

The background features a vertical gradient from light green at the top to a warm orange-brown at the bottom. On the left side, there is a black silhouette of a grass plant with several long, thin blades and a dense root system extending downwards. The title text is overlaid on the upper right portion of the image.

# Investigating soil moisture dynamics for improved applications of radar satellites in agricultural water management

**Coleen Carranza**

## Propositions

1. Decoupling can assess the direct utility of radar-derived surface soil moisture in representing root zone conditions.  
(this thesis)
2. Vegetation backscatter can potentially be used to directly estimate root zone soil moisture.  
(this thesis)
3. Wholeheartedly accepting more members of the minority at every level of the scientific community will reduce the bias still existing in science.
4. Improvement of process-based models cannot be achieved without advancements in measurement techniques.
5. There are no long-term environmental solutions, only present-day solutions which will soon give rise to other problems that require new solutions.
6. Despite improved digital connectivity, modern technology is becoming a tool for widening social disconnect.
7. A two-day weekend is not enough to maintain a work-life balance in today's fast paced living.

Propositions belonging to the thesis, entitled  
Investigating soil moisture dynamics for improved applications of radar  
satellites in agricultural water management

Coleen Carranza  
Wageningen, 25 January 2021



**Investigating soil moisture dynamics for  
improved applications of radar satellites  
in agricultural water management**

Coleen Dorothy Urbano Carranza

## **Thesis committee**

### **Promotor:**

Prof. Dr Coen J. Ritsema  
Professor of Soil Physics and Land Management  
Wageningen University & Research

### **Co-promotor:**

Dr Martine J. van der Ploeg  
Associate Professor, Hydrology and Quantitative Water Management  
Wageningen University & Research

### **Other members:**

Dr Jetse Stoorvogel, Wageningen University & Research  
Prof. Dr Nick van de Giesen, Delft University of Technology  
Dr Theresa Blume, GFZ Helmholtz Centre Potsdam, Germany  
Prof. Dr Teamrat Ghezzehei, University of California, Merced, USA

This research was conducted under the auspices of the Research School for Socio-Economic and Natural Sciences of the Environment (SENSE).

# Investigating soil moisture dynamics for improved applications of radar satellites in agricultural water management

Coleen Dorothy Urbano Carranza

## **Thesis**

submitted in fulfilment of the requirements for the degree of doctor at

Wageningen University

by the authority of the Rector Magnificus

Prof. Dr A.P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Monday 25 January 2021

at 1:30 p.m. in the Aula.

Coleen Dorothy Urbano Carranza  
Investigating soil moisture dynamics for improved applications  
of radar satellites in agricultural water management  
170 pages.

PhD thesis, Wageningen University, Wageningen, NL (2021)  
With references, with summary in English

ISBN: 978-94-6395-584-3  
DOI: 10.18174/532848

*for my little Elis  
you're starting to find your voice  
make it loud  
keep it true  
keep it you*



# Contents

|  | Page       |
|--|------------|
| <b>Contents</b>  | <b>vii</b> |
| <b>Chapter 1 Introduction</b>  | <b>1</b>   |
| <b>Chapter 2 Surface - subsurface soil moisture coupling</b>                       | <b>11</b>  |
| <b>Chapter 3 Subsurface soil moisture dynamics in a cultivated field</b>           | <b>31</b>  |
| <b>Chapter 4 Can radar vegetation backscatter reflect root zone soil moisture?</b> | <b>53</b>  |
| <b>Chapter 5 Monitoring agricultural field trafficability using Sentinel-1</b>     | <b>69</b>  |
| <b>Chapter 6 Root zone soil moisture estimation using Random Forests</b>           | <b>93</b>  |
| <b>Chapter 7 Synthesis</b>   | <b>117</b> |
| <b>Chapter 8 Appendix</b>  | <b>129</b> |
| <b>Bibliography</b>  | <b>135</b> |
| <b>Data Availability</b>   | <b>154</b> |
| <b>Summary</b>   | <b>156</b> |
| <b>Acknowledgements</b>  | <b>160</b> |
| <b>About the author</b>  | <b>165</b> |





*"You can't go back and change the beginning, but you can start where you are and change the ending."*

C.S. Lewis

# 1

## Introduction



## 1.1 Context

Soil moisture dynamics are a fundamental factor in agricultural water management, and are highly variable both spatially and temporally. Accurate estimates of soil moisture status are essential for crop production, soil quality and resource use, but may be difficult to monitor over finer spatio-temporal scales relevant for water management. Radar satellites have potential for providing this information; however translating satellite data into agriculturally useful soil moisture remains a challenge which has prevented full realization of the potential. This PhD project takes a new, innovative look at soil moisture dynamics to discover how radar-derived soil moisture can be directly applied to generate information that support of more efficient and sustainable agriculture.

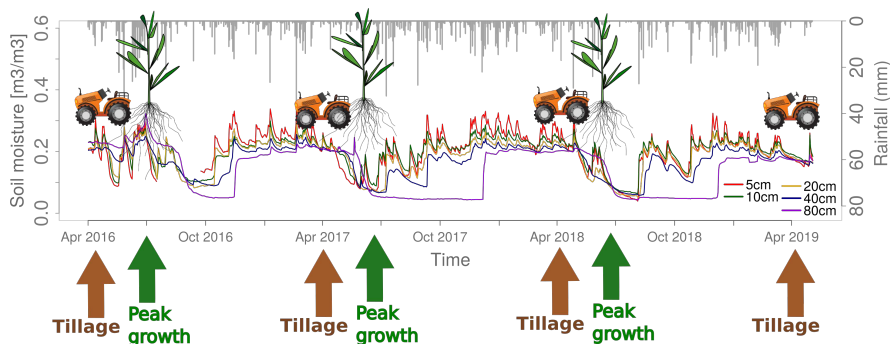
### 1.1.1 Agricultural water management

Soil moisture plays a central role in determining efficient land water management strategies in agricultural areas (Gao et al., 2014; Rivers et al., 2015; Petropoulos et al., 2018). It is a dynamic and highly non-linear environmental variable whose state is controlled by complex interactions of meteorological conditions, vegetation characteristics and responses, and soil properties (Hillel, 1998; Hartge et al., 2016). Because of its dynamic nature, its status should be monitored in order for farmers to efficiently manage their lands throughout the growing season. One of the more obvious and main roles of soil moisture is providing available water for crops (Allen et al., 2005). Aside from this, soil moisture influences the mechanical properties of soils which allows it to support any overlying weight (Ayers and Perumpral, 1982). Therefore, soil moisture information is crucial for selection of optimal tillage practices and for determining irrigation schedules.

In any given area, the general climatology dictates long term or seasonal soil moisture trends, which influence farming schedules and activities. Land and water management practices, in particular tillage and irrigation, are generally adapted to the expected soil moisture status at different stages of a growing season. For example, in temperate areas, a majority of the crops are grown over the summer months (June - August) because of longer day time hours and higher atmospheric temperatures. Since a predefined number of days is already known for crop growth for each specific crop type, tillage is usually initiated over a short span of days in early spring when the soils are still wet and may even be close to saturation. Higher soil moisture levels, however, increase the susceptibility of soil to compaction and lessen mobility of heavy farming machinery (Mosaddeghi et al., 2000). Such conditions ideally require assessment of trafficability to ensure efficient use of the heavy mobile machinery which is now indispensable for tillage and harvesting (Droogers et al., 1996; Müller et al., 2011). In addition, the long term use of heavy machinery leads to a prevalence of subsoil compaction in many agricultural areas (Défossez and Richard, 2002; Raper, 2005). Assessment of trafficability will, therefore, also mitigate compaction rates currently encountered in agricultural fields. After sowing and until crop maturity, ensuring adequate water within the crop's root zone is important for significant productivity levels. Crop growth coincides with increasing evapotranspiration demands, which further translates to increased crop water requirements. Soil moisture throughout the root

zone will determine the amount of water available for uptake. Therefore, monitoring root zone soil moisture will help determine the necessity of irrigation as well as its timing and frequency (Vellidis et al., 2008; Zotarelli et al., 2011; Brocca et al., 2018). Clearly, soil moisture information is essential for farmers throughout the entire year (fig.1.1).

Because present day agriculture is carried out over large areas, optimal and efficient land and water management strategies are needed in order to meet increasing crop productivity and resource efficiency demands. These strategies are geared not only toward high biomass production or business sustainability and profitability, but also toward preservation of soil quality and health (Roger-Estrade et al., 2010; Powlson et al., 2011; Robertson and Swinton, 2005). New and smart technologies which maximize farming activities over agricultural lands and minimize energy and production costs are becoming popular and are continuously being adopted. For instance, remote sensing is becoming an integral part of modern agriculture because of its importance for precision farming (Gebbers and Adamchuk, 2010).



**Figure 1.1:** Matching soil moisture with water requirements throughout a growing season. Soil moisture dynamics are presented by the measurements at several depths. The periods for tillage and crop growth are also indicated.

### 1.1.2 Mapping and monitoring soil moisture with radar satellites

Mapping and monitoring soil moisture conditions in agricultural areas have been achieved using microwave remote sensing (Henderson and Lewis, 1998; Steele-Dunne et al., 2017). Satellites provide a good alternative to in situ soil moisture sensors because installation of spatially distributed measurements can be a tedious task. The microwave or radio portion of the electromagnetic spectrum covers the wavelength range from approximately 1 cm to 1 m. Longer wavelengths of the microwave spectrum are not susceptible to atmospheric scattering which affects shorter optical wavelengths. Microwaves can penetrate through cloud cover, haze, dust, and all but the heaviest rainfall, which allows detection of microwave energy under almost all weather and environmental conditions. However, microwave satellites only measure soil moisture at the the upper surface layer ( $\leq 5$  cm) while most agricultural applications require soil moisture measured over deeper layers, and can be up to 1 m deep (e.g. over the root zone).

There are two forms of microwave remote sensing, the active and passive. In the passive form, energy received by the sensors is related to the temperature and moisture properties of the emitting object or surface. The microwave energy recorded by a passive sensor can be emitted by the atmosphere, reflected from the surface, emitted from the surface, or transmitted from the subsurface. Most passive microwave sensors are characterized by low spatial resolution because of the large fields of view needed to detect sufficient energy to record a signal. Active microwave sensors use a self-contained source of microwave radiation to illuminate the target (Ulaby et al., 1982; Fung, 1994) and are generally divided into imaging and non-imaging categories. At present, the most common forms of imaging active microwave sensors are Synthetic Aperture Radars (SAR). The sensor transmits a microwave signal towards the target and detects the portion scattered back. Upon contact with the surface, the incident energy from the transmitted signal may be reflected, scattered, absorbed (attenuated) or it may penetrate into the medium and be further absorbed or scattered (Moreira et al., 2013). The amount of energy backscattered is dependent on the surface properties and the angle at which the microwave energy strikes the target. The strength of the backscattered signal is used to discriminate between different targets and the time delay between the transmitted and reflected signals determines the distance to the target (e.g. DEMs). SAR utilizes the motion of the antenna along the azimuthal direction to “synthesize” or give the effect of a long antenna, which allows imaging at higher spatial resolutions. It therefore solves the problem of conventional radar systems, that achieve higher spatial resolutions with longer antennas but are prohibitive for flying or outer space.

Using SAR, the amount of energy backscattered from a natural surface is dependent on radar configuration, soil, and vegetation characteristics (Moreira et al., 2013). The radar configuration sets the length and angle of the incoming microwave signal, which determines the type of interactions between the signal and target ground objects. Among the recently launch radar satellites, Sentinel-1 (C-band, 5.405 GHz) has a high revisit time of 1 - 3 days in European regions (Torres et al., 2012) which allows frequent monitoring of surface soil moisture conditions. Furthermore, its 10 m pixel size enables estimation of soil moisture variability at quite high spatial-resolutions. At the C-band, the total backscatter over vegetated agricultural fields is the integrated signal of the vegetation canopy and underlying surface soil layer. Both soil and vegetation components need to be quantified to further estimate surface soil moisture. Surface scattering models over bare soils, which may be process-based (Integral Equation model (IEM, Fung et al., 1992) or semi-empirical (Oh and Kay, 1998 and Dubois et al., 1995), have been applied to quantify the soil backscatter. During vegetated periods, the volume scattering from stalks and leaves and scattering from soil-vegetation interaction are additional components included in the total backscatter. Improved soil moisture estimates are achieved when the vegetation component of the total backscatter is accounted for in most soil moisture retrieval algorithms (e.g. Bindlish and Barros, 2001; Notarnicola et al., 2006; Joseph et al., 2010).

Integration of satellite-derived soil moisture with process-based hydrological models has been achieved through data assimilation (e.g. Das and Mohanty, 2006; Reichle et al., 2004; Draper et al., 2012; Renzullo et al., 2014). However, data-driven and statistical methods have also been applied based on empirical relations between surface and subsurface soil moisture to esti-

mate the latter (e.g. González-Zamora et al., 2016; Gao et al., 2019; Pezij et al., 2020).

### 1.1.3 Soil moisture dynamics

Characterizing soil moisture dynamics across different spatio-temporal scales enable the identification of controls and estimation of parameters for model simulations. At the field or catchment scales, soil moisture variability is influenced by meteorology, topography, vegetation and soil properties. Grayson et al. (1997) grouped soil moisture controls into local and non-local controls. Non-local controls are related to lateral fluxes being dominant under wetter conditions (Famiglietti et al., 1998; Qiu et al., 2001; Zhu and Lin, 2011) whereas local controls are dominant under drier conditions and are related to vertical fluxes (Joshi and Mohanty, 2010; Zhao et al., 2010; Baroni et al., 2013).

