TU WIEN

DEPARTMENT OF GEODESY
AND GEOINFORMATION
MICROWAVE REMOTE SENSING

DISSERTATION

Data-driven soil moisture estimations based on earth observation data and machine learning

Ausgeführt am Department für Geodäsie und Geoinformation,
Forschungsgruppe Mikrowellenfernerkundung,
zum Zwecke der Erlangung des akademischen Grades eines

Doktors der Technischen Wissenschaften (Dr.Techn.)

unter der Leitung von
Univ.Prof. Dipl.-Ing. Dr.Techn. Wolfgang Wagner

eingereicht an der Technischen Universität Wien,
Fakultät für Geoinformation und Mathematik,
von

Felix Greifeneder

Matrikelnummer: 0626383

Bozen, November 2021 ……………………………

TU WIEN

DEPARTMENT OF GEODESY
AND GEOINFORMATION
MICROWAVE REMOTE SENSING

DISSERTATION

Data-driven soil moisture estimations based on earth observation data and machine learning

A thesis submitted at the Department of Geodesy and Geoinformation,
Research Group Microwave Remote Sensing,
in fulfilment of the academic degree of

Doctor of Engineering Sciences (Dr.Techn.)

Under the supervision of
Univ.Prof. Dipl.-Ing. Dr.Techn. Wolfgang Wagner

Research conducted at Technische Universität Wien,
Faculty of Mathematics and Geoinformation,
by

Felix Greifeneder

Matriculation number: 0626383

Bolzano, November 2021 ……………………………………

Supervisor: Prof. Dr. Wolfgang Wagner

TU Wien, Department for Geodesy and Geoinformation

Wiedner Hauptstraße 8, 1040 Vienna, Austria

Email: Wolfgang.wagner@geo.tuwien.ac.at ………………………………

Reviewer: Prof. Dr. Wouter Dorigo

TU Wien, Department for Geodesy and Geoinformation

Wiedner Hauptstraße 8, 1040 Vienna, Austria

Email: Wouter.Dorigo@geo.tuwien.ac.at ………………………………

Reviewer: Dr. Mehrez Zribi

Directeur du Centre d'Etudes Spatiales de la Biosphère (CESBIO)

18, Avenue Edouard Belin, 31401, cedex 9, Toulouse, France

Email: mehrez.zribi@cesbio.cnes.fr ………………………………

Erklärung zur Verfassung der Arbeit / Author's Statement

Hiermit versichere ich, dass die vorliegende Arbeit von mir selbständig angefertigt wurde. Inhalte und Passagen, die aus fremdem Quellen stammen und direkt oder indirekt übernommen worden sind, wurden als solche kenntlich gemacht. Ferner versichere ich, dass ich keine andere, außer der im Literaturverzeichnis angegebenen Literatur verwendet habe. Diese Erklärung bezieht sich sowohl auf Textinhalte sowie alle enthaltenen Abbildungen, Karten und Tabellen.

With this statement, I declare that this thesis was written by me independently. Contents and passages originating from external sources that have been taken over directly or indirectly have been marked as such. Furthermore, I assure that I have not used any literature other than that listed in the bibliography. These statements apply both to the text content and all illustrations, maps, and tables contained therein.

Felix Greifeneder

....

(Unterschrift/Signature)

CONTENTS

Acknowledgement

Throughout the years that I have spent researching and working on my PhD thesis, I have received support from different sides, institutional, academic, and personal. Help without which, I would not have been able to conclude my thesis. The following paragraphs contain an unordered list of some of those people and institutions.

My gratitude goes to Eurac Research to allow me to have a secure position in research over these years and providing the environment to develop personally and professionally. Furthermore, I want to thank the other funders who funded my research and the projects that contributed to this thesis. This includes the province of Bolzano/Bozen, the European Union's Horizon 2020 research and innovation program, the European Fund for Regional Development, Eumetsat, and the European Space Agency.

Without a doubt, I could never have achieved my goals without sound academic support. In this sense, I would like to thank my supervisor Prof. Wolfgang Wagner, first of all, for welcoming and introducing me to the world of science, and for his guidance, his help in finding the red thread and his critical feedback. I also thank Dr. Claudia Notarnicola for allowing me to work for her and her essential support for this thesis, for the many things she taught me and for being a role model scientist. Not to forget her support in my day to day work as a researcher.

I would also like to thank all my colleagues at the Eurac Research Institutes for Earth Observation and Alpine Environment for the many inspiring discussions and conversations and those from different institutions with whom I often collaborated closely through research projects.

Finally, I want to thank my friends and family who provided support, stimulating discussions, and most importantly, the happy distractions to rest my mind outside of my research.

Thesis Abstract

Although its origins lie in military applications, satellite remote sensing was established over the last decades as an essential method for observing our environment through the spatially continuous and global measurement of relevant parameters like land-cover, land-surface temperature, vegetation biomass, ocean salinity, or surface soil moisture. Furthermore, it plays an essential role in observing the atmosphere, mapping land-cover and land-use changes, generating digital elevation models, and many other applications.

This thesis focuses on estimating the surface soil moisture content, which was recognised as an essential climate variable by the Global Climate Observing System of the World Meteorological Organization. It is essential for the understanding of many meteorological and hydrological processes. Spatial and temporal changes of the soil moisture content can help understand and anticipate natural hazards like landslides, floods, or drought. The remote sensing of the soil moisture content, using optical and microwave sensors, has a long history dating back to the 1970s. Different approaches have emerged and established themselves since then. With this thesis, we concentrated on the latest group of approaches, machine learning.

Even though most of the underlying methodologies were already developed during the 1980s and 1990s, machine learning experienced a surge of popularity during the last decade, also for remote sensing and earth observation applications. This increase in popularity was further pushed by the paradigm shift in earth observation, which allows users today to easily access and exploit large quantities of data from different sensors.

The overall goal of this thesis is to use earth observation data combined with machine learning methods to estimate the soil moisture content. To capture the aims of this thesis, we formulated three main questions: *How can we go from site-specific, data-driven machine-learning models to general applicability in large scale applications?; How can we harness the potential of available data and merge data from different sensors and data sources with different spatial and temporal resolutions?; How can we link soil moisture measurements across scales (spatial and temporal)?*

The analysis presented by the thesis in its first part focused on a better understanding of the interactions between microwaves, soil moisture, topography, land-cover, and vegetation, between each other and across spatial scales. By developing a spatial upscaling method for in-situ measurements, we were able to study these interactions and confirm the solid temporal correlation of soil moisture across spatial scales. An essential ambition of this work was also to study the applicability of data-driven methods on a global scale. For this purpose, we performed tests based on different spatial resolutions and used different reference data types. The results demonstrate that accurate estimation is possible, with coarse as well as with high spatial resolutions. The studies also revealed certain limitations related to the potential of retrieval

models relying only on satellite data, the uncertainties of heterogenous reference data, or the validation of high-resolution spatial patterns.

One of the thesis' main outputs is an approach and a model for the high-resolution mapping of surface soil moisture, which we published as part of a software called PYSMM. The practical use of the approach we demonstrated as part of the thesis for mapping soil moisture anomalies. Its relevance was further underlined as it was picked up by scientists at FAO and the USGS to incorporate soil moisture information for wetland detection and assimilation in a hydrological model, respectively.

Kurzfassung

Obwohl ihre Ursprünge im militärischen Bereich liegen, hat sich die Satellitenfernerkundung in den letzten Jahrzehnten als wesentliche Methode zur Beobachtung unserer Umwelt erwiesen, da sie relevante Parameter wie Landbedeckung, Oberflächentemperatur, Biomasse der Vegetation, Salzgehalt der Ozeane oder Bodenfeuchtigkeit räumlich kontinuierlich und global misst. Darüber hinaus spielt sie eine wesentliche Rolle bei der Beobachtung der Atmosphäre, der Kartierung von Landbedeckung und Landnutzungsänderungen, der Erstellung digitaler Höhenmodelle und vielen anderen Anwendungen.

In dieser Arbeit geht es um die Schätzung der Bodenfeuchte, die vom Globalen Klimabeobachtungssystem der Weltorganisation für Meteorologie als eine wesentliche Klimavariable anerkannt wurde. Sie ist für das Verständnis vieler meteorologischer und hydrologischer Prozesse unerlässlich. Räumliche und zeitliche Veränderungen des Bodenfeuchtegehalts können helfen, Naturgefahren wie Erdrutsche, Überschwemmungen oder Dürren zu verstehen und vorherzusehen. Die Fernerkundung des Bodenfeuchtigkeitsgehalts mit optischen und Mikrowellensensoren hat eine lange Geschichte, die bis in die 1970er Jahre zurückreicht. Seitdem haben sich verschiedene Ansätze herausgebildet und etabliert. In dieser Arbeit haben wir uns auf die jüngste Gruppe von Ansätzen, das maschinelle Lernen, konzentriert.

Obwohl die meisten der zugrundeliegenden Methoden bereits in den 1980er und 1990er Jahren entwickelt wurden, erlebte das maschinelle Lernen im letzten Jahrzehnt einen Popularitätsschub, auch für Fernerkundungs- und Erdbeobachtungsanwendungen. Dieser Popularitätsanstieg wurde durch den Paradigmenwechsel in der Erdbeobachtung weiter vorangetrieben, der es den Nutzern heute ermöglicht, problemlos auf große Datenmengen von verschiedenen Sensoren zuzugreifen und diese zu nutzen.

Das übergeordnete Ziel dieser Arbeit ist die Nutzung von Erdbeobachtungsdaten in Kombination mit Methoden des maschinellen Lernens zur Schätzung des Bodenfeuchtegehalts. Um die Absichten dieser Arbeit zu erfassen, haben wir drei Hauptfragen formuliert: Wie können wir von standortspezifischen, datengesteuerten Machine-Learning-Modellen zu einer allgemeinen Anwendbarkeit in großflächigen Anwendungen übergehen?; Wie können wir das Potenzial der verfügbaren Daten nutzen und Daten von verschiedenen Sensoren und Datenquellen mit unterschiedlichen räumlichen und zeitlichen Auflösungen zusammenführen?; Wie können wir Bodenfeuchtemessungen über Skalen hinweg (räumlich und zeitlich) miteinander verknüpfen?

Die im ersten Teil der Arbeit vorgestellte Analyse konzentrierte sich auf ein besseres Verständnis der Wechselwirkungen zwischen Mikrowellen, Bodenfeuchte, Topographie, Landbedeckung und Vegetation, untereinander und über räumliche Skalen hinweg. Durch die Entwicklung einer räumlichen Skalierungsmethode für In-situ-Messungen waren wir in der Lage, diese Wechselwirkungen zu untersuchen und die starke zeitliche Korrelation der Bodenfeuchte über

räumliche Skalen hinweg zu bestätigen. Ein wesentliches Ziel dieser Arbeit war auch die Untersuchung der Anwendbarkeit datengesteuerter Methoden auf globaler Ebene. Zu diesem Zweck haben wir Tests mit unterschiedlichen räumlichen Auflösungen durchgeführt und verschiedene Referenzdatentypen verwendet. Die Ergebnisse zeigen, dass eine genaue Schätzung sowohl bei grober als auch bei hoher räumlicher Auflösung möglich ist. Die Studien zeigten auch bestimmte Einschränkungen in Bezug auf das Potenzial von Modellen, die sich nur auf Satellitendaten stützen, die Unsicherheiten heterogener Referenzdaten oder die Validierung hochaufgelöster räumlicher Muster.

Eines der wichtigsten Ergebnisse dieser Arbeit ist ein Ansatz und ein Modell für die hochauflösende Kartierung der Bodenfeuchte, welches wir als Teil einer Software namens PYSMM veröffentlicht haben. Die praktische Anwendung des Ansatzes haben wir im Rahmen der Arbeit zur Kartierung von Bodenfeuchteanomalien demonstriert. Seine Relevanz wurde weiter unterstrichen, da er von Wissenschaftlern der FAO und des USGS aufgegriffen wurde, um Informationen über die Bodenfeuchte zur Erkennung von Feuchtgebieten bzw. zur Assimilation in ein hydrologisches Modell einzubeziehen.

Relevant other publications and talks

This section contains publications and conference contributions, to which I contributed directly, which proved to be particularly important for shaping the dissertation.

- *Ali, Iftikhar, Felix Greifeneder, Jelena Stamenkovic, Maxim Neumann, and Claudia Notarnicola. "Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data." Remote Sensing 7, no. 12 (2015): 16398–421. https://doi.org/10.3390/rs71215841.*

- *Carvalho-Santos, Claudia, António Monteiro, Salvador Arenas-Castro, Felix Greifeneder, Bruno Marcos, Ana Portela, and João Honrado. "Ecosystem Services in a Protected Mountain Range of Portugal: Satellite-Based Products for State and Trend Analysis." Remote Sensing 10, no. 10 (Oktober 2018): 1573. https://doi.org/10.3390/rs10101573.*

- *Greifeneder, F, C Notarnicola, G Cuozzo, G Bertoldi, S Della Chiesa, G Niedrist, J Stamenkovic, and W Wagner. "Estimation of Surface Soil Moisture in Alpine Areas Based on Medium Spatial Resolution SAR Time-Series and Upscaled in-Situ Measurements." In Proceeding of SPIE Vol. 9243, 9243:92430G-92430G – 10. SPIE, 2014. http://dx.doi.org/10.1117/12.2071252.*

- *Greifeneder, Felix, Claudia Notarnicola, and Wolfgang Wagner. "Using Machine Learning and SAR Data for the Upscaling of Large Scale Modelled Soil Moisture in the Alps." In 11th European Conference on Synthetic Aperture Radar, EUSAR 2016, 1108–11. Hamburg: VDE, 2016.*

- *Gruber, Alexander, Wolfgang Wagner, Alena Hegyiova, Felix Greifeneder, and Stefan Schlaffer. "Potential of Sentinel-1 for High-Resolution Soil Moisture Monitoring." In 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS, 4030–33. IEEE, 2013. https://doi.org/10.1109/IGARSS.2013.6723717.*

- *Lei, Fangni, Wade T. Crow, William P. Kustas, Jianzhi Dong, Yun Yang, Kyle R. Knipper, Martha C. Anderson, et al. "Data Assimilation of High-Resolution Thermal and Radar Remote Sensing Retrievals for Soil Moisture Monitoring in a Drip-Irrigated Vineyard." Remote Sensing of Environment 239 (March 2020): 111622. https://doi.org/10.1016/j.rse.2019.111622.*

- *Notarnicola, C, L Pasolli, G Cuozzo, F Greifeneder, G Bertoldi, S Della Chiesa, G Niedrist, et al. "Temporal and Spatial Soil Moisture Dynamics in Mountain Meadows by Integrating Radarsat 2 Images and Ground Data." In 2014 IEEE Geoscience and Remote Sensing Symposium, 1222–25, 2014. https://doi.org/10.1109/IGARSS.2014.6946652.*

- *Pasolli, Luca, Claudia Notarnicola, Giacomo Bertoldi, Lorenzo Bruzzone, Ruben Remelgado, Felix Greifeneder, Georg Niedrist, Stefano Della Chiesa, Ulrike Tappeiner, and Marc Zebisch. "Estimation of Soil Moisture in Mountain Areas Using SVR Technique*

Applied to Multiscale Active Radar Images at C-Band." Selected Topics in Applied Earth Observations and Remote Sensing 8, no. 1 (2015): 262–83. https://doi.org/JSTARS.2014.2378795.

