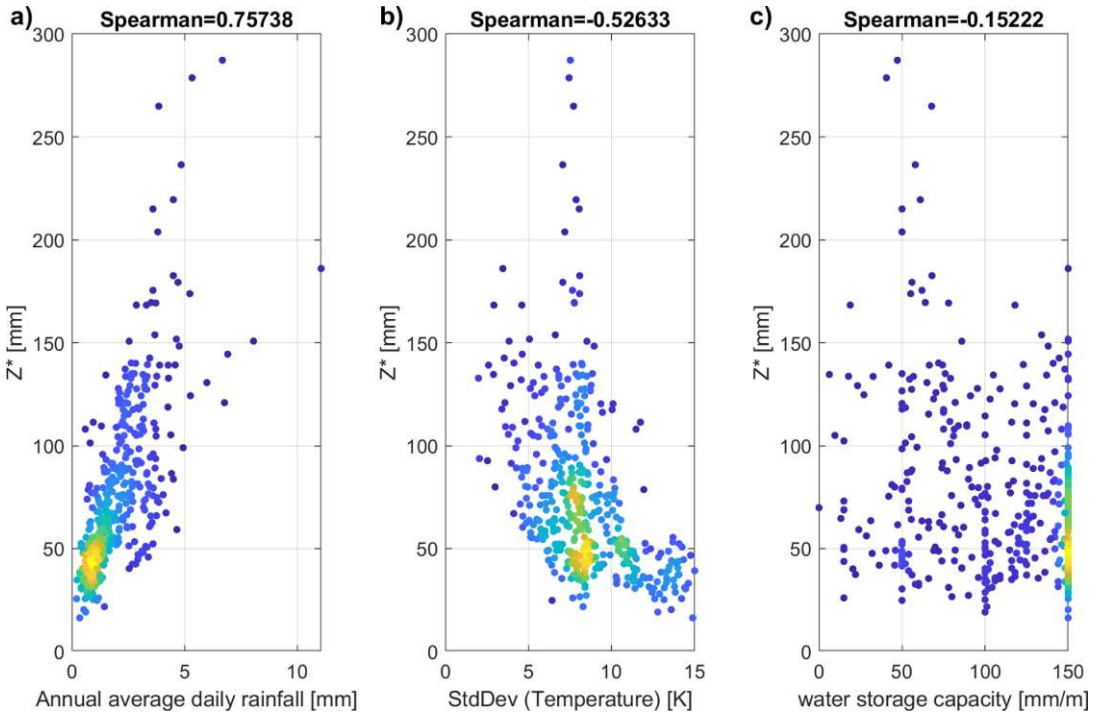


*Figure 3-1: Example of the descriptors selection procedure. In the three panels is shown a scatter density plot of Z* parameter distribution with respect to the annual average daily rainfall (a), the standard deviation of the soil temperature (b) and the s*

These two descriptors were therefore selected to be used in the multilinear regression algorithm, while the soil water storage capacity was discarded (note that the standard deviation of soil temperature was discarded in the successive step). For the sake of brevity, neither the details of the descriptors selection, nor every iteration of the stepwise non-linear backward regression is described here, but the final relationships are directly shown. At the end of the procedure, most of the analyzed descriptors were discarded: the descriptors who resulted more significant for SM2RAIN parameters estimation were only those related to SM, SM noise, rainfall and topography.

### 3.3 *T* parameter

The first obtained relationship was the one relative to the exponential filter parameter $T$. This parameter was the first to be calculated in order to obtain reduced-noise SM from satellite estimates to be used in the following steps. The reference values for $T$ parameter were obtained by applying the exponential filter to ASCAT SM data maximizing R between the filtered SM and the modelled SM from ERA5 (first soil layer 0-7cm). Afterward, the points with R values greater than a fixed threshold of 0.6 were retained and used as reference *T*-values to be compared with the climatic and land descriptors (see paragraph 3.2). The selection of the points with correlation greater than 0.6 was done to avoid fitting the parametric relationship to not representative data. Visual inspection and Spearman's correlation were used to identify which descriptors were better correlated with the reference *T*-values. A non-linear regression model was then iteratively applied to the selected descriptors in order to find the best parametric relationship. The optimal relationship can be written as:

$$T = 0.8788 + 1.7020\ \overline{SMnoise}\ std(|SM_d|) + 0.3555\ {std(|SM_d|)}/{\bar{P}}\ topC \qquad (15)$$

where $\overline{SMnoise}$ is the temporal mean value of the SM noise relative to ASCAT estimates, $std(|SM_d|)$ is the temporal standard deviation of the absolute values of ASCAT SM temporal variations, $\bar{P}$ is the annual average of daily rainfall, and $topC$ is the topographic complexity (spatial standard deviation of elevation values within each pixel boundaries). The reduced-noise SM estimates obtained from the exponential filter application were used in equation (2) and in the calculation of the SM descriptors for the successive SM2RAIN parameters relationships.

### 3.4 *b* parameter

According to Famiglietti and Wood (1994), $b$ can be considered equal to:

$$b = 3 + {2}/{\lambda} \qquad (16)$$

where $\lambda$ is the pore size distribution index. A parametric relationship to estimate $\lambda$ as a function of $a$ parameter was proposed by BROCCA ET AL. (2014B):

$$\lambda = 0.085 \log a + 0.1574 \quad (17)$$

The same relationship was adopted in this study, but the two coefficients were recalibrated using the following procedure. $T$-values from equation (15) were used to obtain filtered ASCAT SM series to which apply SM2RAIN. The three parameters of the balance equation were then calibrated against reference regional rainfall observations (1009 points). The points with R between the observed and estimated rainfall greater than the fixed threshold of 0.6 and with $T$-value less than a threshold fixed to 6, were then selected (as before to avoid fitting the parametric relationship to not representative data) and the two coefficients of equation (17) were calculated by fitting the relationship between the calibrated $a$ and $b$ parameter values, thus obtaining:

$$b = 3 + {}^{2}/_{(\,0.5928 * \log a + 0.3022\,)} \quad (18)$$

### 3.5 $Z^*$ and $a$ parameters

By using equations (15) and (18), the SM2RAIN algorithm was re-applied to ASCAT SM estimates at 1009 points by only calibrating $Z^*$ and $a$ parameters. Again, the points with R between estimated and observed rainfall greater than the fixed threshold of 0.6, and with $T$-value less than 6, were selected to be compared with the climatic and land descriptors. Visual inspection and Spearman's correlation were used to identify which quantities were better related with $Z^*$ parameter, then a linear regression model was applied to them in order to find the $Z^*$ parametric relationship:

$$Z^* = 10.3124 + 0.5186 \, {\bar{P}}/{\overline{|SM_d|}} \quad (19)$$

The same procedure was adopted to find $a$ parametric relationship after recalibrating the SM2RAIN algorithm for only the $a$ parameter and fixing the others through equations (15), (18), and (19). The obtained equations for $a$ was:

$$a = -1.5748 + 13.0324\, Z^* \, \overline{|SM_d|} \tag{20}$$

where $\bar{P}$ is the annual average of daily rainfall and $\overline{|SM_d|}$ is the temporal mean of the absolute values of ASCAT SM temporal variations.

### 3.6 Test of parametric relationships

By using equations (15), (18), (19) and (20), the four SM2RAIN parameters can be obtained from knowing the ASCAT SM timeseries and its noise, the topographic complexity and the mean annual rainfall. To avoid non-physical values for the parameters, the boundaries reported in Table 3-2 were applied, fixing all the parameters that exceed limits to the boundary itself.

*Table 3-2: Upper and Lower boundaries for SM2RAIN parameters*

| **Boundaries** | $Z^*$ **[mm]** | $a$ **[mm/d]** | $b$ **[-]** | $T$ **[d]** |
|---|---|---|---|---|
| **Lower** | 20 | 0.1 | 1 | 0 |
| **Upper** | 800 | 200 | 50 | 8 |

We note that the parametric relationships, obtained through a statistical regression-based approach, show physical reasoning in the expected correlation between SM2RAIN algorithm parameters and climatic and land descriptors. Indeed, equation (15) indicates that the exponential filter $T$ parameter is directly proportional to the mean value of SM noise, to standard deviation of absolute SM variation, to the ratio between the latter and the annual average daily rainfall and to the topographic complexity. All these descriptors increase with either SM measurement error (i.e., SM noise and topographic complexity) or temporal SM variability (i.e., $std(|SM_d|)$ and $1/\bar{P}$); in both cases higher $T$-values are expected, since a higher value of $T$ increases the filtering capacities. Equations (19) and (20) link the estimation of $Z^*$ and $a$ to the value of $\overline{|SM_d|}$ and $\bar{P}$. Indeed $Z^*$ increases with the ratio between $\bar{P}$ and $\overline{|SM_d|}$ because it is a measure of the amount of water stored in the soil, while $a$ is directly correlated with $\bar{P}$.

As mentioned above, to obtain the parametric relationships, SM2RAIN was applied to ASCAT SM for the 1009 points for 5 times, after and before the definition of each parameter relationship, by using the available equations and by calibrating the remaining parameters with the standard approach (i.e., minimization of RMSE). The performances of the obtained rainfall, in terms of R and RMSE are shown in Table 3-3.

A few insights can be deduced from these results. The overall drop in performances is limited, thus demonstrating that the obtained parametric relationships are well suited to estimate the SM2RAIN parameter values. The major drop in correlation can be ascribed to the parameter $T$. This can be easily explained as the parameter $T$ is the only one related to rainfall occurrence, to which the correlation is highly sensitive, while the others parameters are more related to rainfall amount and, hence, to RMSE. A possible reason for the correlation deterioration could be due to error in modelled SM from ERA5. However, different tests with the other soil layers of ERA5 and other modelling approaches were carried out and worse results were obtained (not shown for the sake of brevity). The parametric relationship for $a$ is the one that caused the greatest increase in RMSE (see Table 3-3).

*Table 3-3: Mean value and variation of Pearson's Correlation (R) and Root Mean Square Error (RMSE) for local analysis points, calculated after and before the establishment of each parametric relationship*

| | **Mean R [-]** | **ΔR [-]** | **Mean RMSE [mm/d]** | **ΔRMSE [mm/d]** |
|---|---|---|---|---|
| **Calibrated SM2RAIN** | 0.5951 | | 4.4126 | |
| **T fixed** | 0.5757 | -0.0194 | 4.4909 | 0.0783 |
| **T, b fixed** | 0.5712 | -0.0045 | 4.5226 | 0.0317 |
| **T, b, Z* fixed** | 0.5631 | -0.0081 | 4.6142 | 0.0916 |
| **Parameterized SM2RAIN** | 0.5567 | -0.0064 | 4.7915 | 0.1773 |

Finally, we underline that soil and vegetation descriptors were found not fundamental for obtaining the parametric relationships likely due to the limited accuracy of these datasets at the considered

spatial resolution, particularly for soil information, and the limited influence of vegetation on the analyzed parameters, confirming the findings of SEHGAL ET AL. (2020).

### 3.7 Global analysis

The good results obtained at 1009 points led to the application of the parametric relationships on a global scale. The parameters maps obtained using the parametric relationships on a quasi-global scale (60° S – 60° N) are shown in Figure 3-2. As expected, exponential filter $T$ parameter is greater over desert, forest and mountainous areas (Figure 3-2d), where SM quality is lower, while the distributions of $Z^*$ and $a$ (Figures 3-2a and 3-2b) reflect the known areas where the average rainfall rate is high (equatorial region). ERA5 rainfall was used to obtain the annual average daily rainfall for the parametric relationships and also to calibrate SM2RAIN ASCAT with the standard methodology to verify that the uncalibrated product performances are in line with those of the calibrated SM2RAIN. The performance of the two rainfall datasets were then assessed against the GPM-FR precipitation product, in terms of the categorical indices False Alarm Ratio (FAR), Probability of Detection (POD) and Threat Score (TS), and the continuous indices R, BIASr and RMSEr. It should be noticed that the GPM product contains both the solid and liquid fraction of the precipitation, while SM2RAIN is able to estimate only the liquid fraction. The masking of frozen condition for SM ASCAT product should be able to remove the days of solid precipitation from the comparison, but in case of failure of frozen condition detection, this issue could be a source of error, in particular over high elevation and high latitude regions. Starting from IMERG V05B, full coverage is provided for the latitudes of 60°N-60°S, while the remaining upper and lower latitudes extending to 90° are considered "partial coverage". The current analysis was restrained to the full coverage area (60°N-60°S) to increase the accuracy of the results. From now on, the product obtained from the use of the parametric relationships will be labelled as "*parameterized*", while the one obtained using the standard calibration method will be named as "*calibrated*".

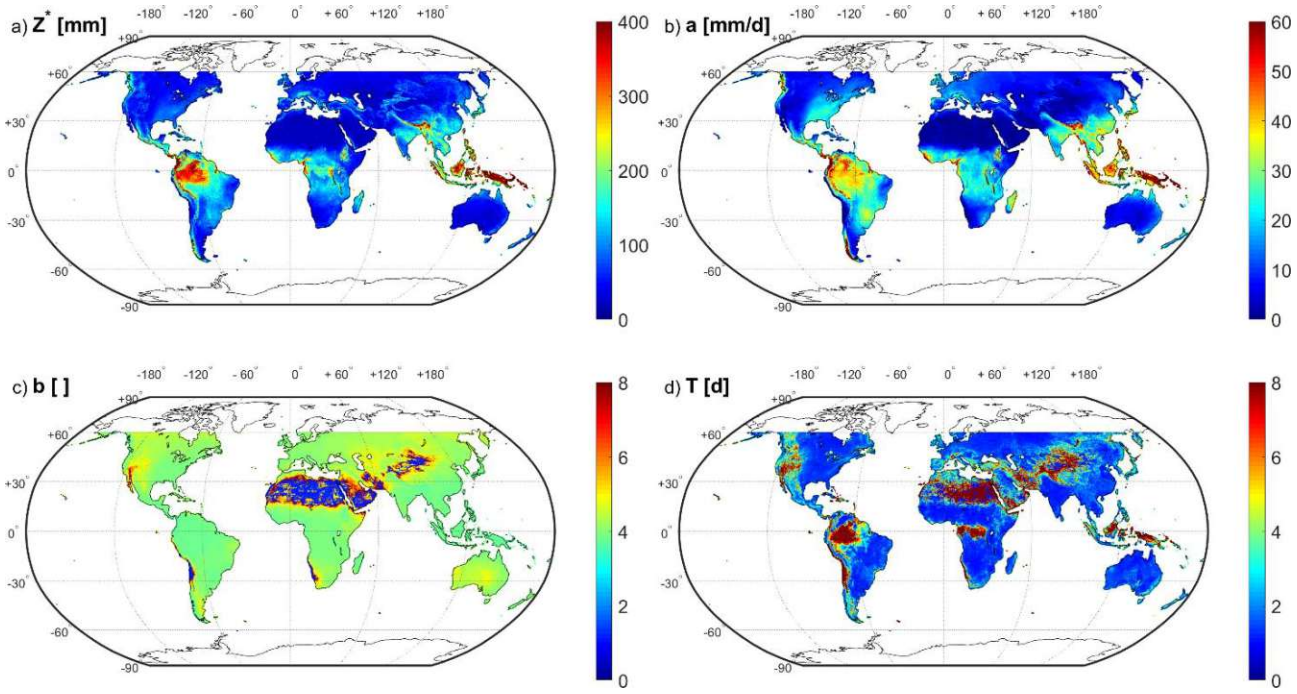


*Figure 3-2: Global Map of SM2RAIN parameter values as obtained from the parametric relationships. Each panel shows: a) parameter Z*, b) parameter a, c) parameter b, d) parameter T.*

The distribution of the obtained categorical indices is shown in Figure 3-3 as boxplots. The indices were calculated for five rainfall classes, respectively the 10$^{th}$, the 30$^{th}$, the 50$^{th}$, the 70$^{th}$ and the 90$^{th}$ percentile of the rainfall calculated for each point. Regarding the FAR, the two products show similar performances for the first two classes, while the parameterized product has a higher percentage of false alarms for the last three classes. Different observations can be done for the POD index: the calibrated product performs slightly better than the parameterized for the first three classes, while it is true the opposite for the others two. From this information, it can be inferred that the parameterized product has greater capability in estimating the major rainfall events. However, the performances of the parameterized product are slightly worse than those of the calibration product for lower percentiles (<50$^{th}$), due to a greater number of false alarms and to a lower detection ability. These results are confirmed by the TS scores, which indicates a slightly better performance of the calibrated product for the first four classes, while for the fifth class the parameterized product performs better.

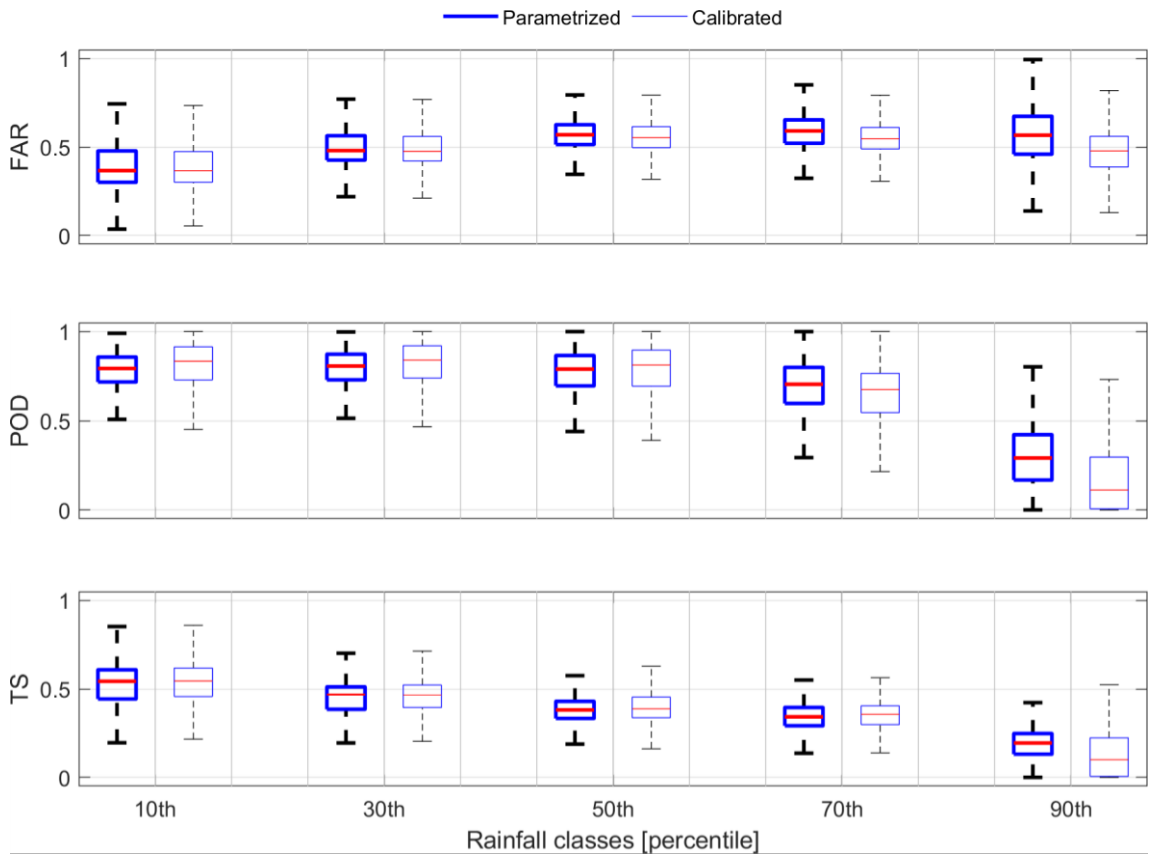


*Figure 3-3: Distribution of False Alarm Ratio, FAR, Probability of Detection, POD and Threat Score, TS categorical indices of the parameterized and calibrated SM2RAIN rainfall products, respectively in dark and light blue, against the benchmark dataset GPM - Final Run, related to the committed area. The indices are calculated for five rainfall classes, according to the intensity of the observed rainfall events being greater than the 10$^{th}$, the 30$^{th}$, the 50$^{th}$, the 70$^{th}$ and the 90$^{th}$ percentiles.*

The differences in the performance are due to the different parameter values adopted by the two products. Figure 3-4a shows the $T$ parameter distribution for the whole area: there are clear differences between the calibrated and parameterized values, in both the median and the range of values. This is probably a consequence of the strategy used to estimate $T$, comparing the satellite SM with a modelled SM, instead of calibrating the $T$-values with respect to reference rainfall. The parameterized product tends therefore to filter the SM data less than the calibrated product (i.e., lower $T$-values), thus increasing the average SM variation, which in turn increase the overall estimated rainfall. This is the cause of the increase of both the FAR and POD indices. $Z^*$ and $a$ (Figures 3-4b and 3-4c) show instead similar behavior between the two products. The values for the parameterized product are slightly greater than the calibrated product that is the reason of the overestimation

tendency noted above. Finally, Figure 3-4d shows the distributions of $b$ parameter, which has similar median value but a very different variability range, due to the relationship, equation (18), used for relating $a$ and $b$ parameters.

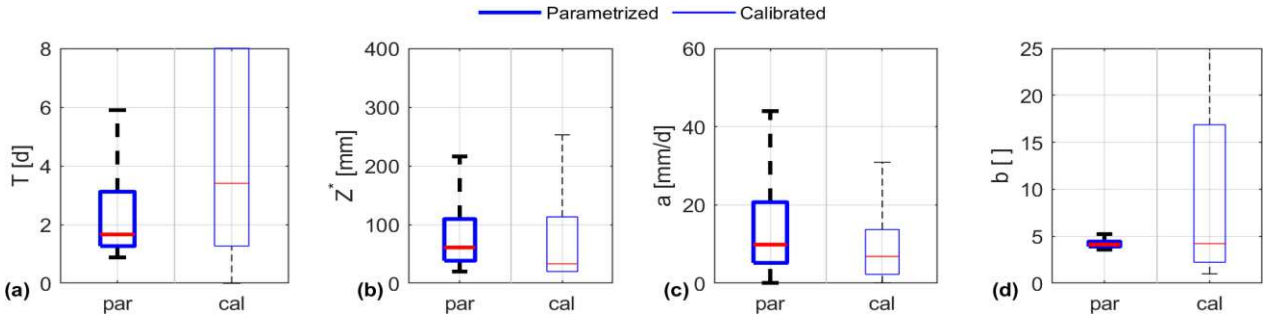


*Figure 3-4: Distribution of SM2RAIN parameters T, Z^*, a and b over the whole area for the parameterized, dark blue, and the calibrated, light blue, SM2RAIN rainfall products.*

R, BIASr and RMSEr were also calculated for the global analysis. The distribution boxplots of these performance indices are shown in Figure 3-5. The obtained results confirm the outcomes of the categorical indices analysis: in terms of R, the calibrated product is slightly better than the parameterized (0.4866 vs 0.4777 in the committed area), but both average value and range of R are overall comparable (Figure 3-5a). BIASr for the parameterized product is around 0.2, confirming the tendency to overestimate rainfall, whereas the calibrated product has a tendency to underestimation (Figure 3-5b). RMSEr values (Figure 3-5c) are very similar between the parameterized and the calibrated product, with differences lower than 5%.

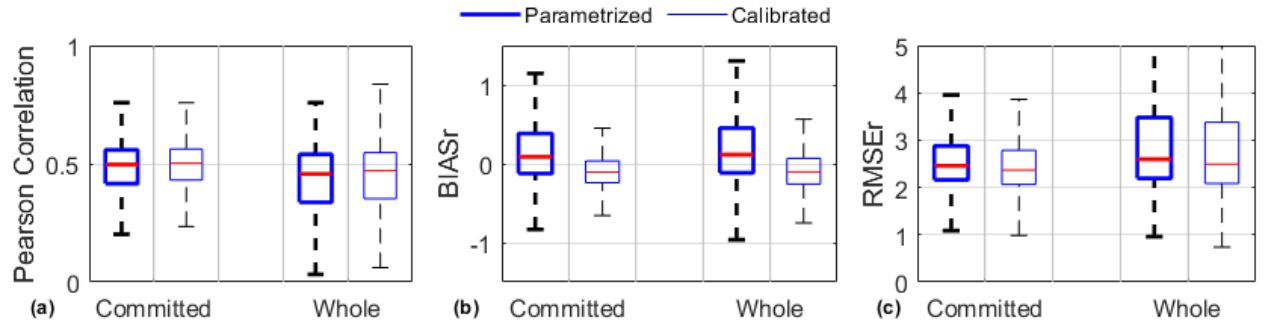


*Figure 3-5: Distribution of Pearson's Correlation, relative BIAS (BIASr) and relative Root Mean Square Error (RMSEr) indices of the parameterized and calibrated SM2RAIN rainfall products, respectively in dark and light blue, against the benchmark dataset GPM - Final Run. In each panel, the results related to the committed area are on the left and those related to the global area are on the right.*

The global map of R and RMSEr difference between parameterized and calibrated products are shown in Figures 3-6 and 3-7, respectively. In the figures, red colour indicates that the parameterized product is better than the calibrated one, while the blue colour indicates the opposite. An overall increase in Pearson's correlation is noticeable in the tropical area, while many mountainous areas show a decrement in R values. The relative error increase in forest, desert and mountainous areas, likely due to the lower filtering of SM values in the parameterized product. One possible cause of the performance deterioration in topographically complex zones could be related to the low spatial resolution of the selected DEM, ETOPO5. For this reason, a different product characterized by higher spatial resolution has been used for the high-resolution rainfall estimation over the Po River basin, descripted in chapter 4.

The mean R values shown in Figure 3-7a are around 0.5 for the committed area and 0.4 globally. These correlation values are likely due to the differences between the dataset used to calibrate SM2RAIN and to obtain the climatology for the parametric relationships, i.e., ERA5, and the product considered as benchmark, i.e., GPM-FR. ERA5 and GPM-FR are indeed not highly correlated by each other; the mean value of R between them equals to 0.5604 in the committed area (0.5412 globally).