Soil hydraulic properties, which are largely determined by physical properties, exert a major control on soil moisture variability. Both textural composition and structural arrangement of soil fragments exert their influence on soil hydraulic properties. Soil structural changes have direct impacts on hydraulic properties and soil moisture estimation (Mapa et al., 1986; Moret and Arrúe, 2007) may persist in both the short- and long-term (Pagliai et al., 2004; Strudley et al., 2008). Land management practices in agricultural fields further contribute to changes in soil physical properties including bulk density (Mallory et al., 2011; Osunbitan et al., 2005), aggregate stability (Paul et al., 2013; Zheng et al., 2018) and macroporosity (Roseberg and McCoy, 1992; Pagliai et al., 2004; Malone et al., 2003). Studies have suggested incorporating agricultural land use and management changes to parameterize hydrological models for a more accurate representation of soil structure and variability in hydraulic properties (e.g. Gonzalez-Sosa et al., 2010; Ndiaye et al., 2007).

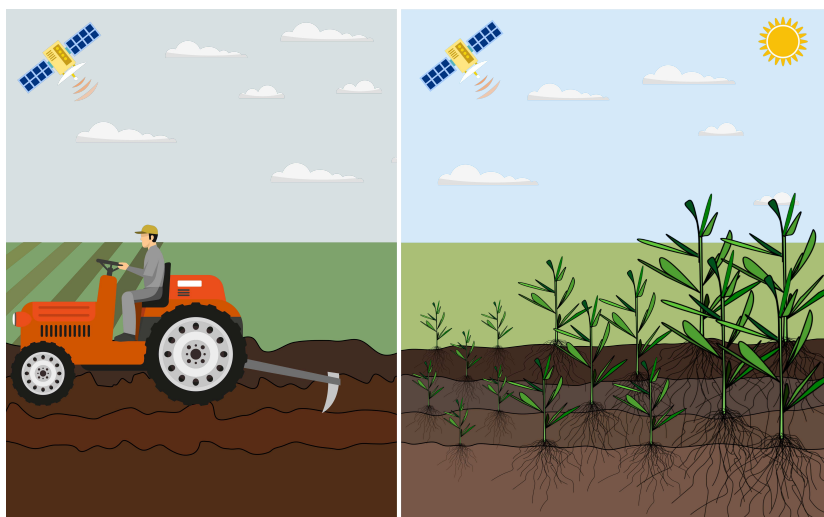
In addition to soil properties, vegetation affects soil moisture dynamics directly through root water uptake and indirectly through root-induced changes to soil structure. The acquisition of water from the soil, a consequence of stomatal opening for photosynthesis, is mostly dictated by transpiration demands. Most studies on vegetation-soil moisture interactions have focused on dry or drying conditions because of the negative consequences for vegetation and agricultural yield (Passioura, 1994; Li et al., 2009). During such periods or in water-controlled environments, vegetation can be both water and nutrient deficient since nutrient availability is limited by water deficits (Porporato et al., 2003; D'odorico et al., 2003). In terms of soil moisture dynamics, Hupet and Vanclooster (2002) showed that vegetation plays a non-negligible role in the observed soil moisture variability within a cultivated field. Furthermore, a hydrogeophysical investigation of root zone soil moisture under two agricultural crops by Srayeddin and Doussan (2009) found that the variability in root water uptake which they attributed to rooting patterns and densities, is influenced by water availability in the root zone. Roots indirectly affect soil moisture dynamics by altering soil hydraulic properties. A recent review by Lu et al. (2020) reports contrasting root effects on soil hydraulic properties depending on which processes are dominant (e.g. root growth (or decay), root density, or diameter). In addition, they noted that the temporal variability arising from agricultural cycles as an important aspect of root-induced modifications to soil hydraulic properties.

## 1.2 Motivation

In the Netherlands, the existing network of canals and waterways over agricultural areas are a potential resource for operational water management. Furthermore, the Netherlands Hydrological Instrument (NHI, De Lange et al., 2014), a national scale modeling framework is already in place for interconnectivity or interoperability of hydrological models that have been developed for specific components of the hydrologic cycle (i.e. surface water, groundwater, unsaturated zone). According to Xu and Tung (2008), operational water management requires decision-making within limited time intervals and involves multiple criteria related to water supply and mitigating water-related hazards.

Radar satellites such as Sentinel-1 can offer soil moisture information at frequent time intervals that could, potentially, be integrated within an operational framework. Several years ago, Wagner et al. (2007) suggested that operational coarse-resolution (25 – 50 km) soil moisture products can be expected in the near future from radiometer and scatterometer systems. True enough, there are already several global coarse soil moisture products available (e.g. SMOS L2/L3 soil moisture, SMAP L2/L3 soil moisture, see Petropoulos et al., 2018). Petropoulos et al. (2018) further note that despite this overwhelming global need, operational earth observation soil moisture products have not yet been widely adopted by the agricultural community. A key technical barrier may be the limited availability of operational soil moisture estimates at high spatio-temporal resolutions. In relation to this, in a study within selected water authorities in the Netherlands, Pezij et al. (2019a) found that operational water managers have indicated the need to access high-resolution spatial data, value-added products and tools for communication to stakeholders. Furthermore, operational water managers may struggle to correctly interpret hydrological model output and therefore need suitable indicators for evidence-based decision-making. More recently, Wagner et al. (2009) and Petropoulos et al. (2018) suggested that the next few years may also see operational soil moisture products at finer (< 1 km) scales based on Sentinel-1 or from downscaled and disaggregated soil moisture products. Therefore, scientific investigations concurrent with efforts to generate operational soil moisture products would be beneficial in determining advantages and limitations of Sentinel-1 in operational (water) management in agricultural areas.

Since Sentinel-1 and other radar satellites only measure surface soil moisture, roughly the upper 5 cm of the soil profile, root zone soil moisture conditions needed in most agricultural applications cannot currently be accurately captured from satellite-derived soil moisture alone. Previous studies have shown that surface soil moisture can be decoupled with deeper layers (e.g. Capehart and Carlson, 1997; Kumar et al., 2009; Hirschi et al., 2014) and may lead to inaccurate estimates of root zone soil moisture. Furthermore, previous studies investigating soil moisture coupling were applied at continental or global scales using land surface models. The coupling behavior of soil moisture in agricultural areas derived from in situ soil moisture measurements are still limited in comparison with model simulated values. However, in situ measurements are valuable for estimating fluxes and for retrieval of spatial and temporal dynamics of soil moisture profiles, especially at the field and catchment scales (Vereecken et al., 2008).



**Figure 1.2:** Monitoring soil moisture status with satellites for operational water management in agricultural areas. The left image portrays Sentinel-1 mapping moisture for tillage to assess mobility and susceptibility to compaction while the right image shows it mapping crop water requirements later in the growing season.

Agricultural water requirements depend on atmospheric conditions and different periods of a growing season. The information contained in radar satellites, such as Sentinel-1, needs to be assessed and matched with agricultural water requirements to determine the its benefits within operational water management framework. Factors controlling field scale soil moisture variability during specific periods within or outside the growing seasons should be determined and linked with water requirements. In order to ultimately obtain measures or indicators for root zone soil moisture that are readily usable for farmers, appropriate methods are needed to balance the information contained in satellite-derived soil moisture and that available in other datasets.

### 1.3 Research questions and thesis outline

In this thesis, the main goal is to investigate the impact of soil moisture dynamics on the ability of radar satellites to provide usable information on agricultural water requirements. With operational water management as one of the primary applications in mind, suitable methods are investigated for matching the soil moisture information contained in radar satellites with the water requirements at different stages of a growing season (fig.1.2).

The following research questions are posed:

- How does soil moisture vary with depth, and under which conditions are surface and subsurface soil moisture coupled?

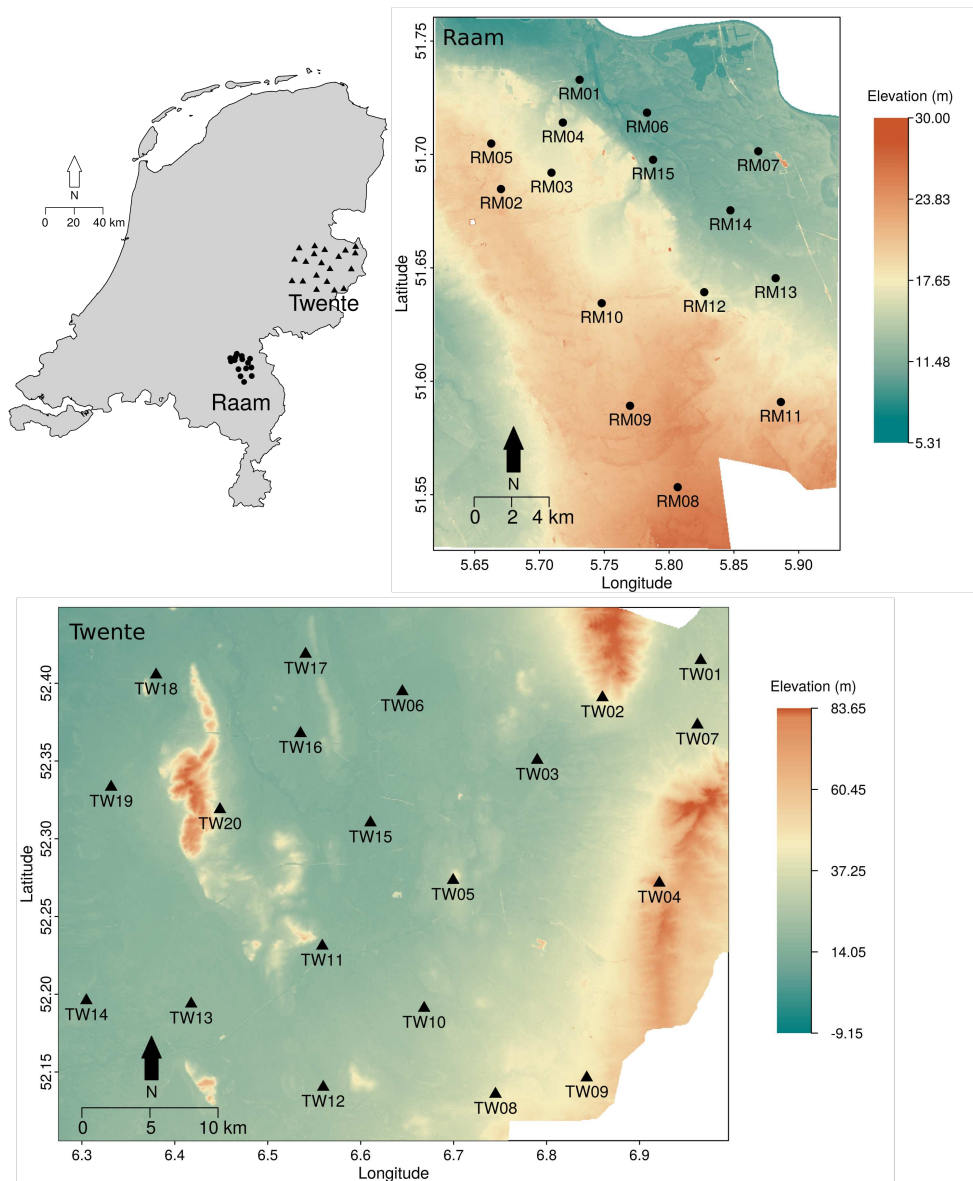
- How does subsurface soil moisture in agricultural areas vary in space and time?
- During dry periods, does vegetation backscatter from Sentinel-1 have the potential to reflect root zone soil moisture?
- During saturated conditions, how can satellite-derived surface soil moisture be used for management of agricultural fields?
- Can a data-driven approach based on surface conditions accurately estimate root zone soil moisture?

In the succeeding sections of this thesis, each chapter addresses one research question (Chapter 2 - 6). Chapter 2 focuses on the vertical variability of soil moisture and how this impacts the ability of satellite-derived soil moisture to represent subsurface soil moisture. In Chapter 3, the field scale soil moisture dynamics over a cultivated field is investigated to quantify variability and describe the controls over space and time. Chapter 4 focuses on the potential of the vegetation component of the total radar backscatter, as an alternative to the soil component, for directly measuring root zone soil moisture. Chapter 5 demonstrates a method to estimate agricultural field trafficability using insights from the results in Chapter 2. Chapter 6 illustrates the capabilities of a purely data-driven method for estimating root zone soil moisture. Each chapter is written as a stand alone journal article format which may result in repetition of some sections in the introduction and methods in the chapters. This is followed by a Synthesis (Chapter 7) where the findings are integrated and discussed in relation to the overall objective of this thesis.

## 1.4 Study areas

Two soil moisture monitoring networks in the Netherlands, Raam and Twente, were used in this thesis. The Twente soil moisture network in the eastern part of the country was established in 2009 (Dente et al., 2011). It was meant primarily as a regional monitoring network for the validation of remote sensing soil moisture products. A total of 20 stations were installed and distributed over an area of approximately 1306 km<sup>2</sup>. The Raam soil moisture network (Benninga et al., 2018) was established in 2016 in the southeastern portion of the country. A total of 15 stations are distributed over the Raam catchment and cover an area of approximately 237 km<sup>2</sup>, hence smaller than the Twente network. In both networks, soil moisture sensors were installed at discrete depths (5, 10, 20, 40, and/or 80 cm). The stations are located in generally flat agricultural fields (fig.1.3) where the most commonly encountered crop is grass, followed by corn, winter wheat, potatoes, sugar beets, and other vegetables. The soils in both networks are mostly sandy, with the stations in Twente network holding slightly higher loam content than those in the Raam network (Benninga et al., 2018; Dente et al., 2011).





**Figure 1.3:** Maps showing the study areas within the Netherlands. The figure indicates the locations of points until 2017. For Twente, some of the stations have been relocated after 2017 because of logistic concerns. Details of the specific changes made for each station are not known. However, the relocated stations were not part of the analysis in the research chapters. Elevation shown is above mean sea level.



"Even with two eyes, you only see half of the picture."