- *Stamenkovic, Jelena, Leila Guerriero, Paolo Ferrazzoli, Claudia Notarnicola, Felix Greifeneder, and Jean Philippe Thiran. "Soil Moisture Estimation by SAR in Alpine Fields Using Gaussian Process Regressor Trained by Model Simulations." IEEE Transactions on Geoscience and Remote Sensing 55, no. 9 (2017): 4899–4912. https://doi.org/10.1109/TGRS.2017.2687421.*

List of Figures

List of Acronyms

ANN	Artificial Neural Network
API	Application Programming Interface
ASAR	Advanced Synthetic Aperture Radar
ASCAT	Advanced Scatterometer
CSSM	Copernicus Surface Soil Moisture
DEM	Digital Elevation Model
ECMWF	European Centre for Medium-Range Weather Forecasts
ESA	European Space Agency
ECV	Essential Climate Variable
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
EVI	Enhanced Vegetation Index
GBRT	Gradient Boosted Regression Trees
GEE	Google Earth Engine
GLDAS	Global Land Data Assimilation System
IEM	Integral Equation Method
ISMN	International Soil Moisture Network
JAXA	Japan Aerospace Exploration Agency
KGE	Kling-Gupta Efficiency
LOGO-CV	Leaf-One-Group-Out-Cross-Validation
LST	land-surface-temperature
LTER	Long Term Ecological Research
MAE	Mean Absolute Error
Metop	Meteorological Operational satellite
Metop-SG	Meteorological Operational Satellite-Second Generation
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NN	Nearest Neighbour
OLM	OpenLandMap
PR	Polarisation Ratio
PYSMM	PYthon Sentinel-1 Soil-Moisture Mapping Toolbox
R	Pearson Correlation Coefficient
RF	Random Forest
RFI	Radio Frequency Interference
RMSE	Root Mean Square Error
SAR	Synthetic Aperture Radar
SCA	Scatterometer
SMAP	Soil Moisture Active-Passive
SMC	Soil Moisture Content
SMOS	Soil Moisture and Ocean Salinity
SMR	Soil Moisture Retrieval
SRTM	The Shuttle Radar Topography Mission

SVR	Support Vector Regression
TUW	TU-Wien
USCRN	US Climate Reference Network
USGS	United States Geological Survey
VWC	Vegetation Water Cloud

I. INTRODUCTION

Per definition, remote sensing is the acquisition of information about an object or phenomenon without making physical contact with the object, in contrast to in-situ observations. Today, the term usually refers to observations made from space- or airborne platforms like satellites or aeroplanes. During the Cold War era, military and espionage applications triggered the development of the first remote sensing satellites and cameras, by the Soviet Union and the United States of America, during the late 1950s and the beginning of the 1960s. Military applications in the following years mainly drove satellite remote sensing and new sensors and technologies. However, soon the potential of satellite observations for civilian purposes was recognised and created a new scientific discipline and a commercial sector to map, monitor, and explore our planet. Scientific and commercial earth observation (EO) applications became a technology driver. Today, a large and growing number of EO systems provide data for countless applications. Two for the scientific community essential programs should be highlighted: the Landsat[1] program by the United Stated Geological Survey (USGS) and the National Aeronautics and Space Administration (NASA) and the Copernicus[2] program by the European Union. The former is noteworthy because it was one of the first dedicated EO missions, and it is still running now with its eighth generation of satellites since the early 1970s. The Copernicus program bundles various environmental datasets – from satellite missions to climate models and in-situ measurements. It not only pushed the boundaries technologically (for example, with its Sentinel satellite missions) but also in terms of data policies and the concept of open data. These two EO programs are essential also for this thesis.

The work presented here focuses on the measurement and mapping the soil moisture content (SMC) based on satellite remote sensing. SMC is a crucial state variable in the complex global water and energy exchange between the land surface and the atmosphere through evaporation and plant transpiration (Bras, 2015). It plays an essential role in understanding critical hydrological and meteorological processes.

I.1 Remote sensing of the soil moisture content

While advanced hydrological models allow deriving detailed information about the spatial (horizontally and vertically) and temporal soil moisture patterns (Silva Ursulino *et al.*, 2019; Andini, Kim and Chun, 2020), they are very complex to set up and require detailed information about the composition of the land surface and meteorological forcings. EO and remote sensing (be it satellite- or aeroplane-borne) provide the only existing technology for the continuous measurement of the SMC over larger areas. The SMC affects the amount of energy emitted or reflected for the Earth's surface and, therefore, directly influences the images captured by different EO sensors. However, because water is not the only contributing parameter, some

[1] https://landsat.gsfc.nasa.gov/

[2] https://www.copernicus.eu/en

modelling is usually required to derive SMC maps from satellite images. We can differentiate the approaches based on the sensor technology into microwave-based and optical-based methods.

The main advantage of the microwave spectrum is that radiation can pass through the atmosphere with little distortion; it can pass through clouds and is independent of the sun as a source of illumination. All active and passive microwave-based retrieval algorithms rely on the dependency between the dielectric properties (e.g. of soil) and their effect on the emitted or reflected energy. The dielectric properties are related to the relative permittivity, which determines the amount of energy lost when microwave radiation travels through a medium—an effect, also known as attenuation. The real part of the dielectric constant (ε') quantifies these properties. For air, ε' equals 1, approximately 4 for dry soil and 80 for water (these values vary depending on the microwave frequency). Besides the SMC, many other factors like vegetation water content, vegetation structure, surface roughness or surface temperature finally determine the total emitted or reflected energy (Ulaby and Batlivala 1978). The difference between active and passive microwave systems is that an active system sends and receives a microwave pulse, whereas passive systems rely on measuring the natural thermal emissions of the Earth's surface. Because they are more relevant for this thesis, we will look at active microwave-based SMC retrieval approaches. We can distinguish three main types:

1. Based on the inversion of a backscattering model, which describes the microwave scattering mechanisms in vegetation and soil. This includes physical models like the Water Cloud Model (Attema and Ulaby 1978) or the Integral Equation Model (Fung, Li, and Chen 1992), and a group of semi-empirical models like the so-called Oh model (Oh, Sarabandi, and Ula 1992) or the Dubois model (Dubois, Zyl, and Engman 1995), usually exploiting ground-based active microwave measurements to characterise some of the scattering mechanisms.
2. Empirical change detection approaches are usually based on temporal normalisation of the backscatter intensities and rely on the fact that the parameters contributing to the measured signal have different temporal change signatures. The approach by Wagner, Lemoine and Rott (1999) or the Normalized Backscatter Moisture Index (Shoshany *et al.*, 2000) are examples in this category.
3. Many machine learning applications for SMC estimation have been presented in recent years (Ali *et al.*, 2015). Section I.3 introduces this group of approaches in detail.

In the optical domain, many different methods to derive the SMC exist. They rely on the relationship between SMC and surface reflectance, vegetation indices or surface temperature (Zhang and Zhou, 2016). The advantage of optical systems, as opposed to microwave systems, is the superior data availability with many existing satellites with optical sensors aboard, covering an extensive range of spatial and spectral resolutions and with long-standing open archives from missions like Landsat, ASTER, or MODIS. Optical retrieval approaches can be coarsely classified based on the used frequency. Some methods rely on visible or near-infrared data. The negative correlation between bare soil's surface reflectance (in some water absorption bands) and the SMC was already described by Ångström (1925). Due to this phenomenon, simple indices like the (Normalized Difference Vegetation Index) NDVI or the normalised-difference-water-index can serve as proxies for the SMC. Land-surface temperature (LST), derived from thermal remote sensing images, is a good proxy for SMC and vegetation water stress (Schmugge, 1978; Chang *et al.*, 2012; Qin *et al.*, 2013). To disentangle contributions by SMC and vegetation, the most commonly used approaches today rely on a combination of LST and vegetation indices. A popular method, which relies on this combination, is the so-called triangle model (Price, 1990). Its name relates to the triangular representation of the LST-vegetation index feature space.

Currently available operational SMC products rely on data from coarse- to medium resolution passive or active microwave sensors like Soil Moisture and Ocean Salinity (SMOS) (Berger *et al.*, 2002) (passive), Soil Moisture Active-Passive (SMAP) (Entekhabi *et al.*, 2010) (passive), or Advance Scatterometer (ASCAT) (Naeimi *et al.*, 2009) (active). As a component of the Copernicus Land Monitoring Service, the Soil Water Index (Bauer-Marschallinger *et al.*, 2018), based on a fusion of Sentinel-1 (S1) and ASCAT data, offers medium resolution (1 km) SMC observations. The advantage of these coarse- to medium-resolution sensors is their high temporal resolution, offering daily observations. With S1, Sentinel-2, and Landsat-8 high resolution (<50 m) remote-sensing data, in the microwave as well as the shortwave and thermal domains are available operationally and on an open-access basis, providing the most crucial foundation for high-resolution SMC mapping (Hornacek *et al.*, 2012; Bauer-Marschallinger *et al.*, 2019).

Due to the long history of the remote-sensing of SMC, many advanced applications emerged over time. For example, in a case study located in Italy, Brocca et al. (2016) integrated ASCAT SMC maps with precipitation and temperature measurements into a proposed operational landslide early warning system. Chaparro, Piles, and Vall-llossera (2016) used data from SMOS and SMAP in a wildfire prevention service. Remotely sensed SMC plays an essential role in improving model estimations through data assimilation (Reichle *et al.*, 2016), and they can even act as a virtual rain gauge to estimate precipitation (Brocca *et al.*, 2014).

I.2 EO Paradigm shift

The past decade has brought considerable changes to the world of satellite remote sensing and EO. It saw the launch of several new satellites and the start of crucial new EO missions. The list below only names a few of the most important ones.

- Landsat-8[3] (USGS), launched in 2013.

---

[3] https://www.usgs.gov/core-science-systems/nli/landsat/landsat-8?qt-science_support_page_related_con=0#qt-science_support_page_related_con

- The Sentinel[4] program (European Commission) with Sentinel-1 A/B launched in 2014 and 2016, Sentinel-2 A/B in 2015 and 2017, Sentinel-3 A/B in 2016 and 2018, and Sentinel-5P in 2017.
- NASA launched the Soil Moisture Active Passive[5] satellite in 2015.
- The Japan Aerospace Exploration Agency (JAXA) launched the ALOS-2[6] (JAXA) in 2014.
- SAOCOM 1 A/B[7] (CONAE) launched in 2018 and 2020.

Beyond the publicly funded missions above, carried out by the national and international space agencies, several commercial companies have launched their satellites and provide imagery and services for several sectors like agriculture or insurances. One example is Planet (founded in 2009), which offers high spatial and temporal resolution global monitoring. Among their fleet of satellites are more than 180 micro-satellites.

Many publicly funded programs now follow an open data policy, meaning that the data are usually free to use for research purposes or grant access to scientists upon request. Even commercial data providers offer some data access opportunities.

These new EO missions that emerged during the last decade illustrate the explosion of data volume, and new missions are being planned and launched in the years to come. Besides the number of new satellites, other reasons for the growing amount of available data are that more and more missions today are operating on a fixed operational schedule, as opposed to more experimental satellites of the past, providing shorter revisit times and higher temporal resolution, and sensors are recording larger areas at once, have higher spatial resolutions, and better sensitivity. The term Big Earth Data emerged to describe this new reality. For example, let us compare the volume of data produced by the European Space Agencies (ESA) Advanced Synthetic Aperture Radar (ASAR) with the Copernicus satellite Sentinel-1. The entire archive of data captured by ASAR in Global-Monitoring-Mode (1 km) has a volume of 1.74 Terra-Byte (TB) and 23.49 TB in Wide-Swath-Mode (150 m). One of the two Sentinel-1 satellites, in one of its acquisition modes (Ground-Range-Detected-High with approximately 20 m spatial resolution), produces 156 TB of data per year. Therefore, in about ten years of operation (2002-2012), ASAR produced less data than one Sentinel-1 satellite in a single year. According to Esch et al. (2018), the amount of data produced in a single day by Sentinel-1, 2, and 3 combined is approximately 20 TB.

Some years ago, the typical EO data workflow looked like this:

[4] https://www.copernicus.eu/en/about-copernicus/infrastructure/discover-our-satellites
[5] https://smap.jpl.nasa.gov/
[6] https://www.eorc.jaxa.jp/ALOS-2/en/about/palsar2.htm
[7] https://saocom.veng.com.ar/en/

1. Identify area and date of interest.
2. Investigate available data.
3. Download data for the required dates and area of interest.
4. Process data locally.
5. Retrieve information.

Due to the reasons discussed above, this approach may often be unfeasible, especially for global analysis. To fully exploit the potential of available data, a paradigm shift in the way they are processed and analysed was inevitable (Sudmanns *et al.*, 2020). Platforms like the Google Earth Engine (GEE) (Gorelick *et al.*, 2017), the Copernicus DIAS[8], the Sentinel-Hub[9], or the currently under development openEO platform[10] provide access to cloud-based processing resources and analysis-ready EO data. These allow a new workflow where the data storage and processing are centralised on a computer cluster. Combined with the availability of pre-processed, "analysis-ready" data on these platforms, this workflow offers another key advantage, besides the reduced requirements for the local processing infrastructure: the pre-processing of large quantities in the past required a high level of specific expertise, with cloud-based platforms like GEE, the barrier for a data analyst to access these data was substantially lowered.

I.3 Machine learning

Machine learning is a subfield of artificial intelligence and is a summary term for various computer algorithms that can learn by example and make predictions. In contrast to other classical modelling approaches, it is entirely empirical, which means that no assumptions or a priori knowledge regarding the relationship between the dependent variable(s) and the predictors are needed. Therefore, machine learning is beneficial for problems where we lack a complete theoretical understanding of the problem or have insufficient data to support the theory. Machine learning can be used for classification (i.e. the classification of data into different categories), for regression (i.e. based on continuous data used to predict unknown data points), or for unsupervised classification or clustering (i.e. the automatic recognition of natural groupings or patterns in data) based on data in different formats, like images, text, sensor measurements, video, or EO data. The history of machine learning goes back to the early 1950s where IBM scientists performed initial experiments, writing a computer program to play the game "Checkers". The term machine learning itself was coined by Arthur Samuel[11], who was a computer scientist at IBM. One of the early key innovations was the introduction of the perceptron by Rosenblatt (1957), a simple algorithm that can be applied to binary classification

[8] https://www.copernicus.eu/sites/default/files/Copernicus_DIAS_Factsheet_June2018.pdf
[9] https://www.sentinel-hub.com/
[10] https://openeo.org/platform/#about
[11] https://en.wikipedia.org/wiki/Arthur_Samuel

problems. It is one of the main concepts behind Artificial Neural Networks (ANN). After the initial excitement in the 1960s and 1970s, the interest began to fade. Other advances in computer-since and the improving performance and availability of PCs led to a resurgence of machine learning in the 1990s when several important algorithms were introduced: In 1995, Tin Kam Ho proposed the first Random Forest (RF) algorithm (Tin Kam Ho, 1995); the Support Vector Machine (SVM) was described first by Cortes and Vapnik (1995); since the early 2010s, Convolutional neural networks (Krizhevsky, Sutskever and Hinton, 2012) and the concept of deep learning caused another significant surge of interest in the topic and the development of new applications in many scientific fields and commercial sectors. These developments were made possible mainly due to the computing power of modern supercomputers and the availability of training datasets like the Imagenet database (Deng *et al.*, 2009).

Fuelled by the developments (see section I.2) of the last 10 to 15 years, machine learning has gained significant popularity also for EO applications. The strong positive trend in the number of publications (based on a Scopus[12] search) containing the keywords "machine learning" and "remote sensing" confirms this growing popularity (Figure I-1). The applications range from land-cover classification to glacier monitoring, from oil-spill mapping to sear surface salinity and soil moisture estimation. In Lary et al. (2018), the authors present a comprehensive overview of machine learning applications in EO. Here we will describe some of them, to give an impression of their large variety.