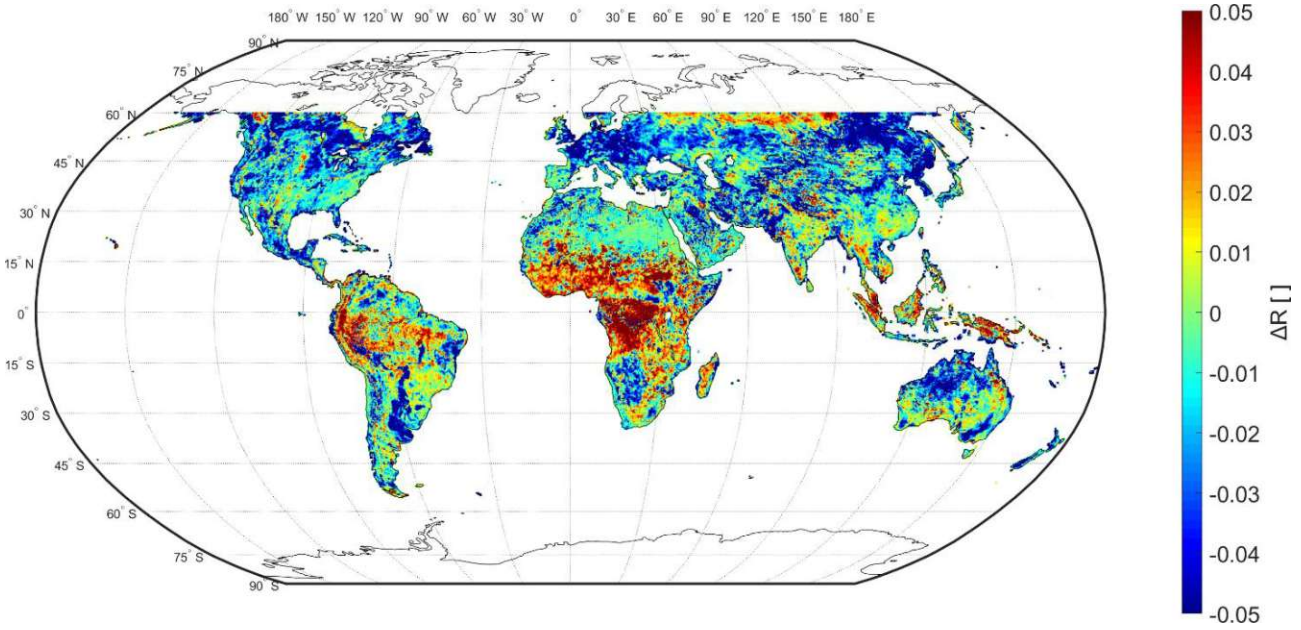


*Figure 3-6: Global map of differences between the parameterized and calibrated SM2RAIN rainfall products for the Pearson's correlation score calculated against GPM – Final Run product. Red areas mean that the parameterized product outperforms the calibrated one, the opposite for blue areas. The parameterized product shows an increase of correlation over dense forest and frozen areas.*

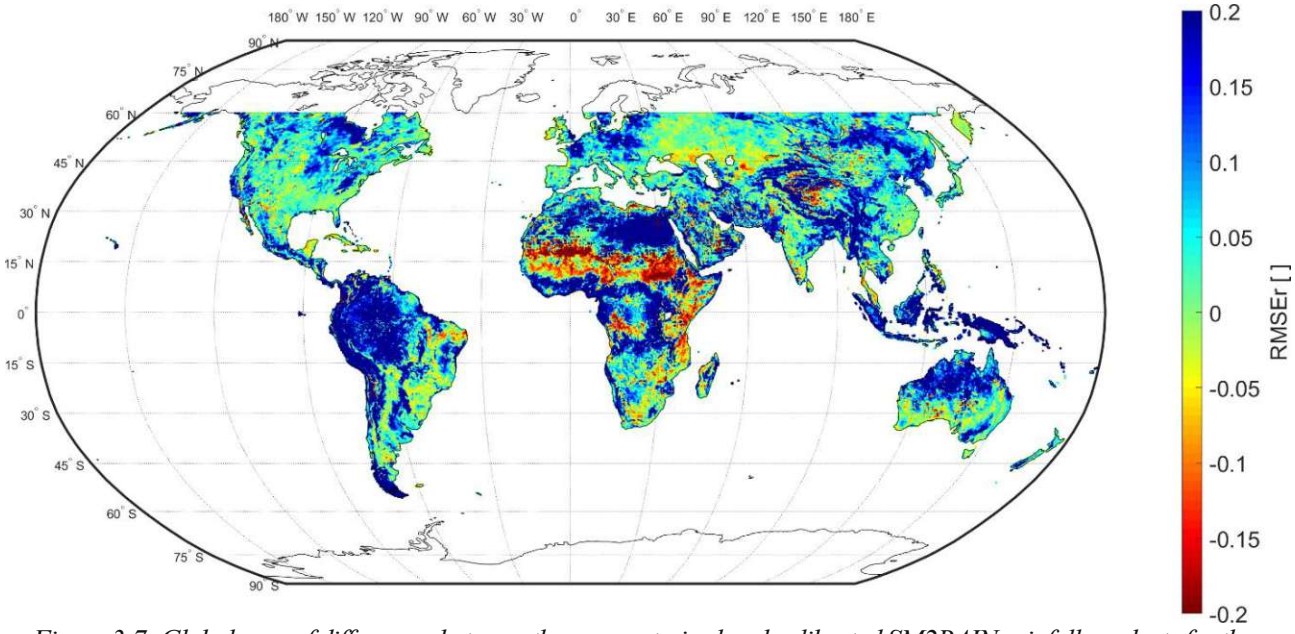


*Figure 3-7: Global map of differences between the parameterized and calibrated SM2RAIN rainfall products for the relative Root Mean Square Error score calculated against GPM – Final Run product. Red areas mean that the parameterized product outperforms the calibrated one, the opposite for blue areas.*

Due to the difficulty to have a reliable rainfall benchmark on a global scale, a TC analysis was performed to assess the capability of SM2RAIN products in rainfall estimation against an unknown true reference. Since TC requires three different products whose errors are uncorrelated, two other global rainfall datasets were selected to be compared with the parameterized and the calibrated products, separately: GPCC and GPM-LR, chosen over GPM-FR as the latter is corrected using GPCC monthly rainfall, and therefore it does not satisfy the condition of uncorrelated error with GPCC. Since GPCC has a low spatial resolution (1-degree), it was interpolated over ASCAT grid using a weighted average, where the weights are the relative inverse of the distance between each ASCAT pixel and the four nearest GPCC points. Figure 3-8 shows the boxplot of the obtained $R_{TC}$.

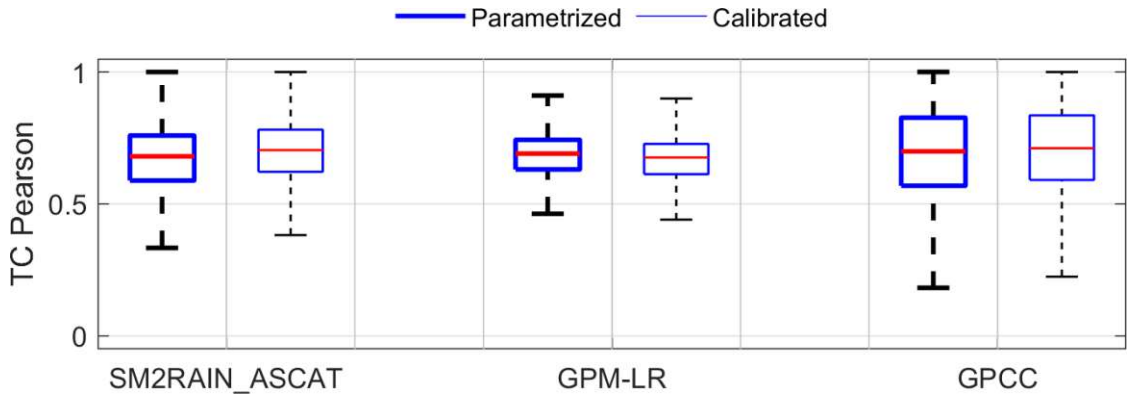


*Figure 3-8: Distribution of the Triple Collocation correlation obtained from the rainfall products triplets composed from SM2RAIN, GPM – Late Run and GPCC, over the committed area. The results of the parameterized products are shown in dark blue, while those of the calibrated product are in light blue.*

SM2RAIN rainfall products have mean correlation values comparable with those of the other products for the committed area, with a slight deterioration when the parametric relationships are used instead of the standard calibration. Moreover, the areas where the parameterized product performs better than the other two is large (red areas in Figure 3-9): GPCC is the best performing product over most of Europe and Asia, half of North America and half of Australia, where there is a high density of gauge stations. GPM-LR performs better over forest and desert areas, where ASCAT SM has large errors (green areas in Figure 3-9). In most of the remaining zones (red areas in Figure 3-9), SM2RAIN ASCAT derived from the parametric relationships performs better than the other products, confirming the capacity of SM2RAIN in estimating rainfall over Africa and South America

(BROCCA ET AL., 2020; MASSARI ET AL., 2020) also when the parameterized product is considered. For further information about the mutual correlation between the products, the individual maps of TC correlation have been added in the appendix (section 7.1). As mentioned above, the TC correlation values should not be taken as absolute measures of accuracy but rather as relative measures between the three datasets.

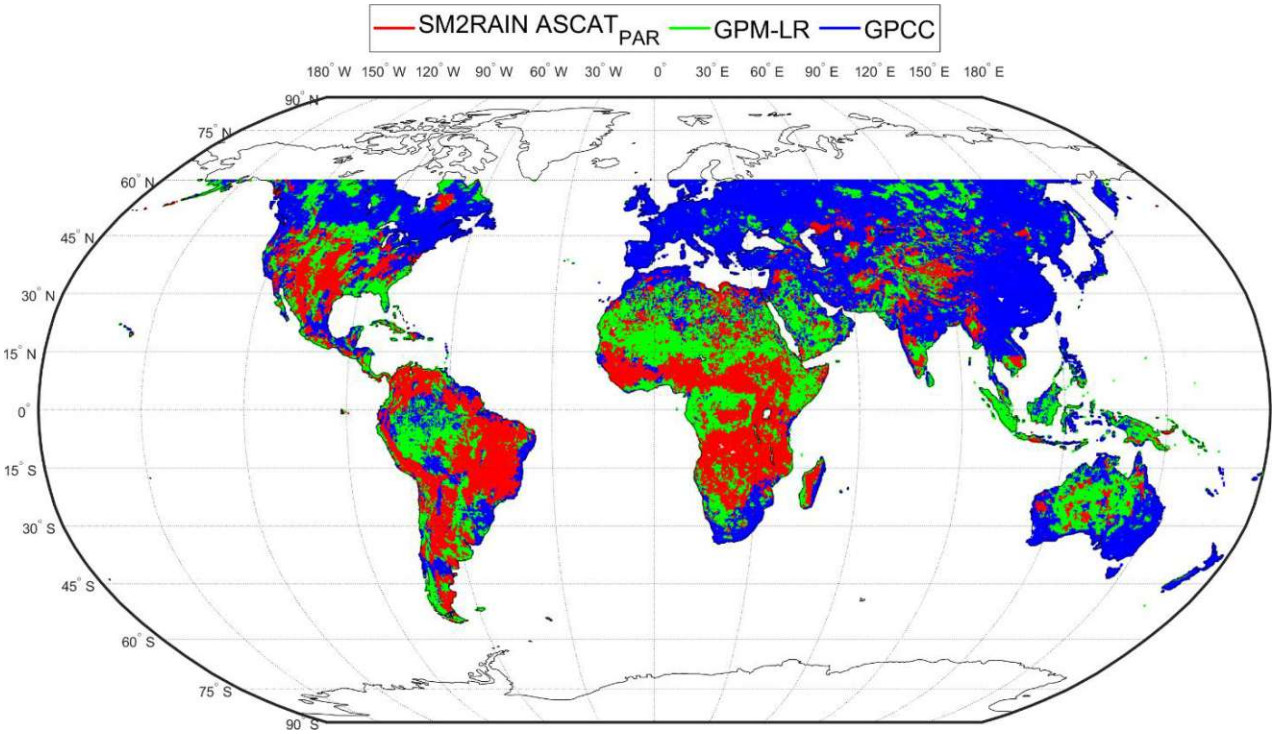


*Figure 3-9: Map of best performing products based on the results of Triple Collocation of the rainfall products triplet SM2RAIN parameterized (red), GPM – Late Run (green) and GPCC (blue). The parameterized SM2RAIN-ASCAT product outperforms the others in those areas characterized by low density of gauge and good quality of the SM retrievals.*

The obtained results confirm the possibility to obtain an estimation of SM2RAIN parameters without calibrate the algorithm by the knowledge of the average SM noise, the topographic complexity and the average rainfall of each analyzed pixel, at the cost of a slight reduction of the performances. The four relationships were therefore applied to estimate SM2RAIN parameters over the Po River basin for its application to high-resolution SM data from S1, as descripted in chapter 4.

# 4. APPLICATION OF SM2RAIN TO HIGH-RESOLUTION SOIL MOISTURE DATA

*In this chapter, the SM2RAIN parametric relationships are applied to high-resolution SM data from Sentinel-1 over the Po River Valley, in Italy. The constants of the four relationships are recalculated to take into consideration the change of the Digital Elevation Model product. The parameterized and calibrated SM2RAIN are then applied to Soil Moisture data from Sentinel-1. The obtained rainfall products are then compared to an observed rainfall dataset obtained from the fusion of gauge and weather radar data. Finally, the advantages of the high spatial resolution are evaluated through a spatial analysis of the obtained product performances.*

This chapter is based on a preprint of the following publication, submitted to Hydrology and Earth System Sciences:

*Filippucci, P., Brocca, L., Quast, R., Ciabatta, L., Saltalippi, C., Wagner, W., and Tarpanelli, A.: High-resolution (1 km) satellite rainfall estimation from SM2RAIN applied to Sentinel-1: Po River Basin as case study, submitted to Hydrology and Earth System Sciences, doi: 10.5194/hess-2021-563, 2021.*

### 4.1 Study area

The analysis was conducted over the Po River Basin, located in Northern Italy (Figure 4-1). The basin extends from the Western Alps to the Adriatic Sea, including Italian and Swiss territories. The region covers an area of around 71000 km$^2$: the Alps outline the boundaries of the basin to the North and West, with altitudes up to 4809 m, while the Apennines mark the South borders. The Po Plain extends to the central part of the basin, broadly divided into a northern and a southern section: the former is generally unsuitable for agriculture, while the latter is more fertile and well irrigated. The average annual rainfall ranges from ~700 to ~1500 mm/year in the analyzed period, 2016-2019, equally distributed during the year, with maximums occurring during autumn and spring seasons. The Po basin area is classified as Cfa (Temperate climate, without dry season and with hot summer) by the Köppen-Geiger climate classification (PEEL ET AL., 2007). In this study, the fraction of the Po River basin external from the Italian boundaries (black line in Figure 4-1) was excluded from the analysis due to the unavailability of rain gauge data.

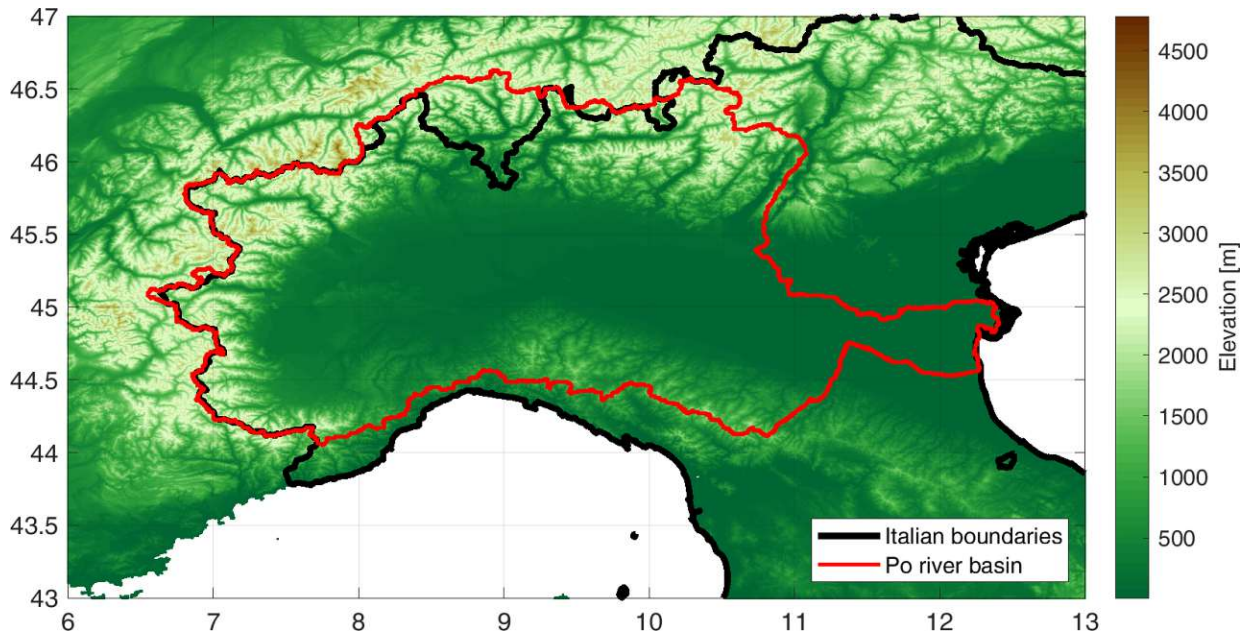


*Figure 4-1: Po River Basin elevation map from ASTGPM. The black line indicates the Italian boundaries, while the red shapes the Po River basin boundaries.*

### 4.2 Data

Several datasets were collected in this study to analyze the feasibility of high-resolution rainfall estimations from SM2RAIN. Specifically, satellite SM products from ASCAT and S1 sensors were analyzed, alongside the selected benchmark rainfall dataset, obtained by the fusion of rain gauges and weather radar measurements through the Modified Conditional Merging (MCM) algorithm (**Bruno et al., 2021**), and the data needed for the parameter estimation within the parameterized SM2RAIN algorithm, i.e., SM noise from ASCAT, topography and rainfall climatology.

*SM measurements*

SM data at 25 km spatial resolution (12.5 km spacing) were obtained from ASCAT (see section 3.1 description), while the high-resolution 1 km estimates (500 m spacing) were derived from the application of S1-RT1 algorithm to S1 data (**Quast et al., 2019**). The spatial sampling was fixed at one-half of the spatial resolution, according to the Nyquist-Shannon sampling theorem, to maximize the details of each SM dataset (Wagner et al., 2013).

S1 mission is composed by a constellation of two polar-orbiting satellites, Sentinel-1A (launched 03/04/2014) and Sentinel-1B (launched 25/04/2016), sharing the same orbital plane 180° apart, each carrying a single C-band Synthetic Aperture Radar (SAR) instrument operating at a center frequency of 5.405 GHz. S1 sensors can operate in four exclusive imaging modes with different spatial resolution (down to 5 m) and swath width (up to 400 km). Particularly, the Interferometric Wide (IW) swath mode, the main sensing mode over land, offers a 20 m x 22 m spatial resolution with a 250-km swath. The revisit time of a single satellite is 12 days, which reduces down to 6 days when considering both sensors. However, since the acquisition strategy prioritizes European landmasses over other regions, the effective temporal resolution over Europe is between 1.5 and 4 days by taking advantage of the overlapping ascending and descending orbits.

SM retrievals at 1 km spatial resolution were obtained by applying a first-order radiative transfer model (RT1) (**Quast et al., 2019**) to a 1 km S1 backscatter ($\sigma_0$) dataset sampled at 500 m pixel

spacing (BAUER-MARSCHALLINGER ET AL., 2019). RT1 is based on a parametric (first-order) solution to the radiative transfer equation (QUAST AND WAGNER, 2016) in conjunction with a timeseries-based non-linear least squares regression to optimize the difference between (incidence-angle dependent) measured and modelled $\sigma_0$. The scattering characteristics of soil- and vegetation are modelled via parametric distribution functions, and the relative SM content (scaled between 0 and 1) is found to be proportional to the nadir hemispherical reflectance (N) of the bidirectional reflectance distribution function used to describe bare-soil scattering characteristics.

To correct for effects induced by seasonal vegetation dynamics, scaled Leaf Area Index (LAI) timeseries provided by ECMWFs ERA5-Land reanalysis dataset have been used to mimic the temporal variability of the vegetation optical depth, accounting for the attenuation of the radiation during propagation through the vegetation layer. Remaining spatial variabilities in soil and vegetation characteristics are accounted for by the model-parameters "single scattering albedo" ($\omega$) and soil-scattering directionality (ts). Within the retrieval-procedure, a unique value for N is obtained for each timestamp, alongside a temporally constant estimate for ts and an orbit-specific estimate for $\omega$ for each pixel individually. A comparison of the obtained RT1 soil-moisture retrievals to ERA5-Land top-layer volumetric water content (swvl1) for a set of ~138 000 pixels over a 4-year time-period from 2016 to 2019 achieves an overall (median) Pearson correlation of 0.55 for areas classified as croplands and 0.65 for areas primarily covered by natural vegetation. A detailed description and performance-analysis of the used soil-moisture dataset will be given in QUAST ET AL., IN PREPARATION.

Due to the presence of systematic differences between S1 acquisitions from different orbits, the obtained soil-moisture timeseries exhibits a periodic disturbance, attributable to unaccounted differences in soil- and vegetation characteristics with respect to the different viewing-geometries. To correct these systematic effects, the timeseries were split with respect to the S1 orbit ID and normalized individually to a range of (0, 1) prior to the incorporation into the SM2RAIN algorithm.

In order to obtain data with the same time spacing, SM data were linearly interpolated at midday and midnight for both datasets. If no data were found within 5 days, each datum in the interval was set to Not a Number (NaN). ASCAT data were resampled on S1-RT1 grid using a weighted average of the four nearest pixels, to allow the inter-comparison of the data. Finally, all the SM products were masked for frozen soil and snow cover conditions, by downloading the Soil Temperature (Tsoil) of the first soil layer (0-7 cm) and Snow Depth data from ERA5-Land (see description below), and excluding the SM estimates obtained over pixels showing a Tsoil < 2 °C or a snow depth > 0.01 m.

*Rainfall measurements*

Two rainfall datasets were considered, to be used as benchmark for the performance assessment and as input for the parameterized version of SM2RAIN, respectively. The first one is a product derived from the integration of ground radar and rain gauge measurements over the Italian territory through the MCM algorithm (**Bruno et al., 2021**). A dense network of rain gauges and weather radars is available over the Italian territory, making it possible to obtain hourly rainfall measurements in real-time. While rain gauges allow a good estimation of punctual rainfall, radar measurements give a good estimation of the general covariance structure of rainfall. MCM uses radar data to condition the spatially limited information of rain gauges, generating a rainfall field with a realistic spatial structure constrained by rain gauge values. The resulting rainfall product is characterized by high spatial (1 km) and temporal (1 h) resolution. These attributes make it a suitable choice for the purpose of comparison with SM2RAIN estimates from high-resolution SM. In this work, the MCM hourly information was resampled to S1 data coordinates. MCM data were temporally cumulated at 12 hours, obtaining two accumulated rainfall measurements per day, respectively at midday and midnight. Rainfall measurements greater than a threshold of 800 mm/day were considered not valid and discarded from the analysis. Even if MCM data were available for the full Po River basin, the territories outside the Italian boundaries were excluded from the analysis due to the absence of rain gauges data.

In order to apply the parameterized version of SM2RAIN, the mean daily rainfall of each pixel in the study area is needed. It was obtained by downloading Total Precipitation and Snowfall daily measurements from the ERA5-Land for the period 1981-2021 (see product description in section 3.1). Daily rainfall data were obtained by subtracting the Snowfall component from ERA5-Land Total Precipitation. The obtained rainfall data were then regridded on S1 grid using a weighted average of the four nearest pixels, as done with ASCAT SM data. The 30-year averaged mean daily rainfall was then calculated for each pixel.

*Topography measurements*

Elevation data from Terra Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) global DEM Version 3 (ASTGTM) were downloaded. The product provides altitude land data at a spatial resolution of 1 arc second (~30 meters resolution at equator). In order to obtain the topographic complexity of each S1 pixel, the standard deviation of the DEM values within each 500m pixel was calculated.

Data interpolation and regridding are expected to introduce small-scale noise in the datasets. Notwithstanding this, the interpolation is unavoidable in order to analyze all the products with the same spatial and temporal sampling.

### 4.3 Rainfall validation

In chapter 3, four parametric relationships were obtained to allow to estimation of the SM2RAIN algorithm parameters along with the $T$ parameter of the original exponential filter, without calibration. It is therefore possible to deduce $T$, $Z^*$, $a$ and $b$ from the knowledge of SM timeseries and its noise, the topographic complexity and the mean daily rainfall of the standard year (obtained by averaging the rainfall in the same DOY). Notwithstanding the good performance obtained, the low resolution (5 arc minutes) of the DEM used in the relationships development, ETOPO5, forbid their use for the 1 km application. For this reason, the procedure descripted in chapter 3 was reapplied in the 1009

points by substituting ETOPO5 with ASTGTM product to recalculate the constant of the parametric relationships. In particular, the recalculated relationships are:

$$T = 0.8351 + 1.2585\,\overline{SMnoise}\,std(|SM_d|) + 0.2777\,{}^{std(|SM_d|)}/_{\bar{P}}\,topC \quad (21)$$

$$Z^* = 10.0678 + 0.5350\,{}^{\bar{P}}/_{\overline{|SM_{fd}|}} \quad (22)$$

$$a = -1.3177 + 13.3579\,Z^*\,\overline{|SM_{fd}|} \quad (23)$$

$$b = 3 + {}^{2}/_{0.4118\ +\ 0.324 * \log a} \quad (24)$$

where $\overline{SMnoise}$ is the average SM noise in the considered pixel, $(|SM_d|)$ is the standard deviation of the absolute values of the SM temporal variations, $\bar{P}$ is the pixel mean daily rainfall, $topC$ is the topographic complexity and $\overline{|SM_{fd}|}$ is the average of the absolute values of the filtered SM temporal variations

In order to obtain rainfall measurements from the SM datasets, SM2RAIN algorithm was applied to both ASCAT and S1-RT1 SM products by using both the calibrated and parameterized versions. In the calibrated SM2RAIN, the algorithm parameters were estimated by minimizing the RMSE between SM2RAIN rainfall estimations and MCM rainfall product at daily time scale for both ASCAT and S1-RT1 SM. For the parameterized SM2RAIN version, the algorithm parameters were obtained through the parametric relationships mentioned above. Since no information regarding S1-RT1 SM noise was available, ASCAT SM noise characteristics were used to calculate S1-RT1 SM2RAIN parameters, assuming that since both ASCAT and S1 sensors operate in C-band, the noises affecting the two SM products are similar. Indeed, the noise level of S1-RT1 is expected to be higher than ASCAT one. This sub-optimal configuration can be therefore considered as a first step to test the data: better results should be obtained when more accurate noise information will be available.

The obtained rainfall can then be accumulated at the desired time step. In order to consider the different temporal resolution of the selected SM products (sub-daily for ASCAT and from 1.5 to 4

days for S1), three accumulation time steps were chosen: 1 day, 10 days and 30 days. The daily rainfall was calculated only for ASCAT product, since the low temporal resolution of S1 prevents to obtain significant results at daily intervals.

Figure 4-2 shows the average 30 days rainfall obtained by the application of the parameterized SM2RAIN to ASCAT and S1-RT1 SM products. By comparing the two figures, the improved resolution of the rainfall obtained from S1-RT1 SM with respect to ASCAT SM is evident: the higher spatial resolution of S1-RT1 allows the generation of detailed features, even if with a granular effect likely due to the uncertainties of the measurements, and with patterns related to the spatial variation of S1 temporal resolution (compare with Figure 4-4e).