Odin, Thor: Ragnarok

# 2

## Surface - subsurface soil moisture coupling



This chapter is based on:

Carranza, C., Van der Ploeg, M. J., and Torfs, P. J. (2018). Using lagged dependence to identify (de)coupled surface and subsurface soil moisture values, *Hydrology and Earth System Sciences* 22, 2255–2267, doi: 10.5194/hess-22-2255-2018

## Abstract

Recent advances in radar remote sensing allow mapping of surface soil moisture at different spatial scales. Furthermore, satellite-derived surface soil moisture measurements can be assimilated into hydrological models for improved estimates of subsurface soil moisture values. However, variability of soil moisture across the soil column is important for accurate estimation of depth-integrated values as decoupling between surface and subsurface can occur. In this study, we employ new methods to investigate the occurrence of (de)coupling between surface and subsurface soil moisture. Using time series datasets, lagged dependence was incorporated in assessing (de)coupling with the idea that surface soil moisture conditions will be reflected at the subsurface after a certain delay. The main approach involves the application of a distributed lag non-linear model (DLNM) to simultaneously represent both the functional relation and the lag structure in the time series. The results of an exploratory analysis using residuals from a fitted loess function serve as a posteriori information to determine (de)coupled values. Both methods allow for a range of (de)coupled soil moisture values to be quantified. Results provide new insights into the decoupled range, especially as we found that its occurrence among the sites investigated is not limited to dry conditions.

## 2.1 Introduction

Although recent decades have seen great advances in remote sensing applications for mapping surface soil moisture (Jackson, 1993; Njoku et al., 2003; Mohanty et al., 2017), most hydrological studies that make use of soil moisture data require integrated values over a certain soil depth (Brocca et al., 2017). Extrapolation of surface soil moisture from remote sensing techniques to depths beyond the sensor's capacity (up to 5 cm) is not a trivial task given the spatio-temporal variability of soil moisture. The vertical distribution of soil moisture, which determines integrated soil moisture content over a soil column, is rarely uniform as more pronounced dynamics are expected closer to the surface compared to deeper in the soil (Hupet and Vanclooster, 2002). Currently, information derived from remote sensing is assimilated into hydrological models to obtain integrated soil moisture values (Houser et al., 1998; Das et al., 2008). However, Kumar et al. (2009) stressed that it is important to assess vertical variability, especially the strength of coupling between surface and subsurface soil moisture for improvement of data assimilation results. Analyses of vertical soil moisture distributions also have important implications for modeling studies, as they could be used for calibration or validation of model parameters (De Lannoy et al., 2006).

The amount of soil moisture at any given time is controlled by factors operating at different time scales. While prevailing atmospheric conditions directly affect surface layers and control the temporal dynamics of soil moisture (Albertson and Montaldo, 2003; Koster et al., 2004), it is the downward movement of water from the surface that dictates the amount of subsurface soil moisture at a given time (Belmans et al., 1983; Rodriguez-Iturbe et al., 1999). Flow rates to the subsurface are driven by hydraulic properties, which are in turn controlled by physical soil characteristics such as texture, bulk density, and structure. Relative to changes in atmospheric conditions, soil physical properties change over longer timescales. Vegetation further modifies vertical soil moisture distribution by root water uptake (Yu et al., 2007) and by changing soil structure (Angers, 1998).

Given the variability along a soil column, under which conditions do surface soil moisture reflect subsurface soil moisture? Several studies have investigated this relation to address the correspondence between surface and subsurface soil moisture. One of the earliest studies was done by Capehart and Carlson (1997) wherein they compared modeling outputs with remote sensing measurements. Using very shallow depths of 5 mm and 5 cm, they referred to decoupling as the deviation from a linear correlation between these depths due to variable drying rates. Further assessments of decoupling from model-generated time series soil moisture data have been investigated using cross-correlation values (Martinez et al., 2008; Mahmood et al., 2012; Ford et al., 2014). High correlation to the subsurface was obtained using lagged values of surface soil moisture. However, cross-correlation is limited to providing a single value throughout the range of soil moisture encountered per lag. Furthermore, cross-correlation generally aims to evaluate the strength of lagged linear dependence between two variables (Shumway and Stoffer, 2010). However, lagged dependence between surface and subsurface soil moisture may not be linear given that non-linear processes determine water flow along the soil profile. Using in situ field measurements, Wilson et al. (2003) investigated spatial

surface (0-6 cm) and subsurface (0-30 cm) soil moisture distribution by calculating statistical metrics and by means of a variogram. Decoupling between the two depths was observed which they suggested to be influenced by vegetation, especially root density at surface soil. Their results were also affected by the dry soil moisture range and emphasized the importance of distinguishing between surface and total soil moisture for future applications of remote sensing to atmospheric studies.

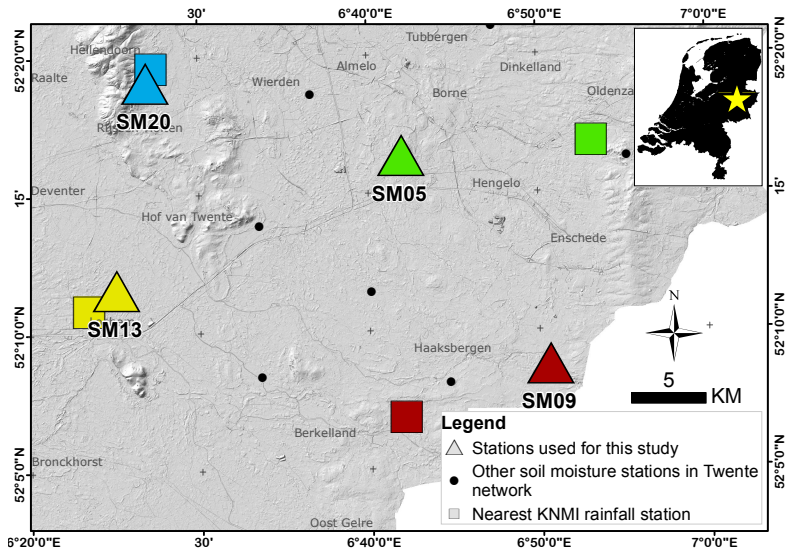
Based on previous studies, the term decoupling refers to a weak dependence between soil moisture contents at the surface and subsurface. Recognition of decoupling is important, however most studies have been limited to providing qualitative characterization of conditions when decoupling occurs (e.g. dry period). Only Capehart and Carlson (1997) identified a mid-range soil moisture ( $\sim 0.3 \text{ cm}^3 \text{ cm}^{-3}$ ) when the surface and very near surface begin to decouple. Their results, however, are limited to a thin layer of the soil column. In this paper, our main objective is to quantitatively identify a range of surface soil moisture values that is decoupled from the subsurface. Furthermore, we consider depths greater than those investigated by Capehart and Carlson (1997). The ability to quantify (de)coupled surface and subsurface soil moisture contents will contribute to more effective estimation of depth-integrated soil moisture data using remote sensing methods and improved data assimilation results in hydrological models.

We utilized in situ time series datasets at depths of 5 cm and 40 cm to represent surface and subsurface, respectively. Values outside the decoupled range are considered coupled since soil moisture is inherently bounded up a maximum value equal to soil's porosity. The investigation of (de)coupling is based on the idea that surface conditions will be reflected at the subsurface after a certain delay indicating strong coupling between the two zones, and vice versa. More focus is given to the decoupled soil moisture range since it has greater implications for extrapolation of surface soil moisture values to deeper soil layers. We applied statistical methods to identify conditions of decoupling with no prior assumptions on the type of functional relation between surface and subsurface. As an exploratory step, we first assessed the dependence without considering lags using regression and residuals analysis. The main approach for assessing decoupling was application of distributed lag non-linear models (Gasparrini et al., 2010) to incorporate both the lag structure and the functional relation between surface and subsurface soil moisture. Applications of distributed lag models to econometrics and environmental epidemiology have been well documented (Almon, 1965; Zanobetti et al., 2002; Bhaskaran et al., 2013; Wu et al., 2013). However, their application to hydrological studies have rarely been explored.

## 2.2 Material and methods

### 2.2.1 Description of datasets and study sites

Four time series datasets from the Twente soil moisture and temperature monitoring network (Dente et al., 2011) were used in this study (fig.2.1). Datasets from 2014 to 2016 are available with only short periods of missing data. The stations are located in agricultural fields with

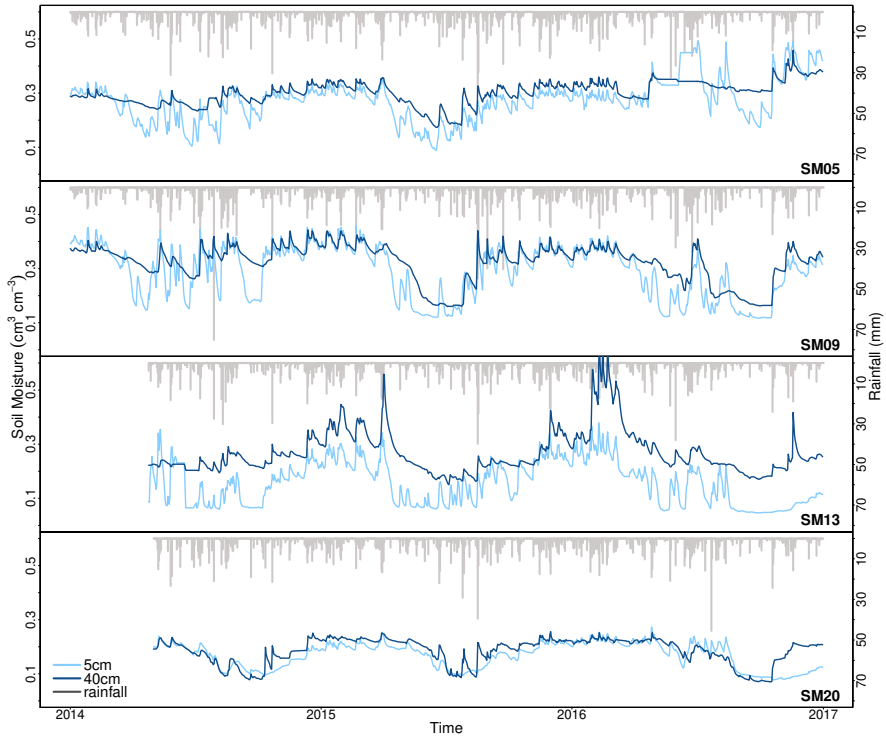


**Figure 2.1:** Location of study site in the eastern part of the Netherlands (inset). Triangles represent stations used within the Twente soil moisture and temperature monitoring network (Dente et al., 2011). Squares represent meteorological stations. Symbols with similar colors indicate the pair of measurements used for the analysis.

sensors installed at 5 cm, 10 cm, 20 cm, and 40 cm depths. To investigate decoupling, only the 5 cm and 40 cm depths were considered because the largest possible distance was desired. Each station consists of EC-TM ECH2O capacitance probes (Decagon Devices, Inc., USA) that logged soil moisture data every 15 minutes. A calibration procedure using gravimetric measurements was applied prior to analysis (Dente et al., 2011).

Land cover in the area varies from corn in one field (SM05), to grass in two fields (SM05 and SM13), to a forest area (SM20). Values at 40 cm capture the root zone of vegetation for each site. In reality, rooting depths vary and depend on species composition, climate, and plant growth rate. However, the depth considered would still allow for approximation of root zone conditions. The landscape is characterized by flat to slightly sloping terrain. It is important to note that SM20 is located at the eastern foot of a small hilly terrain. Throughout the study period, either land cover remained unchanged or the same crop was planted. The soil types for the stations range from coarse sandy soils to weakly silty soils (Wosten et al., 2013). A summary of the land cover and relevant characteristics of the stations are given in table 2.1.

Soil moisture values were averaged into daily values to match the available daily rainfall data from the Dutch national weather service (KNMI). For SM13 and SM20, there are some missing data from the beginning of 2014. The datasets from SM13 begins on April 25, 2014 while SM20 begins on May 2, 2015 (fig.2.2).

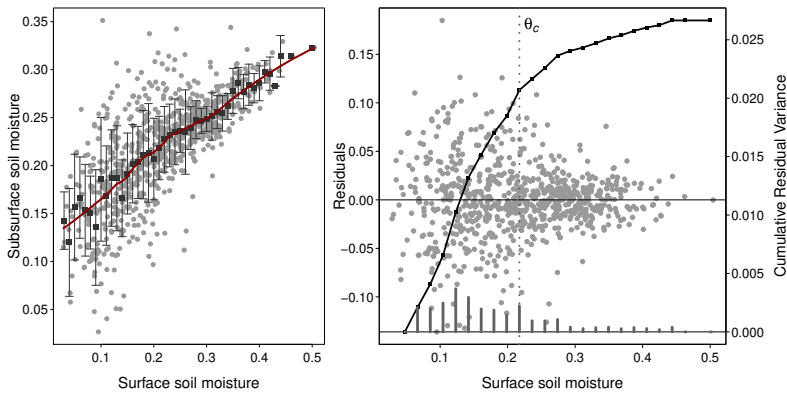


**Figure 2.2:** Time series plots of surface (5 cm in light blue) and subsurface (40 cm in dark blue) soil moisture. Vertical black bars at the top show daily precipitation data from the nearest KNMI station.