A characteristic example of a problem with an incomplete theoretical understanding is bias correction (for example, the bias between measurements from different sensors or model bias), which belongs to the group of regression problems. As an EO-related case, Brown et al. (2008) used machine learning to correct MODIS and AVHRR NDVI time-series biases. The proposed approach relies on a neural network to "predict" the MODIS NDVI based on AVHRR data. In this

[12] https://www.scopus.com/

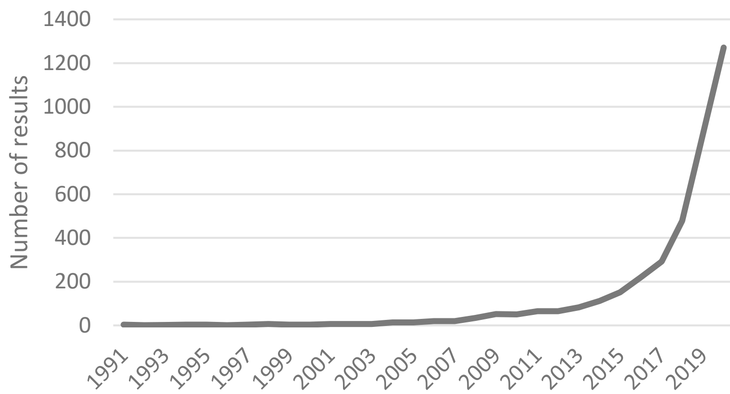

Figure I-1: Numbers of publications in scientific journals, including the terms "machine learning" and "remote sensing" in their keywords.

way, the authors established a long-term data record to serve climatological studies. As an example for unsupervised classification, we can look at Lary et al. (2016), where a clustering algorithm helps to identify dust sources in hype-spectral imagery. One of the most common and mature applications for machine learning in EO is land-use/land-cover (LULC) classification (Talukdar *et al.*, 2020). Numerous scientific studies cover this topic. Adam et al. (2014) demonstrated the potential of machine learning for high accuracy high spatial resolution mapping. They used SVM and RF classifiers to map LULC classes in RapidEye multispectral imagery. Even the operational land-cover monitoring service (Buchhorn, Smets, *et al.*, 2020) of the Copernicus program relies on machine learning (Buchhorn, Lesiv, *et al.*, 2020). This service provides annually updated global LULC maps with a spatial resolution of 100 m and relies on an RF algorithm.

Relatively recently, related to the paradigm shift discussed in section I.2, some applications of deep learning algorithms have found their way into the world of EO (Zhu *et al.*, 2017). Especially for the analysis of hyperspectral imagery (characterised by hundreds of narrow spectral bands), these approaches showed promising results.

In a review of machine learning applications for estimating biomass and soil moisture, Ali et al. (2015) discuss the advantages of machine learning for retrieving biophysical parameters. The theoretical models for these variables are complex and require detailed auxiliary data or assumptions about the present land-cover structure, often challenging to collect, especially for mapping more extensive areas. Machine learning can help to overcome some of these problems; it has two main advantages: 1) it allows to build objective, data driven-models without the necessity of a priori assumptions about the variable dependencies, and it, therefore, is an objective method to test a hypothesis; 2) it is an easy solution for the combination or fusion of various datasets (e.g. optical-, thermal-, and microwave remote sensing, in-situ measurements, and model outputs). Machine learning can fulfil different purposes in the process of parameter estimation. The most common applications are model inversion, empirical or data-driven model creation, and downscaling. Model inversion is necessary to retrieve parameters based on physics-based forward models like the radiative transfer model PROSPECT (describing the optical properties of vegetation canopies) or Integral Equation Method (IEM) (a theoretical radar backscatter model) estimate vegetation properties or the SMC from remote sensing imagery. More traditional approaches like look-up tables are associated with a high computational cost for their generation and parameter retrieval (Hedley, Roelfsema and Phinn, 2009). Machine learning offers a more efficient solution to the problem. To give an example, Moosavi et al. (2016) and Stamenkovic et al. (2017) used Gaussian Process Regression, ANN, and Support-Vector-Regression (SVR) to estimate high-resolution SMC based on ASAR and MODIS data, respectively. Another approach to generating parameter estimation models is empirical modelling, where a machine-learning algorithm uses in-situ measurements to "learn" to estimate the target (e.g. SMC) parameter based on several input features (e.g. satellite imagery combined with auxiliary data). This approach has been successfully applied many times, especially for the high-resolution mapping of SMC, and was subject to many studies (Paloscia *et al.*, 2013; Santi *et al.*, 2014; Pasolli *et al.*, 2015; Li *et al.*, 2020). Most of these studies, focusing on higher spatial resolution data, targeted specific study sites. Generally, these types of approaches are often thought to be site-specific. Some existing studies demonstrated the general applicability of empirical modelling using coarser spatial resolution data. Kolassa et al. (2018) and Greifeneder et al. (2018) tested different machine learning algorithms to derive globally applicable SMC retrieval models.

Besides parameter estimation, machine learning is often used for downscaling. One example was published by Moosavi et al. (2016), based on the downscaling of MODIS LST using Landsat imagery.

As mentioned above, machine learning is the ideal tool for combining data from different sources or data fusion. There is an enormous potential to improve parameter estimations based on a combination of sensors and sensing techniques. Optical, thermal, or microwave remote-sensing all have strengths and weaknesses that can often complement each other. The existing theoretical models often cannot integrate these data types (i.e. we lack theoretical understanding), a prime example of a machine learning application. The point was proven, for example, by Liu, Qian, and Yue (2021), who combined Sentinel-1 (Synthetic Aperture Radar - SAR) and Sentinel-2 (optical) data for the estimation of SMC. They demonstrated that several different machine learning algorithms reliably outperform theoretical modelling approaches by combining these two sensors. SMAP and MODIS data were combined by Bhuiyan et al. (2020) to improve precipitation estimations.

Some of these findings and advantages related to EO and machine learning apply to the machine-learning-based parameter estimation in general. Hence, also for other hydro-meteorological applications, machine learning has gained considerable popularity. Bhuiyan, Anagnostou, and

Kirstetter (2017), Tyralis et al. (2019), Derin et al. (2020) used quantile regression algorithms (quantile regression forests and quantile regression neural networks) to quantify model errors in a hydrological context. Applying similar methods for a different aim, Ahn and Palmer (2016) estimated flood probabilities for ungauged catchments.

I.4 Problem statement

In summary, and concerning the EO based estimation of SMC, we can capture the following issues:

- Physical and semi-empirical backscatter models provide a well-established and reliable base for SMC estimations. They have the advantage that they are well documented and that, due to their physical foundation, their inner workings can be understood and comprehended. However, due to the difficulty to obtain measurements of critical parameters like the surface roughness or detailed information about the LULC type and vegetation structure, they must rely on simplification of the actual scattering mechanisms, especially to map more extensive areas. In many cases, indices have to approximate specific parameters like vegetation status or surface roughness (like the use of a vegetation index for the Triangle-method (Price, 1990) or the Water-Cloud-Model (Attema and Ulaby, 1978)), which means that existing data sources cannot be exploited to their full potential. The complexity of combining data from different sources also hold for change detection approaches.
- Machine learning algorithms allow the establishment of entirely empirical, data-driven models. They are incredibly flexible, making it easy to combine data from different sources (e.g., satellite sensors, in-situ measurements, model outputs) and ordinal scales, which means that it is easier to exploit the available data and the information they contain fully. Together with the development in EO described in section I.2, these approaches constitute great potential. One downside is that machine learning is often considered a black box where it is impossible to understand how the input data contribute to the prediction and what the model learns. This problem is related to the limited generalisation capabilities – i.e. the ability of a model to make predictions outside the training domain, for example, for a different location. The model training requires a large dataset covering feature and target domains as completely as possible to overcome this problem.
- Models with validity for a larger target area require spatially distributed reference datasets that match the scale of target spatial resolution.
- The abundant availability of high-resolution satellite data holds great potential for the high-resolution mapping of SMC. However, it raises the bar for the quality of reference and auxiliary data even further.

- The available high-resolution satellite missions like Sentinel-1, Sentinel-2 or Landsat-8 offer time-series that are too short for some applications like climatological studies or the reliable detection of anomalies for applications like drought mapping and monitoring.

I.5 Objectives

The focus of this thesis is the exploitation of EO datasets, combined with machine learning methods for the estimation of SMC. Based on the problem statement summarised by section I.4, this section defines the three main objectives:

1. *How can we go from site-specific, data-driven machine-learning models to general applicability in large scale applications?*

As the introduction in section I.1 summarises, the estimation of SMC using different types of EO data has a long history (Attema and Ulaby, 1978; Schmugge, 1978; Price, 1990; Fung, Li and Chen, 1992) and is backed up by sound theoretical understanding. Also, machine learning algorithms to derive data-driven models have demonstrated their effectiveness in various studies (Paloscia *et al.*, 2013; Santi *et al.*, 2014; Pasolli *et al.*, 2015, p. 201; Li *et al.*, 2020). The work carried out in the scope of this thesis tries to push the applicability of these approaches further towards global-scale applications, tackling the problems stated in section 1.4. Ultimately, the thesis aims to address the gap that is a globally applicable model for the high-resolution mapping of SMC.

2. *Exploiting the paradigm shift:*

This objective is related to the limitations of the "classic" data processing approach (section I.2) and constitutes a practical necessity and an opportunity. The EO paradigm shift created new means to access analysis-ready data through different online platforms or Application-Programming-Interfaces (APIs), enabling a new level of data exploitation without direct access to ample processing resources. For us, this new potential is a prerequisite to reach objective one and tackle the issues related to generating a spatially distributed training dataset, especially when aiming at high spatial resolutions. The following question summarises these points:

   a. *How can we harness the potential of available data?*

Harnessing this potential also means combining data from different sources (e.g. in-situ, EO, modelling). These data come at different spatial and temporal sampling rates, which means that another related question emerges:

   b. *How can we merge data from different sensors and data sources with different spatial and temporal resolutions?*

Since it renders the necessity for heavy local data processing and sizeable local storage obsolete, the EO paradigm shift also increases the potential for replicability. Therefore, one of the outputs of this work shall be tools to replicate its results and make them accessible for research and practical application.

3. *How can we link soil moisture measurements across scales (spatial and temporal)?*

The third question is relevant from multiple perspectives. It must be considered when linking in-situ data and coarse resolution EO data, such as when we assemble a training dataset to feed a machine learning algorithm or when validating coarse-resolution SMC estimations. We have to consider multiple spatial soil moisture scales to combine SMC estimations from different sources, which may be one way to overcome one of the limitations of available high-resolution EO data, their short time series.

I.6 Study outline

The thesis consists of four main parts. Sections 1, 2 and 3 aim to provide context to the individual articles and integrate them into the overall thesis. This section, section 1, gives an introduction and a general overview of the topics tackled in the thesis, including a discussion of research questions and objectives. The following section, number 2, contains an overview of the four articles and discusses their connections. At its core is section number 3, with four articles published in scientific peer-reviewed journals and presenting this work's scientific contribution. At last, section 4 summarises the results and formulates the conclusions.

I.7 Summary of Publications

Section 2 provides a summary of the four articles, which constitute the main content of this thesis. The numbering, from one to four, is based on the logical flow of the presented work.

I.7.1 Article 1 – "From Point to Pixel Scale: An Upscaling Approach for In Situ Soil Moisture Measurements[13]"

The first article (Greifeneder *et al.*, 2016) was published in the Vadose Zone Journal in 2016. Its main aim was to introduce a method for upscaling in-situ SMC measurements, specifically, to provide reference data for the SMAP Cal/Val activities[14]. In the context of this thesis, it is important because it contributed to a better understanding of the direct influence of terrain, land-cover, or soil type on SMC patterns and the ability of available data to characterise these features. These findings contributed to the model design in article 3 (considering descriptive data is also crucial for overcoming site-specificity – objective 1), and they confirm the temporal

[13] https://doi.org/10.2136/vzj2015.03.0048
[14] https://smap.jpl.nasa.gov/science/validation/

correlation of SMC across different spatial scales. This phenomenon is known as temporal stability, introduced by Vachaud et al. (1985). It is one of the key concepts for anomaly detection in article 4. A more detailed summary of the work carried out in article 1 follows below.

Dealing with medium- to coarse-resolution satellite imagery (like the SMAP data in this study) requires compensating for different measurement scales, especially in mountain areas. We have developed a spatial upscaling method for SMC that combines in situ measurements and remotely sensed data. As mentioned before, the approach relies on correlating spatial patterns of SMC with terrain topography, land-cover, and soil type. The study used data from a small research site in the northern Italian Alps, in Val Mazia. To evaluate the method, we used resampled Envisat ASAR data to reproduce the spatial scale of the SMAP data. Results showed that the representativeness of in situ data could be improved significantly for the 3- by 3-km SMAP pixel scale. Applying the proposed approach improved the correlation (in terms of the Pearson correlation coefficient) between SMC and satellite backscatter from R = 0.05 to 0.28. Another way to test the improvements is to use upscaled in situ measurements to train a machine learning retrieval model, using original versus upscaled ground data for model training. The error of the estimated SMC was improved from Root-Mean-Square-Error (RMSE) = 0.12 to 0.03 m3 m−3.

I.7.2 Article 2 – "The Added Value of the VH/VV Polarization-Ratio for Global Soil Moisture Estimations From Scatterometer Data[15]"

This article (Greifeneder, Notarnicola, *et al.*, 2018) appeared in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. At its core was assessing the radar cross-polarisation channel's value for the compensation of vegetation effects in SMC retrieval models in anticipation of the upcoming second-generation Meteorological-Operation-Satellite (Metop) Scatterometer. The study compared three different machine learning approaches using reanalysis data as an SMC reference for training and testing. The relevance of the presented work concerning this thesis lies in the spatial scope. We demonstrated the applicability of data-driven models for global scale SMC retrievals and analysis (objective 1). Furthermore, the results show the value of the cross-polarisation channel for SMC estimations, and at the same time, they demonstrate the limits of estimations based only on radar data. The work carried out in article 2 is summarised in more detail below.

The successor to the current series of Metop ASCAT, the Metop-SG (second generation) SCA (the launch of Metop-SG-B, which will carry the SCA instrument, is currently scheduled for 2025), will record dual-polarisation, medium spatial resolution, high temporal resolution data at C-band. Taking the current algorithm for the Hydrology Satellite Application Facility (HSAF) medium

[15] https://doi.org/10.1109/JSTARS.2018.2865185

resolution SMC product[16], which is based on single-polarisation ASCAT data, as a starting point, we investigated whether the information contained in the cross-polarisation measurements could improve the vegetation parameterisation to estimate the SMC. The operational HSAF algorithm characterises vegetation dynamics by the relationship between radar backscattering intensity and the incidence angle, the so-called SLOPE parameter. Based on findings from previous studies, we assumed that the polarisation ratio, i.e., VH/VV, could improve this characterisation. Cross-polarisation data from the Scatterometer of NASA's Aquarius mission[17] served as a simulation of an additional ASCAT channel. Machine learning offered the means for an objective assessment of the hypothesis. We employed three different algorithms to avoid biased findings (SVR, ANN and the Bayesian-Regression) and compared four feature configurations: no vegetation compensation, compensation based on SLOPE, compensation based on the polarisation ratio, a combination of SLOPE and the polarisation ratio. The results showed that the information contained in the SLOPE parameter and the polarisation ratio is similar, which confirmed that the cross-polarisation channel is sensitive to changes in vegetation. Based on a global average, the different approaches achieved comparable accuracies. Despite that, analysis of the temporal dynamics of SLOPE and polarisation ratio revealed specific location-specific differences, which affect the spatial distribution of SMC retrieval accuracies. As a result, improvements based on the combination of the two parameters are minor overall, but they can be significant locally.

I.7.3 Article 3 – "A Machine Learning-Based Approach for Surface Soil Moisture Estimations with Google Earth Engine[18]"

Publication number three (Greifeneder, Notarnicola and Wagner, 2021a) appeared in *Remote Sensing* in 2021. This article is the essential publication of the thesis as it builds directly upon the findings of the previous articles, and it touches on all of the objectives discussed in section I.5. The work's main aim was to introduce a system for estimating high spatial resolution SMC applicable independently of the geographic location (objective 1). To exploit a wide range of available satellite and model data, a move into an environment that offers analysis-ready data was a natural choice. In this case, the whole process is based entirely on the GEE (objective 2). Data from the International Soil Moisture Network (ISMN) served as a reference. To achieve these aims, we had to overcome the problem of combining data with different spatial and temporal sampling rates. (objective 3). With this article, we demonstrated the enormous potential that lies in the EO paradigm shift – making the data more accessible – and machine learning methods. However, while the developments of the last decades solved two big problems – access to data and processing infrastructure – the main remaining bottleneck of all machine

[16] https://hsaf.meteoam.it/Products/ProductsList?type=soil_moisture
[17] https://www.nasa.gov/mission_pages/aquarius/overview/index.html
[18] https://doi.org/10.3390/rs13112099.

learning applications is the availability of accurate reference and training data. The following paragraph summarises in more detail the work carried out in the scope of article 3.