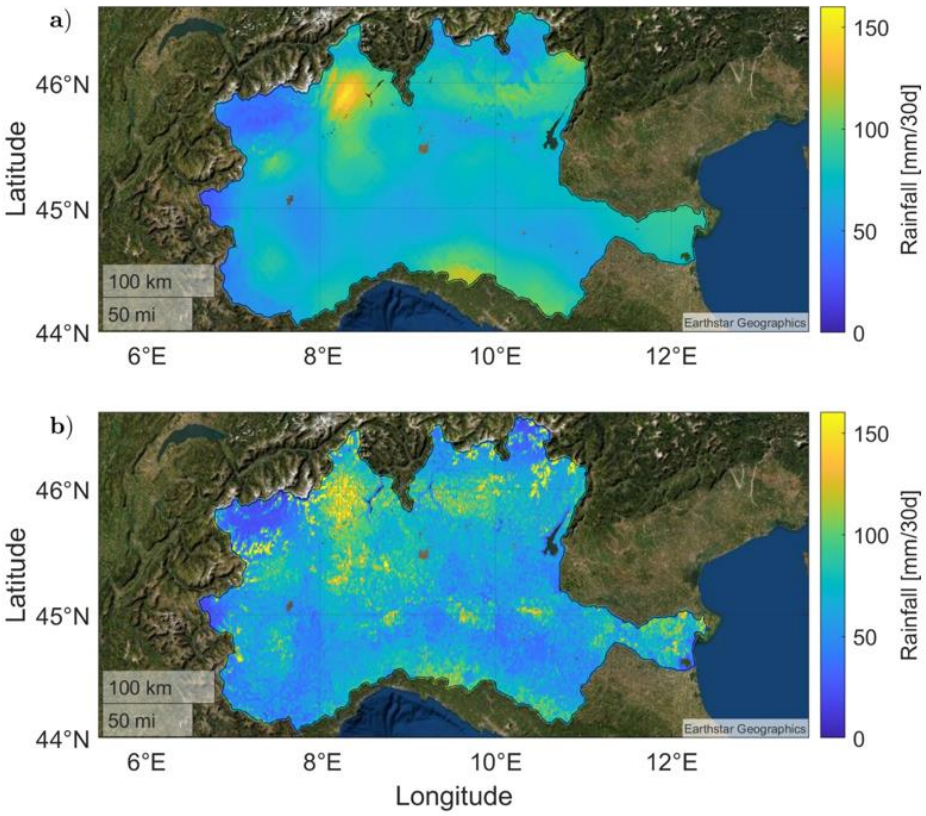


*Figure 4-2: Estimated average 30 days rainfall from the parameterized SM2RAIN applied to ASCAT (Panel a) and S1-RT1 (Panel b) SM product for the period 2016-2019. Map copyright ©2021 GeoBasis-De/BKG (©2009), Google, Inst. Geogr. Nacional Immagini ©2021 TerraMetrics*

The results of R, RMSE and BIAS with respect to the selected time-steps are shown in Figure 4-3. In order to maximize the reliability of the obtained performances, the rainfall accumulation was carried

out by summing up only timestamps available in both the SM2RAIN estimations and the benchmark, for each SM2RAIN product separately. In this way S1-RT1 performances can be better assessed, since a direct accumulation would penalize this product due to the long period of no-data caused by S1 low temporal resolution.

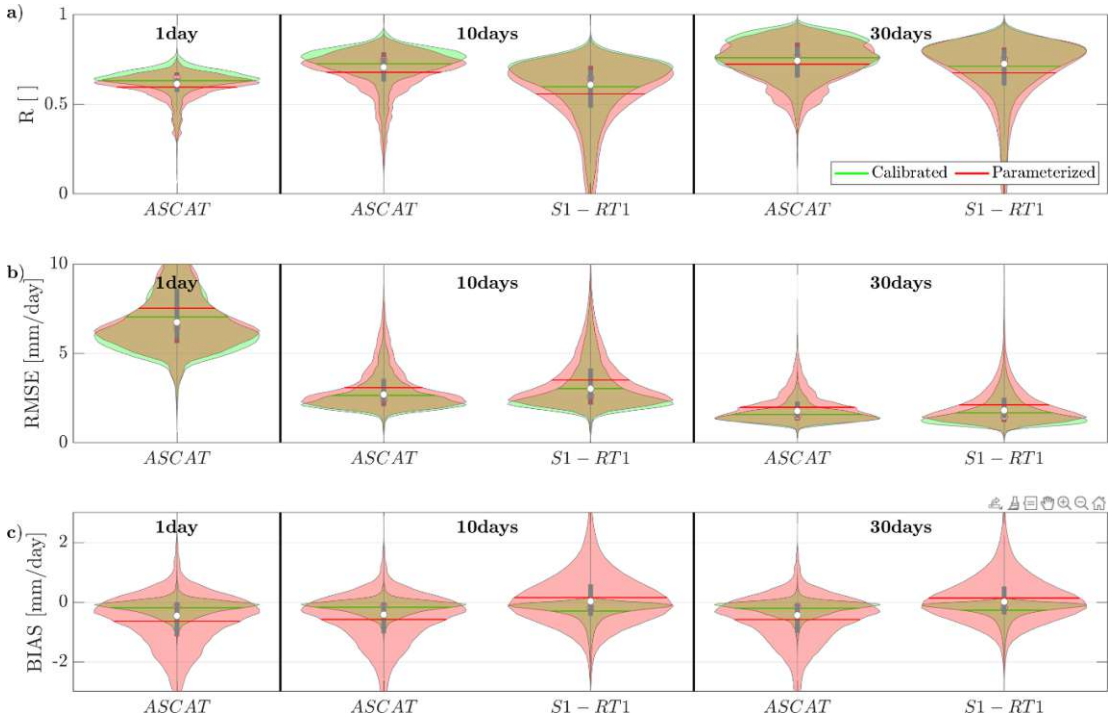


*Figure 4-3: Violin plots of Pearson's Correlation (R, panel a), Root Mean Square Error (RMSE, panel b) and BIAS (panel c) between the rainfall from MCM and from SM2RAIN applied to ASCAT and S1-RT1. ASCAT-derived rainfall was accumulated at 1, 10 and 30 days, while the rainfall from S1-RT1 was accumulated at 10 and 30 days. Each violin shape is obtained by rotating a smoothed kernel density estimator. The green violins are obtained by calibrating SM2RAIN against MCM, while the red violins derived from the parameterized SM2RAIN procedure.*

The SM2RAIN product obtained from ASCAT allows to well reproduce the rainfall of the Po River basin at daily time scale thanks to the high temporal resolution of ASCAT (sub-daily frequency), with a median R of 0.61 for the parameterized product and 0.64 for the calibrated product, confirming the good quality of the data and the importance of its temporal resolution. At higher aggregation time steps, the median R of the parameterized (calibrated) ASCAT-derived rainfall products improve to 0.71 (0.75) for the 10 days accumulation period and to 0.74 (0.77) when 30 days accumulation is considered. Good results are also obtained from the application of SM2RAIN to S1-RT1, with a median R of 0.61 (0.65) and 0.73 (0.75) at 10 and 30 days accumulation time, respectively. Albeit

ASCAT-derived rainfall performs better than the one from S1-RT1 at 10 days, they are equally good for the 30 days accumulated rainfall. The results also confirm the good capabilities of the parameterized SM2RAIN algorithm in rainfall estimation, considering the small differences between the performances obtained by the two algorithm versions. The only exception is the BIAS index, which, as expected, is significantly larger in the parameterized products compared to the calibrated ones. The increased BIAS is due to the ERA5-Land data used to obtain the climatology of the area since its spatial resolution is much lower than the one adopted for this study (i.e., 1 km) and the average spatial pattern of rainfall is quite different from the one measured by MCM.

### 4.4 Spatial validation of rainfall products

Even if the ASCAT product (with lower spatial resolution) is on average the best performing, the spatial comparison of the performances is important to understand the added value of high-resolution SM. In order to better evaluate the differences between the rainfall estimated from ASCAT and S1-RT1, the Pearson's correlation of the 30 days accumulated rainfall derived from the two SM products are analyzed in this section. This temporal step was selected since it is suited for a quality comparison of the two products, being less influenced by the different temporal resolution of the sensors, and because it is optimal for agricultural application.

Generally good performances are obtained from both rainfall products, as shown in Figures 4-4a and 4-4b. Some areas with low R values are shared by both ASCAT and S1-RT1 derived rainfall products. Over mountain areas the errors are mostly related to the lower accuracy of C-band SM data, due to shadowing effects and layover (ULABY ET AL., 1981). The presence of water bodies at the river outlet and over the paddy fields in the western part of the Po basin is also affecting SM, and hence rainfall retrieval accuracy. Finally, the yellow "holes" in the correlation maps resemble the errors caused by low quality gauge data, which affect the rainfall estimation surrounding the gauge sensor. It should also be noticed that many low performing areas are located close to urban centers, which may affect both the SM retrieval quality and the rain gauge measurements (see also appendix, section 7.2).

Notwithstanding this, it is impossible to remove the alleged "bad" gauge stations from the benchmark product, as MCM is an operative product and the clear identification of these stations is often challenging.

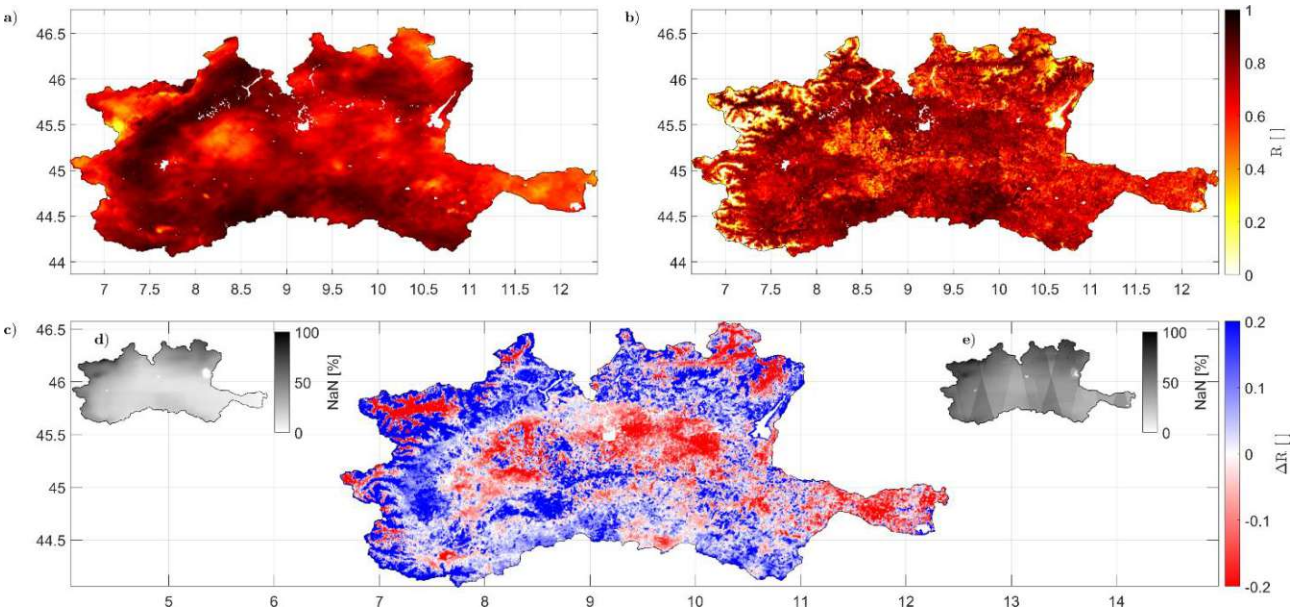


*Figure 4-4: Spatial Pearson's correlation (R) between the 30 days accumulated rainfall derived from MCM and the application of the parameterized SM2RAIN to ASCAT (panel a) and to S1-RT1 (panel b) SM products. Panel c shows the difference between ASCAT and S1-RT1 correlation maps, while panel d) and e) show the percentage of not valid images per pixel respectively for ASCAT and S1-RT1.*

The spatial comparison between the performances of the ASCAT and S1-RT1 derived rainfall is shown on Figure 4-4c, displaying the difference between the correlation values of the two products. Red area means that the S1-RT1 product is performing better, whereas blue areas highlight where ASCAT is providing more accurate rainfall estimates. First of all, it should be noted that while ASCAT derived rainfall product shows average correlation values over the mountainous region in the North and West of the map (see Figure 4-1 for comparison with the DEM map), S1-RT1 correlation are either extremely low or extremely high. This important difference is caused by the high spatial resolution of S1-RT1 product: the improved resolution permits to clearly distinguish the "good" signal originating from the valleys and the "bad" signal coming from the mountain slopes, affected by the noise generated from the aforementioned shadowing and layover effects. This distinction results in areas with respectively very good (valleys) and very bad (mountains) rainfall

estimations. The spatial resolution of ASCAT on the other hand does not permit to distinguish the signals of the two geographical features, causing lower performances over the valleys and higher performances over the slopes in comparison with S1-RT1. The low performances of the pixels located over the mountain slopes are also responsible for the long violin plot tails of S1-RT1 performances that can be noticed in Figure 4-3. S1-RT1 results are particularly lower than those from ASCAT due to the fact that S1-RT1 product calibration was carried out without considering any snow masking, thus reducing the quality of the solution in the pixels affected by snow cover.

A smaller difference in performance can be noticed over the plain, in particular in the north-eastern section, where S1-RT1 rainfall performs overall better than ASCAT. Conversely, in the southern section and specifically over the areas surrounding the Po River and its tributaries, ASCAT derived rainfall is better than S1. An explanation of this behavior can be found in the intensive irrigation practice over this area. Irrigation events cause an increase of the fields SM (Filippucci et al., 2020) that should be sensed by satellites sensors. However, the area surrounding the Po River is composed by many small fields (few hectares each) managed by different farmers, where the irrigation timing is not concurrent. The ASCAT sensor is not able to distinguish the resulting irrigation signal (Brocca et al., 2018) because of its low spatial resolution (25 km) that cause the signals of each field to overlap and average with each other. S1, instead, is more sensitive to the irrigation signal, thanks to its higher spatial resolution.

Considering that the rainfall benchmark product does not contain irrigation information, the drop in Pearson's correlation of the S1-RT1 derived rainfall with respect to ASCAT could be related to the sensitiveness of the former to the aforementioned irrigation events, and not to the SM signal quality. It could be an additional information of great scientific interest but, unfortunately, the absence of detailed irrigation data for the Po Valley makes difficult to verify this hypothesis.

Finally, it should also be noted that this analysis could be biased in the areas characterized by a high presence of missing values (NaN) for one product with respect to the other, which hampers the

statistical significance of the performance indices. Notwithstanding this, the absence of patterns in the maps that resemble the NaN distribution percentage shown in Figure 4-4d and 4-4e, fosters the validity of the analysis.

The performance comparison with respect to RMSE and BIAS and a comparison of the calibrated SM2RAIN products is omitted for the sake of brevity, because no relevant additional information can be obtained from it.

In Figures 4-5 and 4-6, rainfall and SM timeseries of two pixels selected in the north-west of the Po basin are shown, as an example of the increased capacity of S1-RT1 for rainfall retrieval in the mountainous area. Since these pixels are selected in a topographic complex area, they should not be considered representatives of the overall performance and availability of the satellite rainfall products, rather an example of the improved performance derived from the use of S1-RT1 high-resolution SM with respect to ASCAT. Winter and early-spring measurements are masked in both pixels, due to frozen condition or snow cover, according to ERA5-Land information. The pixel in Figure 4-5 is selected over one of the mountain valleys of the Italian territory (7.152°E, 45.710°N), inside the Italian region Valle d'Aosta, in order to show how S1 spatial resolution increases the capabilities in rainfall estimation over such a region. By observing the rainfall timeseries in Figure 4-5a and the standard month distribution in Figure 4-5b, it can be noted how S1-RT1 derived rainfall is in better accordance with the observed one, in particular during autumn months. During late spring and summer, S1-RT1 and ASCAT estimates are more similar, while S1-RT1 often underestimates the observed rainfall, also with respect to ASCAT. In Figure 4-5d, the same behavior can be noted on the averaged SM trends, with the SM sensed by S1-RT1 being on average less than the one from ASCAT during late spring-summer and greater during the autumn season, probably due to the additional vegetation correction operated within S1-RT1.

Figure 4-6 shows the timeseries of a pixel selected over the mountain slopes, in the vicinity of the previous one (7.410°E, 45.824°N). While ASCAT SM estimates (Figure 4-6c and 4-6d) show patterns

that are similar to those in Figure 4-5, S1-RT1 signal is completely different. The SM saturates in the summer period and goes down in autumn, with a strong seasonality that is poorly affected by the rainfall events. This is most probably an issue of the vegetation-correction, since it adds a strong seasonality to pixels that realistically exhibit little vegetation coverage, also due to the low spatial resolution (with respect to S1-RT1) of the LAI product used for correcting the vegetation-seasonalities. This erroneous vegetation seasonality is then counteracted by an erroneous SM seasonality. As expected, the poor quality of SM estimations, greatly affects SM2RAIN capabilities in calculating rainfall in these areas, resulting in very high rainfall rate perceived during summer and very low one during winter, in contrast with the observed data.

Finally, Figure 4-7 shows the timeseries of a pixel selected over the plain (10.684°E, 44.805°N). As can be noted, the period of unavailability of the rainfall datum is greatly reduced in comparison with Figure 4-5 and Figure 4-6, since this area is characterized by higher temperature during the winter and by minor snow cover probability. Overall, S1-RT1 SM shows a greater variability during the summer season with respect to ASCAT (Figure 4-7c, 4-7d), thanks to both the vegetation correction and the higher spatial resolution. This leads to a greater accuracy in the peak rainfall detection of summer 2018 and 2019 (Figure 4-7a) is found. On the other hand, an overestimation of 2017 summer rainfall (potentially due to an error in SM estimation or to an irrigation event) and an underestimation of winter 2019 (probably due to SM saturation). Overall, the rainfall estimate from S1-RT1 is in good accordance with the observed one (Figure 4-7b), proving both the validity of the derived rainfall product and its usefulness for hydrologic modelling.

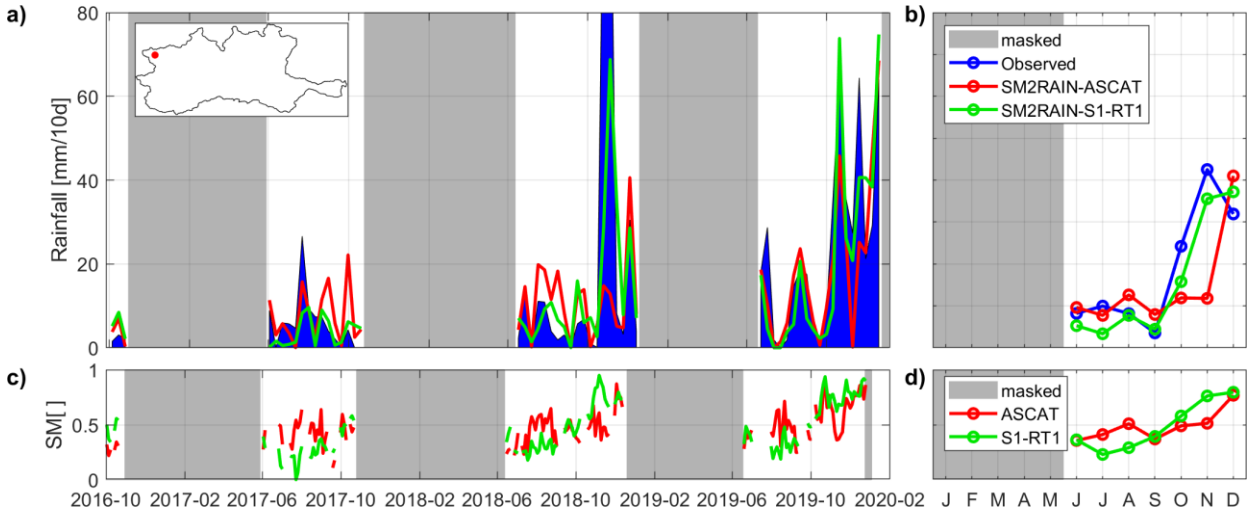


*Figure 4-5: Example of SM and rainfall timeseries over a pixel (7.152°E, 45.710°N) situated within valley inside the mountainous region. In panel a, the timeseries of the observed (blue) and estimated (red SM2RAIN-ASCAT, green SM2RAIN-S1-RT1) 10-days accumulated rainfall products are shown, while panel c displays SM timeseries averaged with a 3 days window. Finally, panel b and d contain the standard month average of the rainfall and SM products, respectively. The periods masked for frozen soil condition or snow cover are highlighted in grey.*

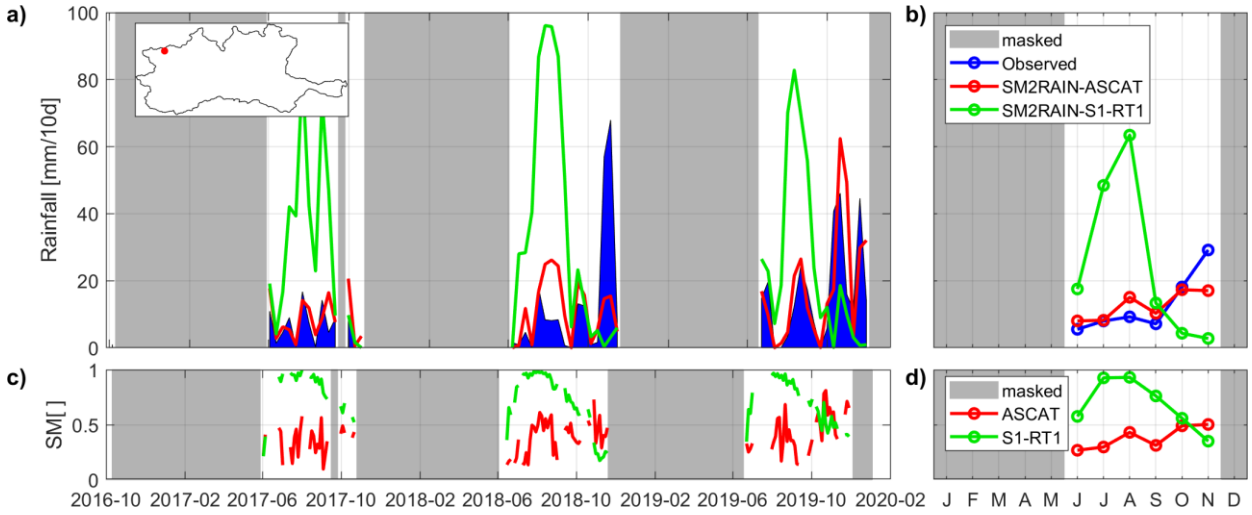


*Figure 4-6: Example of SM and rainfall timeseries over a pixel (7.410°E, 45.824°N) situated over the ridges of the mountainous region. In panel a, the timeseries of the observed (blue) and estimated (red SM2RAIN-ASCAT, green SM2RAIN-S1-RT1) 10-days accumulated rainfall products are shown, while panel c displays SM timeseries averaged with a 3 days window. Finally, panel b and d contain the standard month average of the rainfall and SM products, respectively. The periods masked for frozen soil condition or snow cover are highlighted in grey.*

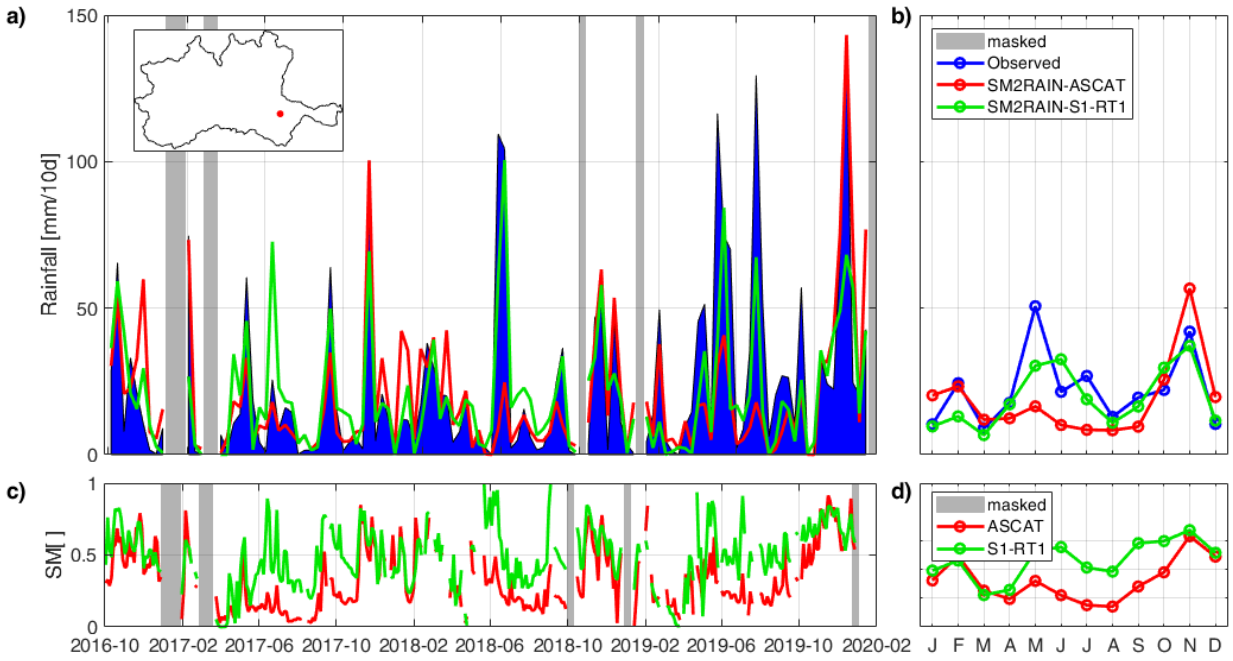


*Figure 4-7: Example of SM and rainfall timeseries over a pixel (10.684°E, 44.805°N) situated over the Po River Plain. In panel a, the timeseries of the observed (blue) and estimated (red SM2RAIN-ASCAT, green SM2RAIN-S1-RT1) 10-days accumulated rainfall products are shown, while panel c displays SM timeseries averaged with a 3 days window. Finally, panel b and d contain the standard month average of the rainfall and SM products, respectively. The periods masked for frozen soil condition or snow cover are highlighted in grey.*

# 5. RIVER DISCHARGE MONITORING

*In this chapter, the river discharge estimation over the Po and Tiber rivers is descripted. Following the data illustration, both the original CM method and the modified approach, CMW, are applied to the two study areas. Four configurations are tested for the selection M: the first three are calibrated over the pixel(s) of maximum Spearman's correlation between the observed data and the analyzed index, averaging a gradually increasing number of pixels to take into consideration the saturation of the information: 1 pixel, 3x3 pixels and 9x9 pixels respectively. The last configuration is not calibrated, rather it is obtained by selecting those pixels that are lowly correlated with C and W. The performances of each index is then evaluated for both the study areas, and the timeseries are analyzed to identify strength and weakness of the method.*

---

This chapter is based on a preprint of the following publication, submitted to Remote Sensing of Environment:

*Filippucci, P., Brocca, L., Bonafoni, S., Saltalippi, C., Wagner, W., and Tarpanelli, A.: Sentinel-2 high-resolution data for river discharge monitoring, submitted to Remote Sensing of Environment, 2021.*

---

### 5.1 Study area

The analysis is focused on two sites located in Italy: the first one along the Po River at the Pontelagoscuro station (44.89 °N, 11.60 °E) and the second one along the Tiber River at the Montemolino station (42.80 °N, 12.40 °E). These study areas were selected on the basis of: 1) the availability of in-situ flow measurements during the analyzed periods (2015-2020); 2) the different regime of flow and 3) the width of the river that is medium-small.