**Table 2.1:** Summary of land cover descriptions at each station covering the period of 2014 to 2016. Soil descriptions and codes are based on BOFEK 2012 (Wosten et al., 2013). Both the slope and distance to nearest ditch were determined from 5m resolution DEM. Datasets are from 2016 and were obtained from the publicly available national topographic database of the Netherlands (TOPNL).

| Station No. | Land cover | BOFEK Soil description                                      | Slope (degrees) | Aerial distance to nearest ditch (m) |
|-------------|------------|---|-----------------|--------------------------------------|
| SM05        | Grass      | Loamy sandy soils with a thick cultivated layer (317)       | 2.22            | 18.97                                |
| SM09        | Corn       | Weakly loamy sand soils with a thick cultivated layer (311) | 2.70            | 1.41                                 |
| SM13        | Grass      | Weak silty soils (podsoils)(304)                            | 1.0             | 17.09                                |
| SM20        | Forest     | Coarse sand (podsoils)(320)                                 | 2.30            | 875.26                               |





**Figure 2.3:** Schematic diagram using hypothetical soil moisture values to show vertical variability. Left: Scatterplot showing the trend with a fitted loess function. The variability can be seen using the standard deviation bars. Right: Scatterplot of the residuals from the fitted function. Soil moisture variability is visible from the variance given by the vertical bars and the cumulative residuals variance given by the black line. A change in variability at an intermediate soil moisture value is marked by a change in the slope of the cumulative variance line, indicated by  $\theta_c$

### 2.2.2 Regression and residuals analysis

As an exploratory step, the dependence between surface and subsurface soil moisture was initially visualized using scatterplots. Conditional means for every  $0.01 \text{ cm}^3 \text{ cm}^{-3}$  interval and vertical bars representing  $\pm$  standard deviation were added to show changes in vertical variability across the soil moisture range. Longer standard deviation bars indicate higher vertical variability (fig.2.3, left). We referred to vertical variability as the unevenness or irregularity in the soil moisture distribution within a certain depth of interest along the vertical profile, in this case up to depths of 40 cm. For the rest of this paper, variability will refer to vertical soil moisture variability, unless otherwise stated. The plotted points were colored per month to show any impacts of seasonality. The effect of rainfall was included by adjusting the sizes of the points proportional to rainfall intensity measured from the nearest KNMI stations. For the overall measure of dependence, Spearman's rank correlation coefficient  $R_s$  was computed for every pair of ranked values in the time series. This was chosen as the assumption of linear dependence was not made.

A flexible non-parametric locally weighted regression function (commonly called a loess function, Cleveland and Devlin, 1988) was fitted along the soil moisture range. This was used to explore and identify trends across the soil moisture range. A linear regression was also fitted only for comparison. Residuals were analyzed further for variability not captured by the fitted function (fig.2.3, right). The residuals variance for every  $0.1 \text{ cm}^3 \text{ cm}^{-3}$  interval as well as the resulting cumulative residuals variance were analyzed to examine variability across the range. The degree of variability was related to the slope of the cumulative variance line, with steep slopes indicating high variability. In addition, a significant change in variance between two points was indicated by a significant change in the slope of the line. The soil moisture value

where a significant change in slope occurred was marked by  $\theta_c$ , this divides the soil moisture range into two groups. The group with a steeper slope was interpreted as the decoupled range, and vice versa. Since the measured variance is sensitive to sample size, a correlation coefficient was calculated to determine if there was significant dependence between the two variables. Residuals variance were first normalized from 0 to 1 because of the varying soil moisture range encountered at each station.

Results of the exploratory methods were considered *a posteriori* knowledge for analysis of lagged dependence and interpretation of results.

### 2.2.3 Analysis of Lagged Dependence

#### *Cross correlation*

Since decoupling is based on the strength of lagged dependence, the existence of lag between surface and subsurface soil moisture values was first determined. Cross-correlation is known to be a quick and easy method to apply for this objective. Lagged values of surface soil moisture were correlated with instantaneous values at the subsurface. A maximum cross-correlation at negative lags indicated that surface soil moisture is leading subsurface soil moisture, and vice versa (Shumway and Stoffer, 2010). A 10-day lag was deemed long enough to show the presence of lag-lead relations in the time series since the maximum correlation occurred within this period.

#### *Distributed lag non-linear model*

We incorporated delayed or lagged effects in evaluating the relation between surface and subsurface values, and eventually in determining the (de)coupled values. It should be emphasized that the analysis was primarily focused on examining the trends and relation between surface and subsurface soil moisture. Moreover, it was not intended to replace other existing models for estimating soil moisture or examining its patterns.

A distributed lag non-linear model (DLNM) developed by Gasparrini et al. (2010) was applied to the 5 cm and 40 cm time series datasets at the study sites. Briefly, the model is capable of simultaneously representing both functional the dependence and delayed response between exposure and response values. We considered surface soil moisture as the exposure values that produced delayed effects to the response values at the subsurface. A non-linear model was selected in order to capture the non-linear dynamics of flow and transport along the soil profile (Mohanty and Skaggs, 2001; Kim and Barros, 2002). Furthermore, DLNM offered enough flexibility to model a variety of dependencies in the time series dataset by selecting a suitable basis function. As an analogy, a DLNM is to a linear time series model (e.g. autoregressive model) just as a generalized linear model is to a linear model, as can be seen in Eq. 2.1.

In assessing lagged dependence, event scale patterns were of interest rather than large scale trends within the time series (Wilson et al., 2004). This required seasonal patterns to be addressed prior to applying the DLNM. This was done by fitting a loess function to the time series and then subtracting it from the original soil moisture values (Cleveland et al., 1990).

Removal of seasonality was further justified by the scatterplot results (see Section 2.3.1). The influence of seasonality on the vertical soil moisture variability is indicated by clustering of observation points occurring within the same months (fig.2.4). De-seasonalized soil moisture values were used for identifying (de)coupled soil moisture conditions.

For consistency in modeling, the range of surface soil moisture values used was from 0-0.50 cm<sup>3</sup> cm<sup>-3</sup>. This was based on the highest surface soil moisture value encountered among the four sites. A lag value of up to 30 days was considered long enough to investigate delayed effects. This period also approximated the recurrence of heavy rainfall within the study sites. A spline function was the basis function chosen to represent the functional dependence and delayed effects as it offered flexibility to capture non-linearities. In addition, contributions from daily rainfall data were used to incorporate current and past meteorological conditions. This was applied as a covariate and was represented with an additional basis function. We only considered delayed effects in vertical flow as lateral movement is deemed negligible in a flat to slightly sloping terrain (table 2.1). The analysis was performed in R software using *dlnm* (Gasparrini, 2011) and *mgcv* (Wood, 2006a) packages.

The following section concisely describes the mathematical formulation of a DLNM. However, the reader may choose to skip this section as the general description of the methods applied have already been given in the text above. For a more detailed explanation, readers are referred to Gasparrini et al. (2010) and Gasparrini et al. (2017).

To more formally describe a DLNM, let us first consider a general time series model, where outcomes  $Y_t$  with  $t = 1, \dots, n$  can be described by:

$$g(\mu_t) = \alpha + \sum_{j=1}^J s_j(x_{tj}; \beta_j) + \sum_{k=1}^K \gamma_k u_{tk} \quad (2.1)$$

where  $\mu \equiv E(Y)$ , assumed to be derived from a Poisson distribution, and  $g$  is a monotonic link function. The functions  $s_j$  denote relationships between the variables  $x_j$  and vector parameters  $\beta_j$ . Other  $u_k$  variables with predictors are included in coefficients  $\gamma_k$  to specify their related effects. The relation between  $x$  and  $g(\mu)$  is represented by  $s(x)$  through a basis function. The complexity of this estimated relationship depends on the type basis function chosen and its dimensions. In the presence of delayed effects, the outcome  $Y$  at any time  $t$  is explained by the past exposures  $x_{t-l}$  with  $l$  as the *lag* representing the elapsed time between exposure and response. The final goal of a DLNM is to simultaneously describe the dependency along both the predictor space and lag dimension. This is achieved by selecting two sets of basis functions that are combined to obtain the cross-basis functions (Gasparrini et al., 2010).

Within the DLNM framework, a response  $Y_t$  at time  $t = 1$  is based on lagged occurrences of predictor  $x_t$ , which is represented by vector  $q_t = [x_{t-l_0}; \dots; x_{t-L}]^T$ . The minimum and maximum lags are given by  $l_0$  and  $L_T$ , respectively. The function represents dependence through:

$$s(q, t) = s(x_{t, t-l_0}, \dots, x_{t-L}) = \sum_{l=l_0}^L f \cdot w(x_{t-L}, l) \quad (2.2)$$

where  $f \cdot w(x_{t-L}, l)$  represents the exposure-lag-response function, which is composed of two marginal functions: the exposure-response function  $f(x)$  and lag-response function  $w(l)$  in the space of the lag. Parameterization of  $f$  and  $w$  is achieved by application of the known basis functions to the vectors  $q_t$  and  $l$ . The result can be expressed as matrices  $\mathbf{R}$  and  $\mathbf{C}$  with dimensions  $(L - l_0 + 1) \times v_x$  and  $(L - l_0 + 1) \times v_l$ , respectively.

The cross basis function  $s$  and parameterized coefficients  $\boldsymbol{\eta}$  are given by:

$$s(x_{t,t-l_0}, \dots, x_{t-L}; \boldsymbol{\eta}) = (\mathbf{1}_{L-l_0+1}^T A_t) \boldsymbol{\eta} = w_t^T \boldsymbol{\eta} \quad (2.3)$$

The values of  $w$  are derived from  $A_t$ , which is computed from the row-wise Kronecker product between matrices  $\mathbf{R}$  and  $\mathbf{C}$ . The dependence is expressed through  $w$  and parameters  $\boldsymbol{\eta}$ . The cross-basis function represents the integral of  $s(x, t)$  over the interval  $[l_0, L]$ , summing the contributions from the exposure history. The estimated dependence to specific exposure values is determined by prediction of  $\hat{\beta}$ , called lag coefficients. The estimated  $\hat{\beta}$  and covariance matrix  $V(\hat{\beta})$  is given by:

$$\hat{\beta} = A_x \hat{\boldsymbol{\eta}} \quad (2.4)$$

$$V(\hat{\beta}) = A_x V(\hat{\boldsymbol{\eta}}) A_x^T \quad (2.5)$$

A further extension to DLNM is the application of penalties for smoothness of the lag structure and shrinkage of lag coefficients to null at very high lags. These penalties were applied in the analysis using a second-order difference (Wood, 2006b) and varying ridge penalties (Obermeier et al., 2015; Gasparrini et al., 2017), respectively. Application of penalties was based on the assumption that, at higher lags, the lag coefficients become smaller and approach the null value.

## 2.2.4 Evaluating (de)coupled soil moisture values

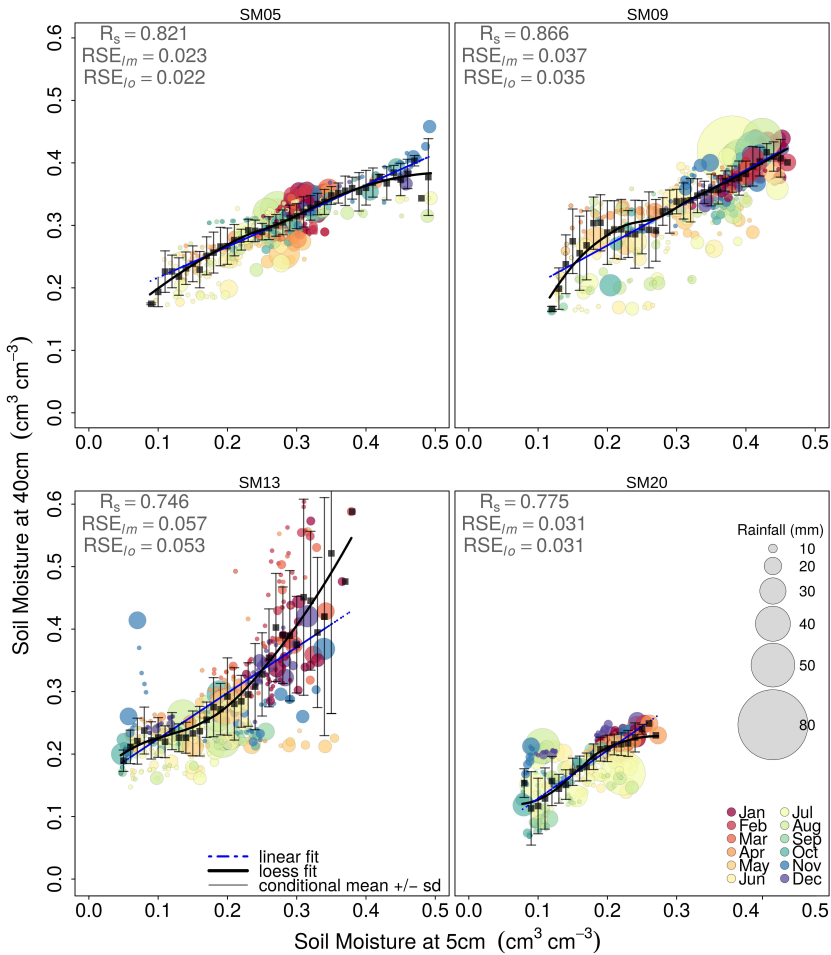
Application of a DLNM resulted in the estimation of parameter  $\hat{\beta}$  for each surface soil moisture value (Eq.2.4 and 2.5). This indicated the strength of dependence between surface and subsurface soil moisture. Higher  $\hat{\beta}$  values indicated stronger dependence or coupling between the two. Hence, we referred to  $\hat{\beta}$  as the relative influence of surface soil moisture on subsurface values.