This study introduces a machine-learning-based approach for high spatial resolution (50 m) mapping of SMC, based on the integration of Landsat-8 optical and thermal images, Copernicus Sentinel-1 C-Band SAR images, and modelled data, executable in the GEE. The novelty of this approach lies in applying an entirely data-driven machine learning concept for global estimation of the surface soil moisture content. Globally distributed in situ data from the ISMN acted as an input for model training. Based on the independent validation dataset, the resulting overall estimation accuracy, in terms of RMSE and R2, was 0.04 m^3m^{-3} and 0.81, respectively. Beyond the retrieval model itself, this article introduces a framework for collecting training data and a stand-alone Python package for soil moisture mapping. The GEE API facilitates data collection and retrieval, which is entirely cloud-based. For soil moisture retrieval, it eliminates the requirement to download or pre-process any input datasets.

I.7.4 Article 4 – "Detection of Soil Moisture Anomalies Based on Sentinel-1[19]"

The last and fourth article (Greifeneder, Khamala, *et al.*, 2018) appeared in *Physics and Chemistry of the Earth* in 2018. It showcases a possible application of high-resolution SMC retrievals to map soil moisture anomalies. For the detection of, anomalies a time series long enough to determine the normal state must be available. High-resolution satellites like Sentinel-1 today cover less than ten years up to now. To extend the temporal coverage of the SMC time-series, we introduced a method to fuse high-resolution estimations with coarse-resolution Global Land Data Assimilation System (GLDAS) model data. In this way, the article contributes to objective number 3 of the thesis and closes the circle back to article one, as the fusion approach relies on the same concept of temporal stability. A more detailed summary of the work carried out follows in the paragraphs below.

One of the applications for SMC measurements is connected to the relationship between SMC anomalies and natural hazards such as droughts, flooding, or landslides. A requirement for detecting and quantifying an anomaly is a long time series (often 10–30 years) to derive a reference value. Herein lies one of the issues of Sentinel-1 based SMC mapping.

We introduced an approach to overcome this problem and enable the Sentinel-1 based SMC anomaly detection. The method built on a cross-calibration between Sentinel-1 SMC estimations and coarse resolution (~30 km) modelled SMC from the GLDAS, covering 1948 to today. As a result, we derived the long-term averages for each Sentinel-1 pixel.

[19] https://doi.org/10.1016/J.PCE.2018.11.009

Results showed that the proposed method allows very accurate reproduction of the average SMC (for any given pixel) – if computed for the Sentinel-1, the RMSE between estimated (GLDAS based) and true average Sentinel-1 SMC is 0.7 %-Vol. Furthermore, the comparison with an in-situ time series showed the correct detection of negative and positive anomalies, respectively. The method presented in article 4 may allow the integration of Sentinel-1 data into, e.g., drought monitoring or flood forecasting applications.

I.8 Author contributions

I.8.1 Article 1 – "From Point to Pixel Scale: An Upscaling Approach for In Situ Soil Moisture Measurements[20]"

Felix Greifeneder, Claudia Notarnicola, and Wolfgang Wagner conceived and designed the study. Felix Greifeneder developed the method and implemented the experiments with guiding comments by Georg Niedrist and Giacomo Bertoldi. Georg Niedrist and Giacomo Bertoldi provided In-measurements. Felix Greifeneder was responsible for the pre-processing of all satellite data. He wrote the first draft of the manuscript and produced all figures. All authors supported the revision and editing of the manuscript.

I.8.2 Article 2 – "The Added Value of the VH/VV Polarization-Ratio for Global Soil Moisture Estimations From Scatterometer Data[21]"

Felix Greifeneder, Claudia Notarnicola, and Wolfgang Wagner conceived and designed the study. Felix Greifeneder conducted the experiments with guiding comments from Sebastian Hahn, Mariette Vreugdenhil, Christoph Reimer, and Simonetta Paloscia. Sebastian Hahn was responsible for pre-processing the Aquarius L-Band and ASCAT data. Emanuele Santi and Simonetta Paloscia provided the soil moisture estimations based on Artificial Neural Networks and Claudia Notarnicola the software to apply Bayesian Regression. Felix Greifeneder wrote the first draft of the manuscript and produced all figures. All authors supported the revision and editing of the manuscript.

I.8.3 Article 3 – "A Machine Learning-Based Approach for Surface Soil Moisture Estimations with Google Earth Engine[22]"

Felix Greifeneder, Claudia Notarnicola, and Wolfgang Wagner conceived and designed the study. Felix Greifeneder developed the method and implemented the experiments with regular exchange between all co-authors. Felix Greifeneder wrote the first draft of the manuscript and

[20] https://doi.org/10.2136/vzj2015.03.0048
[21] https://doi.org/10.1109/JSTARS.2018.2865185
[22] https://doi.org/10.3390/rs13112099.

produced all figures. All authors contributed by editing and organizing the manuscript for the final publication.

I.8.4 Article 4 – "Detection of Soil Moisture Anomalies Based on Sentinel-1[23]"

Felix Greifeneder, Claudia Notarnicola, and Wolfgang Wagner conceived and designed the study. Felix Greifeneder developed the algorithm and carried out the experiments with guiding comments by Marc Zebisch. Erick Khamala and Degelo Sendabo were involved in the field campaigns and provided in-situ and other local datasets. Furthermore, they supported the study through their knowledge of the local climate, land-use, and topography. Felix Greifeneder wrote the first draft of the manuscript and produced all figures. All authors supported the revision and editing for the final manuscript.

[23] https://doi.org/10.1016/J.PCE.2018.11.009

II. RESEARCH ARTICLES

II.1 Article 1 – "From Point to Pixel Scale: An Upscaling Approach for In Situ Soil Moisture Measurements"

written by

Felix Greifeneder, Claudia Notarnicola, Giacomo Bertoldi, Georg Niedrist, and Wolfgang Wagner

Published in the Vadose Zone Journal.

DOI: 10.2136/vzj2015.03.0048

Cited as:
Greifeneder, Felix, Claudia Notarnicola, Giacomo Bertoldi, Georg Niedrist, and Wolfgang Wagner. "From Point to Pixel Scale: An Upscaling Approach for In Situ Soil Moisture Measurements." *Vadose Zone Journal* 15, no. 6 (2016).

This study was supported by the project "HiResAlp" financed by Provincia Autonoma di Bolzano, Alto Adige, Ripartizione Diritto allo Studio, Università e Ricerca Scientifica.

References

Ali, I., F. Greifeneder, J. Stamenkovic, M. Neumann, and C. Notarnicola. 2015. Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. Remote Sens. 7:16398– 16421. doi:10.3390/rs71215841

Barrett, B.W., E. Dwyer, and P. Whelan. 2009. Soil moisture retrieval from active spaceborne microwave observations: An evaluation of current techniques. Remote Sens. 1:210–242. doi:10.3390/rs1030210

Bell, K.R., B.J. Blanchard, M.W. Witczak, and T.J. Schmugge. 1980. Analysis of surface moisture variations within large field sites. Water Resour. Res. 16:796–810. doi:10.1029/WR016i004p00796

Bertoldi, G., S. Della Chiesa, C. Notarnicola, L. Pasolli, G. Niedrist, and U. Tappeiner. 2014. Estimation of soil moisture patterns in mountain grasslands by means of SAR RADARSAT2 images and hydrological modeling. J. Hydrol. 516: 245–257. doi:10.1016/j.jhydrol.2014.02.018

Bindlish, R., T. Jackson, M. Cosh, T. Zhao, and P. O'Neill. 2015. Global soil moisture from the Aquarius/SAC-D satellite: Description and initial assessment. IEEE Geosci. Remote Sens. Lett. 12:923–927. doi:10.1109/LGRS.2014.2364151

Brocca, L., A. Tarpanelli, T. Moramarco, F. Melone, S.M. Ratto, M. Cauduro, et al. 2013. Soil moisture estimation in alpine catchments through modeling and satellite observations. Vadose Zone J. 12(3). doi:10.2136/vzj2012.0102

Bruzzone, L., and F. Melgani. 2005. Robust multiple estimator systems for the analysis of biophysical parameters from remotely sensed data. IEEE Trans. Geosci. Remote Sens. 43:159–174. doi:10.1109/TGRS.2004.839818

Cosh, M.H., T.J. Jackson, R. Bindlish, and J.H. Prueger. 2004. Watershed scale temporal and spatial stability of soil moisture and its role in validating satellite estimates. Remote Sens. Environ. 92:427–435. doi:10.1016/j.rse.2004.02.016

Crow, W.T., D. Ryu, and J.S. Famiglietti. 2005. Upscaling of field-scale soil moisture measurements using distributed land surface modeling. Adv. Water Resour. 28:1–14. doi:10.1016/j.advwatres.2004.10.004

Della Chiesa, S., G. Bertoldi, G. Niedrist, N. Obojes, S. Endrizzi, J.D. Albertson, et al. 2014. Modelling changes in grassland hydrological cycling along an elevational gradient in the Alps. Ecohydrology 7:1453–1473. doi:10.1002/eco.1471

Evgeniou, T., M. Pontil, and T. Poggio. 2000. Regularization networks and support vector machines. Adv. Comput. Math. 13:1–50. doi:10.1023/A:1018946025316

Famiglietti, J.S., J.A. Devereaux, C.A. Laymon, T. Tsegaye, P.R. Houser, T.J. Jackson, et al. 1999. Ground-based investigation of soil moisture variability within remote sensing footprints during the Southern Great Plains 1997 (SGP97) hydrology experiment. Water Resour. Res. 35:1839–1851. doi:10.1029/1999WR900047

Grayson, R.B., and A.W. Western. 1998. Towards areal estimation of soil water content from point measurements: Time and space stability of mean response. J. Hydrol. 207:68–82. doi:10.1016/S0022-1694(98)00096-1

Heathman, G.C., P.J. Starks, L.R. Ahuja, and T.J. Jackson. 2003. Assimilation of surface soil moisture to estimate profile soil water content. J. Hydrol. 279:1–17. doi:10.1016/S0022-1694(03)00088-X

Jones, E.B., M. Owe, and T.J. Schmugge. 1982. Soil moisture variation patterns observed in Hand County, South Dakota. J. Am. Water Resour. Assoc. 18:949–954. doi:10.1111/j.1752-1688.1982.tb00100.x

Kollmann, K. 2012. Klima- und landnutzungsbedingte Bodenverteilung im Matschertal. M.S. thesis. Univ. of Innsbruck, Innsbruck, Austria.

Mohanty, B.P., and T.H. Skaggs. 2001. Spatio-temporal evolution and time-stable characteristics of soil moisture within remote sensing footprints with varying soil, slope, and vegetation. Adv. Water Resour. 24:1051–1067. doi:10.1016/S0309-1708(01)00034-3

Notarnicola, C., L. Pasolli, G. Cuozzo, F. Greifeneder, G. Bertoldi, S. Della Chiesa, et al. 2014. Temporal and spatial soil moisture dynamics in mountain meadows by integrating Radarsat 2 images and ground data. In: Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada. 13–18 July 2014. Inst. Electr. Electron. Eng., Piscataway, NJ. p. 1222–1225.

Pasolli, L., G. Bertoldi, S. Della Chiesa, G. Niedrist, U. Tappeiner, M. Zebisch, et al. 2013. Multi-source and multi-scale soil moisture dynamic modelling. In: 2013 IEEE International Geoscience and Remote Sensing Symposium, Melbourne, VIC, Australia. 21-26 July 2013. Inst. Electr. Electron. Eng., Piscataway, NJ. p. 763–766.

Pasolli, L., C. Notarnicola, G. Bertoldi, L. Bruzzone, R. Remelgado, F. Greifeneder, et al. 2015. Estimation of soil moisture in mountain areas using SVR technique applied to multiscale active radar images at C-band. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8:262–283. doi:10.1109/JSTARS.2014.2378795

Pasolli, L., C. Notarnicola, L. Bruzzone, G. Bertoldi, S. Della Chiesa, V. Hell, et al. 2011. Estimation of soil moisture in an alpine catchment with RADARSAT2 images. Appl. Environ. Soil Sci. 2011:175473. doi:10.1155/2011/175473

Pathe, C., W. Wagner, D. Sabel, M. Doubkova, and J.B. Basara. 2009. Using ENVISAT ASAR global mode data for surface soil moisture retrieval over Oklahoma, USA. IEEE Trans. Geosci. Remote Sens. 47:468–480. doi:10.1109/TGRS.2008.2004711

Richards, J.A., and X. Jia. 1999. Remote sensing digital image analysis. 4th ed. Springer-Verlag, Berlin.

Smola, A.J., and B. Schölkopf. 2004. A tutorial on support vector regression. Stat. Comput. 14:199–222. doi:10.1023/B:STCO.0000035301.49549.88

Vapnik, V.N. 2000. The nature of statistical learning theory. 2nd ed. Springer, New York.

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

Zhan, H., P. Shi, and C. Chuqun. 2003. Retrieval of oceanic chlorophyll concentration using support vector machines. IEEE Trans. Geosci. Remote Sens. 41:2947–2951. doi:10.1109/TGRS.2003.819870

II.2 Article 2 – "The Added Value of the VH/VV Polarization-Ratio for Global Soil Moisture Estimations From Scatterometer Data"

written by

Felix Greifeneder, Claudia Notarnicola, Sebastian Hahn, Mariette Vreugdenhil, Christoph Reimer, Emanuele Santi, Simonetta Paloscia, Wolfgang Wagner

Published in the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

DOI: 10.1109/JSTARS.2018.2865185

Cited as:
Greifeneder, Felix, Claudia Notarnicola, Sebastian Hahn, Mariette Vreugdenhil, Christoph Reimer, Emanuele Santi, Simonetta Paloscia, and Wolfgang Wagner. “The Added Value of the VH/VV Polarization-Ratio for Global Soil Moisture Estimations From Scatterometer Data.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018, 1–12.

The activities were carried out within the project: “Use of SCA cross-polarisation information for soil moisture retrieval,” funded through EUMETSAT ITT 14/208518.

References

Albergel, C., et al., “Skill and global trend analysis of soil moisture from reanalyses and microwave remote sensing,” J. Hydrometeorol., vol. 14, no. 4, pp. 1259–1277, 2013.

Albergel, C., P. de Rosnay, G. Balsamo, L. Isaksen, and J.Mu˜noz-Sabater, “Soil moisture analyses at ECMWF: Evaluation using global groundbased in situ observations,” J. Hydrometeorol., vol. 13, pp. 1442–1460, 2012.

Attema, E. W. P., and Ulaby, F. T, “Vegetation modelled as a water cloud,” Radio Sci., vol. 13, no. 2, pp. 357–364, 1978.

Berger, M., et al., “Measuring ocean salinity with ESA’s SMOS mission - advancing the science,” Esa Bull. Sp. Agency, no. 111, pp. 113– 121, 2002.

Berrisford, P., et al., “The ERA-Interim archive Version 2.0,” Tech. Rep. ECMWF, 2011. [Online]. Available: http://www.ecmwf.int/publications/library/do/references/show?id=90276

Bindish, R., T. Jackson, M. Cosh, T. Zhao, and P. O’Neill, “Global soil moisture from the Aquarius/SAC-D satellite: Description and initial assessment,” IEEE Geosci. Remote Sens. Lett., vol. 12, no. 5, pp. 923–927, May 2015.

Bindlish, R., and A. P. Barros, “Including vegetation scattering effects in a radar based soil moisture estimation model,” Remote Sens. Hydrol., vol. 2000, no. 267, pp. 354–361, 2000.

Chang, C., and C. Lin, “LIBSVM: A library for Support Vector Machines,” ACM Trans. Intell. Syst. Technol., vol. 2, pp. 1–39, 2013.

Del Frate, F., P. Ferrazzoli, and G. Schiavon, “Retrieving soil moisture and agricultural variables by microwave radiometry using neural networks,” Remote Sens. Environ., vol. 84, no. 2, pp. 174–183, 2003.