The Po River is located in Northern Italy and it has a drainage area of 74'000 $km^2$: it is the longest Italian river and the fifth European river for river discharge. During the analyzed period the average discharge value at Pontelagoscuro station was equal to 1215 $m^3/s$, with the minimum and the maximum values being equal to 406 and 5905 $m^3/s$, respectively. The maximum cross-section width is around 350 m, during the seasonal peaks in spring (May-June) and autumns (November). The river discharge data are available online from the website of the Regional Agency for Environmental Protection (www.arpae.it/documenti.asp).

The Tiber River in central Italy has a drainage area of 17'375 $km^2$ and its average river discharge at Montemolino station is much smaller than the one of the Po River: during the analyzed period it is around 30 $m^3/s$, with the minimum and maximum extremes being equal to 1.6 and 804 $m^3/s$, respectively. The river width is about 50 meters, surrounded by riparian vegetation. The data of Montemolino are provided by the Department of Environment, Planning and Infrastructure of the Umbria Region, which is in charge for the collection of hydro meteorological data in the Umbria Region.

### 5.2 S2 data in Near-InfraRed band

The S2 mission is composed of two satellites, Sentinel-2A (launched on the 23 June 2015) and Sentinel-2B (launched on 7 March 2017). The satellites are placed in the same sun-synchronous orbit, phased 180° to each other. This strategy permits to increase the revisit time of S2 from 10 days (using one satellite) to 5 days (using both of them), with higher revisit frequency in the areas where adjacent

orbits overlap. Each satellite carries a MultiSpectral Instrument (MSI) sensor, capable of sensing electromagnetic radiations in 13 spectral bands, from visible to short-wave infrared. In this analysis, data from band 8 (central wavelength ~833 nm, bandwidth=106 nm) are selected, characterized by a spatial resolution of 10 meters. Two products are available: the Level-1C product with orthorectified TOA (Top-Of-Atmosphere) reflectance and the Level-2A product with surface reflectance. Generally, the atmospheric correction applied to obtain the Level-2 product over land is different from the one applied over ocean/sea. The river is a fairly complex and mixed environment to be managed. In the absence of detailed studies on this, and because the effects of the atmosphere in clear-sky conditions can be assumed negligible when computing the ratio of neighboring pixels in the same area (as used in this study, see section 2.2), we selected Level-1C product to be used in this analysis.

The dimension of each image (~600 MB) and the number of images to be downloaded for a temporal analysis, make the process of collecting and reading the data not easy and computationally challenging. In order to show the method and illustrate the new algorithm, we referred to the S2 data Level-1C for bands 8 dataset, provided by the Meteorological Environmental Earth Observation (MEEO – www.meeo.it) company, together with the relative cloud mask probability product. This collection includes a subset of Pontelagoscuro area in the period January 2018 – September 2019 and it was used to generate the figures shown in the section 2.2. However, Google Earth Engine platform was used for the analysis relative to the period 2015-2020 over both Pontelagoscuro and Montemolino stations, in order to lighten the computational load: the results and conclusions at chapters 5 and 6 are referring to the cloud data.

### 5.3 Google Earth Engine cloud platform

Following the algorithm definition, the analysis at the study sites was carried out through Google Earth Engine, GEE (Gorelick et al., 2017), a cloud computing platform for processing satellite imagery and many kinds of geospatial and observation data. GEE comprehends a catalogue of satellite imagery and geospatial datasets of over twenty petabytes at planetary scale, together with the analysis

capabilities needed to process them. The data are freely available for scientist, researcher and developers, who can access and process them through the Earth Engine API both using Javascript or Python. The development of Web-Application is also possible from the code editor interface. For this study, we developed a code to read, extract and analyze S2 data for the period 2015-2020 at both the considered gauge stations, Pontelagoscuro and Montemolino.

### 5.4 Po river: Pontelagoscuro analysis

A total of 238 images of S2 Level 1C band 8 were available on GEE (accessed on 17/06/2021) for the Pontelagoscuro area over the Po River during the period January 2015 – December 2020. After the cloud masking procedure, 168 cloud free images were retained and analyzed through the CM and the CMW procedure. The timeseries for the four configurations are shown in Figure 5-1, together with the hydrograph of the observed river discharge at the Pontelagoscuro station. All the configurations were successful in detecting the river dynamic. $CMW$ approach shows a general better agreement with the observed discharge, in particular at the beginning of the flood events in March 2016, May and November 2018, and November 2019, when the high river discharge triggered a high level of sediment transportation. This can also be noticed by observing the scatter plots on the right of Figure 5: the green dots in the top-left angles obtained through the $CM$ approach, characterized by low reflectance ratios and high discharge values due to the high sediment transportation condition, are avoided with $CMW$ approach.

The results of the different configurations show that by increasing the number of pixels for the calculation of $M$ (from the first to the fourth configuration, respectively on the top and the bottom of the Figure 5-1) $CM$ approach gradually agrees with $CMW$ approach, with the exception of the flood events above mentioned. This can be attributed to the fact that by selecting multiple pixels, the reflectance ratio is more sensitive to river discharge variation and less sensitive to residual sources of error, e.g., residual clouds and cloud shadows, and therefore the detection capacities of both approaches increase. Even if the new formulation is in good agreement with the observed river

discharge and improve the previous methodology, the peak floods experienced in November 2018 and November 2019 are still underestimated for all the configurations. This is probably due to the limited river width variation for very high flow. The pixels sensitive to the variation for very high flow are extremely few, due to the river embankments, and none of them was selected by the *M* configurations. Such a discharge peak is in fact encountered just two-times during the study period and therefore it has too low importance in terms of Spearman's correlation to be selected by the calibrated approaches and too low statistical significance to be selected by the uncalibrated approach. Finally, there are also a few peaks that are still not detected, due to the low temporal resolution of S2 (e.g., see the flood event in December 2016). Notwithstanding this, the reproduction of the river discharge follows quite well the high and low flows.

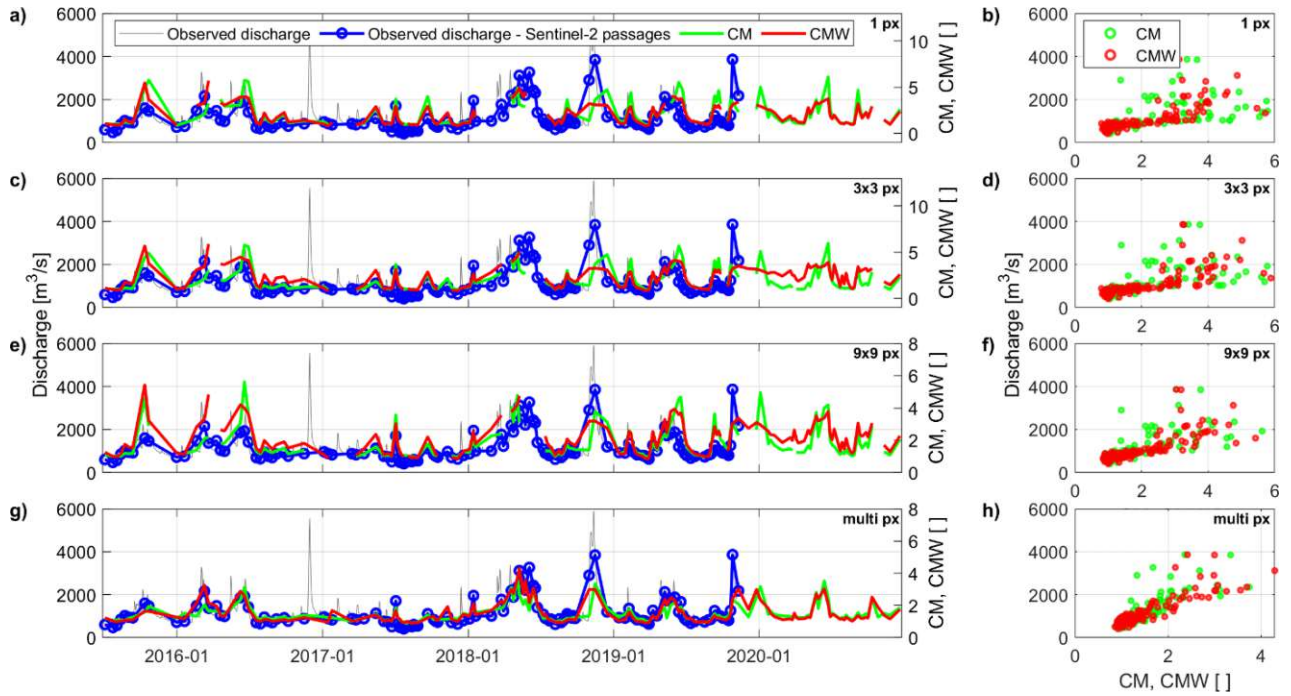


*Figure 5-1: Po River at the Pontelagoscuro station: comparison in terms of temporal series (a, c, e, g) and scatter plots (b, d, f, h) between the ground observed discharge and the CM (green) or CMW (red) timeseries for the configuration at single pixel (a and b), at 3x3 averaged pixel (c and d), at 9x9 averaged pixel (e and f) and for multiple uncalibrated pixels (g and h).*

The obtained improvement in river discharge estimation was confirmed by the correlations S and R shown in Table 5-1. In terms of Pearson's correlation, an improvement of the *CMW* approach with respect to the *CM* approach can be noticed for all the configurations, with increment that goes from

+0.06 to +0.18. Similarly, it can be observed an improvement by the use of multiple pixels rather than a single one that goes from +0.03 to +0.25 for the two approaches, again in terms of R. Spearman's correlation also improved, but the variations were slighter. In terms of river discharge estimated by applying the cubic law of Eq. 7, the performance scores of Table 5-1 underline similar results, with a globally decrease of rRMSE up to 16% and an increase of NS up to 0.32, considering both the multiple pixels and the different methodology.

*Table 5-1: Performance scores for the different configurations (1 for single pixel, 2 for 3x3 pixels, 3 for 9x9 pixels and 4 for multiple pixels) for the period 2015-2020 at Pontelagoscuro station. Spearman's (S) and Pearson's (R) correlations are calculated between the CMW or CM timeseries and the ground observed river discharge. Root Mean Square Error (RMSE), relative Root Mean Square Error (rRMSE) and Nash-Sutcliffe efficiency, NS, refer to the simulated discharge calculated according Eq. 7 and the ground observed discharges.*

| Configuration | CM | | | | | CMW | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | S [-] | R [-] | RMSE [m³/s] | rRMSE [-] | NS [-] | S [-] | R [-] | RMSE [m³/s] | rRMSE [-] | NS [-] |
| 1 | 0.82 | 0.55 | 515 | 0.45 | 0.38 | 0.88 | 0.73 | 444 | 0.39 | 0.54 |
| 2 | 0.84 | 0.58 | 544 | 0.48 | 0.30 | 0.88 | 0.72 | 436 | 0.39 | 0.53 |
| 3 | 0.86 | 0.68 | 475 | 0.42 | 0.44 | 0.89 | 0.77 | 406 | 0.36 | 0.59 |
| 4 | 0.86 | 0.80 | 444 | 0.38 | 0.56 | 0.84 | 0.87 | 366 | 0.32 | 0.70 |

By moving from the single to aggregated pixels, the performance increased, reaching the best values for configuration 4. This is justified by the S2's pixel dimensions, which are too small with respect to the Po River width: it is convenient to aggregate multiple pixels for a proper representation of the river discharge. Details on the pixels selected from each mask are provided in this section (Figure 5-2), along with the correlation maps used for their selection. In particular, Figure 5-2a shows the pixels of the box with lower (less than 5th percentile) temporal coefficient of variation, selected to obtain the $C$ timeseries. As expected, they were mostly concentrated over roads and bare soil areas, due to the temporal stability of the reflectances measured in these areas. Figure 5-2b shows the pixels within the water mask selected for the $W$ timeseries. They represent the lower (less than 5th percentile)

product between the average reflectance and its standard deviation, represented in Figure 5-2d. Figure 5-2c represents the selected $M$ pixels: the red area shows the pixels averaged for the 9x9 pixels calibration (configuration 3); the blue area represent the pixels selected by the 3x3 calibration (configuration 2); the green dot shows the single pixel showing the highest Spearman's correlation with the observed discharge (configuration 1); and, finally, the light grey area shows the $M$ pixels selected by the configuration 4 (without using any ground observation for calibration) and coming from the evaluation of the S between the reflectance timeseries of each pixel included in the water mask (except those pixels already used for $W$ and $C$) and the timeseries already calculated for $C$ (or $W$) represented in Figure 5-2e (or 5-2f). Specifically, the blue pixels in Figure 5-2e are lowly correlated with $C$, and coincide with those showing greater correlation with low flow (compare with Figure 2-3c) and to those always inside the water. Conversely, the pixels in yellow in Figure 5-2f are highly correlated with $W$ timeseries (stable water pixels) and therefore identify the inner river, while the blue pixels coincide with those showing greater correlation the high flow (compare with Figure 2-3d). Permanently flooded pixels (highly correlated with $W$) are also poorly correlated with $C$ (Figure 5-2e), therefore the condition of S with $W < 0.7$ was needed to avoid their selection for the $M$ mask in configuration 4. This condition helps reaching a tradeoff between the optimal number of pixels selected for $M$ and the sensitiveness of the procedure to the high river discharge, by limiting the pixels selected because the low correlation with $C$, sensitive to low flow condition. In fact, the pixels sensitive to high flow condition are expected to be less than those sensitive to low flow condition due to the nature of the phenomenon: therefore, by limiting the selection of the pixels sensitive to low flow, the sensitiveness to different water regimes should be obtained.

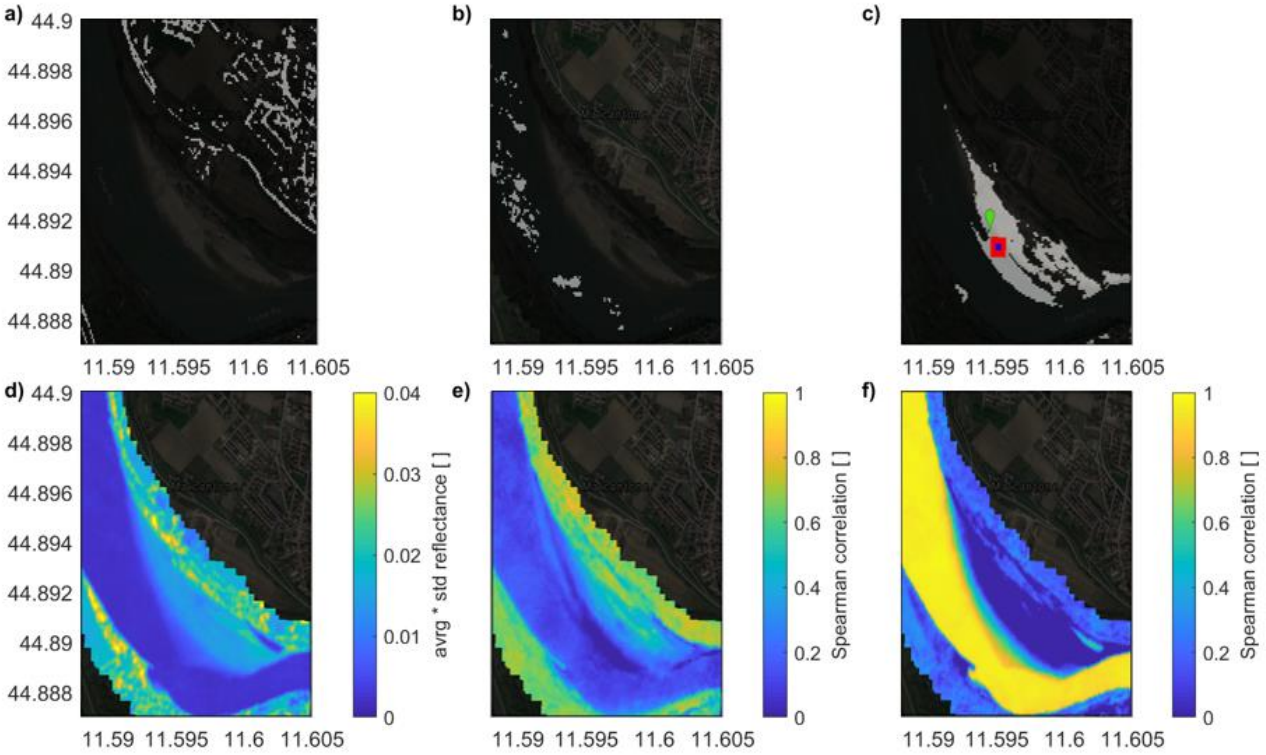


*Figure 5-2: Po River at Pontelagoscuro: in light gray the pixels used for calculating C (panel a), W (panel b). Panel c shows the location of the pixel M obtained from configuration 1 (1 pixel, greed dot), configuration 2 (3x3 pixels, blue square), configuration 3 (9x9 pixels, red square) and configuration 4 (multiple uncalibrated pixels, light gray). Panel d) shows the product between the average and standard deviation of the NIR reflectance for each pixel within the water mask; panel e) and f) represent the Spearman's correlation map between the reflectance timeseries of each pixel in the water mask and the average reflectance calculated for the C (panel e) and W timeseries (panel f), respectively. Background Copyright ©2021 Immagini ©2021, CNES / Airbus, Maxar Technologies.*

In order to better asses the source of errors, the S2 images of the days when estimated river discharge were very different from the observed one were analyzed and two main causes of error were detected. The first one is related to the clouds: even if the S2 cloud probability mask allows to remove the majority of the them, some residual clouds were undetected and unmasked, whereas in some cases the mask fails in correctly identifying the clouds and filter out valid pixels. An example of these issues is shown in Figures 5-3a, 5-3c and 5-3b, 5-3d. The first couple of panels show the S2 1c band 8 image collected on 24/05/2018 over Pontelagoscuro area, before and after the application of the cloud probability mask at 50%. Residual clouds can be clearly distinguished over the river. Furthermore, cloud shadows were not considered in the masking process, and the pixels affected by those were not removed. Clouds are very bright in the NIR, therefore their presence cause river discharge

overestimation when they overlap with $C$ or $W$ and underestimation when they coincide with $M$ pixels. In contrast, cloud shadows reduce the reflectance of the affected pixels, causing underestimation when they coincide with $C$ or $W$ and overestimation when they coincide with $M$. Figures 5-3b-5-3d, instead, show the image collected on 10/06/2017: few clouds were present on the west area of the river, but the algorithm masked out a large number of pixels that were confused with clouds, hampering the river discharge estimation. The second issue is related to the reflectance values of $C$. It was noticed that in some days the reflectance values of the terrain pixels were significantly lower/higher than in other days, potentially due to e.g., the effect of soil moisture variations or bidirectional reflectance effects. These conditions led to a lower/higher $C$ value, which cause the under/overestimation of $CMW$ and consequently of the flow peaks, worsening the effect of the pixels' saturation.

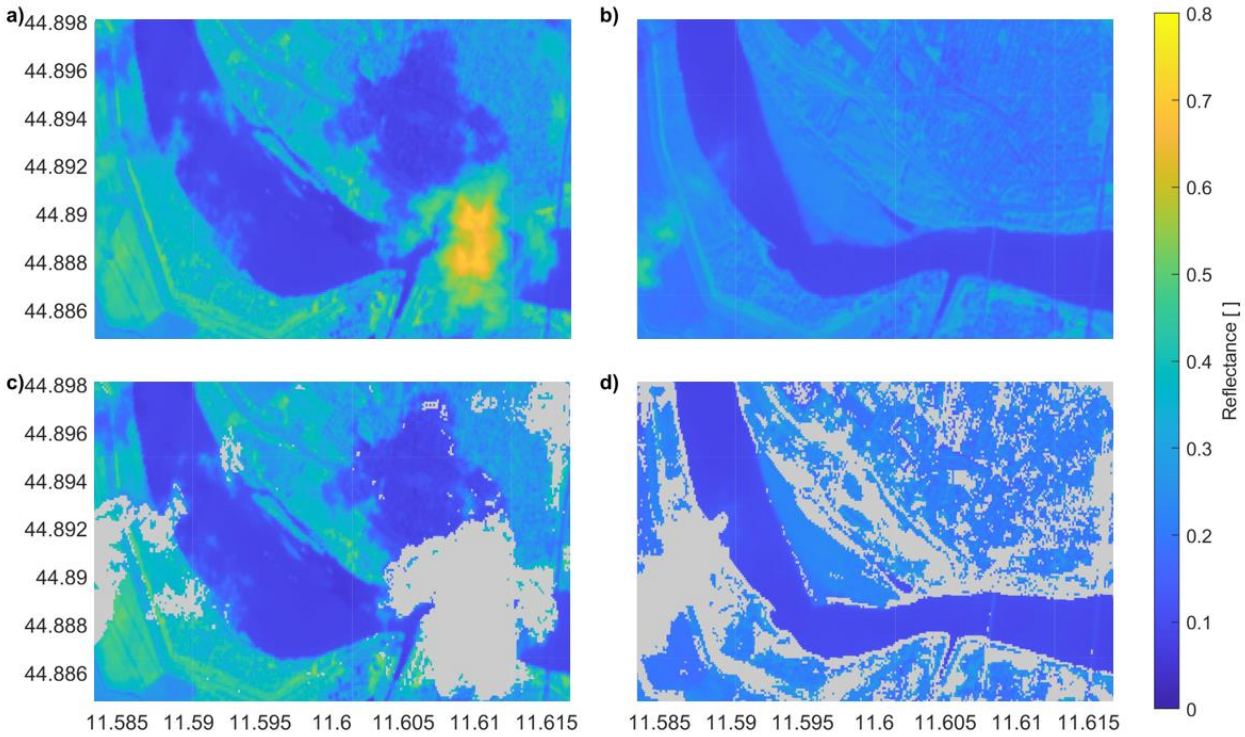


*Figure 5-3: Po River at Pontelagoscuro: S2 1c Band 8 images acquired on 24/05/2018 (Panel a and c) and on 06/10/2017 (Panel b-d), prior (a and b) and after (c and d) the application of the cloud mask equal or above 50% of cloud probability.*

### 5.5 Tiber river: Montemolino analysis

With the purpose to investigate the limit (or the benefits) of the procedure when applied to narrow rivers, the same analysis was implemented at Montemolino station, along the Tiber River. The total number of available images for the period 2015-2020 was 529, obtained in 446 different days. After the cloud masking, 374 images with valid data were analyzed (based on the access on 17/06/2021).

Figure 5-4 shows the results for the uncalibrated procedure (configuration 4). The pixels selected for $C$ (Figure 5-4a) coincided also in this case with roads and parking. The river bed was again successfully identified by the product of the average reflectance and its standard deviation (Figure 5-4d). The holes within the river course were caused by river zones that are not detected from the JRC Global Surface water map. In Figure 5-4b the pixels selected for $W$ are shown. It can be noticed that they do not fall within the center of the river but prevalently on its southern borders. A possible cause of this phenomena is the shadow created by the riparian vegetation, which decreases the average reflectance of the pixels affected by it and facilitates their selection. Figures 5-4e and 5-4f show the S between the reflectance timeseries of each pixel in the water mask and the $C$ and $W$ timeseries, respectively. Figure 5-4e is particularly interesting because it shows many pixels lowly correlated with $C$, over the northern river boundaries, which are then selected to obtain $M$ timeseries by the uncalibrated configuration, as shown in Figure 5-4c. In contrast with the results obtained at Pontelagoscuro, by comparing Figure 5-4e with Figure 5-4f, it can be noticed that many pixels show also a low correlation with both $W$, while for the Po River the pixels lowly correlated with $C$ were better correlated with $W$ and vice versa. By further analyzing the selected area, a failure of the image co-registration procedure was detected: the coordinates of each image often not correspond to the real coordinates of the area, but there was often a misregistration of a few pixels.

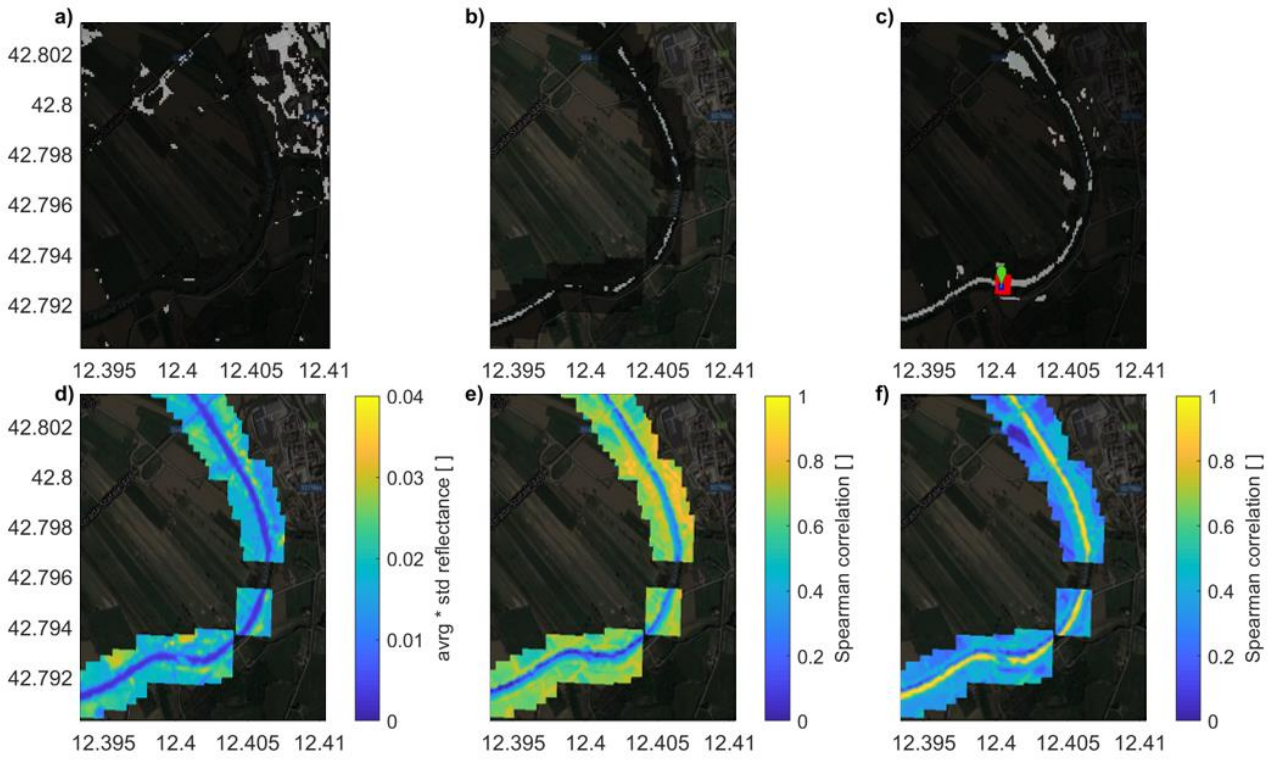


*Figure 5-4: Tiber River at Montemolino: in light gray the pixels used for calculating C (panel a), W (panel b). Panel c shows the location of the pixel M obtained from configuration 1 (1 pixel, greed dot), configuration 2 (3x3 pixels, blue square), configuration 3 (9x9 pixels, red square) and configuration 4 (multiple uncalibrated pixels, light gray). Panel d) shows the product between the average and standard deviation of the NIR reflectance for each pixel within the water mask; panel e) and f) represent the Spearman's correlation map between the reflectance timeseries of each pixel in the water mask and the average reflectance calculated for the C (panel e) and W timeseries (panel f), respectively. Background Copyright ©2021 Immagini ©2021, CNES / Airbus, European Space Imaging, Maxar Technologies.*

An example of this issue is shown in Figure 5-5, where the S2 1c band 8 image acquired on 22/08/2017 is shown (panel a) along with a clipping of the image acquired on 27/08/2017 (panel b): a miss-registration of ~1-2 pixels can be noticed by comparing the two images. This apparent movement of the full river section caused many pixels vary "artificially" between soil and water that were in turn selected for *M*, even being insensitive to river discharge. The error was also found in Pontelagoscuro images, but it was less impacting for river discharge estimation due to the different extent of the river, even if it could be responsible of the missing information for the greater river discharge peaks, since as above mentioned those affect very few pixels. To correct the issue, several attempts were done using GEE *"register"* function without satisfying results, mainly because of the presence of geometric distortions between the analyzed S2 images. The correction of this effect

should be therefore entrusted to the data provider, in order to increase the multi-temporal consistency of S2 dataset.