## 2.3 Results

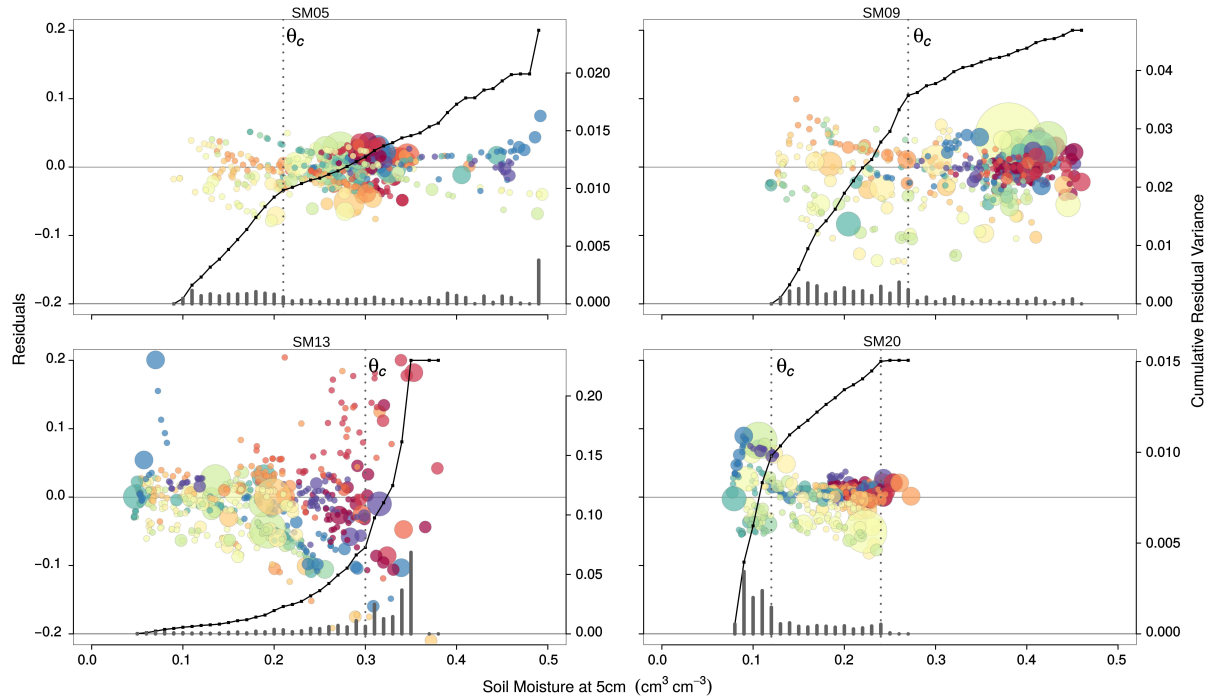
### 2.3.1 Regression and Residuals analysis

The overall dependence between surface and subsurface given by the Spearman's rank coefficient ( $R_s$ ) range from 0.746 to 0.866 (fig.2.4). However, even with a high overall dependence, variability is not uniform across the soil moisture range (fig.2.4). Except for SM13, increased variability is observed towards drier soil moisture values. Furthermore, the degree of variability also differs among the four sites. The most pronounced variability is observed at SM13 and the

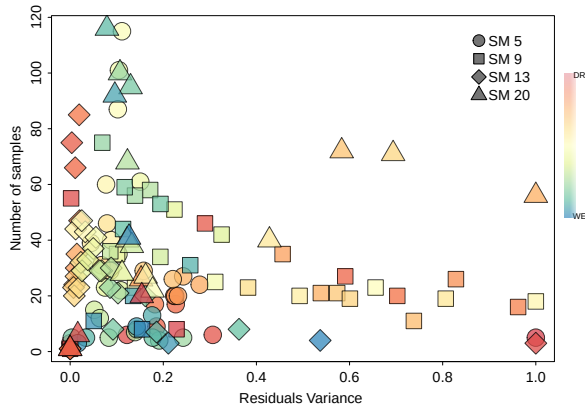
least at SM05. Clustering of observation points occurring within the same months indicate that seasonality dictates soil moisture values and impacts soil moisture variability. Rainfall events measured on the same day do not show a clear effect on surface and subsurface soil moisture dependence. Observations with higher rainfall intensities appear scattered in the plots (fig.2.4). In addition, the said observation points do not necessarily fall along the fitted functions or at the wet soil moisture region of the scatterplots. As lag is not considered, the impact of rainfall on variability is not fully captured in the scatterplots alone.



**Figure 2.4:** Scatter plots of 5 cm vs 40 cm soil moisture values at  $lag = 0$ . Colors correspond to the months in a year and sizes of points are proportional to rainfall intensity. Trends along the soil moisture range shown with the fitted loess function (black line). A linear function (blue line) is also fitted for comparison. The overall dependence using Spearman's rank correlation  $R_s$  is given in the upper left corner each plot. Residual standard errors (RSE) for loess (lo) and linear (lm) fits are also shown.



**Figure 2.5:** Residual variance plots from the fitted loess function. Vertical bars at the bottom of each plot represent the variance for every 0.1 cm<sup>3</sup> cm<sup>-3</sup> interval. The cumulative residual variance (gray line) shows a change in slope at  $\theta_c$ , indicated by the vertical black dashed line. This separates the soil moisture values into a range with higher variance (steeper slope) and another range with lower variance (gentler slopes). The range with higher variance is considered decoupled, and vice versa.

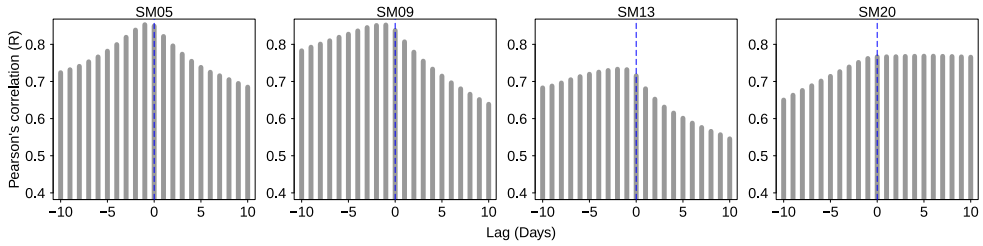


**Figure 2.6:** Scatterplot of sample size vs. normalized residual variance calculated for each  $0.01 \text{ cm}^3 \text{ cm}^{-3}$  interval. Colors indicate soil moisture conditions at each point. The plot of points indicate very weak linear dependence, which is further confirmed by a  $-0.24$  correlation coefficient.

Assessment of the regression fit quality was performed by comparison using residual standard errors (RSE). The results for both linear and loess functions show highly similar values (fig.2.4). This indicates that, in this case, a linear function captures the relation between surface and subsurface values. Nevertheless, the more flexible loess function was preferred for further residuals analysis because of its slightly better model fit and, using only visual inspection of fig.2.4, it more closely approximates the calculated conditional mean.

Figure 2.5 shows the residual plots with lines of the cumulative residuals variance. The change in slope of the line is a feature consistent for all sites regardless of the magnitude of residual variance. The changes in variability are more clearly observed from the residuals than from the standard deviation bars in the scatterplots. The change in slope at  $\theta_c$  is highlighted by the vertical dashed line. The decoupled soil moisture range corresponds to the section of cumulative residuals variance line with a steeper slope. Specifically, the ranges of decoupled surface soil moisture values (in  $\text{cm}^3 \text{ cm}^{-3}$ ) were 0.08-0.21 for SM05, 0.12-0.27 for SM09, 0.30-0.39 for SM13, and 0.08-0.12 for SM20. Except for SM13, the decoupled values are within the dry to intermediate soil moisture range. The cumulative residuals variance line for SM13 appears to increase exponentially with increasing surface soil moisture. This differs from the other three sites which show a distinct decrease in slope at increasing soil moisture values. For SM20, a second point is identified with a change in slope. The flat line starting from  $0.24\text{-}0.28 \text{ cm}^3 \text{ cm}^{-3}$  indicates there is still lowered variance at the very wet soil moisture range.

The correlation between normalized variance and sample size yielded a value of  $-0.24$  (fig.2.6). This low correlation magnitude confirms that the variance obtained for the soil surface moisture intervals was not strongly influenced by the sample size used.



**Figure 2.7:** Cross-correlation plots of soil moisture values. The lagged surface soil moisture values at 5 cm are correlated with subsurface values at 40 cm. A 1 to 2 day lead of surface soil moisture is observed, except for SM20. This is indicated by having maximum the correlation values at lags of -1 to -2 days. At SM20, the maximum correlation occurs at positive lags.

### 2.3.2 Cross-correlation

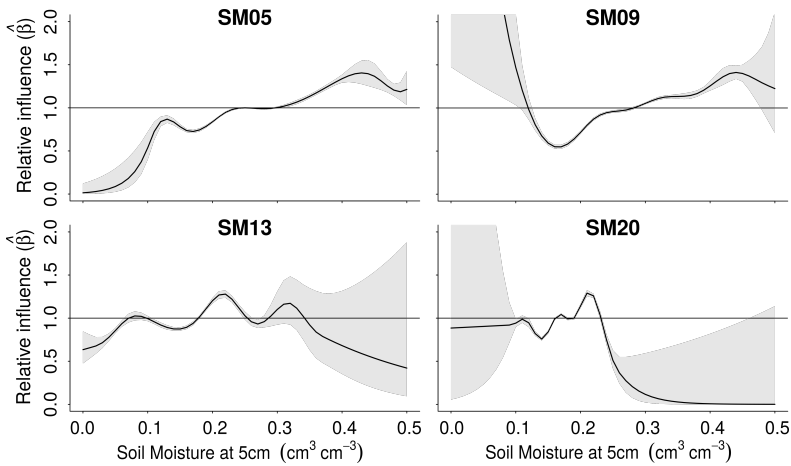
Figure 2.7 shows cross-correlation values at the four sites. Maximum correlation occurs at -1 to -2 days lag, except at SM20. This translates to a 1 to 2 day lead of surface soil moisture values. For SM20, the maximum correlation occurs at positive lags. Correlation values from lag = 0 to lag = 10 are almost equal at SM20. Although this indicates leading subsurface values, it does not eliminate the possibility of having a lag between surface and subsurface values (see Section 2.4.2). Other factors may play a role in having leading subsurface values in the cross-correlation plots. Hence, SM20 was still analyzed for decoupling using DLNM.

### 2.3.3 Distributed lag model

Figure 2.8 shows the overall  $\hat{\beta}$  for each surface soil moisture value with 5% and 95% confidence intervals in shaded gray regions. In order to identify a range that is decoupled, a threshold value ( $\hat{\beta}_c$ ) must be specified. This value is comparable to the intermediate soil moisture  $\theta_c$  identified from Section 2.3.1. The values of  $\theta_c$  provided a suitable guide for identifying a threshold common to all four sites (table 2.2). The corresponding  $\hat{\beta}$  values obtained at  $\theta_c$  were very close to 1, therefore, setting the threshold  $\hat{\beta}_c = 1$  seemed a reasonable choice. This was preferred over the exact  $\hat{\beta}$  at each  $\theta_c$  since the latter was defined using exploratory methods at  $lag = 0$ . Using the chosen  $\hat{\beta}_c = 1$ , surface soil moisture values with  $\hat{\beta} < 1$  are considered decoupled while those with  $\hat{\beta} \geq 1$  are coupled.

Based on  $\hat{\beta}_c$ , the identified decoupled values are generally in the dry to intermediate soil moisture range (fig.2.8), except for SM13 where decouple values are at the wet range. Table 2.2 shows the decoupled values identified based on the selected  $\hat{\beta}_c$ . The behavior and trends of  $\hat{\beta}$  also differ for each station. For instance, at SM05 and SM09, there is a general increase in  $\hat{\beta}$  from dry towards wet surface soil moisture values. SM20 also shows increasing  $\hat{\beta}$  over a limited soil moisture range ( $0.1 - 0.25 \text{ cm}^3 \text{ cm}^{-3}$ ). Outside this range, the estimated  $\hat{\beta}$  values for SM20 were less than one and have very broad confidence intervals. Recall that the range used for DLNM was only for uniformity among the four study sites. The lack of or very few observations for very dry or very wet soil moisture conditions led to wider confidence intervals not only for SM20 but also for the other three sites. Compared to the three sites,





**Figure 2.8:** The relative influence of surface soil moisture on subsurface values obtained by summing the predicted  $\hat{\beta}$  along the 30-day lag. The threshold value ( $\hat{\beta}_c$ ) used to identify the decoupled range is indicated by the horizontal line. Surface soil moisture values below  $\hat{\beta}_c$  are considered decoupled. The 5% and 95% confidence intervals of the predicted values are shown as shaded regions.

**Table 2.2:** Summary of coupling analysis showing the surface soil moisture with significant change in variability ( $\theta_c$ ) from residuals analysis and representing significant lagged influence ( $\hat{\beta}_c$ ) using a threshold of  $\hat{\beta}_c = 1$  from the distributed lag non-linear model. Values of  $\hat{\beta}$  at  $\theta_c$  are presented for assessment of the suitability of the  $\hat{\beta}_c = 1$  threshold.

| Station No. | SSM at $\theta_c$ | $\hat{\beta}$ at $\theta_c$ | Decoupled values using threshold $\hat{\beta}_c = 1$ |
|-------------|-------------------|-----------------------------|--|
| SM05        | 0.21              | 0.90                        | <0.24  |
| SM09        | 0.27              | 0.97                        | <0.28  |
| SM13        | 0.30              | 1.18                        | >0.34  |
| SM20        | 0.12              | 0.94                        | 0.16 > SSM > 0.23                                    |

the estimated  $\hat{\beta}$  values for SM13 show decreasing values towards the wet soil moisture range ( $> 0.3 \text{ cm}^3 \text{ cm}^{-3}$ ). From the intermediate to dry soil moisture conditions, the values fluctuate around the designated  $\hat{\beta}_c$ .

## 2.4 Discussion

### 2.4.1 Decoupled soil moisture values

Regression and residuals analyses show that there is an inherent vertical variability between surface and subsurface soil moisture values based on the lack of 1:1 correspondence between the two (fig.2.4). This inherent variability is also not uniform as higher variability is observed

at certain soil moisture ranges. The cumulative residual variance plots (fig.2.5) clearly indicate the soil moisture values where vertical variability starts to become consistently larger. The increase in variability further translates to weak lagged dependence which we observe as low  $\hat{\beta}$  values from DLNM. The increase in vertical variability and weakening of lagged dependence is what we considered as decoupling between the surface and subsurface soil moisture.