Entekhabi, D., et al., “The soil moisture active passive (SMAP) mission,” Proc. IEEE, vol. 98, no. 5, pp. 704–716, May 2010.

Evans, D. L., T. G. Farr, J. P. Ford, T. W. Thompson, and C. L. Werner, “Multipolarization radar images for geologic mapping and vegetation discrimination,” IEEE Trans. Geosci. Remote Sens., vol. GE-24, no. 2, pp. 246–257, Mar. 1986.

Evgeniou, T., M. Pontil, and T. Poggio, “Regularization networks and support vector machines,” Adv. Comput. Math., vol. 13, no. 1, pp. 1–50, 2000.

Fois, F., C. C. Lin, M. Betto, H. Barr´e, M. Loiselet, and G. Mason, “The MetOp second generation scatterometer instrument,” in Proc. EUMETSAT Meteorolog. Satell. Conf., 2012, no. 1, pp. 8–11, 2012.

Gelman, A., Bayesian Data Analysis, 2nd ed. London, U.K.: Chapman & Hall, 2004.

Gherboudj, I., R. Magagi, A. A. Berg, and B. Toth, “Soil moisture retrieval over agricultural fields from multi-polarized and multi-angular RADARSAT-2 SAR data,” Remote Sens. Environ., vol. 115, no. 1, pp. 33–43, 2011.

Gruber, A., et al., “Performance inter-comparison of soil moisture retrieval models for the MetOp-A

Hollmann, R., et al., “The ESA climate change initiative: Satellite data records for essential climate variables,” Bull. Amer.Meteorol. Soc., vol. 94, no. 10, pp. 1541–1552, 2013.

Hornik, K., M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” Neural Netw., vol. 2, no. 5, pp. 359–366, 1989.

Jacome, A., M. Bernier, K. Chokmani, Y. Gauthier, J. Poulin, and D. De S`eve, “Monitoring volumetric surface soil moisture content at the La Grande basin boreal wetland by radar multi polarization data,” Remote Sens., vol. 5, no. 10, pp. 4919–4941, Oct. 2013.

Kim, Y., T. Jackson, R. Bindlish, H. Lee, and S. Hong, “Radar vegetation indices for estimating the vegetation water content of rice and soybean,” Geosci. Remote Sens. Lett., vol. 9, no. 4, pp. 564–568, 2012.

Le Vine, D. M., G. S. E. Lagerloef, F. R. Colomb, S. H. Yueh, and F. A. Pellerano, “AQUARIUS: An instrument to monitor sea surface salinity from space,” IEEE Trans. Geosci. Remote Sens., vol. 45, no. 7, pp. 2040–2050, Jul. 2007.

Linden, A., and J. Kindermann, “Inversion of multilayer nets,” in Proc. Int. Joint Conf. Neural Netw., 1989, vol. 2, pp. 425–430.

Moran, S., L. Alonso, J. F. Moreno, M. P. Cendrero Mateo, D. Fernando De La Cruz, and A. Montoro, “A RADARSAT-2 quadpolarized time series for monitoring crop and soil conditions in Barrax, Spain,” IEEE Trans. Geosci. Remote Sens., vol. 50, no. 4, pp. 1057–1070, Apr. 2012.

Naeimi, V., K. Scipal, Z. Bartalis, S. Hasenauer, and W. Wagner, “An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations,” IEEE Trans. Geosci. Remote Sens., vol. 47, no. 7, pp. 1999–2013, Jul. 2009.

Nelder, J. A., and R.Mead, “A simplex method for function minimization,” Comput. J., vol. 7, no. 4, pp. 308–313, 1965.

Njoku, E., T. Jackson, and S. Chan, “NRT AMSR2 unified L2B half-orbit 25 km EASE-GRID surface soil moisture,” Dataset available online from the NASA Global Hydrology Center DAAC, Huntsville, AL, USA, 2016.

Notarnicola, C., and F. Posa, “Inferring vegetation water content from C- and L-band SAR images,” IEEE Trans. Geosci. Remote Sens., vol. 45, no. 10, pp. 3165–3171, Oct. 2007.

Notarnicola, C., M. Angiulli, and F. Posa, “Use of radar and optical remotely sensed data for soil moisture retrieval over vegetated areas,” IEEE Trans. Geosci. Remote Sens., vol. 44, no. 4, pp. 925–935, Apr. 2006.

Oh, Y., K. Sarabandi, and F. T. Ulaby, “An empirical model and an inversion technique for radar scattering from bare soil surfaces,” IEEE Trans.Geosci. Remote Sens., vol. 30, no. 2, pp. 370–381, 1992.

Paloscia, S., P. Pampaloni, S. Pettinato, and E. Santi, “A comparison of algorithms for retrieving soil moisture from ENVISAT/ASAR images,” IEEE Trans. Geosci. Remote Sens., vol. 46, no. 10, pp. 3274–3284, Oct. 2008.

Paloscia, S., S. Pettinato, E. Santi, C. Notarnicola, L. Pasolli, and A. Reppucci, “Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation,” Remote Sens. Environ., vol. 134, pp. 234–248, Jul. 2013.

Pasolli, L., et al., “Estimation of soil moisture in an Alpine catchment with RADARSAT2 images,” Appl. Environ. Soil Sci., vol. 2011, pp. 1–12, 2011.

Pasolli, L., et al., “Estimation of soil moisture in mountain areas using SVR technique applied to multiscale active radar images at C-Band,” Sel. Top. Appl. Earth Observ. Remote Sens., vol. 8, no. 1, pp. 262–283, 2015.

Pathe, C., W. Wagner, D. Sabel, M. Doubkova, and J. B. Basara, “Using ENVISAT ASAR Global Mode data for surface soil moisture retrieval over Oklahoma, USA,” IEEE Trans. Geosci. Remote Sens., vol. 47, no. 2, pp. 468–480, Feb. 2009.

Santi, E., S. Paloscia, S. Pettinato, and G. Fontanelli, “A prototype ANN based algorithm for the soil moisture retrieval from L-Band in view of the incoming SMAP mission,” in Proc. IEEE 13th Spec. Meeting Microw. Radiometry Remote Sens. Environ., 2014, pp. 5–9.

Santi, E., S. Paloscia, S. Pettinato, and G. Fontanelli, “Application of artificial neural networks for the soil moisture retrieval from active and passive microwave spaceborne sensors,” Int. J. Appl. Earth Observ. Geoinf., vol. 48, pp. 61–73, 2015.

Santi, E., S. Pettinato, S. Paloscia, P. Pampaloni, G. Macelloni, and M. Brogioni, “An algorithm for generating soilmoisture and snowdepthmaps from microwave spaceborne radiometers: HydroAlgo,” Hydrol. Earth Syst. Sci., vol. 16, no. 10, pp. 3659–3676, 2012.

Smola, A. J., and B. Schölkopf, “A tutorial on support vector regression,” Stat. Comput., vol. 14, no. 3, pp. 199–222, Aug. 2004.

Srivastava, P. K., P. O. Neill, M. Cosh, R. Lang, and A. Joseph, “Evaluation of radar vegetation indices for vegetation water content estimation using data from a ground-based SMAP simulator,” in Proc. IEEE Int. Geosci. Remote Sens. Symp., 2015, pp. 1296–1299.

Steiger, J. H., “Tests for comparing elements of a correlation matrix.,” Psychol. Bull., vol. 87, no. 2, pp. 245–251, 1980.

Ulaby, F. T., P. C. Dubois, and J. Van Zyl, “Radar mapping of surface soil moisture,” J. Hydrol., vol. 184, no. 1–2, pp. 57–84, 1996.

Ulaby, F. T., R. K. Moore, and A. K. Fung, Microwave Remote Sensing Active and Passive-Volume III: From Theory to Applications. Norwood, MA, USA: Artech House, 1986.

Wagner, W., et al., “The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications,” Meteorol. Zeitschrift, vol. 22, no. 1, pp. 5–33, 2013.

Wagner, W., G. Lemoine, and H. Rott, “A method for estimating soil moisture from ERS scatterometer and soil data,” Remote Sens. Environ., vol. 70, no. 2, pp. 191–207, 1999.

Wagner, W., G. Lemoine, M. Borgeaud, S. Member, and H. Rott, “A study of vegetation cover effects on ERS scatterometer data,” IEEE Trans. Geosci. Remote Sens., vol. 37, no. 2, pp. 938–948, Mar. 1999.

Wagner, W., J. Noll, M. Borgeaud, S. Member, and H. Rott, “Monitoring soil moisture over the Canadian prairies with the ERS Scatterometer,” IEEE Trans. Geosci. Remote Sens., vol. 37, no. 1, pp. 206–216, Jan. 1999.

Willett, P., “Dissimilarity-based algorithms for selecting structurally diverse sets of compounds,” J. Comput. Biol., vol. 6, no. 3–4, pp. 447–457, 1999.

Zhan, H., P. Shi, and C. Chuqun, “Retrieval of oceanic chlorophyll concentration using support vector machines,” IEEE Trans. Geosci. Remote Sens., vol. 41, no. 12, pp. 2947–2951, Dec. 2003.

II.3 Article 3 – "A Machine Learning-Based Approach for Surface Soil Moisture Estimations with Google Earth Engine"

written by

Felix Greifeneder, Claudia Notarnicola, Wolfgang Wagner

Publication in the MDPI journal Remote Sensing

DOI: 10.3390/rs13112099

Cited as:
Greifeneder, Felix, Claudia Notarnicola, and Wolfgang Wagner. "A Machine Learning-Based Approach for Surface Soil Moisture Estimations with Google Earth Engine." Remote Sensing 13, no. 11 (2021). https://doi.org/10.3390/rs13112099.

This work was partially funded by the Horizon 2020 project "Ecopotential—Improving Future Ecosystem Benefits through Earth Observation, which has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n 641762) and the European Fund for Regional Development project "DPS4ESLAB".

References

Aboutalebi, M.; Allen, N.; Torres-Rua, A.F.; Coopmans, C. Estimation of soil moisture at different soil levels using machine learning techniques and unmanned aerial vehicle (UAV) multispectral imagery. In Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping IV; International Society for Optical Engineering: Baltimore, MD, USA, 2019; p. 26.

Ahn, K.-H.; Palmer, R. Regional flood frequency analysis using spatial proximity and basin characteristics: Quantile regression vs. parameter regression technique. J. Hydrol. 2016, 540, 515–526.

Albergel, C.; Rüdiger, C.; Pellarin, T.; Calvet, J.-C.; Fritz, N.; Froissard, F.; Suquia, D.; Petitpa, A.; Piguet, B.; 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. Hydrol. Earth Syst. Sci. 2008, 12, 1323–1337.

Ali, I.; Greifeneder, F.; Stamenkovic, J.; Neumann, M.; Notarnicola, C. Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data. Remote Sens. 2015, 7, 16398–16421.

Ångström, A. The Albedo of Various Surfaces of Ground. Geogr. Ann. 1925, 7, 323–342.

Armijo, L. Minimization of functions having Lipschitz continuous first partial derivatives. Pac. J. Math. 1966, 16, 1–3.

Barzilai, J.; Borwein, J.M. Two-Point Step Size Gradient Methods. Ima J. Numer. Anal. 1988, 8, 141–148.

Bauer-Marschallinger, B.; Freeman, V.; Cao, S.; Paulik, C.; Schaufler, S.; Stachl, T.; Modanesi, S.; Massari, C.; Ciabatta, L.; Brocca, L.; et al. Toward Global Soil Moisture MonitoringWith Sentinel-1: Harnessing Assets and Overcoming Obstacles. IEEE Trans. Geosci. Remote Sens. 2018, 57, 520–539.

Bauer-Marschallinger, B.; Paulik, C.; Hochstöger, S.; Mistelbauer, T.; Modanesi, S.; Ciabatta, L.; Massari, C.; Brocca, L.; Wagner, W. Soil Moisture from Fusion of Scatterometer and SAR: Closing the Scale Gap with Temporal Filtering. Remote Sens. 2018, 10, 1030.

Bauer-marschallinger, B.; Schaufler, S.; Navacchi, C. Validation Report Surface Soil Moisture Collection 1KM Version 1. 2018. Available online: https://land.copernicus.eu/global/sites/cgls.vito.be/files/products/CGLOPS1_VR_SSM1km-V1_I1.20.pdf (accessed on 3 November 2020).

Beaudoin, A.; Le Toan, T.; Gwyn, Q. SAR observations and modeling of the C-band backscatter variability due to multiscale geometry and soil moisture. IEEE Trans. Geosci. Remote Sens. 1990, 28, 886–895.

Beck, H.E.; Zimmermann, N.E.; McVicar, T.R.; Vergopolan, N.; Berg, A.;Wood, E.F. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci. Data 2018, 5, 180214.

Bell, J.E.; Palecki, M.A.; Baker, C.B.; Collins, W.G.; Lawrimore, J.H.; Leeper, R.; Hall, M.E.; Kochendorfer, J.; Meyers, T.P.; Wilson, T.; et al. U.S. Climate Reference Network Soil Moisture and Temperature Observations. J. Hydrometeorol. 2013, 14, 977–988.

Berger, M.; Camps, A.; Font, J.; Kerr, Y.; Miller, J.; Johannessen, J.A.; Boutin, J.; Drinkwater, M.R.; Skou, N.; Floury, N.; et al. Measuring ocean salinity with ESA's SMOS mission—Advancing the science. ESA Bull. Eur. Space Agency 2002, 111, 113–121.

Bhuiyan, M.A.E.; Anagnostou, E.N.; Kirstetter, P.-E. A Nonparametric Statistical Technique for Modeling Overland TMI (2A12) Rainfall Retrieval Error. IEEE Geosci. Remote Sens. Lett. 2017, 14, 1898–1902.

Bhuiyan, M.A.E.; Yang, F.; Biswas, N.K.; Rahat, S.H.; Neelam, T.J. Machine Learning-Based Error Modeling to Improve GPM IMERG Precipitation Product over the Brahmaputra River Basin. Forecasting 2020, 2, 248–266.

Bircher, S.; Skou, N.; Jensen, K.H.; Walker, J.P.; Rasmussen, L. A soil moisture and temperature network for SMOS validation in Western Denmark. Hydrol. Earth Syst. Sci. 2012, 16, 1445–1463.

Bras, R.L. Complexity and organization in hydrology: A personal view. Water Resour. Res. 2015, 51, 6532–6548.

Buchhorn, M.; Smets, B.; Bertels, L.; Lesiv, M.; Tsendbazar, N.-E.; Herold, M.; Fritz, S. Copernicus Global Land Service: Land Cover 100m: Collection 3: Epoch 2015: Globe. Zenodo 2020. [CrossRef]

Calvet, J.-C.; Fritz, N.; Froissard, F.; Suquia, D.; Petitpa, A.; Piguet, B. In situ soil moisture observations for the CAL/VAL of SMOS: The SMOSMANIA network. In 2007 IEEE International Geoscience and Remote Sensing Symposium; IEEE: Barcelona, Spain, 2007; pp. 1196–1199.

Chander, G.; Markham, B.L.; Helder, D.L. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Remote Sens. Environ. 2009, 113, 893–903.

Chang, T.-Y.; Wang, Y.-C.; Feng, C.-C.; Ziegler, A.D.; Giambelluca, T.W.; Liou, Y.-A. Estimation of Root Zone Soil Moisture Using Apparent Thermal InertiaWith MODIS Imagery Over a Tropical Catchment in Northern Thailand. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2012, 5, 752–761.

Chatterjee, S.; Huang, J.; Hartemink, A.E. Establishing an Empirical Model for Surface Soil Moisture Retrieval at the U.S. Climate Reference Network Using Sentinel-1 Backscatter and Ancillary Data. Remote Sens. 2020, 12, 1242.

Cho, E.; Jacobs, J.M.; Jia, X.; Kraatz, S. Identifying Subsurface Drainage using Satellite Big Data and Machine Learning via Google Earth Engine. Water Resour. Res. 2019, 55, 8028–8045.