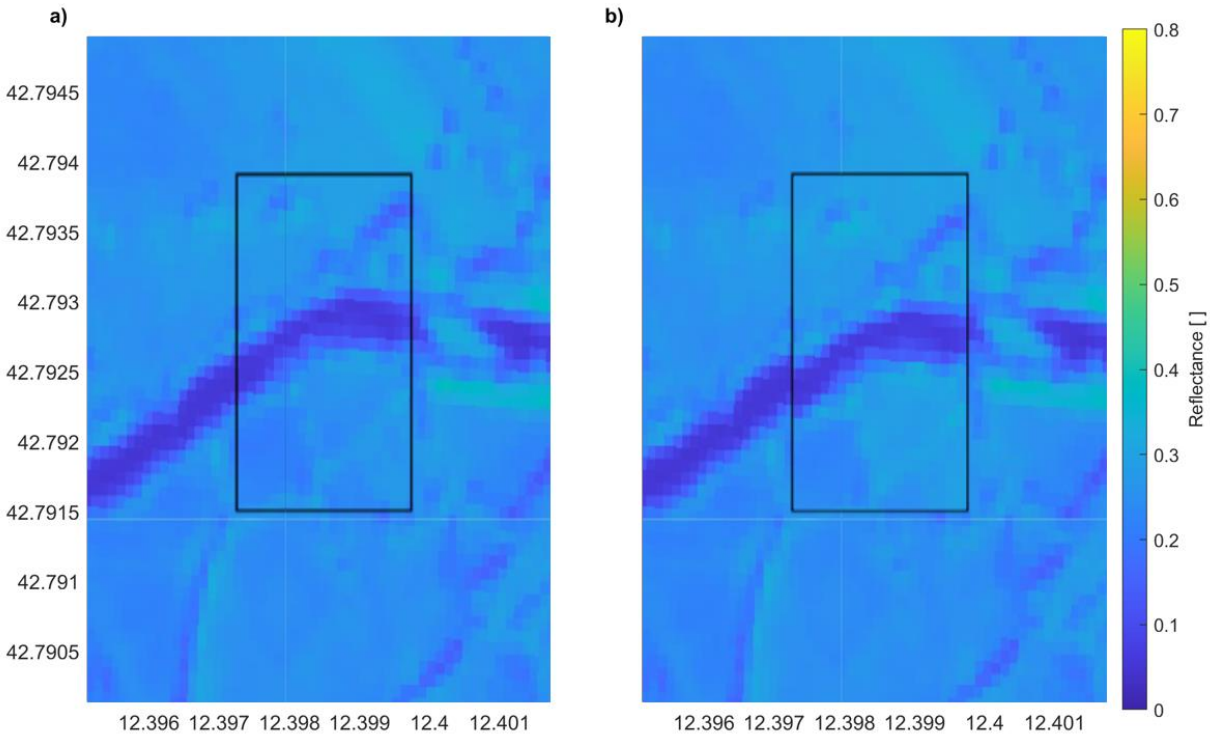


*Figure 5-5: Tiber River at Montemolino: detail of S2 1c Band 8 image acquired on 22/08/2017 (Panel a and b) and on 27/08/2017 (central patch Panel b), showing a miss-registration between the two images of about 1-2 pixels.*

The $M$ pixels chosen over the fields, in the north of the Figure 5-4c, highlight two additional limitations of the automatic procedure. First, the vegetated pixels, which are lowly correlated with both $C$ and $W$, are not taken into account by the current approach, and could be potentially wrongly selected for $M$ with configuration 4, as it happens in the example. Second, the use of the percentile thresholds to obtain $M$ pixels can lead to select too many of them, especially for small river, where the sensitive area could be narrow.

The $M$ pixels selected by the calibrated procedures (configurations 1-3: green dot, blue area and red area, respectively, in Figure 5-4c) fall in the same area inside the river that is periodically submerged (it is shown in light blue color in Figure 5-4d: in some periods it dries, therefore the average reflectance and the standard deviation are higher with respect to the rest of the river area) and it is therefore sensitive to river discharge variation.

Once again, the obtained $C$ and $W$ timeseries were used to estimate the river discharge with the $CMW$ and the $CM$ approaches, using both the uncalibrated and calibrated $M$ timeseries. The performances of the obtained approaches for the river discharge estimation are shown in Table 5-2.

*Table 5-2: Performance scores for the different configurations (1 for single pixel, 2 for 3x3 pixels, 3 for 9x9 pixels and 4 for multiple pixels) for the period 2015-2020 at Montemolino station. Spearman's (S) and Pearson's (R) correlations are calculated between the CMW or CM timeseries and the ground observed river discharge. Root Mean Square Error (RMSE), relative Root Mean Square Error (rRMSE) and Nash-Sutcliffe efficiency, NS, refer to the simulated discharge calculated according Eq. 7 and the ground observed discharges*

| Configuration | *CM* | | | | | *CMW* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ***S*** **[-]** | ***R*** **[-]** | ***RMSE*** **[m³/s]** | ***rRMSE*** **[-]** | ***NS*** **[-]** | ***S*** **[-]** | ***R*** **[-]** | ***RMSE*** **[m³/s]** | ***rRMSE*** **[-]** | ***NS*** **[-]** |
| **1** | 0.88 | 0.46 | 26 | 1.57 | 0.18 | 0.91 | 0.63 | 24 | 1.44 | 0.30 |
| **2** | 0.86 | 0.42 | 27 | 1.61 | 0.13 | 0.91 | 0.62 | 23 | 1.39 | 0.36 |
| **3** | 0.77 | 0.40 | 26 | 1.55 | 0.21 | 0.84 | 0.46 | 26 | 1.52 | 0.23 |
| **4** | 0.42 | 0.26 | 30 | 1.72 | 0.05 | 0.55 | 0.33 | 29 | 1.66 | 0.12 |

The misregistration issue affects particularly the uncalibrated procedure, which shows the lowest performance in terms of S, R, RMSE and NS. The calibrated methods are also affected, with score values worse than those obtained at Pontelagoscuro. Nevertheless, the relevance of considering sediment transport is again confirmed by the stable improvement in performances comparing the $CMW$ and the $CM$ results. Among the calibrated results for , the first two configurations, where the average spatial window was not applied or applied for a limited area (3x3), show the best performance,. This is probably due to both the moderate dimensions of the river, which reduce the number of pixels sensitive to river discharge variation, and the image misregistration issue.

Figure 5-6 shows the timeseries of the ground observed discharge and those derived by $CMW$ and $CM$ approaches at Montemolino station for all the configurations. The quality of the uncalibrated procedure (Figure 5-6g-h) is relatively poor: even if flood periods were correctly detected, the surplus of $M$ pixels deteriorates the information. The calibrated procedure shows instead better results, even

if all the river discharge peaks were underestimated, probably due to saturation issues. Low flow conditions were better represented here, but the chosen fitting function is not well suited for this location. A possible cause could be the different river section shape, which can greatly shift the relationship between river discharge and $CMW$ timeseries.

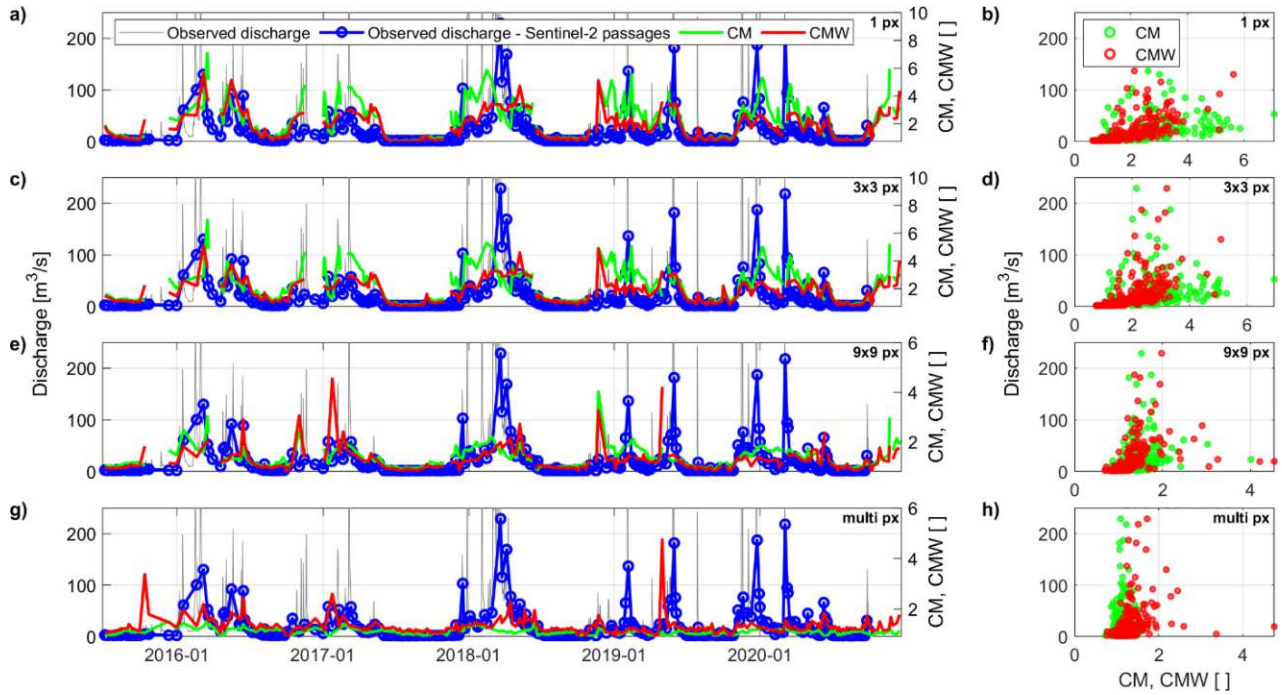


*Figure 5-6: Tiber River at the Montemolino station: comparison in terms of temporal series (a, c, e, g) and scatter plots (b, d, f, h) between the ground observed discharge and the $CM$ (green) or $CMW$ (red) timeseries for the configuration at single pixel (a and b), at 3x3 averaged pixel (c and d), at 9x9 averaged pixel (e and f) and for multiple uncalibrated pixels (g and h).*

# 6. CONCLUSIONS

*This work of thesis investigated the use of high spatial resolution data from the Sentinel Missions of the European Union Copernicus Earth Observation program for the estimation of water cycle variables, in order to assess whether high quality data can improve our ability to observe this important natural system from space. Specifically, the estimation of rainfall and river discharge from Sentinel-1 and Sentinel-2, respectively, was considered, due to their great impact in influencing human activities. As expected, these data can improve the accuracy of estimates in many areas where coarser resolution data are not accurate enough to provide useful information. In particular, the higher resolution products give the opportunity to observe features that would otherwise be averaged with surrounding pixels. Notwithstanding this, the low temporal resolution and some issues in georeferencing the data are still major problems that cause many events of both rainfall and river discharge to be missed.*

## 6.1 Final remarks

The number of natural disasters is expected to increase in the next decades due to the ongoing climate change, worsening the living conditions of the affected areas inhabitants. High-resolution data are fundamental to assist the authorities to prevent or mitigate these events, therefore it is of primary importance to increase their accuracy, availability and use, together with our understanding of the physic behind the phenomena we observe. For this reason, European Union developed Copernicus, its own Earth observation program. It aims at achieving a global, continuous, autonomous, high quality, wide range Earth observation capacity. This can be obtained by exploiting Sentinel Missions, a family of innovative satellites characterized by the most advanced technologies that ESA is developing and launching to fulfil Copernicus objectives. These sensors are capable of providing accurate, timely and open access information to any interested user, with the final objective, among other things, to improve the management of the environment, to understand and mitigate the effects of climate change, and to ensure civil security. Sentinel missions produce everyday more than 15 Terabytes of data, available to users and scientist worldwide. This enormous amount of information makes necessary dedicated research efforts to develop new approaches for observing Earth system from space and/or adapt the existing methodologies to the new high-resolution data.

With this scope, the information obtained from S1 and S2 missions from the Copernicus program have been used in this thesis to develop two methodologies aimed at the estimation of rainfall and river discharge, two variables that critically contribute to the characterization of Earth's climate. These variables were chosen among the others that contribute to the water cycle due to their clear impact on Earth, which allows to observe them from space, and their important impact on human lives. Existing algorithms for the estimation of rainfall and river discharge were developed and adapted to be applied to Copernicus data. Specifically, in order to demonstrate the usefulness of Sentinel data, this thesis has a twofold objective:

1) adapt and developing SM2RAIN algorithm for the rainfall estimation to high-resolution data from S1, by also considering the absence of high-resolution calibration data;
2) adapt and developing CM approach for the river discharge assessment to S2 data, exploiting the possibility to investigate new detailed feature such as the sediment transportation.

Improvement in the estimation of the variables were obtained with both the developed methodologies, thanks to the additional information contained in the high-resolution data from Copernicus missions. The main results of the analysis are summed up in the following paragraphs.

### 6.2 Parameterized SM2RAIN validation

The first objective of this thesis was to develop four parametric relationships capable of estimating SM2RAIN parameters values from climatic and land descriptors, in order to allow SM2RAIN application at high spatial resolution (1 km) and in areas where there is no availability of a benchmark rainfall product to be used for calibration. A local analysis was performed over 1009 points characterized by a spatial resolution of 25 km and uniformly distributed over areas covered by rain gauges in Australia, India, Italy and US, for which high quality observed rainfall data were available. Several climatic and land descriptor datasets were analyzed to obtain an inclusive description of each point and to find the descriptors better related to the four SM2RAIN parameters. The four parametric relationships were developed, obtaining the parameters estimation from the knowledge of the SM timeseries and their noise, the topographic complexity and the mean annual rainfall. The major drops in correlation due to the use of the parametric relationships are caused by the estimation of the $T$ parameter, responsible of the SM data filtering, while the major increase in RMSE is caused by the $a$ parameter, related to the drainage rate estimation.

To validate the obtained results, a global application of SM2RAIN on ASCAT SM was performed using the parametric relationships (called parameterized SM2RAIN-ASCAT) and the standard calibration methodology (called calibrated SM2RAIN-ASCAT). ERA5 rainfall was used to assess mean annual rainfall and as calibration dataset, while GPM-FR rainfall was used as benchmark to

calculate performance indices. From the analysis of the categorical and continuous scores, an overall similar capacity in rainfall estimation between the calibrated and parameterized product is found. In particular, even if the calibrated product has slightly better performances both in terms of correlation and RMSE, the parameterized product resulted more capable in the detection of larger rainfall events.

Finally, the triple collocation method was performed by using GPM-LR, GPCC and the two SM2RAIN-ASCAT products. The analysis revealed that, even if the parameterized version of SM2RAIN-ASCAT has a slightly lower average correlations than the others, there are several regions (e.g., in Africa and South America) in which its performance is better than both GPM-LR and GPCC, suggesting the utility of this product for rainfall estimation.

It is worth noticing that the parametric relationships have been linked to the mean annual rainfall and average SM variation, even though they are physically based on soil characteristics. This was not found probably due to the coarse spatial resolution of the SM product selected for this analysis and the low accuracy of the soil composition dataset. This point may cause a lack of robustness in the parameters estimation, since they are related to statistical quantities that are subjected to fluctuation due to, e.g., climatic variation (climatic change for rainfall) or anthropic effects (like the effect of irrigation on the average SM variation). Nevertheless, the results are still satisfactory and represent a step forward with respect to the simple calibration of SM2RAIN, particularly with respect to the high-resolution rainfall application. Indeed, the quality of SM2RAIN products is limited in specific regions, such as desert, densely vegetated and mountainous areas, but their capabilities in estimating accumulated rainfall (rather than an instantaneous rainfall rate measured by the majority of the satellite-based products), make them a precious source of information to study rainfall process and variability worldwide, also in conjunction with other precipitation products derived from EO as noted by Massari et al. (2020).

### 6.3 High-resolution rainfall for Po River basin

Once the parametric relationships were obtained, the high-resolution rainfall estimation analysis was conducted. Specifically, the SM2RAIN algorithm was applied to a 1 km spatial resolution SM product from S1, obtained through an algorithm based on a first order solution of the Radiative Transfer equation, RT1, over the Italian fraction of the Po River Basin, to retrieve a 1 km high-resolution rainfall product from satellite remote sensing. This region was selected due to the availability of a benchmark dataset from radar and rain gauge data and the presence of many geographical features, such as mountains, hills, lakes, rivers and plains, which make it a good test area for this analysis. Two versions of SM2RAIN were applied in this analysis to compare the resulting performances: one uncalibrated, to foster the high-resolution rainfall estimation in other regions where benchmarks data are unavailable, and one calibrated against the observed data. In order to assess the improvements related to the high spatial resolution of S1, SM2RAIN was also applied to ASCAT SM (25 km spatial resolution), resampled to S1-RT1 grid for comparison. The analysis was carried out at different temporal accumulation steps (1 day, 10 days and 30 days) to take the different temporal resolution of the two SM products, 1.5 to 4 days for S1-RT1 and sub-daily for ASCAT, into account.

The results show that it is indeed possible to obtain high-resolution rainfall data from S1, even if the low temporal resolution of the data does not allow to calculate daily rainfall. It is instead possible to calculate it with ASCAT data due to the higher temporal resolution, with good results (median R, of 0.61 and 0.64 for the parameterized and calibrated SM2RAIN). When 10 days accumulated rainfall is considered, S1-RT1 derived rainfall from the parameterized (calibrated) SM2RAIN performs quite well, with a median R of 0.61 (0.65), but ASCAT performances are higher, with a median R of 0.71 (0.75). At higher accumulation steps, the performance differences reduce, until ASCAT and S1-RT1 derived rainfall reach almost equal R for the 30 days accumulated rainfall (around 0.75). Similar conclusion can be deduced by analyzing RMSE values, while for the BIAS index the differences

between the calibrated and the parameterized SM2RAIN results are larger, probably due to the low spatial resolution of the product used to obtain the Po River Basin climatology (ERA5-Land).

Even if on average the rainfall from ASCAT seems to be slightly better performing than the one from S1, the analysis of the spatial distribution of R shows instead the true benefits of the high-resolution SM. In the complex mountain regions, S1 obtains extremely good performance over the valleys and bad performance over the ridges, unsuited for SM remote sensing, whereas ASCAT's R always represents an average of the two features signals due to the lower spatial resolution. S1 derived rainfall is generally better performing than the one from ASCAT also in the northern section of the Po Valley plain, while the latter is better in the southern section, where irrigation is widely practiced. The fragmentary nature of the irrigation in this area could be the cause of this phenomena: S1-RT1 should be more sensitive than ASCAT to the signal generated by various small fields, where irrigation in not concurrent, thanks to its higher spatial resolution, but since irrigation is not considered in the benchmark product, the resulting R is reduced.

Summing up, high-resolution rainfall from satellite remote sensing is possible and it is able to observe features that are averaged in products with lower spatial resolution, like the rainfall within mountain valleys and potentially the fields' irrigation. Notwithstanding this, the low temporal resolution is currently a limitation for its application in many fields, even if high spatial resolution rainfall at monthly temporal resolution is still useful for agriculture, water resource management and index-based insurances. Nevertheless, the estimation of SM at 1 km spatial resolution is still a brand-new field and several improvements are expected in the next years, which will allow to further improve also the performance of the derived high-resolution rainfall. Moreover, such high-resolution rainfall products are rarely available to the users, and most of them are just based on spatial interpolation or other downscaling techniques, with many limitations, as mentioned in chapter 1. SM2RAIN allows instead to derive rainfall directly from SM variations, exploiting a physical relationship. This allows to obtain reliable rainfall estimates, particularly over the areas where the SM quality is high.

### 6.4 River discharge estimation over Po and Tiber rivers

The second target of this thesis was to take advantage of the high-resolution of S2 data and their availability on Google Earth Engine platform to increase the capabilities and the accuracy of the $CM$ approach in estimating river discharge from NIR data over medium/narrow rivers. The method was enriched with the additional concept of exploiting a water ($W$) area to take into account the variation of sediment load in the river, responsible of a water reflectance increase that caused river discharge underestimations. The averaging of multiple pixels to obtain $M$ timeseries was also proposed, to increase the methodology sensitiveness to river discharge by avoiding as much as possible the condition of completely flooded or dry $M$. Four different configurations were applied for the selection of $M$ pixels to test this concept: three calibrated (1 pixel, 3x3 pixels and 9x9 pixels) against the observed river discharge, and one uncalibrated, based on selecting for $M$ the pixels with low Spearman's correlation with $W$ or $C$.

To test its efficacy with different water regime, the methodology was applied over two study areas in Italy: at Pontelagoscuro along the Po River, and at Montemolino along the Tiber River, for the period 2015-2020, exploiting Google Earth Engine cloud platform. Results indicated a stable performance improvement of the $CMW$ approach with respect to the $CM$ approach for almost all the areas and configurations, proving the relevance of taking into consideration the sediments to correctly estimate the river discharge. The uncalibrated configuration has proved the most effective at Pontelagoscuro station with respect to the calibrated ones. Over the Tiber River, instead, the uncalibrated configuration failed to detect the pixels sensitive to the variation of river discharge due to the presence of heavily vegetated pixels and the misregistration issues of the S2 images that introduces uncertainties in the identification of $C$, $M$ and $W$ pixels. The issue was present also at Pontelagoscuro station, but with a less severe impact due to the greater river width. Finally, two additional sources of error were identified: the presence of cloud shadows and residual clouds, not masked by the cloud

masking procedure, and the variability of the value of $C$, due e.g., to the effect of soil moisture variations or bidirectional reflectance effects.

This work represents a new step toward a non-calibrated methodology capable of obtaining a river discharge proxy over ungauged river, an extremely important information, especially due to the fostered water scarcity related to the climate change. The sediment contribution proved to be necessary to obtain reliable river discharge estimation. Despite this, the methodology is still unsuitable for ungauged rivers: indeed, the proposed approach is able to obtain a good proxy of water flow without using any gauge data, but river discharge observations are still needed for calibration purposes. Moreover, the methodology needs to be tested in different environments to evaluate its skills. Finally, well registered and more frequent NIR high-resolution data are also fundamental to increase the accuracy of the retrievals and to avoid the missing of river discharge peaks. Indeed, the CMW approach still has different source of noises that can potentially impact the estimated index value. The use of optical sensors, anyhow, ensures higher temporal resolution with respect to altimetry data and it is less sensitive to cloud coverage than the approaches based on the river width estimation. Moreover, the river width evaluation is limited by the use of a binary state water/not-water to retrieve the length of the flooded river section, which may reduce its precision. Indeed, the proposed methodology showed promising results, with potential to further improve the river discharge estimation also considering future integrations of the existing techniques.

### 6.5 Further steps

This thesis demonstrated how the high-resolution information from the Sentinel Missions of the Copernicus program can increase our capacity in observing the water cycle from space, and specifically to increase the quality of rainfall and river discharge estimations, two of the most important variables that contribute to the water cycle. The described analysis highlighted how the high-resolution of Sentinel data can be used to study features and objectives that are not observable though coarser resolution data, leading the way to new researches to further improve remote sensing

observation capabilities. Although the performance of the rainfall and river discharge estimations obtained in this thesis are satisfactory, a detailed analysis of errors and noises related to the high-resolution products (e.g., the co-registration errors for S2 imagery or the SM calibration over mountain regions for S1-RT1) is required, to increase the methods accuracy and applicability, and to provide useful satellite data for many applications.

Further research will be carried out by using the results obtained in this work of thesis as starting point. Specifically, regarding SM2RAIN application for rainfall estimation, the following research topics requires detailed analysis:

1. Improve the automatic estimation of SM2RAIN parameters values by including additional descriptors (e.g., Radio Frequency Interference indicators, Land Cover, high-resolution topography), with the objective of reducing the BIAS error resulting from the application of the parameterized version of SM2RAIN;
2. Test machine learning approaches for self-calibrated SM2RAIN in order to assess their capacity in estimating the desired SM2RAIN parameters;
3. Evaluate the integration of multiple SM products (e.g., from ASCAT; S1; Global Navigation Satellite Systems, GNSS; Soil Moisture Active Passive, SMAP) and/or top-down products (e.g., GPM IMERG) to obtain high spatial and temporal resolution rainfall estimates, i.e., 1 km and 1 hour, in order to increase the fields of applicability of the obtained rainfall datasets;
4. Use S1-RT1 information for detection and quantification of irrigation, in order to provide to the policy makers information regarding this important anthropic alteration of the water cycle;
5. Use the high spatial resolution products obtained in this thesis for hydrological applications such as flood forecasting, landslide prediction and water resource management, to assess whether the high-resolution input data effectively improves the models' forecasting and estimation capabilities;
6. Analyze the error propagation and uncertainties assessment of SM2RAIN-derived rainfall;

7. Improve SM2RAIN formulation by including additional processes (e.g., evaporation, runoff and interception) and by improving the description of drainage.