Both residuals analysis and DLNM were successful in identifying a decoupled soil moisture range and there is good agreement between the results from both. Three out of four sites show decoupled values in the dry to intermediate soil moisture range (fig.2.5 and table 2.2). These results agree with the known range where decoupling is expected (Capehart and Carlson, 1997; Hirschi et al., 2014; Wilson et al., 2003). For SM05 and SM09, the intermediate soil moisture value,  $\theta_c$  that marks when decoupling begins (table 2.2) is close to that identified by Capehart and Carlson (1997). They obtained a value of  $0.3 \text{ cm}^3 \text{ cm}^{-3}$  as the point below which decoupling begins. However, results for SM13 do not conform to the traditional concept of decoupling. This result is significant as it implies that decoupling may occur at any value and is not confined to dry soil moisture range.

The vegetation type at each site exerts some influence on the soil moisture variability and the resulting (de)coupled values. First, the vegetation type affects how much ground surface is directly exposed to atmospheric conditions. Forested areas and grass fields are almost fully covered by vegetation compared to a corn field where the crops are organized in equidistant rows. Vegetation or canopy cover will determine how atmospheric conditions affect the soil moisture values. For instance, the amount of intercepted precipitation and evaporation are both dependent on vegetation cover. This in turn will have direct impacts to the surface soil moisture dynamics at each of the sites. For comparison, the variability given by the standard deviation bars in fig.2.4 and variance in fig.2.5 at the cornfield (SM09) is higher compared to that of the grass field (SM05) or the forested area (SM20). In addition, the forested area (SM20) has the smallest range of soil moisture values among the four sites. This may be due to the large intercepted rainfall by the forest canopy. Root water uptake (RWU) is another way by which vegetation affects soil moisture variability. RWU can have significant influence on the subsurface dynamics. The influence of RWU may vary for different vegetation types as it can be exerted over a range of depths, leading to differences in the resulting (de)coupled values.

Among the four sites, the subsurface trends observed for the 40 cm values at SM13 show consistently high values, which can be more pronounced during winter months. This resulted in decoupling during wet soil moisture conditions in fig.2.8. This trend is different from the other three sites which only show a slight increase in the subsurface values. Further inspection of the time series data at SM13 reveals no sudden disturbance in the signal which could be attributed to errors in the sensor. Field investigation confirmed an increase in silt content at 40 cm compared to the upper layers. The increase in silt content promotes a decrease of hydraulic conductivity over depth that results in a slower vertical flow towards deeper layers. The presence of burrowing and hibernating animals was also observed at the site during winter. These create macropores which eventually alter the hydraulic properties of the soil (Kodešová et al., 2006; Beven and Germann, 2013). We infer that, at the measurement domain

of the sensor, these burrows or macropores facilitated faster vertical flow to the subsurface. Alternatively, if the burrows produced voids around the measurement domain, this would result in lowered soil moisture or data gaps due to the loss of sensor to soil media contact. However, there were no gaps observed that coincided with the burrowing animals' period of hibernation. During precipitation events, soil moisture flowing from upper layers arrived more rapidly at 40 cm depths due to the presence of macropores. There it accumulated and flowed more slowly to deeper layers because of the low hydraulic conductivity promoted by the increase in silt content. The overall effect of these factors was the pronounced increase in soil moisture values at 40 cm compared to those at 5 cm during winter periods as observed from the time series dataset in fig.2.2.

Site-specific characteristics at each station control the magnitude of variability as well as the range at which decoupling is observed. However, the occurrence of decoupling is independent of the magnitude of variability since it was observed from SM05 where variability is least up to SM13 where it is greatest. The methods applied in this study only identify conditions when decoupling occurs but do not explicitly determine its controls. Identification of controls for decoupling requires a separate analysis where mechanistic models or statistical approaches can be applied.

#### 2.4.2 Assessing the use of lagged dependence for identifying decoupled conditions

To assess the applicability of the methods applied, we further discuss their strengths and weaknesses. We also present opportunities for further studies as well as foreseen limitations for other sites.

*Strengths:* The residuals analysis and DLNM methods allow quantification of a range of soil moisture values where decoupling occurs. This provides further extension to previous studies where decoupling is only described qualitatively. As seen from the results at the four sites, decoupling can occur at any soil moisture value, and is not confined to dry periods or ranges. Furthermore, by making no initial assumptions on data distributions and the type of functional relation and lag structure, the methods applied were considered robust. Non-linear functions were applied as they conform to the nonlinearity of water flow in the unsaturated zone. They can also handle a variety of bivariate dependence, even in cases where the relation is linear, as shown by the highly similar fit of the loess and linear functions in Section 2.3.1.

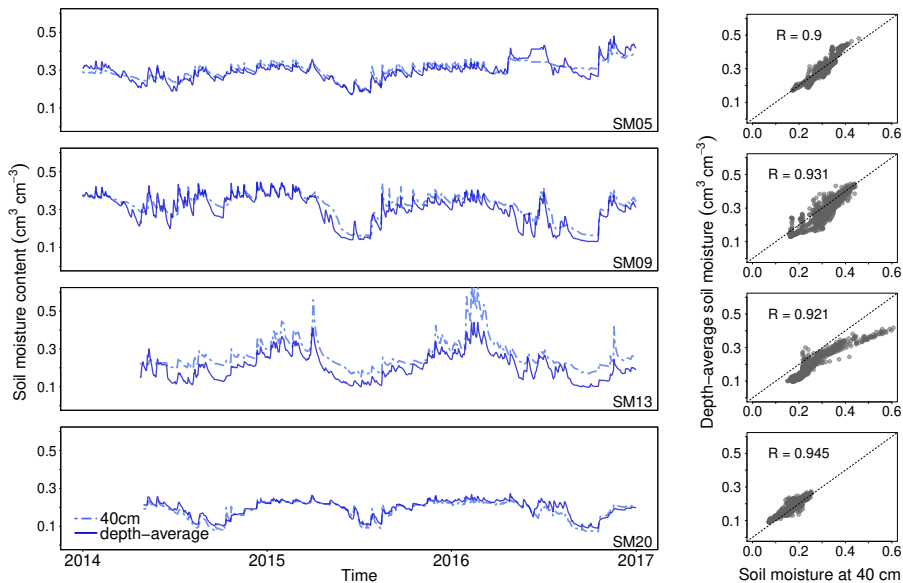
*Weaknesses:* The first aspect that needs to be further investigated is the selected  $\hat{\beta}_c$  value for identifying the decoupled soil moisture range. Although the selection in this study was based on trends identified from time series datasets, the methods applied should be tested further using other datasets to confirm the suitability of  $\hat{\beta}_c = 1$  for other depths and soil types. The choice of  $\beta_c$  is crucial as it dictates which soil moisture values are expected to be decoupled. For instance, at the sites where decoupling occurs during dry conditions, a higher  $\beta_c$  value would enlarge the decoupled range. A similar effect would be expected for the site with decoupling during wet conditions. However, a lower  $\beta_c$  value could result to decoupling only during extreme soil moisture conditions (e.g very wet or very dry).

Another aspect to further examine is the use of cross-correlation for confirming the presence of leading surface soil moisture values. Results from SM20 show maximum correlation at positive lags which indicate leading subsurface values (fig.2.7). The weakness of using cross-correlation as a test for the presence of lag can be two-fold. First, cross-correlation can also capture the effect of subsurface dynamics such as groundwater influence and lateral flow. We infer that in SM20, subsurface dynamics dominates and masks the lag relation sought. An additional covariate representing subsurface dynamics was not included in the DLNM analysis since a dominant downward vertical flow was assumed. This assumption was based on the flat slopes encountered at SM20 (table 2.1). Therefore, the occurrence of subsurface lateral flow or groundwater influence pose limitations to the applicability of DLNM for assessing decoupling. Second, cross-correlation is limited to evaluating linear lagged dependence and in incorporating non-linear lagged dependence can make the test more robust. Equivalent methods exist (e.g. mutual information content, Qiu et al., 2014) but they are much more computationally demanding when the goal is simply to check for the existence of lag-lead relation.

*Opportunities:* In relation to utilizing remote sensing techniques, our results imply that the accuracy of estimating subsurface values from surface soil moisture can be greatly affected by vertical coupling. Lower variability and hence lower uncertainties are expected in the coupled soil moisture range. Assessment of decoupling can be used in combination with modeling studies as a preliminary method to determine the range where variability is expected to be higher. Furthermore, it can be helpful in assessing whether simulation results capture the variabilities observed in both the coupled and decoupled ranges. Taking decoupling into account can also assist in evaluating the necessity of complex models for simulating vertical soil moisture content.

For data assimilation (DA) applications, (de)coupling methods can be used for cross-comparison of the vertical coupling derived from DA model outputs with those observed from long term in situ measurements. This can aid in examining the adequacy of the assumed inherent connection between surface and subsurface values. As Kumar et al. (2009) pointed out, land surface models vary in their representation of the strength of this connection (e.g. weak or strong connection) which contributes the degree in which modeling results are improved. They also suggested that strong coupling is a more robust choice unless independent information suggests that a more decoupled surface-subsurface representation is more realistic. In this aspect, the analysis applied in this study could be a valuable tool in determining which type of surface-subsurface coupling is the more optimal choice. Furthermore, the assumed connection strength is adopted for the whole range of soil moisture values. The results of our analysis show that at any given site, decoupling will occur regardless of degree of soil moisture variability. A variable coupling strength could be adopted based on the soil moisture range where decoupling is likely to occur as an alternative to the single value for the whole range.

Although the study focused on vertically discrete values, the results are also applicable for depth-averaged values commonly used in remote sensing and DA applications. This requires that the vertically discrete values adequately capture the overall dynamics within zone being



**Figure 2.9:** Subsurface soil moisture dynamics for vertically-discrete (40 cm) and depth-average value. Left: Time series of soil moisture at 40 cm and depth-averaged values. The dynamics observed for depth-average values are highly similar to those at 40 cm. Right: Scatterplot showing that these two sets of values are highly correlated.

investigated. In such a case, we infer that the translation to depth-averaged values would result in (de)coupled values that are close, but not identical, to the values obtained when only comparing two discrete depths. As an illustration, we calculated the depth-average values using all the available measurements at each site (i.e. 5, 10, 20 and 40 cm depth) following the formula from Qiu et al. (2014). Figure 2.9 (left) reveals highly similar dynamics for both discrete and depth-average values. Therefore, it can be expected that the results from a regression and DLNM analyses using depth-average values would be highly similar to the original results in fig.2.5 and fig.2.8. However, if the vertically discrete values insufficiently represent the subsurface dynamics, larger deviations in the resulting decoupled values can be expected.

*Limitations:* In this study, only meteorological factors were incorporated in the DLNM analysis since vertical movement was assumed to be the dominant flow mechanism. However, the subsurface can also be influenced by lateral movement or groundwater by capillary rise. In such scenarios, decoupling will not be limited to changes in surface conditions. For this, SM20 provides an excellent example. This station is located at the foot of a small hill (fig.2.2) where the occurrence of lateral subsurface movement is highly probable. This shows that although the analysis would be limited to smaller scales, or even a single point, recognition of regional setting is important for interpretation of results. In addition, subsurface dynamics can also be affected by capillary rise in areas with shallow groundwater. For future applications, the

effect of both capillary rise and lateral movements to subsurface dynamics should be assessed and included in the DLNM analysis, but caution should be exercised when interpreting the results. Assessment of decoupling with DLNM is deemed more applicable to areas where the subsurface has insignificant groundwater influence and where vertical downward movement is the dominant flow mechanism.

## 2.5 Conclusions

The methods applied in this study allow for investigation of vertical soil moisture variability. More importantly, application of DLNM allowed for decoupled soil moisture range to be quantitatively identified. The results also reveal that decoupling is not confined to dry soil moisture range as implied by previous studies. The reasons for decoupling are manifold and controls for the dry soil moisture range may differ from those for the wet range. The results of this study have implications for remote sensing and data assimilation methods, especially for uncertainties related to the use of surface soil moisture to obtain integrated soil moisture values.

"Everything is hard before it is easy."

Johann Wolfgang von Goethe

# 3

## Subsurface soil moisture dynamics in a cultivated field



This chapter is based on:

Carranza, C. and van der Ploeg, M. J. (2020). Characterizing root zone soil moisture dynamics in a cultivated field, *in prep.*

## Abstract

Cultivated fields are dynamic environments because of cyclical land management and vegetation growth. Soil physical properties change consistently due to farming which impacts soil moisture dynamics at seasonal or annual scales. In this study, the contributions of tillage and vegetation growth to subsurface soil moisture dynamics over three seasons (summer 2017, winter 2017, summer 2018) were investigated using in situ measurement stations distributed in a cultivated field. From the temporal stability analysis applied, root-induced variability was estimated to be up to nine times larger compared to those due to tillage. During the 2018 summer drought, the variability in root water uptake (RWU) depends on soil moisture and atmospheric demands. For energy-limited periods, RWU variability strongly correlates with potential evapotranspiration. Once water-limited conditions occurred, RWU variability dramatically decreases, and further implies an adaptive water uptake strategy. The impact of these seasonal changes on soil hydraulic properties was analyzed through inverse hydrological modeling using a pore-flow model. For stations with reasonable modelling accuracy, significant differences for some of the hydraulic parameters were observed. Among them was the much lower saturated water contents ( $\theta_s$ ) estimated for winter 2017 when the soil was bare compared to the two summers when crops were grown. However, a pore-flow model was not sufficient to capture the soil moisture dynamics for almost half of the stations, especially during the two summer periods. These combined modeling results further demonstrate that cultivated fields have temporally dynamic soil hydraulic properties and may systematically require the application of a flow regime other than pore-flow to accurately capture soil moisture dynamics.



## 3.1 Introduction

In agricultural lands, field scale soil moisture variability is important for selection of optimal water management strategies that maximize crop growth (Zhang and Oweis, 1999; Liao et al., 2008; Geerts and Raes, 2009), minimize the harmful impacts of pesticides on groundwater resources (Shipitalo et al., 2000; Ghafoor et al., 2013) as well as erosion (Keesstra et al., 2016). Local scale knowledge of soil moisture dynamics and its spatial variability is also vital to improve our understanding of biogeochemical processes and subsurface flow processes (Wang et al., 2015).