Colliander, A.; Jackson, T.J.; Bindlish, R.; Chan, S.; Das, N.; Kim, S.B.; Cosh, M.H.; Dunbar, R.S.; Dang, L.; Pashaian, L.; et al. Validation of SMAP surface soil moisture products with core validation sites. Remote Sens. Environ. 2017, 191, 215–231.

Derin, Y.; Bhuiyan, A.E.; Anagnostou, E.; Kalogiros, J.; Anagnostou, M.N. Modeling Level 2 Passive Microwave Precipitation Retrieval Error Over Complex Terrain Using a Nonparametric Statistical Technique. IEEE Trans. Geosci. Remote Sens. 2020, 1–12.

Didan, K. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006, V006 ed.; NASA EOSDIS LP DAAC: Sioux Falls, SD, USA, 2015.

Dorigo,W.;Wagner,W.; Hohensinn, R.; Hahn, S.; Paulik, C.; Xaver, A.; Gruber, A.; Drusch, M.; Mecklenburg, S.; Van Oevelen, P.; et al. The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. Hydrol. Earth Syst. Sci. 2011, 15, 1675–1698.

Dorigo,W.A.; Xaver, A.; Vreugdenhil, M.; Gruber, A.; Hegyiová, A.; Sanchis-Dufau, A.D.; Zamojski, D.; Cordes, C.;Wagner,W.; Drusch, M. Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network. Vadose Zone J. 2013, 12, vzj2012.0097.

Dubois, P.; Van Zyl, J.; Engman, T. Measuring soil moisture with imaging radars. IEEE Trans. Geosci. Remote Sens. 1995, 33, 916–926.

Duffy, N.; Helmbold, D. Boosting Methods for Regression. Mach. Learn. 2002, 47, 153–200.

Elith, J.; Leathwick, J.R.; Hastie, T. A working guide to boosted regression trees. J. Anim. Ecol. 2008, 77, 802–813.

Entekhabi, D.; Njoku, E.G.; O'Neill, P.E.; Kellogg, K.H.; Crow, W.T.; Edelstein, W.N.; Entin, J.K.; Goodman, S.D.; Jackson, T.J.;

European Space Agency. Sentinel-1 User Handbook. European Space Agency, European Comission. 2013. Available online: https://sentinel.esa.int (accessed on 16 December 2020).

Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The Shuttle Radar Topography Mission. Rev. Geophys. 2007, 45, 1–33.

Friedman, J.H. Greedy function approximation: A gradient boosting machine. Ann. Stat. 2001, 29, 1189–1232.

Fung, A.K.; Li, Z.; Chen, K.S. Backscatterin from randomly rough dielectric surface. IEEE Trans. Geosci. Remote Sens. 1992, 30, 356–369.

Google. Google Earth Engine: Sentinel-1 Pre-Processing. 2016. Available online: https://developers.google.com/earth-engine/sentinel1 (accessed on 18 December 2017).

Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 2017, 202, 18–27.

Greifeneder, F. PYSMM. Zenodo 2021.

Greifeneder, F.; Khamala, E.; Sendabo, D.;Wagner,W.; Zebisch, M.; Farah, H.; Notarnicola, C. Detection of soil moisture anomalies based on Sentinel-1. Phys. Chem. Earth Parts ABC 2018, 1–8.

Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 2009, 377, 80–91.

Hengl, T. Clay content in % (kg/kg) at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution. Zenodo 2018.

Hengl, T. Sand content in % (kg/kg) at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution. Zenodo 2018.

Hengl, T. Soil bulk density (fine earth) 10 x kg/m-cubic at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution. Zenodo 2018.

Hengl, T. Soil texture classes (USDA system) for 6 soil depths (0, 10, 30, 60, 100 and 200 cm) at 250 m. Zenodo 2018.

Hird, J.N.; DeLancey, E.R.; McDermid, G.J.; Kariyeva, J. Google Earth Engine, Open-Access Satellite Data, and Machine Learning in Support of Large-Area Probabilistic Wetland Mapping. Remote Sens. 2017, 9, 1315.

Hoekman, D.H. Radar Remote Sensing Data for Applications in Forestry. Dissertation, Internally Prepared, Laboratory of Geo-Information Science and Remote Sensing,Wageningen University,Wageningen, The Netherlands. 1990. Available online: https://library.wur.nl/WebQuery/wurpubs/12860 (accessed on 12 January 2021).

Hornacek, M.; Wagner, W.; Sabel, D.; Truong, H.-L.; Snoeij, P.; Hahmann, T.; Diedrich, E.; Doubkova, M. Potential for High Resolution Systematic Global Surface Soil Moisture Retrieval via

Change Detection Using Sentinel-1. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2012, 5, 1303–1311.

Hrisko, J.; Ramamurthy, P.; Gonzalez, J.E. Estimating heat storage in urban areas using multispectral satellite data and machine learning. Remote Sens. Environ. 2021, 252, 112125.

Johnson, J.; et al. The Soil Moisture Active Passive (SMAP) Mission. Proc. IEEE 2010, 98, 704–716.

Kohavi, R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. Int. Jt. Conf. Artif. Intell. 1995, 14, 1137–1143.

Kolassa, J.; Reichle, R.H.; Liu, Q.; Alemohammad, S.H.; Gentine, P.; Aida, K.; Asanuma, J.; Bircher, S.; Caldwell, T.; Colliander, A.; et al. Estimating surface soil moisture from SMAP observations using a Neural Network technique. Remote Sens. Environ. 2018, 204, 43–59.

Larson, K.M.; Small, E.E.; Gutmann, E.D.; Bilich, A.L.; Braun, J.J.; Zavorotny, V.U. Use of GPS receivers as a soil moisture network for water cycle studies. Geophys. Res. Lett. 2008, 35, 24405.

Lei, F.; Crow,W.T.; Kustas, W.P.; Dong, J.; Yang, Y.; Knipper, K.R.; Anderson, M.C.; Gao, F.; Notarnicola, C.; Greifeneder, F.; et al. Data assimilation of high-resolution thermal and radar remote sensing retrievals for soil moisture monitoring in a drip-irrigated vineyard. Remote Sens. Environ. 2020, 239, 111622.

Li, Y.; Yan, S.; Chen, N.; Gong, J. Performance Evaluation of a Neural Network Model and Two Empirical Models for Estimating Soil Moisture Based on Sentinel-1 SAR Data. Prog. Electromagn. Res. C 2020, 105, 85–99.

Liu, D.;Wang, G.; Mei, R.; Yu, Z.; Yu, M. Impact of initial soil moisture anomalies on climate mean and extremes over Asia. J. Geophys. Res. Atmos. 2014, 119, 529–545.

Liu, Y.; Qian, J.; Yue, H. Combined Sentinel-1AWith Sentinel-2A to Estimate Soil Moisture in Farmland. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2021, 14, 1292–1310.

Moghaddam, M.; Entekhabi, D.; Goykhman, Y.; Li, K.; Liu, M.; Mahajan, A.; Nayyar, A.; Shuman, D.; Teneketzis, D. A Wireless Soil Moisture Smart Sensor Web Using Physics-Based Optimal Control: Concept and Initial Demonstrations. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2010, 3, 522–535.

Moghaddam, M.; Silva, A.; Clewley, D.; Akbar, R.; Hussaini, S.A.; Whitcomb, J.; Devarakonda, R.; Shrestha, R.; Cook, R.B.; Prakash, G.; et al. Soil Moisture Profiles and Temperature Data from SoilSCAPE Sites, USA; ORNL DAAC: Oak Ridge, TN, USA, 2017.

Moosavi, V.; Talebi, A.; Mokhtari, M.H.; Hadian, M.R. Estimation of spatially enhanced soil moisture combining remote sensing and artificial intelligence approaches. Int. J. Remote Sens. 2016, 37, 5605–5631.

Morrison, K.; Wagner, W. Explaining Anomalies in SAR and Scatterometer Soil Moisture Retrievals From Dry Soils With Subsurface Scattering. IEEE Trans. Geosci. Remote Sens. 2019, 58, 2190–2197.

Naeimi, V.; Scipal, K.; Bartalis, Z.; Hasenauer, S.;Wagner,W. An Improved Soil Moisture Retrieval Algorithm for ERS and METOP Scatterometer Observations. IEEE Trans. Geosci. Remote Sens. 2009, 47, 1999–2013.

Notarnicola, C.; Angiulli, M.; Posa, F. Use of radar and optical remotely sensed data for soil moisture retrieval over vegetated areas. IEEE Trans. Geosci. Remote Sens. 2006, 44, 925–935.

Oh, Y.; Sarabandi, K.; Ulaby, F.T. An empirical model and an inversion technique for radar scattering from bare soil surfaces. IEEE Trans. Geosci. Remote Sens. 1992, 30, 370–381.

Ojo, E.R.; Bullock, P.R.; L'Heureux, J.; Powers, J.; McNairn, H.; Pacheco, A. Calibration and evaluation of a frequency domain reflectometry sensor for real-time soil moisture monitoring. Vadose Zone J. 2015, 14, 1–15.

Osenga, E.C.; Arnott, J.C.; Endsley, K.A.; Katzenberger, J.W. Bioclimatic and Soil Moisture Monitoring Across Elevation in a Mountain Watershed: Opportunities for Research and Resource Management. Water Resour. Res. 2019, 55, 2493–2503.

Paloscia, S.; Pettinato, S.; Santi, E.; Notarnicola, C.; Pasolli, L.; Reppucci, A. Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation. Remote Sens. Environ. 2013, 134, 234–248.

Pasolli, L.; Notarnicola, C.; Bertoldi, G.; Bruzzone, L.; Remelgado, R.; Greifeneder, F.; Niedrist, G.; Della Chiesa, S.; Tappeiner, U.; Zebisch, M. Estimation of Soil Moisture in Mountain Areas Using SVR Technique Applied to Multiscale Active Radar Images at C-Band. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2015, 8, 262–283.

Paulik, C.; Plocon, A.; Hahn, S.; Mistelbauer, T.; Reimer, C. TUW-GEO/pytesmo. Pytesmo 2017.

Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.;Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 2011, 12, 2825–2830.

Pham, T.D.; Yokoya, N.; Xia, J.; Ha, N.T.; Le, N.N.; Nguyen, T.T.T.; Dao, T.H.; Vu, T.T.P.; Takeuchi,W. Comparison of Machine Learning Methods for Estimating Mangrove Above-Ground Biomass

Using Multiple Source Remote Sensing Data in the Red River Delta Biosphere Reserve, Vietnam. Remote Sens. 2020, 12, 1334.

Price, J. Using spatial context in satellite data to infer regional scale evapotranspiration. IEEE Trans. Geosci. Remote Sens. 1990, 28, 940–948.

Qin, J.; Yang, K.; Lu, N.; Chen, Y.; Zhao, L.; Han, M. Spatial upscaling of in-situ soil moisture measurements based on MODIS derived apparent thermal inertia. Remote Sens. Environ. 2013, 138, 1–9.

Quegan, S.; Le Toan, T.; Yu, J.; Ribbes, F.; Floury, N. Multitemporal ERS SAR analysis applied to forest mapping. IEEE Trans. Geosci. Remote Sens. 2000, 38, 741–753.

Rodell, M.; Houser, P.R.; Jambor, U.; Gottschalck, J.; Mitchell, K.; Meng, C.-J.; Arsenault, K.; Cosgrove, B.; Radakovich, J.; Bosilovich, M.; et al. The Global Land Data Assimilation System. Bull. Am. Meteorol. Soc. 2004, 85, 381–394.

Santi, E.; Paloscia, S.; Pettinato, S.; Fontanelli, G. A prototype ann based algorithm for the soil moisture retrieval from l- band in view of the incoming SMAP mission. In 2014 13th Specialist Meeting on Microwave Radiometry and Remote Sensing of the Environment (MicroRad); IEEE: Pasadena, CA, USA, 2014; pp. 5–9.

Schmugge, T. Remote Sensing of Surface Soil Moisture. J. Appl. Meteorol. Climatol. 1978, 17, 1549–1557.

Smith, A.B.; Walker, J.P.; Western, A.W.; Young, R.I.; Ellett, K.M.; Pipunic, R.C.; Grayson, R.B.; Siriwardena, L.; Chiew, F.H.S.; Richter, H.G. The Murrumbidgee soil moisture monitoring network data set. Water Resour. Res. 2012, 48, 1–6.

Srivastava, P.K.; Han, D.; Ramirez, M.R.; Islam, T. Machine Learning Techniques for Downscaling SMOS Satellite Soil Moisture Using MODIS Land Surface Temperature for Hydrological Application. Water Resour. Manag. 2013, 27, 3127–3144.

Stamenkovic, J.; Guerriero, L.; Ferrazzoli, P.; Notarnicola, C.; Greifeneder, F.; Thiran, J.-P. Soil Moisture Estimation by SAR in Alpine Fields Using Gaussian Process Regressor Trained by Model Simulations. IEEE Trans. Geosci. Remote Sens. 2017, 55, 4899–4912.

Stendardi, L.; Karlsen, S.R.; Niedrist, G.; Gerdol, R.; Zebisch, M.; Rossi, M.; Notarnicola, C. Exploiting Time Series of Sentinel-1 and Sentinel-2 Imagery to Detect Meadow Phenology in Mountain Regions. Remote Sens. 2019, 11, 542.

Systematic Observation Requirements for Satellite-Based Data Products for Climate; GCOS Secretariat: Geneva, Switzerland, 2011; Available online: https://climate.esa.int/sites/default/files/gcos-154.pdf (accessed on 3 November 2020).

Tagesson, T.; Fensholt, R.; Guiro, I.; Rasmussen, M.O.; Huber, S.; Mbow, C.; Garcia, M.; Horion, S.; Sandholt, I.; Holm-Rasmussen, B.; et al. Ecosystem properties of semiarid savanna grassland inWest Africa and its relationship with environmental variability. Glob. Chang. Biol. 2015, 21, 250–264.

Traganos, D.; Aggarwal, B.; Poursanidis, D.; Topouzelis, K.; Chrysoulakis, N.; Reinartz, P. Towards Global-Scale Seagrass Mapping and Monitoring Using Sentinel-2 on Google Earth Engine: The Case Study of the Aegean and Ionian Seas. Remote Sens. 2018, 10, 1227.

Tyralis, H.; Papacharalampous, G.; Burnetas, A.; Langousis, A. Hydrological post-processing using stacked generalization of quantile regression algorithms: Large-scale application over CONUS. J. Hydrol. 2019, 577, 123957.

Ulaby, F.T.; Batlivala, P.P.; Dobson, M.C. Microwave Backscatter Dependence on Surface Roughness, Soil Moisture, and Soil Texture: Part I-Bare Soil. IEEE Trans. Geosci. Electron. 1978, 16, 286–295.

Ulander, L. Radiometric slope correction of synthetic-aperture radar images. IEEE Trans. Geosci. Remote Sens. 1996, 34, 1115–1122.

Upreti, D.; Huang, W.; Kong, W.; Pascucci, S.; Pignatti, S.; Zhou, X.; Ye, H.; Casa, R. A Comparison of Hybrid Machine Learning Algorithms for the Retrieval of Wheat Biophysical Variables from Sentinel-2. Remote Sens. 2019, 11, 481.

Verrelst, J.; Muñoz, J.; Alonso, L.; Delegido, J.; Rivera, J.P.; Camps-Valls, G.; Moreno, J. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. Remote Sens. Environ. 2012, 118, 127–139.

Vollrath, A.; Mullissa, A.; Reiche, J. Angular-Based Radiometric Slope Correction for Sentinel-1 on Google Earth Engine. Remote Sens. 2020, 12, 1867.

Wagner,W. Big Data infrastructures for processing Sentinel data. In Photogramm. Week; University of Stuttgart: Stuttgart, Germany, 2015; pp. 93–104.

Wagner,W.; Lemoine, G.; Borgeaud, M.; Member, S.; Rott, H. A Study of Vegetation Cover Effects on ERS Scatterometer Data. Geosci. Remote Sens. IEEE Trans. 1999, 37, 938–948.

Wagner,W.; Lemoine, G.; Rott, H. A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data. Remote Sens. Environ. 1999, 70, 191–207.