Similarly, the research related to the river discharge estimation should be focused on the following topics:

1. Test the developed CMW approach to different satellite products and study areas over multiple climates, in order to evaluate the replicability of the obtained results and the limitations of the selected approach;
2. Investigate the role of vegetation, to understand its influence in the definition of the C pixels and correct its effect on the selected M reflectances;
3. Improve the cloud masking algorithm, to reduce the number of unmasked clouds and to develop a procedure to mask clouds' shadow;
4. Analyze river clusters to find the most suitable relationship between the CMW index and the observed river discharge;
5. Further develop the uncalibrated procedure, to foster its application on multiple environments where the ground reference is missing and the river discharge estimation through the traditional approaches is impossible due to the absence of calibration data.

# 7. APPENDIX

*In this section, additional material is reported. The global maps of Triple Collocation correlation for the parametrized SM2RAIN applied to ASCAT, GPM-LR and GPCC are firstly shown. Then, a more detailed analysis of the areas characterized by stable low accuracy of rainfall estimation through SM2RAIN over the Po River basin is reported. Lastly, the results of testing a different time accumulation concept for the high-resolution rainfall products obtained over the Po River basin are shown.*

## 7.1 Triple collocation results

This section shows the global maps of TC correlation for the parametrized SM2RAIN applied to ASCAT, GPM-LR and GPCC. It should be noted that TC correlation values should not be taken as absolute measures of accuracy but rather as relative measures between the three datasets. The section is based on the following publication, published by Journal of Hydrology:

*Filippucci, P., Brocca, L., Massari, C., Saltalippi, C., Wagner, W., and Tarpanelli, A.: Toward a self-calibrated and independent SM2RAIN rainfall product, Journal of Hydrology, 603(A), 126837, doi: 10.1016/j.jhydrol.2021.126837, 2021.*

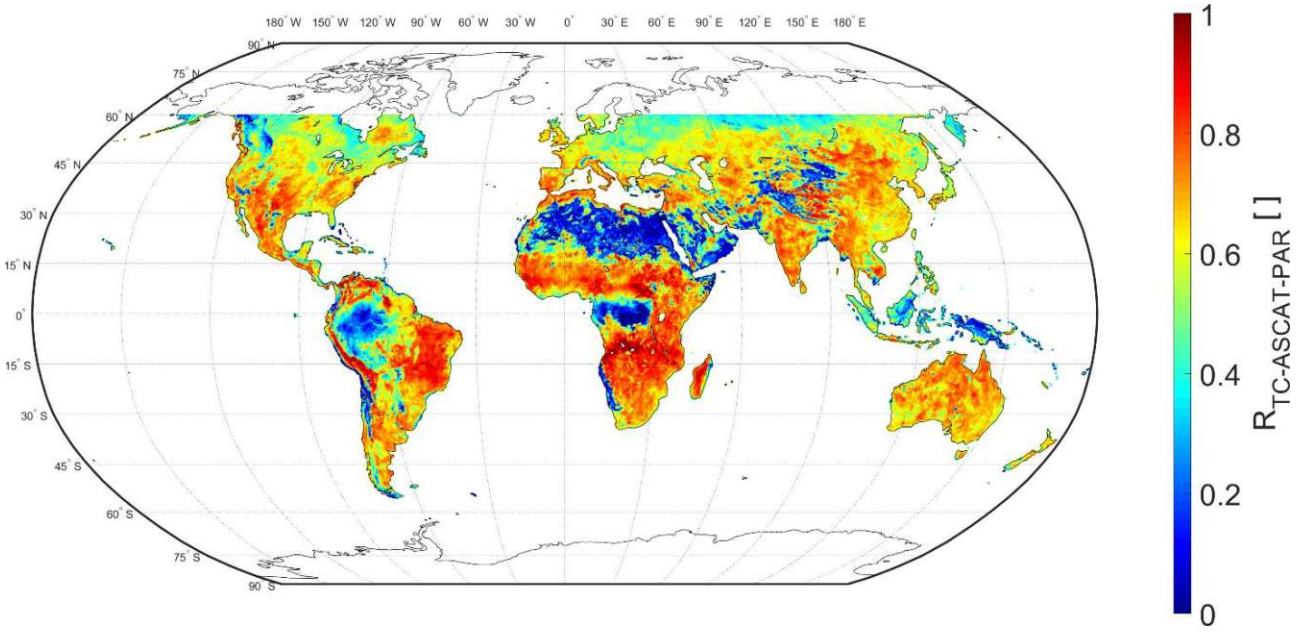


*Figure 7-1: Map of TC correlation of the parameterized SM2RAIN-ASCAT based on the results of Triple Collocation of the rainfall products triplet SM2RAIN parameterized, GPM – Late Run and GPCC.*

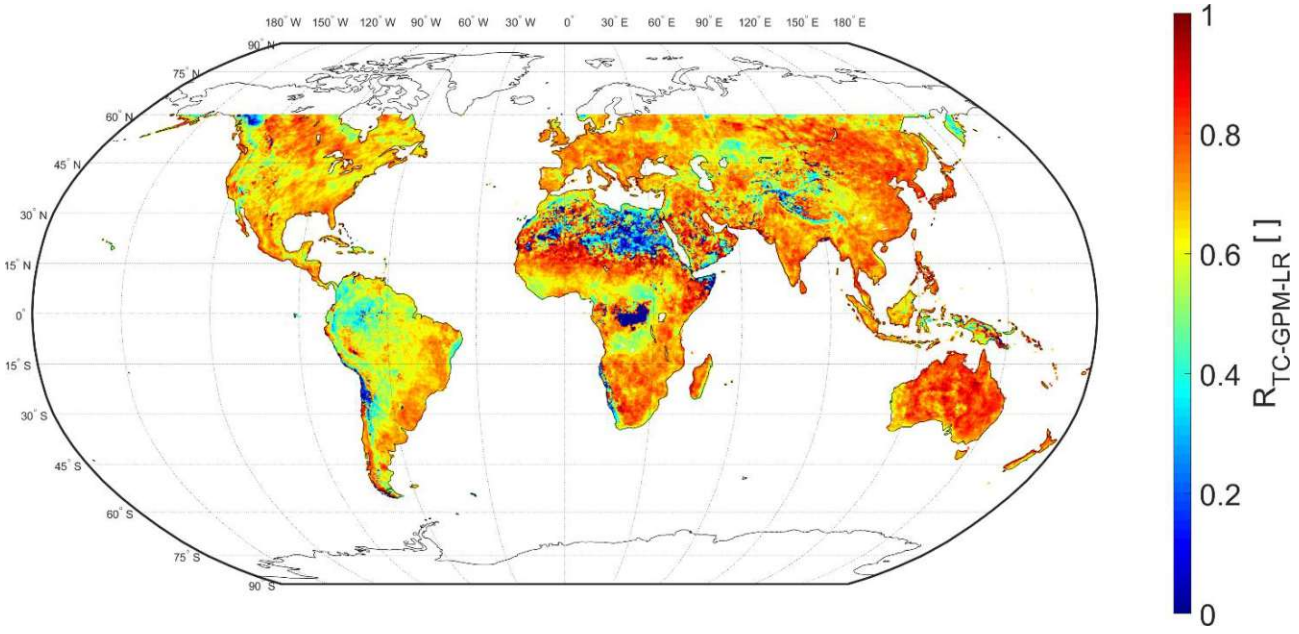


*Figure 7-2: Map of TC correlation of the GPM – Late Run based on the results of Triple Collocation of the rainfall products triplet SM2RAIN parameterized, GPM – Late Run and GPCC.*

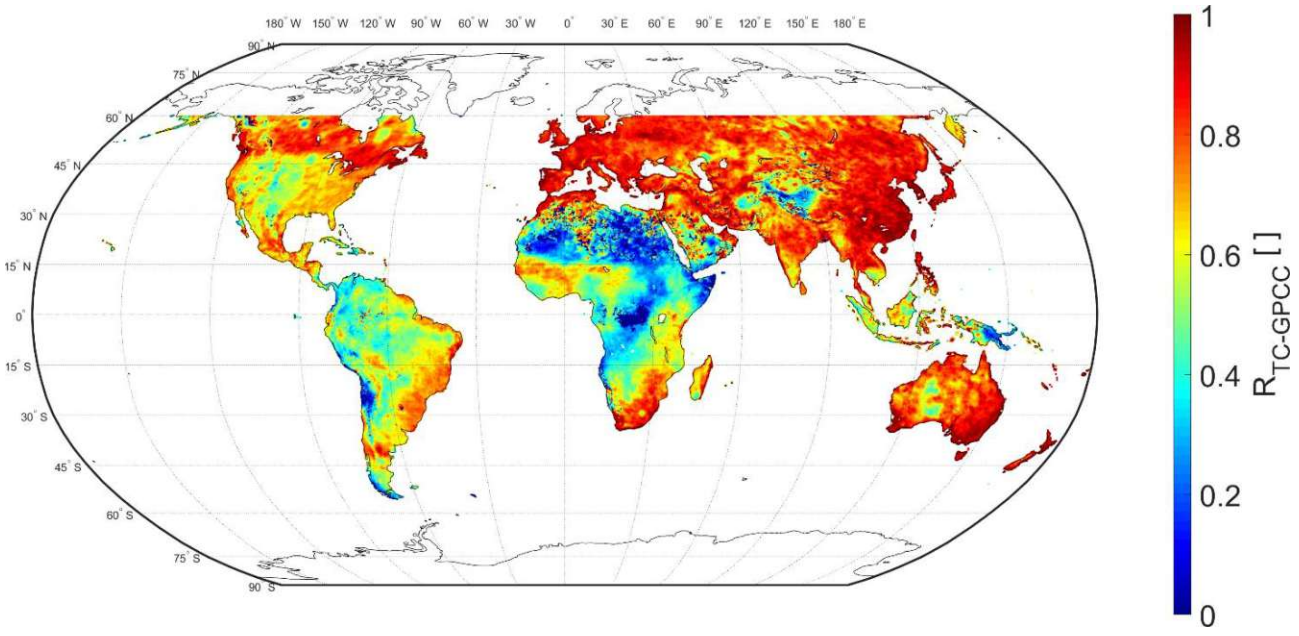


*Figure 7-3: Map of TC correlation of the GPCC based on the results of Triple Collocation of the rainfall products triplet SM2RAIN parameterized, GPM – Late Run and GPCC*

## 7.2 Parameterized SM2RAIN over Po River basin: issues with rainfall estimation

---

This section will present an analysis of the areas characterized by stable low accuracy of rainfall estimation through SM2RAIN over the Po River basin. First the probable causes of the issue are identified by looking at the performance patterns, then three potential "bad performing" gauges are analyzed. The section is based on a preprint of the following publication, submitted to Hydrology and Earth System Sciences:

*Filippucci, P., Brocca, L., Quast, R., Ciabatta, L., Saltalippi, C., Wagner, W., and Tarpanelli, A.: High-resolution (1 km) satellite rainfall estimation from SM2RAIN applied to Sentinel-1: Po River Basin as case study, submitted to Hydrology and Earth System Sciences, doi: 10.5194/hess-2021-563, 2021.*

---

The obtained results show that the high-resolution information from S1 sensors allows to increase the accuracy of SM (and thus of rainfall) in areas where coarse resolution data are not able to obtain reliable estimates. Conversely, over some region the rainfall obtained from the application of SM2RAIN to S1-RT1 SM shows worse performance with respect to the one obtained when the algorithm is applied to ASCAT data, as it happens on many mountainous areas. Finally, there exist some areas in which the accuracy of the rainfall obtained from the application of both the calibrated and parameterized SM2RAIN to ASCAT or S1-RT1 SM products is stably low, as discussed in section 4.4. This issue can depend by multiple factors, as SM signal quality, failure of the models (for rainfall or SM estimation) or accuracy of the benchmark rainfall product. In this section, an attempt to identify those area is made, by highlighting the pixels in which the Pearson's correlation between the 30 days accumulated rainfall from MCM and the four SM2RAIN (from ASCAT and from S1, calibrated or parameterized) derived products is always less than a threshold, fixed at 0.65, as shown in Figure 7-4. Multiple areas of stable low performances can be distinguished in the figure, highlighted in blue. Two main reasons of this behavior can be identified: issues with the SM sensing and issues with the benchmark product. In particular, the blue areas located in mountainous region in the North and the West of Figure 7-4 map, should be affected by both the source of error, since on topographically complex areas SM retrieval is difficult and weather radar accuracy drops. Notwithstanding this, ASCAT performance are still higher than those of S1-RT1 in these areas (compare with Fig. 4-4). This fact has a threefold explanation: first, S1-RT1 SM estimations are obtained without considering any snow masking, thus their accuracy over mountain region regularly affected by snow cover is limited; second the low quality of ASCAT SM retrieval over topographically complex area is mitigated by the presence in each ASCAT pixel of valleys and/or plateau in which SM accuracy is higher; third, SM2RAIN algorithm hypothesis could be not valid over these areas since the runoff rate should be not negligible.

Instead, the areas within the light blue rectangles in Figure 7-4, are characterized by the presence of paddies and water bodies: here the low performance should be caused by low SM quality, due to the

impossibility of retrieve SM information over flooded areas with active microwave sensors. Finally, the remaining blue regions should be affected by low quality of the benchmark product. This can be related either to "bad" performing gauge stations, recognizable through the central position of a gauge with respect to the low performing area (e.g., the two regions in the Center-East black rectangles), or to issues with weather radar and rain gauges measurements, where the blue patterns are concentrated between two or more rain gauges (e.g., the region within the black rectangles on the South-West).

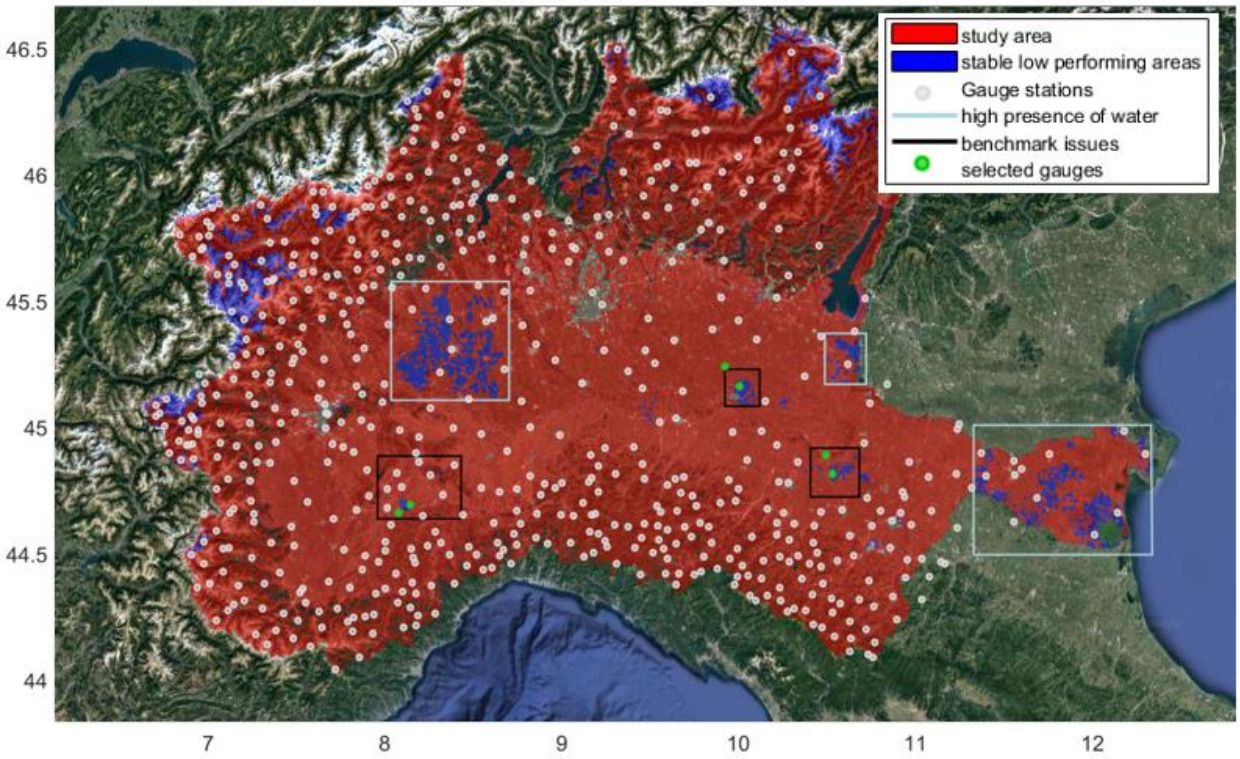


*Figure 7-4: Map of Po River basin. The blue pixels indicate the areas where Pearson's correlation between the 30 days accumulated rainfall from MCM and the calibrated and parameterized SM2RAIN applied to ASCAT or S1-RT1 is stably less than a threshold of 0.65. The light blue rectangles surround the areas with paddy areas or abundant water bodies, while black rectangles outline areas with alleged "bad" performing gauge station. Finally, the white dots show the gauge stations location and the green dots the rain gauge selected to be further analyzed. Map copyright ©2021 GeoBasis-De/BKG (©2009), Google, Inst. Geogr. Nacional Immagini ©2021 TerraMetrics.*

In order to better analyze this aspect, three stations located in within the three black rectangles in Figure 7-4 were selected, together with the nearest neighbor stations. The MCM timeseries of the pixels that includes the stations were extracted, in order to compare them and assess the quality of the considered rain gauges. The qualitative comparison of the stations is displayed in Figure 7-5, where the scatter plots for each pair of rain gauges is shown together with their position in the map (Figure

7-5a). In particular, a clear issue with the rain gauge named A1 can be appreciated in Figure 7-5b, with this sensor measuring rainfall peaks up to 300 mm/day, absent from the nearest gauges. The issue can be confirmed by the low Pearson's correlation between its timeseries and the one of the nearest rain gauge, equal to 0.53, which is significantly lower than the mean Pearson's correlation calculated between each couple of nearest stations within the study area, equal to 0.87 (standard deviation equal to 0.1). Also Figure 7-5c shows strange patterns of rainfall: even if there are no large peaks, several rainfall events are sensed with different magnitude by the two stations named B1 and B2, as can be noticed by looking at the number of points that tends to the x and y axis of the figure which indicate severe over- or underestimation. Also in this case, the measured Pearson's correlation is lower than the average, equal to 0.71. Finally, the station C1 measures several peaks of rainfall that are higher than those sensed by the nearest rain gauge, C2 (Figure 7-5d). Notwithstanding this, in this case the variation between the two sensors seems to be caused by the natural spatial variability of the rainfall, as demonstrated by the high Pearson's correlation between the two timeseries, equal to 0.88. This was expected since the low performing region is not located around one of the stations, but in between them, over a hilly area that could affect the weather radar measurements.

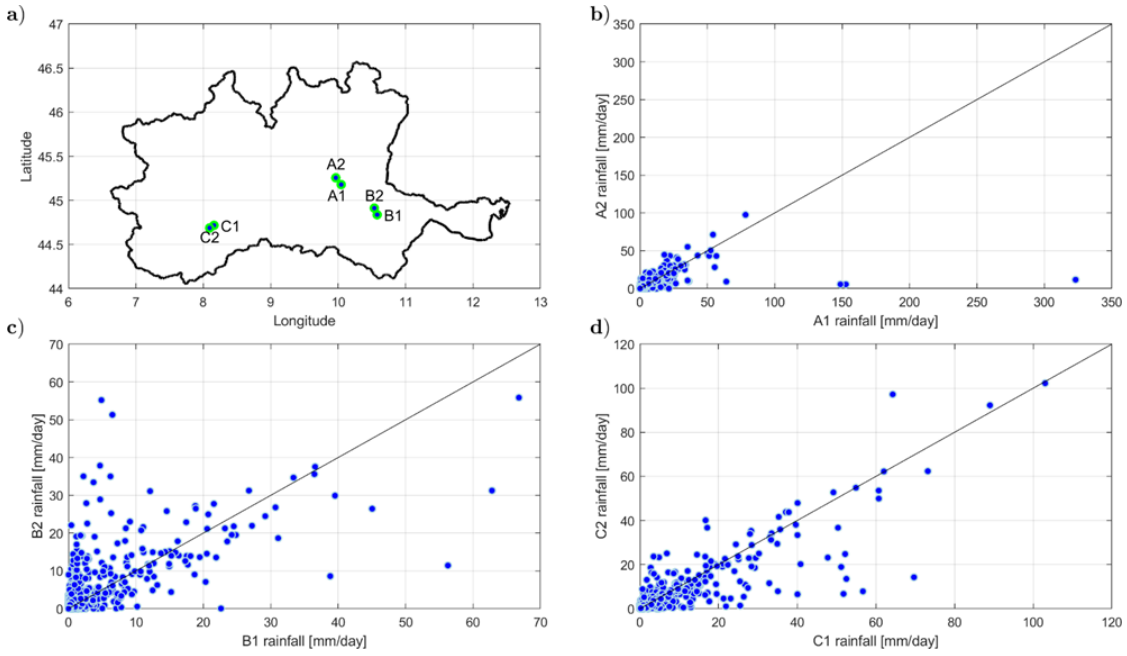


*Figure 7-5: Panel a shows the boundary of the Po River basin, together with the position of three couple of stations (A1-A2, B1-B2 and C1-C2) with alleged "bad" MCM performance. The scatter plots of the daily rainfall measured by each couple of stations is then shown in Panel b, c and d.*

Errors in the selected benchmark product are a known limitation of the direct validation of rainfall datasets. Notwithstanding this, this fact is also the proof of the need of further researches in the rainfall measurement fields, since the merging of different rainfall products, each with its limitation, often complementary, can be beneficial to obtain a measurement as close as possible to reality.

## 7.3 Parameterized SM2RAIN over Po River basin: time accumulation

---

A different time accumulation concept for the high-resolution rainfall products obtained over the Po River basin is tested in this section. The performances of the SM2RAIN derived rainfall products are recalculated to confirm the results obtained in chapter 4. The section is based on a preprint of the following publication, submitted to Hydrology and Earth System Sciences:

*Filippucci, P., Brocca, L., Quast, R., Ciabatta, L., Saltalippi, C., Wagner, W., and Tarpanelli, A.: High-resolution (1 km) satellite rainfall estimation from SM2RAIN applied to Sentinel-1: Po River Basin as case study, submitted to Hydrology and Earth System Sciences, doi: 10.5194/hess-2021-563, 2021.*

---

The performance indexes of the high-resolution rainfall products of chapter 4 were calculated at three different temporal steps: 1 day, 10 days and 30 days. In order to obtain them, the timeseries of each estimated product and of the observed one were accumulated according to the selected period by considering only the intervals in which the data was available in both the datasets (estimated and benchmark). This choice was made to obtain the best accurate assessment of each product, by calculating its potential in estimating rainfall against a concurrent dataset. Notwithstanding this, the comparison of ASCAT and S1-RT1 based on such performances could be biased, because in this way the analyzed indexes are calculated against two different benchmark datasets, each resulting by the selected product's valid data periods. In this section, we decided therefore to calculate again the performance indexes by accumulating the rainfall of the observed and estimated datasets only over the periods in which the three datasets (i.e., MCM, ASCAT and S1-RT1) are available together, and to insert in this appendix the corresponding of Figures 4-3 (Figure 7-6) and 4-4 (Figure 7-7) with the newly calculated indexes. To further increase the comparison quality and to avoid the period in which just one S1 sensor was in orbit and thus the associated drop in performance, only the data subsequent to 01/10/2016 were considered for the new indexes calculation.

In comparison with chapter 4 results, here ASCAT performances slightly increase (Figure 7-6), due evidently to the removal of some low performing data. This is confirmed by the appearance of some patterns within the ASCAT correlation map in Figure 7-7a that resemble the invalid pixel percentage distributions map of Figure 7-7d. Notwithstanding this, the areas in which S1-RT1 outperforms ASCAT are almost confirmed (Figure 7-7c), although shrinked in the plain, confirming the obtained results.