At the field scale, soil moisture dynamics is determined by complex interaction of processes, which have been broadly categorized into local and non-local controls (Grayson et al., 1997). Local controls such as soil properties (Famiglietti et al., 2008; Penna et al., 2013) and vegetation (Baroni et al., 2013) influence vertical drainage and evapotranspiration fluxes. Non local controls including topography leads to surface and subsurface lateral flow (Joshi and Mohanty, 2010). The variability in precipitation might, however, be more important for scales larger than the field scale (Rosenbaum et al., 2012). At deeper soil layers, the direct impact of meteorological variables on soil moisture dynamics diminishes but is primarily modulated by vegetation activity through root water uptake (RWU). Hupet and Vanclooster (2002) showed that vegetation growth, and subsequently evapotranspiration and root water uptake play a non-negligible role in the temporal dynamics of the observed soil moisture patterns in a cultivated (maize) field. Soil-root interactions and their responses to different soil moisture conditions determine water transport processes in the soil-plant-air continuum. Acquisition of water from the soil results in variable root and hydraulic conductivity which depends on water shortage, soil salinity and on water demands from transpiring the above-ground shoot (Steudle, 2000). Recent studies have shown that the rhizosphere has a different hydraulic property than the bulk soil (Carminati et al., 2010; Tardieu et al., 2017). Carminati et al. (2011) found differences in the hydraulic properties of the rhizosphere and bulk soil during a drying phase and immediately after rewetting.

Compared to other landscapes, moisture dynamics in agricultural fields is further affected by land management practice (Kováč et al., 2005; Govaerts et al., 2007; Fabrizzi et al., 2005). Physical soil properties such as bulk density (Mallory et al., 2011; Osunbitan et al., 2005), aggregate stability (Paul et al., 2013; Zheng et al., 2018) and macroporosity (Roseberg and McCoy, 1992; Pagliai et al., 2004; Malone et al., 2003) are altered by land management both in the short- and long-term (Pagliai et al., 2004; Strudley et al., 2008). Soil structural changes have direct impacts to hydraulic properties and soil moisture estimation (Mapa et al., 1986; Moret and Arrúe, 2007). Studies have suggested incorporating land use changes to parameterize hydrological models for a more accurate representation of soil structure and variability in hydraulic properties (e.g Gonzalez-Sosa et al. 2010; Ndiaye et al. 2007). Furthermore, Schwen et al. (2011a) and Alletto and Coquet (2009) demonstrated that temporally varying soil hydraulic parameters improved simulation results.

To elucidate the processes that act as controls for soil moisture variability, temporal stability analysis (TSA, Vachaud et al. 1985) has been applied at regional and catchment scales (e.g.

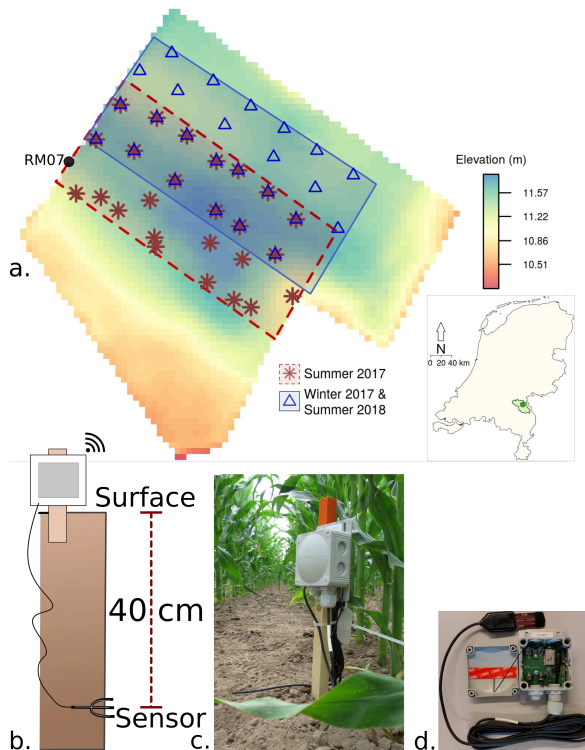
Guber et al. 2008; Vanderlinden et al. 2012; Fry and Guber 2020). In TSA, soil moisture variability at a certain location is characterized in terms of its mean and standard deviation relative to the trends derived from all measurements combined. Based on this, the presence of locations exhibiting persistent bias or those which are consistently wetter or consistently dryer than the average are highlighted. Using TSA in grass fields, Wang et al. (2012) demonstrated that the temporal stability in the root zone is mainly affected by vegetation phenology. Furthermore, it may control soil moisture variability at both seasonal and annual time scales. Martinez et al. (2013) analyzed TSA using stochastic simulations of soil moisture values derived from known statistical distributions of soil hydraulic properties. Focusing on the impact of soil hydraulic properties on soil moisture temporal stability, their main findings indicate that coarser soil textures have higher mean relative differences than finer textures. In addition, RWU was found to decrease the mean relative differences variability in the root zone and increased variability below it. In recent years, TSA has been relatively popular for optimizing sampling locations, especially for large areal campaigns, because it allows identification of sites that are representative of the spatial mean (e.g. Brocca et al. 2009; Schneider et al. 2008; Gao et al. 2019).

In this study, our main goal is to quantify subsurface soil moisture variability using TSA from almost 12 months of subsurface in situ measurements and to explore its impacts on soil hydraulic properties. Field scale subsurface soil moisture dynamics from in situ measurements are limited in literature, perhaps due to the difficulty of obtaining spatially-distributed subsurface measurements in a single field. However, in situ measurements are essential since they capture conditions that may otherwise be difficult to replicate in process-based models. Installed at 40 cm depth, the measurements from 25 sampling points captured the active root zone of the crops encountered in the field. Furthermore, the sensors were integrated into a wireless LoRaWAN architecture for real-time monitoring. Since meteorological conditions over a field is assumed to be uniform, only the contributions of soil heterogeneity and vegetation activity were related to the variability observed over two growing seasons (2017 and 2018) and the winter season in between. Inverse modeling was applied to the in situ measurements to determine how the dynamics observed from the in situ measurements impacted the soil hydraulic properties.

## 3.2 Materials and Methods

### 3.2.1 Site Description

The study area is a cultivated field situated at the southeastern portion of the Netherlands (fig.3.1). The terrain is generally flat and the soil is dominantly sandy (82.1% sand, 10.5% silt, 5.2% clay, and 2.2% organic matter). The crops grown every year rotates from corn, chicory, potato, and grass. The field is part of the Raam soil moisture network (station RM07, Benninga et al., 2018). Measurements at RM07 were used to provide information on the surface moisture conditions, yet are not further included in the analysis. For every growing season, the field is tilled before the seeds are sown and after the crops are harvested.

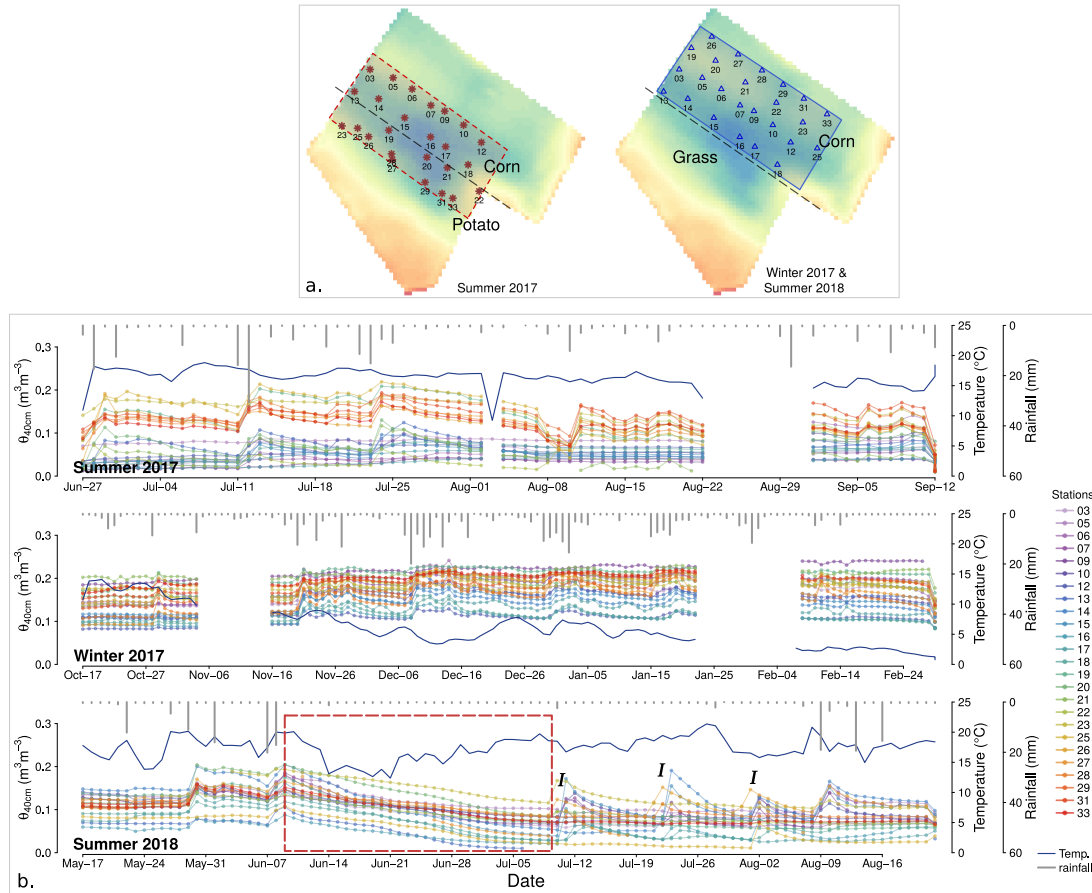


**Figure 3.1:** Map of the soil moisture sensors installed in the cultivated field. a) Location of sensors for over the measurement period from 2017 to 2018. Within the same field, the location of RM07, a sensor in the Raam soil moisture and monitoring network (Benninga et al., 2018) is also shown. b) A cross-sectional diagram of the installation configuration at each location. c) Photo of the above-ground part of one sensor. d) Components at each installation site, which include the soil moisture sensor with a customized logger and signal transmitter in a white waterproof enclosure.

### 3.2.2 Soil moisture network with LoRaWAN

Twenty-five soil moisture sensors were installed within the field at 40 cm depth during the growing seasons of 2017 and 2018, as well as the winter period in between. There were two sets of sampling locations for the duration of the measurement campaign because a portion of the field was converted into grass after the 2017 harvest (fig.3.1 and 3.2). Installation in grass was not considered since there is high likelihood that the sensors will be destroyed by cattle.

For the growing season of 2017, the crops encountered at the sites were corn and potato. The soil moisture locations were almost equally (13 in corn, 12 in potato) divided between the two crops. The sampling campaign started on June 26 and were uninstalled by Sept 12, just before the harvest period. The measurement campaign resumed after harvest in the autumn of the same year. Since the southwestern portion of the field was converted into grass, 12



**Figure 3.2:** Soil moisture measurements at 40 cm depth ( $\theta_{40cm}$ ) for 2017 - 2018. a) Maps showing the station locations and crop types encountered in the field. b) Time series plots of  $\theta_{40cm}$  over the study period. For summer 2018, the red box highlights a continuous drying period at the beginning of the drought. *I*'s indicate the days when irrigation was applied.

out of 25 stations were relocated on the northeastern side. The second field measurement campaign then started from Oct 10 2017 up to March 01 2018 just before the start of tillage activities. The third field campaign started from May 17 2018, after the seeds have been sown, up to August 22 2018, before harvesting. For 2018, only corn was planted in the area where the soil moisture sensors were re-installed. Measurements from the soil moisture sensors were logged every 15 minutes for each station over the duration of the sampling campaigns. However, continuous measurements across the seasons were not possible in the field because of tillage.

The soil moisture sensors were integrated into a wireless LoRaWAN framework that enabled real-time monitoring and facilitated early detection of malfunctions to minimize data losses. Briefly, LoRa (short for Long Range) is a radio frequency carrier signal of telecommunications devices. LoRa allows the conversion of measured data into a signal that can be transmitted from the sensor location to the back-end user elsewhere. It is considered as a part of the Internet of Things (IoT) technology and is owned by Semtech in 2012 (Alippi et al., 2010). One advantage of LoRa is its increased communication range for transmitting data, which can be up to 20 km on good atmospheric conditions (non-rainy days). LoRa Physical Layer Protocol works on sub-GHz frequencies of frequency bands 433-, 868-, 915-, 923-MHz depending on country-specific regulations. LoRaWAN is an open standard technology that links the LoRa signal to a certain application. Thus, LoRaWAN is a communication protocol and network architecture, while LoRa is the physical layer that enables the long range link LoRaWAN (2021). LoRaWAN contains the data transfer layer that allows transmission of data to any device already connected to the cloud. For the soil moisture network, each sensor was connected to a customized logger and wireless data transmitter (fig.3.1b-d). All 25 data transmitters were then connected to a wireless LoRa network gateway that was installed inside the barn next to the cultivated field. Aside from receiving all the soil moisture measurements, the gateway communicates with other LoRa gateways distributed across the Netherlands, which allows transmission of the data to the end-user.

#### *Sampling optimization using Spatial Simulated Annealing (SSA)*

Spatial simulated annealing (SSA, Van Groenigen and Stein, 1998), a geostatistical method for sampling optimizations, was applied to optimize soil moisture sampling locations. To solve spatial optimization problems, SSA algorithm creates slight perturbations in the sampling design and applies a random search technique. The suitability of a sampling design is evaluated by minimizing the values of an objective function, which in this case is the average kriging prediction error variance (Brus and Heuvelink, 2007; Melles et al., 2011). SSA optimization runs by randomly moving each of the sampling stations, calculating the objective function, and accepting improved designs (i.e. smaller values for the objective function) over a fixed number of iterations. In some instances, worse designs are accepted with a decreasing probability ( $P$ ) to avoid being trapped in a local minima.  $P$  was set to  $P \leq 20\%$  as suggested by Brus and Heuvelink (2007). The cooling factor in SSA, which dictates the rate at which  $P$  decreases to zero, whereby  $P$  was set to exponentially decrease as a function of number of iterations to ensure convergence. The number of iterations was set to 10000 and the cooling factor was

set at 1000.