Wang, Y.; Jiang, B.; Liang, S.; Wang, D.; He, T.; Wang, Q.; Zhao, X.; Xu, J. Surface Shortwave Net Radiation Estimation from Landsat TM/ETM+ Data Using Four Machine Learning Algorithms. Remote Sens. 2019, 11, 2847.

Young, R.; Walker, J.P.; Yeoh, N.; Smith, A.; Ellett, K.M.; Merlin, O.; Western, A. Soil Moisture and Meteorological Observations from the Murrumbidgee Catchment. 2008. Available online: http://www.oznet.org.au/documentation/Soil_Moisture_Meteorological_Observation_of_Murrumbidgee_Catchment.pdf (accessed on 28 February 2020).

Zacharias, S.; Bogena, H.; Samaniego, L.; Mauder, M.; Fuß, R.; Pütz, T.; Frenzel, M.; Schwank, M.; Baessler, C.; Butterbach-Bahl, K.; et al. A Network of Terrestrial Environmental Observatories in Germany. Vadose Zone J. 2011, 10, 955–973.

Zhang, D.; Zhou, G. Estimation of Soil Moisture from Optical and Thermal Remote Sensing: A Review. Sensors 2016, 16, 1308.

II.4 Article 4 – "Detection of Soil Moisture Anomalies Based on Sentinel-1"

written by

Felix Greifeneder, Erick Khamala, Degeleo Sendabo, Wolfgang Wagner, Marc Zebisch, Hussein Farah, Claudia Notarnicola

Published in the Elsevier Journal Physics and Chemistry of the Earth.

DOI: 10.1016/j.pce.2018.11.009

Cited as:
Greifeneder, Felix, Erick Khamala, Degelo Sendabo, Wolfgang Wagner, Marc Zebisch, Hussein Farah, and Claudia Notarnicola. "Detection of Soil Moisture Anomalies Based on Sentinel-1." *Physics and Chemistry of the Earth, Parts A/B/C*, no. November (2018): 1–8. https://doi.org/10.1016/J.PCE.2018.11.009.

This study was executed in the scope of the project "Synergy of multi-temporal Sentinel 1 and Sentinel 2 image for soil water resources monitoring in Africa", funded within the European Space Agency "Tiger Bridge – Water for Agriculture" framework.

References

Alexakis, D.D., Mexis, F.D.K., Vozinaki, A.E.K., Daliakopoulos, I.N., Tsanis, I.K., Jun. 2017. Soil moisture content estimation based on sentinel-1 and auxiliary earth observation products. A hydrological approach. Sensors 17 (6), 1455.

Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M., Notarnicola, C., 2015. Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. Rem. Sens. 7 (12), 16398–16421.

Baghdadi, N., El Hajj, M., Zribi, M., 2016. Coupling SAR C-band and optical data for soil moisture and leaf area index retrieval over irrigated grasslands. In: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2016. pp. 3551–3554 Novem, no. 3.

Berger, M., Camps, a, Font, J., Kerr, Y., Miller, J., Johannessen, J., Boutin, J., Drinkwater, M.R., Skou, N., Floury, N., Rast, M., Rebhan, H., Attema, E., 2002. Measuring ocean salinity with ESA's SMOS mission - advancing the science. Esa Bull. Sp. Agency 111, 113–121.

F. Chen, K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q. Y. Duan, M. Ek, and A. Betts, "Modeling of land surface evaporation by four schemes and comparison with FIFE observations," J. Geophys. Res. Atmos., vol. 101, no. D3, pp. 7251–7268, Mar. 1996.

Dorigo, W. a., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, a., Gruber, a., Drusch, M., Mecklenburg, S., Van Oevelen, P., Robock, a., Jackson, T., 2011. The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. Hydrol. Earth Syst. Sci. 15 (5), 1675–1698.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P.D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y.Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S.I., Smolander, T., Lecomte, P., Dec. 2017. ESA CCI Soil Moisture for improved Earth system understanding: state-of-the art and future directions. Remote Sens. Environ. 203, 185–215.

Doubková, M., Van Dijk, A.I.J.M., Sabel, D., Wagner, W., Blöschl, G., 2012. Evaluation of the predicted error of the soil moisture retrieval from C-band SAR by comparison against modelled soil moisture estimates over Australia. Remote Sens. Environ. 120, 188–196.

Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A., Vapnik, V., 1997. Support vector regression machines. Adv. Neural Inf. Process. Syst. 9, 155–161.

Entekhabi, D., Njoku, E.G., O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Van Zyl, J., 2010. The soil moisture active passive (SMAP) mission. Proc. IEEE 98 (5), 704–716.

European Space Agency, 2013. Sentinel-1 User Handbook.

Greifeneder, F., Notarnicola, C., Wagner, W., 2016. Using machine learning and SAR data for the upscaling of large scale modelled soil moisture in the Alps. In: 11th European Conference on Synthetic Aperture Radar 2016. EUSAR, pp. 1108–1111.

Gruber, A., Wagner, W., Hegyiova, A., Greifeneder, F., Schlaffer, S., 2013. Potential of Sentinel-1 for high-resolution soil moisture monitoring. In: 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS, pp. 4030–4033.

Kenya Meteorological Department, 2017. The Outlook For The October-November- December (OND) 2017 Season And Review Of Rainfall During The 'Long Rains' (March To May) 2017 & June-July-August (JJA) 2017 Seasons. [Online]. Available: http://www.meteo.go.ke/pdf/seasonal.pdf, Accessed date: 10 October 2017.

Liu, D., Wang, G., Mei, R., Yu, Z., Yu, M., 2014. Impact of initial soil moisture anomalies on climate mean and extremes over Asia. J. Geophys. Res. 119 (2), 529–545.

Paloscia, S., Pettinato, S., Santi, E., Notarnicola, C., Pasolli, L., Reppucci, A., Jul. 2013. Soil moisture mapping using Sentinel-1 images: algorithm and preliminary validation. Remote Sens. Environ. 134, 234–248.

Pasolli, L., Notarnicola, C., Bertoldi, G., Bruzzone, L., Remelgado, R., Greifeneder, F., Niedrist, G., Della Chiesa, S., Tappeiner, U., Zebisch, M., 2015. Estimation of soil moisture in mountain areas using SVR technique applied to multiscale Active radar images at C-band. Sel. Top. Appl. Earth Obs. Remote Sens. 8 (1), 262–283.

Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin*, J.K., Walker, J.P., Lohmann, D., Toll, D., Mar. 2004. The global land data assimilation system. Bull. Am. Meteorol. Soc. 85 (3), 381–394.

Santi, E., Paloscia, S., Pettinato, S., Fontanelli, G., 2015. Application of artificial neural networks for the soil moisture retrieval from active and passive microwave spaceborne sensors. Int. J. Appl. Earth Obs. Geoinf. 48, 61–73.

Sepulcre-Canto, G., Horion, S., Singleton, A., Carrao, H., Vogt, J., 2012. Development of a combined drought indicator to detect agricultural drought in Europe. Nat. Hazards Earth Syst. Sci. 12 (11), 3519–3531.

Ulaby, F.T., Member, S., Batlivala, P.P., 1978. Microwave backscatter dependence on surface roughness , soil moisture , and soil Texture : Part 1I-bare soil. IEEE Trans. Geosci. Electron. 16 (4), 286–295.

Vachaud, G., Silans, A.P.D.E., Balabanis, P., Vauclin, M., 1985. Temporal stability of spatially measured soil water probability density function. Soil Sci. Soc. Am. J. 49 (3), 822–828.

Wagner, W., Pathe, C., Doubkova, M., Sabel, D., Bartsch, A., Hasenauer, S., Blöschl, G., Scipal, K., Martínez-Fernández, J., Löw, A., 2008. Temporal stability of soil moisture and radar backscatter observed by the advanced synthetic aperture radar (ASAR). Sensors (Peterboroug) 8 (2), 1174–1197.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., De Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Blöschl, G., Eitzinger, J., Steinnocher, K., Zeil, P., Rubel, F., 2013. The ASCAT soil moisture product: a review of its specifications, validation results, and emerging applications. Meteorol. Z. 22 (1), 5–33.

Walker, J.P., Willgoose, G.R., Kalma, J.D., Jun. 2004. In situ measurement of soil moisture: a comparison of techniques. J. Hydrol. 293 (1–4), 85–99.

Wilks, D.S., 1995. Statistical Methods in the Atmospheric Science. Academic Press.

Zreda, M., Desilets, D., Ferré, T.P.A., Scott, R.L., Nov. 2008. Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons. Geophys. Res. Lett. 35 (21), L21402.

Zreda, M., Shuttleworth, W.J., Zeng, X., Zweck, C., Desilets, D., Franz, T., Rosolem, R., Nov. 2012. COSMOS: the cosmic-ray soil moisture observing system. Hydrol. Earth Syst. Sci. 16 (11), 4079–4099.

III. THESIS CONCLUSIONS AND OUTLOOK

III.1 Summary and Conclusion

The thesis aimed to explore the potential of machine learning for the global scale analysis of SMC. The developments of the last decades – high spatial and temporal resolution satellite missions, the shift towards open data policies, better availability of analysis-ready data and processing infrastructure – have created immense new potential for many applications of EO data. While the four articles making up this thesis already include detailed discussions of the results and conclusions, this section focuses on their interrelationships and their contribution to the general aims and problems formulated in Part I.

Article 1 presented a method for bridging the gap between the different scales of in situ SMC measurements and satellite pixels for both validation and retrieval approaches. The method was designed specifically for the Long-Term-Socio-Ecological-Research site in Val Mazia[24] in northern Italy, which served as one of the candidate validation sites for the SMAP cal/val project[25]. However, the same approach may be helpful in other areas with highly complex terrain or land cover. The results and performance analysis of the proposed method show how it successfully reduces the effects of local SMC variability and helps strengthen the relationship between satellite and in situ measurements. Furthermore, the study produced valuable findings in the context of this thesis' general aims by analysing the correlation of SMC across scales and the impact of features like land-cover, soil-type, or topography. We were able to build on these insights for the model design in articles 3 and 4. In this way, article 1 made a valuable contribution towards the objectives formulated in Part I, especially objectives number one and three.

We presented the first attempt at machine-learning-based SMC estimations at the global scale in article 2. For this purpose, the study had to tackle several problems related to objective one (Part I, section 1.4 and 1.5). To tackle the issues related to the properties of the reference SMC dataset, namely, good coverage of both target and feature domains requiring a good distribution, we used the ERA-Interim dataset for training and algorithm testing. The results further demonstrated that with the appropriate reference dataset, the problem of site-specificity was reduced. For the article's main aim, the analysis of the added value of the cross-polarisation channel, we exploited one quality of machine learning that sets it apart from process-driven models – the general flexibility of these methods makes an excellent tool for hypothesis testing.

Confirming the results of previous studies (Gruber *et al.*, 2014), different machine learning algorithms performed similarly in this study, which is in line with analysis performed in the scope of article 3 where other qualities, besides the estimation accuracy, like the flexibility to handle different ordinal scales of input data or the training efficiency led to the choice of algorithm.

[24] https://ltereurac.wimuu.com/it/
[25] https://smap.jpl.nasa.gov/science/validation/

The results of article 2 gave a practical example for the radar cross-polarisation channel's sensitivity towards vegetation dynamics, described in an experimental setting by (Ulaby, Moore and Fung, 1986). We demonstrated the potential improvements, but the study also demonstrates the limitations of an approach that relies only on radar backscatter information, which shows in the relatively high error values for some regions.

The third article is the main work of this thesis. It picked up findings from articles 1 and 2 and other related work (Pasolli *et al.*, 2015; Greifeneder, Notarnicola and Wagner, 2016; Greifeneder *et al.*, 2017; Stamenkovic *et al.*, 2017) to tackle all of the main objectives formulated in Part I and estimate SMC at a high spatial resolution on a quasi-global scale. The novelty of this approach is the application of a data-driven model in a large-scale context.

With our work in the first two articles, we concentrated mainly on microwave remote sensing as a data source. In Article 3, data from various sources with different properties in terms of spatial- and temporal-resolution were integrated: point measurements from the ISMN (as training and validation dataset), high-resolution satellite data from Sentinel-1 and Landsat-8, medium resolution satellite data from MODIS, coarse resolution model data, static data like land-cover and soil properties. It turned out that machine learning, notably the GBRT algorithm, effectively merges these different data types. The validation performed very well. The overall accuracy was in line with requirements set by the Global Observing System[26], which means that the method performs very well, also compared to existing operational products and previous studies. As discussed in the article (Part II) the main limitations of the approach can be related to the irregular coverage of the feature and target space by the SMC reference dataset.

The combination of many datasets was enabled by consistently relying on GEE for the data collection. It offered the capabilities to perform the necessary interpolation operations on the fly, and we could avoid downloading large amounts of data. In the end, only the pre-processed and completely assembled training datasets were downloaded to perform the machine learning algorithm training offline. The use of GEE brought another critical advantage; We were able to implement the estimation of SMC to be executed directly online, which means that no base data has to be downloaded the perform estimations. The estimation software was published as a python package PYSMM (Greifeneder, Notarnicola and Wagner, 2021b). Besides the obvious advantages of the EO paradigm shift, related to the simplified access to data and processing resources, a significant advantage for science is that it significantly increases the reproducibility of results.

Being the main work of the thesis, it also generated its biggest impact. Numerous scientific and general users picked up the results, especially in the shape of PYSMM. The software enables us

[26] https://climate.esa.int/sites/default/files/gcos-154.pdf

to deliver, in a semi-operational way, soil moisture maps for the GRAPEX[27] project by USGS, which led to a joint publication (Lei *et al.*, 2020). The publication of PYSMM helped to increase the reach of our work significantly. According to the PyPI download statistics [28], the software was downloaded more than 13,000 times when writing this text. A group also picked it up at FAO who used PYSMM to map wetlands in Southeast Asia.

The final article proposed a method for detecting SMC anomalies based on remotely sensed, high spatial resolution SMC estimations. The approach exploited the correlation of SMC across spatial scales for the downscaling of average SMC climatologies extracted from coarse resolution GLDAS simulations, and thus, it contributed to a better understanding of the questions raised through objective 3 (part I). We successfully demonstrated how high-resolution SMC estimations could be used for practical applications like drought monitoring or climate forecasting with the presented solution.

Certain limitations apply to the methods introduced in this thesis (as discussed in the conclusions above and part II). These are the areas where future work could continue the development. The list below summarises the points with the most significant potential for improvement in a very condensed way:

1. Training dataset: Article 3 analysed the limitations related to a sparse training dataset. The results suggest that estimation accuracies could be improved through a better representation of the feature-target space. A possible way to achieve this could be integrating in-situ data as a reference with alternative reference datasets like modelled SMC, following the idea explored in (Greifeneder, Notarnicola and Wagner, 2016)
2. Validation: This refers primarily to the validation of high-resolution spatial SMC patterns. Articles 3 and 4 perform validation based on time-series of in-situ point measurements, which allows a good understanding of the sensitivity of temporal SMC variabilities. Continuous data with a similar sampling rate and soil depth would be necessary to validate spatial patterns.
3. Deep learning: Throughout the last ten years, deep learning has been an incredible trend in the world of machine learning, initiated by the incredible successes in image classification (Krizhevsky, Sutskever and Hinton, 2012). Various recent studies suggest that the method holds significant potential for estimating SMC (Lee *et al.*, 2019; Ahmed *et al.*, 2021; Li *et al.*, 2021). The advantages of deep learning compared to traditional machine learning approaches could be improved efficiency and the ability to extract

[27] https://www.ars.usda.gov/northeast-area/beltsville-md-barc/beltsville-agricultural-research-center/hydrology-and-remote-sensing-laboratory/docs/grapex/grapex-overview/

[28] https://pepy.tech/project/pysmm

relevant information when working with a large number of features (i.e. exploiting a large number of data sources).

IV. BIBLIOGRAPHY

Adam, E. *et al.* (2014) 'Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers', *International Journal of Remote Sensing*, 35(10), pp. 3440–3458. doi:10.1080/01431161.2014.903435.