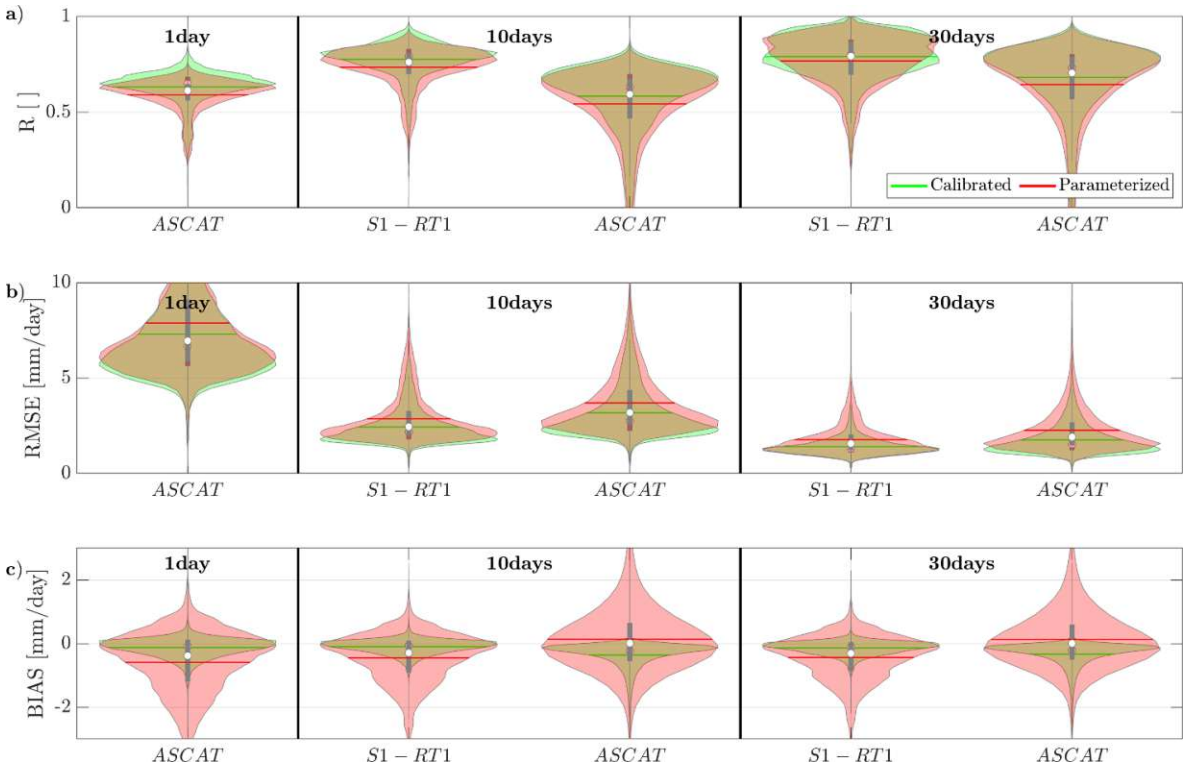


*Figure 7-6 Violin plots of Pearson's Correlation (R, panel a), Root Mean Square Error (RMSE, panel b) and BIAS (panel c) between the rainfall from MCM and from SM2RAIN applied to ASCAT and S1-RT1. ASCAT-derived rainfall was accumulated at 1, 10 and 30 days, while the rainfall from S1-RT1 was accumulated at 10 and 30 days. Only the periods in which all three products are available are considered in the accumulation. Each violin shape is obtained by rotating a smoothed kernel density estimator. The green violins are obtained by calibrating SM2RAIN against MCM, while the red violins derived from the parameterized SM2RAIN procedure.*

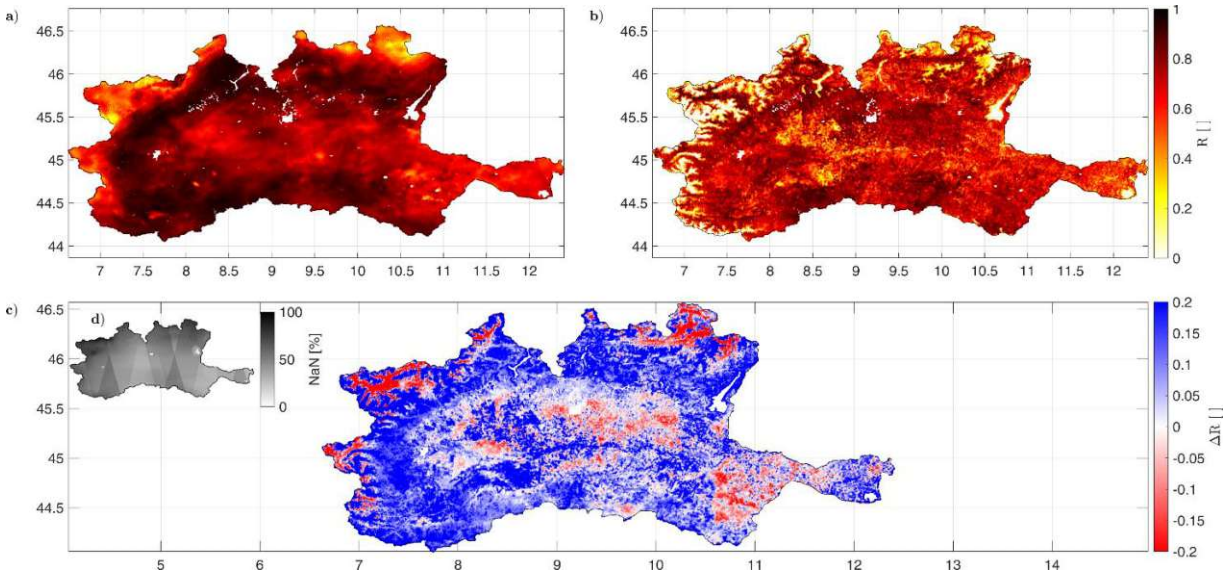


*Figure 7-7: Spatial Pearson's correlation (R) between the 30 days accumulated rainfall derived from MCM and the application of the parameterized SM2RAIN to ASCAT (panel a) and to S1-RT1 (panel b) SM products, considering only for the periods in which all three products are available. Panel c shows the difference between ASCAT and S1-RT1 correlation maps, while panel d shows the percentage of not valid images per pixel.*

# 8. REFERENCES

Abdalla, S., Abdeh Kolahchi, A., Adusumilli, S., Aich Bhowmick, S., Alou-Font, E., Amarouche, L., …: Altimetry for the future: Building on 25 years of progress, Advances in Space Research, 68(2). 319–363, doi: 10.1016/j.asr.2021.01.022, 2021.

Adler, R. F., and Negri, A. J.: A satellite infrared technique to estimate tropical convective and stratiform rainfall, Journal of Applied Meteorology, 27(1), 30–51, doi: 10.1175/1520-0450(1988)027<0030:ASITTE>2.0.CO;2, 1998.

Ahmed, R., Prowse, T., Dibike, Y., Bonsal, B., and O'Neil, H.: Recent Trends in Freshwater Influx to the Arctic Ocean from Four Major Arctic-Draining Rivers, Water, 12, 1189, doi: 10.3390/w12041189, 2020.

Ahn J. H. and Park, Y. J.: Estimating Water Reflectance at Near-Infrared Wavelengths for Turbid Water Atmospheric Correction: A Preliminary Study for GOCI-II, Remote Sensing, 12(22), 3791, doi: 10.3390/rs12223791, 2020.

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J. C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., and Martin, E.: From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations, Hydrology and Earth System Sciences, 12, 1323–1337, doi: 10.5194/hess-12-1323-2008, 2008.

Andreadis, K. M., Clark, E. A., Lettenmaier, D. P., and Alsdorf, D. E.: Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model, Geophysical Research Letters, 2007, 34(10), L10403, doi: 10.1029/2007GL029721, 2007.

Ba, M. B., and Gruber, A.: GOES multispectral rainfall algorithm, Journal of Applied Meteorology, 40 (8), 1500–1514, doi: 10.1175/1520-0450(2001)040<1500:GMRAG>2.0.CO;2, 2001.

Barrett, E. C. and Beaumont, M. J.: Satellite rainfall monitoring: An overview, Remote Sensing Reviews, 11(1–4), 23–48, doi: 10.1080/02757259409532257, 1994.

Bauer, P., Amayenc, P., Kummerow, C. D., and Smith, E. A.: Over-ocean rainfall retrieval from multisensor data of the Tropical Rainfall Measuring Mission. Part II: Algorithm implementation, Journal of Atmospheric and Oceanic Technology, 18(11), 1838–1855, doi: 10.1175/1520-0426(2001)018<1838:OORRFM>2.0.CO;2, 2001.

Bauer-Marschallinger, B., Paulik, C., Hochstöger, S., Mistelbaue, T., Modanesi, S., Ciabatta, L., Massari, C., Brocca, L., and Wagner, W.: Soil Moisture from Fusion of Scatterometer and SAR: Closing the Scale Gap with Temporal Filtering, Remote Sensing, 10(7), 1030, doi: 10.3390/rs10071030, 2018.

Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., Modanesi, S., Massari, C., Ciabatta, L., Brocca, L., and Wolfgang, W.: Towards Global Soil Moisture Monitoring with Sentinel-1: Harnessing Assets and Overcoming Obstacles, IEEE Transactions on Geoscience and Remote Sensing, 57(1), 520–539, doi: 10.1109/TGRS.2018.2858004, 2019.

Belloni, R., Camici, S., and Tarpanelli, A.: Towards the continuous monitoring of the extreme events through satellite radar altimetry observations, Journal of Hydrology, 603(A), 126870, doi: 10.1016/j.jhydrol.2021.126870, 2021.

Bjerklie, D. M., Lawrence Dingman, S., Vorosmarty, C. J., Bolster, C. H., and Congalton, R. G.: Evaluating the potential for measuring river discharge from space, Journal of Hydrology, 278(1-4), 17–38, doi: 10.1016/S0022-1694(03)00129-X, 2003.

Bjerklie, D. M., Moller, D., Smith, L. C., and Dingman, S. L.: Estimating discharge in rivers using remotely sensed hydraulic information, Journal of Hydrology, 309(1-4), 191–209, doi: 10.1016/j.jhydrol.2004.11.022, 2005.

Bjerklie, D. M.: Estimating the bankfull velocity and discharge for rivers using remotely sensed river morphology information, Journal of Hydrology, 341(3-4), 144–155, doi: 10.1016/j.jhydrol.2007.04.011, 2007.

Black, E., Tarnavsky, E., Maidment, R., Greatrex, H., Mookerjee, A., Quaife, T., and Brown, M.: The Use of Remotely Sensed Rainfall for Managing Drought Risk: A Case Study of Weather Index Insurance in Zambia, Remote Sensing, 8(4), 342, doi: 10.3390/rs8040342, 2016.

Boergens, E., Buhl, S., Dettmering, D., Klüppelberg, C., and Seitz, F.: Combination of multi-mission altimetry data along the Mekong River with spatio-temporal kriging, Journal of Geodesy, 91(5), 519–534, doi: 10.1007/s00190-016-0980-z, 2017.

Brakenridge, G. R., Nghiem, S. V., Anderson, E., and Mic, R.: Orbital microwave measurement of river discharge and ice status, Water Resources Research, 43, W04405, doi: 10.1029/2006WR005238, 2007.

Brocca, L., Melone, F., Moramarco, T., and Wagner, W.: A new method for rainfall estimation through soil moisture observations, Geophysical Research Letters, 40, 853–858, doi: 10.1002/grl.50173, 2013.

Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., and Levizzani, V.: Soil as a natural rain gauge: estimating global rainfall from satellite soil moisture data, Journal of Geophysical Research, 119(9), 5128–5141, doi: 10.1002/2014JD021489, 2014a.

Brocca, L., Camici, S., Melone, F., Moramarco, T., Martínez-Fernández, J., Didon-Lescot, J. F., and Morbidelli, R.: Improving the representation of soil moisture by using a semi-analytical infiltration model, Hydrological Processes, 28, 2103–2115, doi: 10.1002/hyp.9766, 2014b.

Brocca, L., Massari, C., Ciabatta, L., Moramarco, T., Penna, D., Zucco, G., Pianezzola, L., Borga, M., Matgen, P., and Martínez-Fernández, J.: Rainfall estimation from in situ soil moisture observations at several sites in Europe: an evaluation of SM2RAIN algorithm, Journal of Hydrology and Hydromechanics, 63(3), 201–209, doi: 10.1515/johh-2015-0016, 2015.

Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A., and Fernàndez-Prieto, D.: How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products, International Journal of Applied Earth Observation and Geoinformation, 73(C), 752–766, doi: 10.1016/j.jag.2018.08.023, 2018.

Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Massari, C., Camici, S., Schüller, L., Bojkov, B., and Wagner, W.: SM2RAIN-ASCAT (2007-2018): global daily satellite rainfall from ASCAT

soil moisture, Earth System Science Data, 11, 1583–1601, doi: 10.5194/essd-11-1583-2019, 2019.

Brocca, L., Massari, C., Pellarin, T., Filippucci, P., Ciabatta, L., Camici, S., Kerr, Y. H., and Fernández-Prieto, D.: River flow prediction in data scarce regions: soil moisture integrated satellite rainfall products outperform rain gauge observations in West Africa, Scientific Reports, 10, 12517, doi: 10.1038/s41598-020-69343-x, 2020.

Brunetti, M. T., Melillo, M., Peruccacci, S., Ciabatta, L., and Brocca, L.: How far are we from the use of satellite data in landslide forecasting?, Remote Sensing of Environment, 210, 65–75, doi: 10.1016/j.rse.2018.03.016, 2018.

Bruno, G., Pignone, F., Silvestro, F., Gabellani, S., Schiavi, F., Rebora, N., Giordano P., and Falzacappa, M.: Performing Hydrological Monitoring at a National Scale by Exploiting Rain-Gauge and Radar Networks: The Italian Case, Atmosphere, 12(6), 771, doi: 10.3390/atmos12060771, 2021.

Camici, S., Ciabatta, L., Massari, C., and Brocca, L.: How reliable are satellite precipitation estimates for driving hydrological models: a verification study over the Mediterranean area, Journal of Hydrology, 563, 950–961, doi: 10.1016/j.jhydrol.2018.06.067, 2018.

Casella, D., Panegrossi, G., Sanò, P., Milani, L., Petracca, M., and Dietrich S.: A novel algorithm for detection of precipitation in tropical regions using PMW radiometers, Atmospheric Measurement Techniques, 8 1217-1232, doi:10.5194/amt-8-1217-2015, 2015.

Cenci, L., Pulvirenti, L., Boni, G., Chini, M., Matgen, P., Gabellani, S., Squicciarino, G., and Pierdicca, N.: An evaluation of the potential of Sentinel 1 for improving flash flood predictions via soil moisture-data assimilation, Advances in Geosciences, 44, 89–100, 2017, doi: 10.5194/adgeo-44-89-2017, 2017.

Chandanpurkar, H. A., Reager, J. T., Famiglietti, J. S., and Syed, T. H.: Satellite- and Reanalysis-Based Mass Balance Estimates of Global Continental Discharge (1993–2015), Journal of Climate, 30, 8481–8495, doi: 10.1175/JCLI-D-16-0708.1, 2017. Chen, F., Crow, W. T., Ciabatta, L., Filippucci, P., Panegrossi, G., Marra, A. C., Puca, S., and Massari, C.: Enhanced large-scale validation of satellite-based land rainfall products, Journal of Hydrometeorology, 22(2), 245–257, doi: 10.1175/JHM-D-20-0056.1, 2020.

Ciabatta, L., Brocca, L., Massari, C., Moramarco, T., Puca, S., Rinollo, A., Gabellani, S., and Wagner, W.: Integration of satellite soil moisture and rainfall observations over the Italian territory, Journal of Hydrometeorology, 16(3), 1341–1355, doi: 10.1175/JHM-D-14-0108.1, 2015.

Ciabatta, L., Camici, S., Brocca, L., Ponziani, F., Stelluti, M., Berni, N., and Moramarco, T.: Assessing the impact of climate-change scenarios on landslide occurrence in Umbria Region, Italy, Journal of Hydrology, 541(A), 258–295, doi: 10.1016/j.jhydrol.2016.02.007, 2016.

Ciabatta, L., Marra, A.C., Panegrossi, G., Casella, D., Sanò, P., Dietrich, S., Massari, C., and Brocca, L.: Daily precipitation estimation through different microwave sensors: verification study over Italy, Journal of Hydrology, 545, 436-450, doi: 10.1016/j.jhydrol.2016.12.057, 2017

Ciabatta, L., Massari, C., Brocca, L., Gruber, A., Reimer, C., Hahn, S., Paulik, C., Dorigo, W., Kidd, R., and Wagner, W.: SM2RAIN-CCI: a new global long-term rainfall data set derived from ESA CCI soil moisture, Earth System Science Data, 10, 267–280, doi: 10.5194/essd-10-267-2018, 2018.

Crochemore, L., Isberg, K., Pimentel, R., Pineda, L., Hasan, A., and Arheimer, B.: Lessons learnt from checking the quality of openly accessible river flow data worldwide, Hydrological Sciences Journal, 65(5), 699–711, doi: 10.1080/02626667.2019.1659509, 2020.

Crow, W. T., Huffman, G. F., Bindlish, R., and Jackson, T. J.: Improving satellite rainfall accumulation estimates using spaceborne soil moisture retrievals, Journal of Hydrometeorology, 10, 199–212, doi: 10.1175/2008JHM986.1, 2009.

Dari, J., Brocca, L., Quintana-Seguí, P., Escorihuela, M. J., Stefan, V., and Morbidelli, R.: Exploiting High-Resolution Remote Sensing Soil Moisture to Estimate Irrigation Water Amounts over a Mediterranean Region, Remote Sensing, 12(16), 2593, doi: 10.3390/rs12162593, 2020.

De Frasson, R. P. M., Pavelsky, T. M., Fonstad, M. A., Durand, M. T., Allen, G. H., Schumann, G., Lion, C., Beighley, R. E., and Yang, X.: Global Relationships Between River Width, Slope, Catchment Area, Meander Wavelength, Sinuosity, and Discharge: Geophysical Research Letters, 46(6), 3252–3262, doi: 10.1029/2019GL082027, 2019.

Di Baldassarre, G., Schumann, G., and Bates, P. D.: A technique for the calibration of hydraulic models using uncertain satellite observations of flood extent, Journal of Hydrology, 367(3-4), 276–282, doi: 10.1016/j.jhydrol.2009.01.020, 2009.

Dinku, T.: Challenges with availability and quality of climate data in Africa, In: Melesse, A., Abtew, W., and Senay, G.: Extreme Hydrology and Climate Variability, Elsevier, 71–80. ISBN: 9780128159989, doi: 10.1016/B978-0-12-815998-9.00007-5, 2019.

Durand, M., Andreadis, K. M., Alsdorf, D. E., Lettenmaier, D. P., Moller, D., and Wilson, M.: Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. Geophysical Research Letters, 35(20), L20401, doi: 10.1029/2008GL034150, 2008.Dziubanski, D. J., and Franz, K. J.: Assimilation of AMSR-E snow water equivalent data in a spatially-lumped snow model, Journal of .Hydrology, 540, 26–39, doi: 10.1016/j.jhydrol.2016.05.046, 2016.

Enenkel, M., Osgood, D., Anderson, M., Powell, B., McCarty, J., Neigh, C., Carroll, M., Wooten, M., Husak, G., Hain, C., and Brown, M.: Exploiting the Convergence of Evidence in Satellite Data for Advanced Weather Index Insurance Design, Weather, Climate, and Society, 11(1), 65–93, doi: 10.1175/WCAS-D-17-0111.1, 2019.

Famiglietti, J. and Wood, E. F.: Multiscale modeling of spatially variable water and energy balance processes, Water Resources Research, 30, 3061–3078, doi: 10.1029/94WR01498, 1994.

Fekete, B. and Vörösmarty, C.: The current status of global river discharge monitoring and potential new technologies complementing traditional discharge measurements. Proceedings of the Predictions in Ungauged Basins: PUB Kick-off (Proceedings of the PUB Kick-off meeting held in Brasilia, 20-22 November 2002), IAHS Publ. 309, doi: unavailable, 2007.

Fekete, B. M., Robarts, R. D., Kumagai, M., Nachtnebel, H. P., Odada, E., and Zhulidov, A. V.: Time for in situ renaissance, Science, 349, 685–686, doi: 10.1126/science.aac7358, 2015.

Feng D., Gleason C. J., Yang X., and Pavelsky T. M.: Comparing discharge estimates made via the BAM algorithm in high-order Arctic rivers derived solely from optical CubeSat, Landsat, and Sentinel-2 data, Water Resources Research, 55(9), 7753-7771, doi: 10.1029/2019wr025599, 2019.

Filippucci, P., Tarpanelli, A., Massari, C., Serafini, A., Strati, V., Alberi, M., Raptis, K. G. C., Mantovani, F., and Brocca, L.: Soil moisture as a potential variable for tracking and quantifying irrigation: A case study with proximal gamma-ray spectroscopy data, Advances in Water Resources, 136, 103502, doi: 10.1016/j.advwatres.2019.103502, 2020.

Gentile, F., Bisantino, T., Corbino, R., Milillo, F., Romano, G., and Liuzzi, G. T.: Monitoring and analysis of suspended sediment transport dynamics in the Carapelle torrent (southern Italy), Catena, 80(1), 1–8, doi: 10.1016/j.catena.2009.08.004, 2010.

Gleason C. J., and Durand M. T.: Remote Sensing of River Discharge: A Review and a Framing for the Discipline, Remote Sensing, 12(7), 1107, doi: 10.3390/rs12071107, 2020

Elmi O., Tourian M. J., Bárdossy A., and Sneeuw N.: Spaceborne River Discharge From a Nonparametric Stochastic Quantile Mapping Function, Water Resources Research, 57(12), e2021WR030277, doi: 10.1029/2021WR030277, 2021.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote sensing of Environment, 202, 18–27, doi: 10.1016/j.rse.2017.06.031, 2017.

Gravelle, R.: Discharge Estimation: Techniques and Equipment, chapter in book: Geomorphological Techniques (Online Edition), British Society for Geomorphology, Clarke L.E., Nield J.M, ISSN 2047-0371, doi: unavailable, 2015.

Griffith, C. G., Woodley, W. L., Grube, P. G., Martin, D. W., Stout, J., and Sikdar, D. N.: Rain estimation from geosynchronous satellite imagery—Visible and infrared studies, Monthly Weather Review, 106(8), 1153–1171, doi: 10.1175/1520-0493(1978)106<1153:REFGSI>2.0.CO;2,1978.

Habib, E., Haile, A. T., Tian, Y., and Joyce, R. J.: Evaluation of the high-resolution CMORPH satellite rainfall product using dense rain gauge observations and radar-based estimates, Journal of Hydrometeorology, 13(6), 1784–1798, doi: 10.1175/JHM-D-12-017.1, 2012.

Hahn, S., Delogu, F., Gabellani, S., and Brocca L.: Product Validation Report (PVR) H113, Tech. rep., H SAF, Doc. No: SAF/HSAF/CDOP3/PVR/H113, available at: http://hsaf.meteoam.it/CaseStudy/GetDocumentUserDocument?fileName=H113_ASCAT_SSM_CDR_PVR_v0.3.pdf&tipo=PVR, 2018.

Hannah, D. M., Demuth, S., Van Lanen, H. A. J., Looser, U., Prudhomme, C., Rees, G., Stahl, K., and Tallaksen, L. M.: Large-scale river flow archives: importance, current status and future needs, Hydrological Processes, 25 (7), 1191-1200, doi: 10.1002/hyp.7794, 2011.

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

Herschy, R. W.: Streamflow Measurement (2nd Edition), Taylor and Francis, London, ISBN: 9780367865184, doi: unavailable, 1995.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and Iguchi, T.: The global precipitation measurement mission, Bulletin of the American Meteorological Society, 95(5), 701–722, doi: 10.1175/BAMS-D-13-00164.1, 2014.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E. J., Sorooshian, S., Stocker, E. F., Tan, J., Wolff, D. B., and Xie, P.: Integrated Multi-satellitE Retrievals for the Global Precipitation Measurement (GPM) mission (IMERG), Chapter 19 in Adv. Global Change Res., Vol. 67, Satellite Precipitation Measurement, Levizzani, V., Kidd, C., Kirschbaum, D., Kummerow, C., Nakamura, K., and Turk F. J. (Ed.), Springer Nature, Dordrecht, ISBN 978-3-030-24567-2 / 978-3-030-24568-9 (eBook), 343-353, doi: 10.1007/978-3-030-24568-9_19, 2020.

Huo J., Qu X., Zhu D., Yuan Z., and Zeng Z.: Runoff monitoring in the Lhasa River Basin using passive microwave data, International Journal of Applied Earth Observation and Geoinformation, 103, 102486, doi: 10.1016/j.jag.2021.102486, 2021.

IPCC: The Physical Science Basis, Summary for Policymakers. International Panel of Climate Change, Working Group 1 contribution, Climate Change 2013, Cambridge University Press, 1–30, doi: 10.1017/CBO9781107415324.004, 2013

Jakeman, A. J., Hornberger, G. M., Littlewood, I. G., Whitehead, P. G., Harvey, J. W., and Bencala, K. E.: A systematic approach to modelling the dynamic linkage of climate, physical catchment descriptors and hydrological response components, Math. Comp. Sim., 33, 359–366, doi: 10.1016/0378-4754(92)90122-W, 1992.

Jodar, J., Carpintero, E., Martos-Rosillo, S., Ruiz-Constan, A., Marin-Lechado, C., Cabrera-Arrabal, J. A., Navarrete-Mazariegos, E., Gonzalez-Ramon, A., Lamban, L. J., Herrera, C., and González-Dugo, M.P.: Combination of lumped hydrological and remote-sensing models to evaluate water resources in a semi-arid high altitude ungauged watershed of Sierra Nevada (Southern Spain), Science of the Total Environment, 625, 285–300, doi: 10.1016/j.scitotenv.2017.12.300, 2018.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P.: CMORPH: A method that produces global precipitation estimates from passive nicrowave and infrared data at high spatial and temporal resolution, Journal of Hydrometeorology, 5(3), 487–503, doi: 10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2, 2004.

Keesstra, S. D., Davis, J., Masselink, R. H., Casalí, J., Peeters, E. T., and Dijksma, R.: Coupling hysteresis analysis with sediment and hydrological connectivity in three agricultural catchments in Navarre, Spain, Journal of Soils and Sediments, 19(3), 1598–1612, doi: 10.1007/s11368-018-02223-0, 2019.

Kidd, C. and Huffman, G.: Global precipitation measurement, Meteorological Application, 18, 334–353, doi: 10.1002/met.284, 2011.

Kidd, C. and Levizzani, V.: Status of satellite precipitation retrievals, Hydrology and Earth System Science, 15, 1109–1116, doi: 10.5194/hess-15-1109-2011, 2011.

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., and Kirschbaum, D. B.: So, How Much of the Earth's Surface is Covered by Rain Gauges?, Bulletin of the American Meteorological Society, 98, 69–78, doi: 10.1175/BAMS-D-14-00283.1, 2017.

Kremezi, M. and Karathanassi, V.: Correcting the BRDF effects on Sentinel-2 ocean images, Proc. SPIE 11174, Seventh International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2019), 111741C (27 June 2019), doi: 10.1117/12.2533653, 2019.

Kummerow, C., Simpson, J., Thiele, O., Barnes, W., Chang, A. T. C., Stocker, E., Adler, R. F., Hou, A., Kakar, R., Wentz, F., Ashcroft, P., Kozu, T., Hong, Y., Okamoto, K., Iguchi, T., Kuroiwa, H., Im, E., Haddad, Z., Huffman, G., Ferrier, B., Olson, W. S., Zipser, E., Smith, E. A., Wilheit, T. T., North, G., Krishnamurti, T., and Nakamura, K.: The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit, Journal of Applied Meteorology, 39(12), 1965–1982 doi: 10.1175/1520-0450(2001)040%3C1965:TSOTTR%3E2.0.CO;2, 2000.

Kummu, M., Guillaume, J., de Moel, H., Eisner, S., Flörke, M., Porkka, M., Siebert, S., Veldkamp T. I. E., and Ward, P. J.: The world's road to water scarcity: shortage and stress in the 20th century and pathways towards sustainability, Scientific Reports, 6, 38495, doi: 10.1038/srep38495, 2016.

La Barbera, P., Lanza, L. G., and Stagi, L.: Tipping bucket mechanical errors and their influence on rainfall statistics and extremes, Water Science & Technology, 45(2), 1–9, doi: 10.2166/wst.2002.0020, 2002.

Lin, P., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., David, C. H., Durand, M., Pavelsky, T. M., Allen, G. H., Gleason, C. J., and Wood, E. F.: Global Reconstruction of Naturalized River Flows at 2.94 Million Reaches, Water Resources Research, 55, 6499–6516, doi: 10.1029/2019WR025287, 2019.

Lopez, P. L., Immerzeel, W. W. Sandoval, E. A. R., Sterk, G., and Schellekens, J.: Spatial Downscaling of Satellite-Based Precipitation and Its Impact on Discharge Simulations in the Magdalena River Basin in Colombia, Frontiers in Earth Science, 6, 68, doi: 10.3389/feart.2018.00068, 2018.

Maggioni, V., and Massari, C.: On the performance of satellite precipitation products in riverine flood modeling: a review, Journal of Hydrology, 558, 214–224, doi: 10.1016/j.jhydrol.2018.01.039, 2018.

Malinowski, R., Groom, G., Schwanghart, W., and Goswin, H.: Detection and Delineation of Localized Flooding from WorldView-2 Multispectral Data, Remote Sensing, 7, 14853–14875, doi: 10.3390/rs71114853, 2015.

Massari, C., Crow, W., and Brocca, L.: An assessment of the performance of global rainfall estimates without ground-based observations, Hydrology and Earth System Sciences, 21, 4347–4361, doi: 10.5194/hess-21-4347-2017, 2017.