Ahmed, A.A.M. *et al.* (2021) 'Deep Learning Forecasts of Soil Moisture: Convolutional Neural Network and Gated Recurrent Unit Models Coupled with Satellite-Derived MODIS, Observations and Synoptic-Scale Climate Index Data', *Remote Sensing*, 13(4), p. 554. doi:10.3390/rs13040554.

Ahn, K.-H. and Palmer, R. (2016) 'Regional flood frequency analysis using spatial proximity and basin characteristics: Quantile regression vs. parameter regression technique', *Journal of Hydrology*, 540, pp. 515–526. doi:10.1016/j.jhydrol.2016.06.047.

Ali, I. *et al.* (2015) 'Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data', *Remote Sensing*, 7(12), pp. 16398–16421. doi:10.3390/rs71215841.

Andini, N., Kim, D. and Chun, J.A. (2020) 'Operational soil moisture modeling using a multi-stage approach based on the generalized complementary principle', *Agricultural Water Management*, 231, p. 106026. doi:10.1016/j.agwat.2020.106026.

Ångström, A. (1925) 'The Albedo of Various Surfaces of Ground', *Geografiska Annaler*, 7(4), pp. 323–342. doi:10.1080/20014422.1925.11881121.

Attema, E. and Ulaby, F.T. (1978) 'Vegetation modelled as a water cloud', *Radio Science*, 13(2), pp. 357–364. doi:http://dx.doi.org/10.1029/RS013i002p00357.

Bauer-Marschallinger, B. *et al.* (2018) 'Soil Moisture from Fusion of Scatterometer and SAR: Closing the Scale Gap with Temporal Filtering', *Remote Sensing*, 10(7), p. 1030. doi:10.3390/rs10071030.

Bauer-Marschallinger, B. *et al.* (2019) 'Toward Global Soil Moisture Monitoring With Sentinel-1: Harnessing Assets and Overcoming Obstacles', *IEEE Transactions on Geoscience and Remote Sensing*, 57(1), pp. 520–539. doi:10.1109/TGRS.2018.2858004.

Berger, M. *et al.* (2002) 'Measuring ocean salinity with ESA's SMOS mission - Advancing the science', *Esa Bulletin-European Space Agency*, (111), pp. 113–121.

Bhuiyan, M.A.E. *et al.* (2020) 'Machine Learning-Based Error Modeling to Improve GPM IMERG Precipitation Product over the Brahmaputra River Basin', *Forecasting*, 2(3), pp. 248–266. doi:10.3390/forecast2030014.

Bhuiyan, M.A.E., Anagnostou, E.N. and Kirstetter, P.-E. (2017) 'A Nonparametric Statistical Technique for Modeling Overland TMI (2A12) Rainfall Retrieval Error', *IEEE Geoscience and Remote Sensing Letters*, 14(11), pp. 1898–1902. doi:10.1109/LGRS.2017.2728658.

Bras, R.L. (2015) ‘Complexity and organization in hydrology: A personal view: HYDROLOGIC COMPLEXITY AND ORGANIZATION’, *Water Resources Research*, 51(8), pp. 6532–6548. doi:10.1002/2015WR016958.

Brocca, L. *et al.* (2014) ‘Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data’, pp. 5128–5141. doi:10.1002/2014JD021489.

Brocca, L. *et al.* (2016) ‘Use of Satellite Soil Moisture Products for the Operational Mitigation of Landslides Risk in Central Italy’, in *Satellite Soil Moisture Retrieval*. Elsevier, pp. 231–247. doi:10.1016/B978-0-12-803388-3.00012-7.

Brown, M.E. *et al.* (2008) ‘Neural networks as a tool for constructing continuous NDVI time series from AVHRR and MODIS’, *International Journal of Remote Sensing*, 29(24), pp. 7141–7158. doi:10.1080/01431160802238435.

Buchhorn, M., Lesiv, M., *et al.* (2020) ‘Copernicus Global Land Cover Layers—Collection 2’, *Remote Sensing*, 12(6), p. 1044. doi:10.3390/rs12061044.

Buchhorn, M., Smets, B., *et al.* (2020) ‘Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe’. Zenodo. doi:10.5281/ZENODO.3939050.

Chang, T.-Y. *et al.* (2012) ‘Estimation of Root Zone Soil Moisture Using Apparent Thermal Inertia With MODIS Imagery Over a Tropical Catchment in Northern Thailand’, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(3), pp. 752–761. doi:10.1109/JSTARS.2012.2190588.

Chaparro, D., Piles, M. and Vall-llossera, M. (2016) ‘Remotely Sensed Soil Moisture as a Key Variable in Wildfires Prevention Services’, in *Satellite Soil Moisture Retrieval*. Elsevier, pp. 249–269. doi:10.1016/B978-0-12-803388-3.00013-9.

Cortes, C. and Vapnik, V. (1995) ‘Support-vector networks’, *Machine Learning*, 20(3), pp. 273–297. doi:10.1007/BF00994018.

Deng, J. *et al.* (2009) ‘ImageNet: A large-scale hierarchical image database’, in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops)*, Miami, FL: IEEE, pp. 248–255. doi:10.1109/CVPR.2009.5206848.

Derin, Y. *et al.* (2020) ‘Modeling Level 2 Passive Microwave Precipitation Retrieval Error Over Complex Terrain Using a Nonparametric Statistical Technique’, *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–12. doi:10.1109/TGRS.2020.3038343.

Entekhabi, D. *et al.* (2010) ‘The soil moisture active passive (SMAP) mission’, *Proceedings of the IEEE*, 98(5), pp. 704–716. doi:10.1109/JPROC.2010.2043918.

Esch, T. *et al.* (2018) ‘Exploiting big earth data from space – first experiences with the timescan processing chain’, *Big Earth Data*, 2(1), pp. 36–55. doi:10.1080/20964471.2018.1433790.

Fung, A.K., Li, Z. and Chen, K.S. (1992) ‘Backscatterin from randomly rough dielectric surface’, *IEEE Transactions on Geoscience and Remote Sensing*, 30(2), pp. 356–369.

Gorelick, N. *et al.* (2017) ‘Remote Sensing of Environment Google Earth Engine : Planetary-scale geospatial analysis for everyone’, *Remote Sensing of Environment*, 202, pp. 18–27. doi:10.1016/j.rse.2017.06.031.

Greifeneder, F. *et al.* (2016) ‘From Point to Pixel Scale: An Upscaling Approach for In Situ Soil Moisture Measurements’, *Vadose Zone Journal*, 15(6). doi:10.2136/vzj2015.03.0048.

Greifeneder, F. *et al.* (2017) ‘Mapping Soil Moisture Content Using Sentinel-1 and Sentinel-2: Case Studies From Kenya’, in *ISRSE-37*. Tshwane, South Africa, p. 263. Available at: http://isrse37.org/wp-content/uploads/2017/05/ISRSE-37-Abstract-Book-as-at-8-May-2017.pdf.

Greifeneder, F., Khamala, E., *et al.* (2018) ‘Detection of soil moisture anomalies based on Sentinel-1’, *Physics and Chemistry of the Earth, Parts A/B/C*, (November), pp. 1–8. doi:10.1016/J.PCE.2018.11.009.

Greifeneder, F., Notarnicola, C., *et al.* (2018) ‘The Added Value of the VH/VV Polarization-Ratio for Global Soil Moisture Estimations From Scatterometer Data’, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 1–12. doi:10.1109/JSTARS.2018.2865185.

Greifeneder, F., Notarnicola, C. and Wagner, W. (2016) ‘Using machine learning and SAR data for the upscaling of large scale modelled soil moisture in the Alps’, in *11th European Conference on Synthetic Aperture Radar, EUSAR 2016*. Hamburg: VDE, pp. 1108–1111.

Greifeneder, F., Notarnicola, C. and Wagner, W. (2021a) ‘A Machine Learning-Based Approach for Surface Soil Moisture Estimations with Google Earth Engine’, *Remote Sensing*, 13(11). doi:10.3390/rs13112099.

Greifeneder, F., Notarnicola, C. and Wagner, W. (2021b) *PYSMM*. Zenodo. doi:10.5281/ZENODO.4552813.

Gruber, A. *et al.* (2014) ‘Performance inter-comparison of soil moisture retrieval models for the MetOp-A ASCAT instrument’, in *Geoscience and Remote Sensing Symposium (IGARSS)*. doi:10.1109/IGARSS.2014.6946969.

Hedley, J., Roelfsema, C. and Phinn, S.R. (2009) ‘Efficient radiative transfer model inversion for remote sensing applications’, *Remote Sensing of Environment*, 113(11), pp. 2527–2532. doi:10.1016/j.rse.2009.07.008.

Hornacek, M. *et al.* (2012) 'Potential for High Resolution Systematic Global Surface Soil Moisture Retrieval via Change Detection Using Sentinel-1', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(4), pp. 1303–1311. doi:10.1109/JSTARS.2012.2190136.

Kolassa, J. *et al.* (2018) 'Estimating surface soil moisture from SMAP observations using a Neural Network technique', *Remote Sensing of Environment*, 204, pp. 43–59. doi:10.1016/j.rse.2017.10.045.

Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) 'ImageNet classification with deep convolutional neural networks', *Communications of the ACM*, 60(6), pp. 84–90. doi:10.1145/3065386.

Lary, D.J. *et al.* (2016) 'Machine learning in geosciences and remote sensing', *Geoscience Frontiers*, 7(1), pp. 3–10. doi:10.1016/j.gsf.2015.07.003.

Lary, D.J. *et al.* (2018) 'Machine Learning Applications for Earth Observation', in Mathieu, P.-P. and Aubrecht, C. (eds) *Earth Observation Open Science and Innovation*. Cham: Springer International Publishing, pp. 165–218. doi:10.1007/978-3-319-65633-5_8.

Lee, C. suk *et al.* (2019) 'Estimation of soil moisture using deep learning based on satellite data: a case study of South Korea', *GIScience & Remote Sensing*, 56(1), pp. 43–67. doi:10.1080/15481603.2018.1489943.

Lei, F. *et al.* (2020) 'Data assimilation of high-resolution thermal and radar remote sensing retrievals for soil moisture monitoring in a drip-irrigated vineyard', *Remote Sensing of Environment*, 239, p. 111622. doi:10.1016/j.rse.2019.111622.

Li, Q. *et al.* (2021) 'Improved daily SMAP satellite soil moisture prediction over China using deep learning model with transfer learning', *Journal of Hydrology*, 600, p. 126698. doi:10.1016/j.jhydrol.2021.126698.

Li, Y. *et al.* (2020) 'Performance Evaluation of a Neural Network Model and Two Empirical Models for Estimating Soil Moisture Based on Sentinel-1 SAR Data', *Progress In Electromagnetics Research C*, 105, p. 15.

Liu, Y., Qian, J. and Yue, H. (2021) 'Combined Sentinel-1A With Sentinel-2A to Estimate Soil Moisture in Farmland', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, pp. 1292–1310. doi:10.1109/JSTARS.2020.3043628.

Moosavi, V. *et al.* (2016) 'Estimation of spatially enhanced soil moisture combining remote sensing and artificial intelligence approaches', *International Journal of Remote Sensing*, 37(23), pp. 5605–5631. doi:10.1080/01431161.2016.1244366.

Naeimi, V. *et al.* (2009) 'An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations', *IEEE Transactions on Geoscience and Remote Sensing*, 47(7), pp. 1999–2013. doi:10.1109/TGRS.2008.2011617.

Paloscia, S. *et al.* (2013) ‘Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation’, *Remote Sensing of Environment*, 134, pp. 234–248. doi:10.1016/j.rse.2013.02.027.

Pasolli, L. *et al.* (2015) ‘Estimation of Soil Moisture in Mountain Areas Using SVR Technique Applied to Multiscale Active Radar Images at C-Band’, *Selected Topics in Applied Earth Observations and Remote Sensing*, 8(1), pp. 262–283. doi:JSTARS.2014.2378795.

Price, J.C. (1990) ‘Using spatial context in satellite data to infer regional scale evapotranspiration’, *IEEE Transactions on Geoscience and Remote Sensing*, 28(5), pp. 940–948. doi:10.1109/36.58983.

Qin, J. *et al.* (2013) ‘Spatial upscaling of in-situ soil moisture measurements based on MODIS-derived apparent thermal inertia’, *Remote Sensing of Environment*, 138, pp. 1–9. doi:10.1016/j.rse.2013.07.003.

Reichle, R. *et al.* (2016) *SMAP L4 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data*. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center. Available at: http://dx.doi.org/10.5067/YK70EPDHNF0L.

Rosenblatt, F. (1957) *The Perceptron-a preceiving and recognizing automation*. 85-460–1. Cornell Aeronautical Laboratory.

Santi, E. *et al.* (2014) ‘A prototype ann based algorithm for the soil moisture retrieval from l- band in view of the incoming SMAP mission’, *2014 13th Specialist Meeting on Microwave Radiometry and Remote Sensing of the Environment (MicroRad)*, (November 2015), pp. 5–9. doi:10.1109/MicroRad.2014.6878897.

Schmugge, T. (1978) ‘Remote Sensing of Surface Soil Moisture’, *Journal of Applied Meteorology and Climatology*, 17(10), pp. 1549–1557. doi:10.1175/1520-0450(1978)017<1549:RSOSSM>2.0.CO;2.

Shoshany, M. *et al.* (2000) ‘The relationship between ERS-2 SAR backscatter and soil moisture: Generalization from a humid to semi-arid transect’, *International Journal of Remote Sensing*, 21(11), pp. 2337–2343. doi:10.1080/01431160050029620.

Silva Ursulino, B. *et al.* (2019) ‘Modelling Soil Water Dynamics from Soil Hydraulic Parameters Estimated by an Alternative Method in a Tropical Experimental Basin’, *Water*, 11(5), p. 1007. doi:10.3390/w11051007.

Stamenkovic, J. *et al.* (2017) ‘Soil Moisture Estimation by SAR in Alpine Fields Using Gaussian Process Regressor Trained by Model Simulations’, *IEEE Transactions on Geoscience and Remote Sensing*, 55(9), pp. 4899–4912. doi:10.1109/TGRS.2017.2687421.

Sudmanns, M. *et al.* (2020) ‘Big Earth data: disruptive changes in Earth observation data management and analysis?’, *International Journal of Digital Earth*, 13(7), pp. 832–850. doi:10.1080/17538947.2019.1585976.

Talukdar, S. *et al.* (2020) 'Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review', *Remote Sensing*, 12(7), p. 1135. doi:10.3390/rs12071135.

Tin Kam Ho (1995) 'Random decision forests', in *Proceedings of 3rd International Conference on Document Analysis and Recognition*. *3rd International Conference on Document Analysis and Recognition*, Montreal, Que., Canada: IEEE Comput. Soc. Press, pp. 278–282. doi:10.1109/ICDAR.1995.598994.

Tyralis, H. *et al.* (2019) 'Hydrological post-processing using stacked generalization of quantile regression algorithms: Large-scale application over CONUS', *Journal of Hydrology*, 577, p. 123957. doi:10.1016/j.jhydrol.2019.123957.

Ulaby, F.T. and Batlivala, P. (1978) 'Microwave Backscatter Dependence on Surface Roughness , Soil Moisture , and Soil Texture : Part 1I-Bare Soil', *IEEE Transactions on Geoscience Electronics*, 16(4), pp. 286–295.

Ulaby, F.T., Moore, R.K. and Fung, A.K. (1986) *Microwave Remote Sensing Active and Passive-Volume III: From Theory to Applications*. Artech House.

Wagner, W., Lemoine, G. and Rott, H. (1999) 'A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data', *Remote Sensing of Environment*, 70(2), pp. 191–207. doi:10.1016/S0034-4257(99)00036-X.

Zhang, D. and Zhou, G. (2016) 'Estimation of Soil Moisture from Optical and Thermal Remote Sensing: A Review', *Sensors*, 16(8), p. 1308. doi:10.3390/s16081308.

Zhu, X.X. *et al.* (2017) 'Deep learning in remote sensing: a review', *IEEE Geoscience and Remote Sensing Magazine*, 5(4), pp. 8–36. doi:10.1109/MGRS.2017.2762307.