Massari, C., Brocca, L., Pellarin, T., Abramowitz, G., Filippucci, P., Ciabatta, L., Maggioni, V., Kerr, Y., and Fernández-Prieto, D.: A daily/25km short-latency rainfall product for data scarce regions based on the integration of the GPM IMERG Early Run with multiple satellite soil moisture products, Hydrology and Earth System Sciences, 24, 2687–2710, doi: 10.5194/hess-24-2687-2020, 2020.

McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., and Stoffelen, A.: Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown targetd Geophysical Research Letters, 41, 6229–6236, doi: 10.1002/2014GL061322, 2014.

Merlin, O., Chehbouni, A., Walker, J. P., Panciera, R., and Kerr, Y. H.: A simple method to disaggregate passive microwave-based soil moisture, IEEE Transactions on Geoscience and Remote Sensing, 46(3), 786–796, doi: 10.1109/TGRS.2007.914807, 2008.

Michelson, D. B.: Systematic correction of precipitation gauge observations using analyzed meteorological variables, Journal of Hydrology, 290(3–4), 161–177, doi: 10.1016/j.jhydrol.2003.10.005, 2004.

Morrison, K., and Wagner, W.: Explaining Anomalies in SAR and Scatterometer Soil Moisture Retrievals From Dry Soils With Subsurface Scattering, IEEE Transactions on Geoscience and Remote Sensing, 58 (3), 2190–2197, doi: 10.1109/TGRS.2019.2954771, 2020.

Mueller N., Lewis A., Roberts D., Ring S., Melrose R., Sixsmith J., Lymburner L., McIntyre A., Tan P., Curnow S., and Ip A.: Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia, Remote Sensing of Environment, 174, 341-352, doi: 10.1016/j.rse.2015.11.003, 2016.

Nash, J. E. and Sutcliffe, J. V.: River Flow Forecasting through Conceptual Model. Part 1 - A Discussion of Principles, Journal of Hydrology, 10, 282–290, doi: 10.1016/0022-1694(70)90255-6, 1970.

Neal, J., Schumann, G., Bates, P., Buytaert, W., Matgen, P., and Pappenberger, F.: A data assimilation approach to discharge estimation from space, Hydrological Processes, 23(25), 3641–3649, doi: 10.1002/hyp.7518, 2009.

Nguyen, P., Shearer, E. J., Tran, H., Ombadi, M., Hayatbini, N., Palacios, T., Huynh, P., Updegraff, G., Hsu, K., Kuligowski, B., Logan, W. S., and Sorooshian, S.: The CHRS Data Portal, an easily accessible public repository for PERSIANN global satellite precipitation data, Nature Scientific Data, 6, Article 180296, doi: 10.1038/sdata.2018.296, 2019

Nohara, D., Kitoh, A., Hosaka, M., and Oki, T.: Impact of Climate Change on River Discharge Projected by Multimodel Ensemble, Journal of Hydrometeorology, 7(5), 1076–1089, doi: 10.1175/JHM531.1, 2006.

Pai, D., Sridhar, L., Rajeevan, M., Sreejith, O., Satbhai, N., and Mukhopadhyay, B.: Development of a new high spatial resolution (0.25×0.25) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region, Mausam, 65, 1–18, doi: 10.54302/mausam.v65i1.851, 2014

Paris, A., Dias de Paiva, R., Santos da Silva, J., Medeiros Moreira, D., Calmant, S., Garambois, P. A., Collischonn, W., Bonnet, M. P., and Seyler, F.: Stage-discharge rating curves based on satellite altimetry and modeled discharge in the Amazon basin, Water Resources Research, 52(5), 3787–3814, doi: 10.1002/2014WR016618, 2016.

Parr, D., Wang, G., and Bjerklie, D.: Integrating Remote Sensing Data on Evapotranspiration and Leaf Area Index with Hydrological Modeling: Impacts on Model Performance and Future Predictions, Journal of Hydrometeorology, 16, 2086–2100, doi: 10.1175/JHM-D-15-0009.1, 2015.

Pavelsky T. M.: Using width-based rating curves from spatially discontinuous satellite imagery to monitor river discharge, Hydrological Processes, 28, 3035-3040, doi: 10.1002/hyp.10157, 2014.

Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen-Geiger climate classification, Hydrology and Earth System Sciences, 11, 1633–1644, doi: 10.5194/hess-11-1633-2007, 2007.

Pekel, J. F., Cottam, A., Gorelick, N., and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, Nature, 540, 418–422, doi: 10.1038/nature20584, 2016.

Pellarin, T., Louvet, S., Gruhier, C., Quantin, G., and Legout, C.: A simple and effective method for correcting soil moisture and precipitation estimates using AMSR-E measurements, Remote Sensing of Environment, 136, 28–36, doi: 10.1016/j.rse.2013.04.011, 2013.

Pelletier, P. M.: Uncertainties in the single determination of river discharge: a literature review, Canadian Journal of Civil Engineering, 15, 834-850, doi: 10.1139/l89-116, 1988.

Peng, J., Loew, A., Merlin, O., and Verhoest, N. E.: A review of spatial downscaling of satellite remotely sensed soil moisture, Reviews of Geophysics, 55, 341–366, doi: 10.1002/2016RG000543, 2017.

Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gul-lett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Førland, E. J., Hanssen-Bauer, I., Alexandersson, H., Jones, P., and Parker, D.: Homogeneity adjustments of in situ atmospheric climate data: a review, International Journal of Climatology, 18, 1493–1517. doi: 10.1002/(SICI)1097-0088(19981115)18:13<1493::AID-JOC329>3.0.CO;2-T, 1998.

Piecuch, C. G., Bittermann, K., Kemp, A. C., Ponte, R. M., Little, C. M., Engelhart, S. E., and Lentz, S. J.: River-discharge effects on United States Atlantic and Gulf coast sea-level changes,

Proceedings of the National Academy of Sciences, 115(30), 7729–7734, doi: 10.1073/pnas.1805428115, 2018.

Pignone, F., Rebora, N., Silvestro, F., and Castelli, F.: GRISO – Rain, CIMA Research Foundation, Savona, Italy, Operational Agreement 778/2009 DPC-CIMA, Year-1 Activity Report 272/2010. doi: unavailable, 2010.

Ponziani, F., Pandolfo, C., Stelluti, M., Berni, N., Brocca, L., and Moramarco, T.: Assessment of rainfall thresholds and soil moisture modeling for operational hydrogeological risk prevention in the Umbria region (central Italy), Landslides, 9, 229-237, doi: 10.1007/s10346-011-0287-3, 2012.

Post, D. A. and Jakeman, A. J.: Predicting the daily streamflow of ungauged catchments in S.E. Australia by regionalising the parameters of a lumped conceptual rainfall-runoff model, Ecological Modeling, 123, 91 – 104, doi: 10.1016/S0304-3800(99)00125-8, 1999.

Quast, R. and Wagner, W.: Analytical solution for first-order scattering in bistatic radiative transfer interaction problems of layered media, Applied Optics, 55(20), 5379–5386, doi: 10.1364/ao.55.005379, 2016.

Quast, R., Albergel, C., Calvet, J. C., and Wagner, W.: A Generic First-Order Radiative Transfer Modelling Approach for the Inversion of Soil and Vegetation Parameters from Scatterometer Observations, Remote Sensing, 11(3), 285, doi: 10.3390/rs11030285, 2019.

Quast, R., Wagner, W., Bauer-Marschallinger, B., and Vreugdenhil, M.: Soil moisture retrieval from Sentinel-1 using a first-order radiative transfer model - a case-study over the Po-Valley, in preparation

Ragettli, S., Cortés, G., Mcphee, J., and Pellicciotti, F.: An evaluation of approaches for modelling hydrological processes in high-elevation, glacierized Andean watersheds, Hydrological Processes, 28, 5674–5695, doi: 10.1002/hyp.10055, 2013.

Rockström, J., Falkenmark, M., Lannerstad, M., and Karlberg, L.: The planetary water drama: Dual task of feeding humanity and curbing climate change, Geophysical Research Letters, 39, L15401, doi: 10.1029/2012GL051688, 2012.

Sauvageot, H.: Rainfall measurement by radar: A review, Atmospheric Research, 35, 27–54, doi: 10.1016/0169-8095(94)90071-X, 1994

Schamm, K., Ziese, M., Becker, A., Finger, P., Meyer-Christoffer, A., Schneider, U., Schröder, M., and Stender, P.: Global gridded precipitation over land: a description of the new GPCC First Guess Daily product, Earth System Science Data, 6, 49–60, doi: 10.5194/essd-6-49-2014, 2014.

Schumann, G., Matgen, P., Ho_mann, L., Hostache, R., Pappenberger, F., and Pfister, L.: Deriving distributed roughness values from satellite radar data for flood inundation modelling, Journal of Hydrology, 344(1-2), 96–111, doi: 10.1016/j.jhydrol.2007.06.024, 2007.

Schwatke, C., Dettmering, D., Bosch, W., and Seitz, F.: DAHITI–an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry, Hydrology and Earth System Sciences, 19(10), 4345–4364, doi: 10.5194/hess-19-4345-2015, 2015.

Sefton, C. E. M., and Howarth, S. M.: Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales, J. Hydrol. 211, 1–16, doi: 10.1016/S0022-1694(98)00163-2, 1998.

Sehgal, V., Gaur, N., and Mohanty, B. P., Global Surface Soil Moisture Drydown Patterns, Water Resources Research, 56, e2020WR027588, doi: 10.1029/2020WR027588, 2020.

Seibert, J.: Regionalisation of parameters for a conceptual rainfall-runoff model, Agric. Forest Meteor. 98–99, 279–293, doi: 10.1016/s0168-1923(99)00105-7, 1999.

Shi, Z., Chen, Y., Liu, Q., and Huang, C.: Discharge Estimation Using Harmonized Landsat and Sentinel-2 Product: Case Studies in the Murray Darling Basin, Remote Sensing, 12(17), 2810, doi: 10.3390/rs12172810, 2020.

Sichangi A. W., Wang L., Yang K., Chen D., Wang Z., Li X., Zhou J., Liu W., and Kuria D.: Estimating continental river basin discharges using multiple remote sensing data sets, Remote Sensing of Environment 179, 36-53, doi: 10.1016/j.rse.2016.03.019, 2016.

Silberstein, R. P.: Hydrological models are so good, do we still need data?, Environmental Modelling & Software, 21, 1340–1352, doi: 10.1016/j.envsoft.2005.04.019, 2006.

Stoffelen, A., Toward the true near-surface wind speed: Error modeling and calibration using triple collocation, Journal of Geophysical Research: Oceans, 103, 7755–7766, 1998, doi: 10.1029/97JC03180, 1998.

Strangeways, I.: Precipitation: Theory, measurement and distribution. Cambridge: Cambridge University Press., ISBN: 9780521851176, doi: unavailable, 2006.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K.-L.: A review of global precipitation data sets: Data sources, estimation, and intercomparisons, Reviews of Geophysics, 56, 79–107, doi: 10.1002/2017RG000574, 2018.

Sun, W., Ishidaira, H., Bastola, S., and Yu, J.: Estimating daily time series of streamflow using hydrological model calibrated based on satellite observations of river water surface width: Toward real world applications. Environmental Research, 139, 36–45, doi: 10.1016/j.envres.2015.01.002, 2015.

Syed, T. H., Famiglietti, J. S., Chen, J., Rodell, M., Seneviratne, S. I., Viterbo, P., and Wilson, C. R.: Total basin discharge for the Amazon and Mississippi River basins from GRACE and a land-atmosphere water balance, Geophysical Research Letters, 32(24), L24404, doi: 10.1029/2005GL024851, 2005.

Tapiador, F. J., Turk, F. J., Petersen, W., Hou, A. Y., Garcia-Ortega, E., Machado, L., Angelis, C. F., Salio, P., Kidd, C., Huffman, G. J. and de Castro, M.: Global precipitation measurement: Methods, datasets and applications, Atmospheric Research, 104, 70–97, doi: 10.1016/j.atmosres.2011.10.021, 2012.

Tarpanelli, A., Brocca, L., Lacava, T., Melone, F., Moramarco, T., Faruolo, M., Pergola, N., and Tramutoli, V.: Toward the estimation of river discharge variations using MODIS data in ungauged basins, Remote Sensing of Environment, 136, 47–55, doi: 10.1016/j.rse.2013.04.010, 2013.

Tarpanelli, A., Amarnath, G., Brocca, L., Massari, C., and Moramarco, T.: Discharge estimation and forecasting by MODIS and altimetry data in Niger-Benue River, Remote Sensing of Environment, 195, 96–106, doi: 10.1016/j.rse.2017.04.015, 2017a.

Tarpanelli, A., Massari, C., Ciabatta, L., Filippucci, P., Amarnath, G., and Brocca, L.: Exploiting a constellation of satellite soil moisture sensors for accurate rainfall estimation, Advances in Water Resources, 108, 249–255, doi: 10.1016/j.advwatres.2017.08.010, 2017b.

Tarpanelli, A., Iodice, F., Brocca, L., Restano, M., and Benveniste, J.: River Flow Monitoring by Sentinel-3 OLCI and MODIS: Comparison and Combination, Remote Sensing, 12, 3867, doi: 10.3390/rs12233867, 2020.

Tarpanelli, A., Camici, S., Nielsen, K., Brocca, L., Moramarco, T., and Benveniste, J.: Potentials and limitations of Sentinel-3 for river discharge assessment, Advances in Space Research, 68(2), 593–606, doi: 10.1016/j.asr.2019.08.005, 2021.

Tourian, M. J., Sneeuw, N., and Bardossy, A.: A quantile function approach to discharge estimation from satellite altimetry (ENVISAT), Water Resources Research, 49(7), 4174–4186, doi: 10.1002/wrcr.20348, 2013.

Tourian, M. J., Tarpanelli, A., Elmi, O., Qin, T., Brocca, L., Moramarco, T., and Sneeuw, N.: Spatiotemporal densification of river water level time series by multimission satellite altimetry, Water Resources Research, 52, 1140–1159, doi: 10.1002/2015WR017654, 2016.

Tourian, M. J., Schwatke, C., and Sneeuw, N.: River discharge estimation at daily resolution from satellite altimetry over an entire river basin, Journal of Hydrology, 546, 230–247, doi: 10.1016/j.jhydrol.2017.01.009, 2017.

Trenberth, K. E.: Changes in Precipitation with Climate Change. Climate Change Research, Climate Research, 47(1), 123–138, doi: 10.3354/cr00953, 2011.

Ulaby, F. T., Moore, R. K., and Fung, A. K.: Microwave Remote Sensing, Active and Passive. vol. 1, 2, 3, Artech House, ISBN-13: 978-0890061909, 978-0201107609, 978-0890061923, doi: unavailable, 1981.

Villarini, G., Mandapaka, P. V., Krajewski, W. F., and Moore, R. J.: Rainfall and sampling uncertainties: A rain gauge perspective, Journal of Geophysical Research, 113, D11102, doi: 10.1029/2007JD009214, 2008.

Vörösmarty, C., Askew, A., Grabs, W., Barry, R. G., Birkett, C., Döll, P., Goodison, B., Hall, A., Jenne, R., Kitaev, L., Landwehr, J., Keeler, M., Leavesley, G., Schaake, J., Strzepek, K., Sundarvel, S. S., Takeuchi, K., and Webster, F.: Global water data: A newly endangered species, Eos Transactions American Geophysical Union, 82(5), 54–58, doi: 10.1029/01EO00031, 2001.

Wagener, T., Wheater, H. S., and Gupta, H. V.: Rainfall-Runoff Modelling in Gauged and Ungauged Catchments, Imperial College Press, London p. 300, doi: 10.1142/p335, 2004.

Wagner, W., Lemoine, G., and Rott, H.: A method for estimating soil moisture from ERS scatterometer and soil data, Remote Sensing of Environment, 70, 191–207, doi: 10.1016/S0034-4257(99)00036-X, 1999.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Blöschl, G., Eitzinger, J., and Steinnocher, K.: The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications, Meteorologische Zeitschrift, 22, 5–33, doi: 10.1127/0941-2948/2013/0399, 2013.

Wang, S., Liu, S., Mo, X., Peng, B., Qiu, J., Li, M., Liu, C., Wang, Z., and Bauer-Gottwein, P.: Evaluation of Remotely Sensed Precipitation and Its Performance for Streamflow Simulations in Basins of the Southeast Tibetan Plateau, Journal of Hydrometeorology, 16(6), 2577–2594, doi: 10.1175/JHM-D-14-0166.1, 2015.

Wentz, F. J., and Spencer, R. W.: SSM/I rain retrievals within a unified all-weather ocean algorithm, Journal of the Atmospheric Sciences, 55(9), 1613–1627, doi: 10.1175/1520-0469(1998)055<1613:SIRRWA>2.0.CO;2, 1998.

Wieder, W. R., Boehnert, J., and Bonan, G. B.: Evaluating soil biogeochemistry parameterizations in Earth system models with observations, Global Biogeochemical Cycles, 28, 211–222, doi: 10.1002/2013GB004665, 2014.

Xie, P., Yatagai, A., Chen, M., Hayasaka, T., Fukushima, Y., Liu, C., and Yang S.: A gauge-based analysis of daily precipitation over East Asia, Journal of Hydrometeorology, 8, 607. 626, doi: 10.1175/JHM583.1, 2007.

Yang J., Huang X., and Tang Q.: Satellite-derived river width and its spatiotemporal patterns in China during 1990–2015, Remote Sensing of Environment, 247, 111918, doi: 10.1016/j.rse.2020.111918, 2020.

Zakharova E., Nielsen K., Kamenev G., and Kouraev A.: River discharge estimation from radar altimetry: Assessment of satellite performance, river scales and methods, Journal of Hydrology, 583, 124561, doi: 10.1016/j.jhydrol.2020.124561, 2020.

Zhang, Y., Pan, M., and Wood, E. F.: On Creating Global Gridded Terrestrial Water Budget Estimates from Satellite Remote Sensing, Surveys on Geophysics, 37, 249–268, doi: 1 0.1007/s10712-015-9354-y, 2016.

# 9. PUBLICATIONS

## Under review

**Filippucci, P.,** Brocca, L., Quast, R., Ciabatta, L., Saltalippi, C., Wagner, W., and Tarpanelli, A.: High resolution (1 km) satellite rainfall estimation from SM2RAIN applied to Sentinel-1: Po River Basin as case study, submitted to Hydrology and Earth System Sciences, 2021.

**Filippucci, P.,** Brocca, L., Bonafoni, S., Saltalippi, C., Wagner, W., and Tarpanelli, A.: Sentinel-2 high-resolution data for river discharge monitoring, submitted to Remote Sensing of Environment, 2021.

## Published in peer-reviewed journals

Abdalla S., Abdeh Kolahchi A., Adusumilli S., Aich Bhowmick S., Alou-Font E., Amarouche L., … **Filippucci P.,** …, Altimetry for the future: Building on 25 years of progress, Advances in Space Research, 68(2), 319-363, doi: 10.1016/j.asr.2021.01.022, 2021.

Brocca, L., Tarpanelli, A., **Filippucci, P.,** Dorigo, W., Zaussinger, F., Gruber, A., and Fernández-Prieto, D.: How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products, International Journal of Applied Earth Observation and Geoinformation, 73(C), 752–766, doi: 10.1016/j.jag.2018.08.023, 2018.

Brocca, L., **Filippucci, P.,** Hahn, S., Ciabatta, L., Massari, C., Camici, S., Schüller, L., Bojkov, B., and Wagner, W.: SM2RAIN-ASCAT (2007-2018): global daily satellite rainfall from ASCAT soil moisture, Earth System Science Data, 11, 1583–1601, doi: 10.5194/essd-11-1583-2019, 2019.

Brocca, L., Massari, C., Pellarin, T., **Filippucci, P.,** Ciabatta, L., Camici, S., Kerr, Y.H., and Fernández-Prieto, D.: River flow prediction in data scarce regions: soil moisture integrated satellite rainfall products outperform rain gauge observations in West Africa, Scientific Reports, 10(1), 12517, doi: 10.1038/s41598-020-69343-x, 2020.

Chen, F., Crow, W. T., Ciabatta, L., **Filippucci, P.,** Panegrossi, G., Marra, A. C., Puca, S., and Massari, C.: Enhanced Large-Scale Validation of Satellite-Based Land Rainfall Products, Journal of Hydrometeorology, 22(2), 245–257, doi: 10.1175/JHM-D-20-0056.1, 2021.

**Filippucci, P.,** Tarpanelli, A., Massari, C., Serafini, A., Strati, V., Alberi, M., Raptis, K. G. C., Mantovani, F., and Brocca, L.: Soil moisture as a potential variable for tracking and quantifying irrigation: A case study with proximal gamma-ray spectroscopy data, Advances in Water Resources, 136, 103502, doi: 10.1016/j.advwatres.2019.103502, 2020.

**Filippucci, P.,** Brocca, L., Massari, C., Saltalippi, C., Wagner, W., and Tarpanelli, A.: Toward a self-calibrated and independent SM2RAIN rainfall product, Journal of Hydrology, 603(A), 126837, doi: 10.1016/j.jhydrol.2021.126837, 2021.

Massari, C., Brocca, L., Pellarin, T., Abramowitz, G., **Filippucci, P.,** Ciabatta, L., Maggioni, V., Kerr, Y., and Fernández-Prieto, D.: A daily/25km short-latency rainfall product for data scarce regions based on the integration of the GPM IMERG Early Run with multiple satellite soil moisture products, Hydrology and Earth System Sciences, 24, 2687–2710, doi:10.5194/hess-24-2687-2020, 2020.

Tarpanelli, A., Santi, E., Tourian, M.J., **Filippucci, P.,** Amarnath, G., Brocca, L.: Daily river discharge estimates by merging satellite optical sensors and radar altimetry through artificial neural network, IEEE Transactions on Geoscience and Remote Sensing, 57(1), 329–341, doi: 10.1109/TGRS.2018.2854625, 2019.

Zaussinger, F., Dorigo, W., Gruber, A., Tarpanelli, A., **Filippucci, P.,** and Brocca, L.: Estimating irrigation water use over the contiguous United States by combining satellite and reanalysis soil moisture data, Hydrology and Earth System Sciences, 23, 897–923, doi:10.5194/hess-23-897-2019, 2019.

## Book chapters

Ciabatta, L., Camici, S., Massari, C., **Filippucci, P.,** Hahn, S., Wagner, W., and Brocca, L.: Soil moisture and precipitation: the SM2RAIN algorithm for rainfall retrieval from satellite soil moisture, In: Levizzani V., Kidd C., Kirschbaum D., Kummerow C., Nakamura K., and Turk F. J. (eds.) Satellite Precipitation Measurement, Advances in Global Change Research, vol 69. Springer, Cham, 1013-1027, ISBN: 978-3-030-35798-6 doi: 10.1007/978-3-030-35798-6_27, 2020.

## Conference contributions

Brocca, L., Camici, S., Ciabatta, L., Tarpanelli, A., Modanesi, S., **Filippucci, P.,** Massari, C., Brunetti, M. T., Peruccacci, S., Gariano, S. L., and Melillo, M.: Recent advances in using satellite soil moisture and precipitation for flood and landslide prediction in the Mediterranean basin, EGU 2019, 07-12/04/2019

Brocca, L., Dari, J., **Filippucci, P.,** Tarpanelli, A., Quintana-Segui, P., Escorihuela, M. J., and Morbidelli, R.: The missing information for hydrological modelling in agricultural areas: irrigation, EGU 2019, 07-12/04/2019

Ciabatta, L., Massari, C., Panegrossi, G., Marra, A. C., **Filippucci, P.,** Casella, D., Sanò, P., Dietrich, S., Melfi, D., and Brocca, L.: The HSAF H64 soil moisture-precipitation integrated product: development and preliminary results, EGU 2019, 07-12/04/2019

Tarpanelli, A., **Filippucci, P.,** Brocca, L., and Moramarco, T.: River discharge monitoring with multi-mission satellite sensor, EGU 2019, 07-12/04/2019

Tarpanelli, A., Camici, S., **Filippucci, P.,** Brocca, L., and Moramarco, T.: Towards river discharge estimation for ungauged sites, EGU 2019, 07-12/04/2019

Massari, C., Brocca, L., Pellarin, T., **Filippucci, P.,** Ciabatta, L., Maggioni, V., Kerr, Y., and Fernàndez-Prieto, D.: A Derived Optimal Linear Interpolation approach for merging multiple satellite soil moisture-based rainfall products with IMERG Early Run, EGU 2019, 07-12/04/2019

Zaussinger, F., Dorigo, W., Gruber, A., Tarpanelli, A., **Filippucci, P.,** and Brocca, L.: Estimating irrigation water use over the contiguous United States by combining satellite and reanalysis soil moisture data, EGU 2019, 07-12/04/2019

Bauer-Marschallinger, B., Reimer, C., Tuan, L., Vreugdenhil, M., Freeman, V., Massari, C., Modanesi, S., Ciabatta, L., **Filippucci, P.,** Brocca, L., and Wagner, W.: Sentinel-1 and MetOp ASCAT Driving Two New 1km Soil Moisture Products in Copernicus: First Data over Europe 2015-2018, EGU 2019, 07-12/04/2019

Brocca, L., Camici, S., Modanesi, S., **Filippucci, P.,** Ciabatta, L., Massari, C., Hahn, S., Wagner, W., Campione, E., Puca, S., and Cacciamani, C.: Operational services of satellite soil moisture products for flood, landslide, drought and precipitation, EMS Annual Meeting 2019, 09-13/09/2019

Quast, R., Wagner, W., Calvet, J. C., Albergel, C., Bertrand, B., Brocca, L., **Filippucci, P.,** and Hobbs, S.: Assessing prospects of sub-daily radar-observations to improve the understanding of soil- and vegetation dynamics, EGU 2020, 04-08/05/2020

Brocca, L., Camici, S., Massari, C., Ciabatta, L., **Filippucci, P.,** Villarini, G., and Tramblay, Y.: Satellite soil moisture improves rainfall just where needed, EGU 2020, 04-08/05/2020

Tarpanelli, A., Nielsen, K., **Filippucci, P.,** Belloni, R., Camici, S., Brocca, L., Moramarco, T., Restano, M., and Benveniste, J.: Monitoring of river discharge through the combination of multiple satellite data: RIDESAT project, EGU 2020, 04-08/05/2020

**Filippucci, P.,** Brocca, L., Tarpanelli, A., Massari, C., Ciabatta, L., Wagner, W., Bauer-Marschallinger, C., and Saltalippi, C.: Remote sensing of rainfall at high spatial-temporal resolution through soil moisture, EGU 2020, 04-08/05/2020

**Filippucci, P.,** Brocca, L., Tarpanelli, A., Massari, C., Wagner, W., and Saltalippi, C.: Toward Self-calibrated SM2RAIN-based rainfall product, EGU 2021, 19-30/04/2021

**Filippucci, P.,** Brocca, L., Saltalippi, C., Wagner, W., and Tarpanelli, A.: River discharge estimation for narrow rivers: benefits of Sentinel-2 high resolution data, 4th Hydrospace-GEOGloWS 2021, 7-11/06/2021