A set of soil moisture measurements was obtained on June 08 2017 prior to the installation of sensors at 40 cm to determine the kriging prediction error variance required needed for SSA. However, collection of spatially distributed samples at 40 cm depth was considered impractical for the given sampling optimization task. Therefore, only surface soil moisture were measured since they influence and may be indicative of subsurface conditions. Thirty surface soil moisture measurements at 5 cm using a TDR were obtained along three transects that were at least 20 m apart. This was deemed as a reasonable guide for identifying the subsurface sampling locations based on the assumption that soil moisture values encountered at the 5 cm depth will have equal probability of being encountered at the 40 cm depth, although not necessarily because of the same processes. A spherical model was applied to estimate the semi-variogram parameters using the 30 surface soil moisture samples. The modelled variogram was then used to estimate the kriging variance required in SSA. Nine out of the 30 measurement locations were used as the initial input for SSA. Sixteen sampling locations were optimized for the 2017 growing season while 12 locations were optimized for the 2017 winter and 2018 growing season.

### 3.2.3 Field scale heterogeneity using Temporal Stability analysis

Using the relative differences (RD) approach, the temporal stability (TS) at each station were evaluated with respect to the mean value from all of the stations. The daily mean soil moisture measurements were applied for the TS analysis. The relative difference ( $\delta_{ij}$ ), is defined as:

$$\delta_{ij} = \frac{\theta_{ij} - \bar{\theta}_j}{\bar{\theta}_j} \quad (3.1)$$

where  $\theta_{ij}$  is the measurement at location  $i$  of sampling time  $j$  and  $\bar{\theta}_j$  is the mean from all the measurements ( $N$ ) of each sampling time  $j$ :

$$\bar{\theta}_j = \frac{1}{N} \sum_{i=1}^N \theta_{ij} \quad (3.2)$$

For each location  $i$ , the mean and the standard deviation of the relative differences (MRD and SDRD) are given by:

$$MRD (\bar{\delta}_i) = \frac{1}{M} \sum_{j=1}^M \delta_{ij} \quad (3.3)$$

$$SDRD (\sigma(\delta_i)) = \sqrt{\frac{1}{M-1} \sum_{j=1}^M (\delta_{ij} - \bar{\delta}_i)^2} \quad (3.4)$$

where  $M$  is the number of samples for each station  $i$ .

Due to the bounded nature of soil moisture, the datasets obtained may have skewed distribution, rendering the mean a poor measure of central tendency. In such cases, median is a better measure than the mean. In this study, TSA was carried out also using the median but produced highly similar values and trends with those using the mean (fig.8.1, Appendix). In this case, the mean and the median may be used interchangeably since the distributions are not highly skewed. Most, if not all, studies encountered use the mean for TSA (e.g. Brocca et al., 2009; Martinez et al., 2013) but often lacks assessment of its suitability as a measure of central tendency.

### 3.2.4 Diurnal soil moisture dynamics during the 2018 summer drought

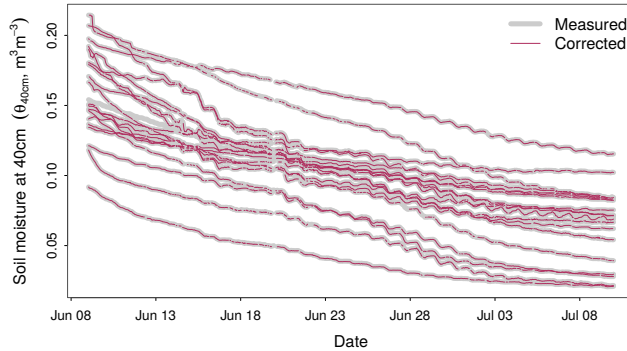
In 2018, western Europe was hit by a summer drought from June until early August. In the study area, a drying trend coinciding with the drought was observed until the start of July (fig.3.2). Irrigation was applied several times to the corn fields starting early July because of the prolonged dry period. Diurnal soil moisture dynamics were investigated over the month-long drying period (June 9 - July 9) to determine root water uptake variability. It is assumed that during this period, the influence of meteorological conditions at 40 cm depth were primarily modulated by vegetation activity. Atmospheric conditions influence daytime transpiration rates by dictating the amount of water needed by vegetation and, consequently, determining the decrease in soil water content during a drying period.

#### *Temperature sensitivity correction*

The sensitivity of soil moisture sensors to temperature was addressed by applying a correction method outlined in Cobos and Campbell (2007). Daily variations in temperature can cause sinusoidal patterns to be overprinted onto the 'true' soil moisture values. If the effect of temperature is negligible, a linearly decreasing trend in soil moisture would be expected for several consecutive dry days. For the sensors at 40 cm, the soil moisture values were more or less identical before and after applying the temperature correction, indicating a negligible effect of diurnal temperature variations on the soil moisture measurements (fig.3.3). These findings agree with Cobos and Campbell (2007), who noted that at soil depth greater than 15 cm or under a full vegetative canopy, the temperature sensitivity due to diurnal fluctuations in soil temperature would be small or unnoticeable.

#### *Diurnal soil moisture dynamics*

The impacts of root water uptake on soil moisture variability were investigated based on the diurnal changes in soil moisture from June 9 to July 9, 2018. This drying period facilitated the estimation of root water uptake up since the changes in soil moisture was mainly attributed to vegetation activity because of the absence of rainfall over the said month-long period. Using the 15-minute soil moisture measurements, subsets were obtained for specific daily time intervals. Diurnal changes (thereafter referred to as RWU) were calculated based on differences in soil moisture between sunrise and sunset. Furthermore, the afternoon and morning



**Figure 3.3:** Measured and temperature-corrected soil moisture at 40 cm depth. For a few observations, corrected values were not obtained because of missing temperature measurements.

changes were also calculated to determine whether differences occur at different daytime periods. Afternoon and morning changes were based on differences against measurements at 12 noon.

The temporal RWU dynamics were compared with daily potential evapotranspiration ( $ET_{pot}$ ), which were measured from the nearest KNMI weather station (Volkel). To facilitate comparison, the volumetric soil moisture content in  $m^3 m^{-3}$  was converted to millimeters to match the units of the  $ET_{pot}$  obtained. The measured volumetric soil moisture was assumed to be representative over a 5.2 cm (52 mm) thickness. This is based on the average of the values found by Benninga et al. (2018) (6 cm) and Vaz et al. (2013) (4.4 cm), who investigated sampling diameters of theta probe sensors. The corresponding values in millimeters were obtained by multiplying the volumetric soil moisture content by 52. The spatial patterns between RWU and soil moisture states were compared using the relative differences approach primarily to determine whether a correlation exists between the two. In other words, do locations which are relatively wetter (or drier) show relatively larger (or smaller) RWU?

RWU dynamics was further investigated based on transpiration regimes that relate soil moisture and  $ET_{pot}$ . Assuming that precipitation is the only source of water in the soil at 40 cm, periods with sufficient or surplus of precipitation are, on the one hand, considered energy-limited since  $ET_{pot}$  is limited by energy and not water. On the other hand, precipitation or soil moisture deficit during droughts are considered water-limited periods. In land surface models, a critical soil moisture content ( $\theta_{crit}$ , Seneviratne et al. 2010) separates the two regimes and has been found to influence land-atmosphere interactions and feedbacks.

### 3.2.5 Estimation of soil hydraulic parameters using Inverse modeling

Inverse numerical simulations for a 1-dimensional (1D) vertical flow domain was applied in Hydrus-1D to estimate the soil hydraulic parameters at each station. Soil moisture measurements were used as inputs in the pore flow equations to numerically solve for water retention



parameters based on van Genuchten (1980). The vertical water flow in an unsaturated porous media is solved numerically using Richards' equation:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial Z} \left[ K(h) \left( \frac{\partial h}{\partial Z} + 1 \right) \right] - S \quad (3.5)$$

where  $t$  is the time (days),  $\theta$  is the volumetric water content ( $\text{cm}^3 \text{cm}^{-3}$ ),  $h$  is the soil water pressure head (cm),  $Z$  is the spatial coordinate (cm) defined as positive upward,  $K(h)$  is the unsaturated hydraulic conductivity function ( $\text{cm d}^{-1}$ ) and  $S$  is a sink term representing water uptake by plant roots ( $\text{cm d}^{-1}$ ).  $K(h)$  is derived from a water retention curve, given by van Genuchten (1980):

$$\theta(h) = \frac{\theta_s - \theta_r}{[1 + (\alpha h)^n]^m}, h \leq 0 \quad (3.6)$$

$$K(h) = K_s S_e^l (1 - (1 - S_e^{1/m})^m)^2$$

$$m = 1 - \frac{1}{n} \quad (3.7)$$

$$S_e = \frac{\theta - \theta_r}{\theta_s - \theta_r}$$

where  $\theta_r$  and  $\theta_s$  denote residual and saturated volumetric water contents ( $\text{cm}^3 \text{cm}^{-3}$ ), respectively;  $\alpha$  ( $\text{cm}^{-1}$ ) and  $n$  (-) are fitting parameters of soil water characteristic curve;  $K_s$  is the saturated hydraulic conductivity ( $\text{cm d}^{-1}$ );  $l$  (-) is the pore connectivity parameter ; and  $S_e$  (-) is the relative saturation.

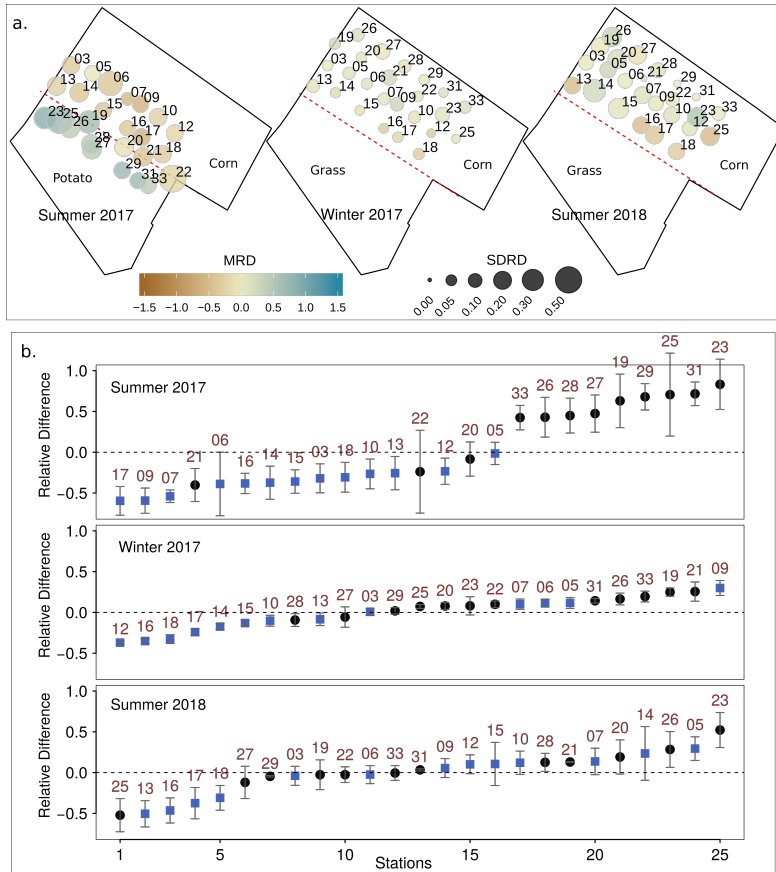
Inverse numerical simulations were carried out over a vertical length of 1 m to cover the rooting depths of the crops. The initial profile distribution of pressure heads was made uniform over the flow domain under the assumption of a saturated state during initial conditions. The initial pressure head value for the profile was set equal to -100 cm. Variable atmospheric conditions was chosen for the upper boundary while a free drainage was set for the lower boundary condition. Data on meteorological conditions were obtained the nearest meteorological station (Volkel) from the Royal Netherlands Meteorological Institute (KNMI). The simulation was started at the first day of a year (Jan 1) for a spin up period to minimize the impact of initial conditions. For the growing seasons, root growth and root water uptake models were included in the simulation considering the growth properties of corn and potato.

## 3.3 Results and Discussion

### 3.3.1 Time stability and local controls

Based on fig.3.4, soil moisture values for the 2017 growing season shows distinct patterns over corn and potatoes, with higher mean relative difference (MRD) in the latter than the former. In contrast to MRD, the relative difference standard deviation (SDRD) appear to be

comparable for both crops. The range of SDRD values ( $\sim 0.30$ - $0.5$ ) are also larger during the 2016 growing season compared to those calculated for other seasons. This suggests that soil moisture at 40 cm vary for both corn and potato. Because there is only a single soil textural class in the field, the variability at  $\theta_{40\text{cm}}$  could be related to structural characteristics that may be due crop-specific management (e.g. Mudgal et al. 2010; Çerçioglu et al. 2019). It is inferred that the creation of mounds and furrows for the potato field resulted in increased water retention compared to the corn fields based on higher MRDs in the former.



**Figure 3.4:** Plots of relative differences (RD) over different seasons. a) Spatial plots of mean and standard deviations in RD (MRD and SDRD). b) The stations are ranks based on increasing MRD for the different seasons. The dots indicate MRD while the bars indicate SDRD for each station. Blue square dots are stations kept in the same locations while the black circles were changed after 2017 harvest. MRD and SDRD are in  $\text{m}^3 \text{m}^{-3}$  units

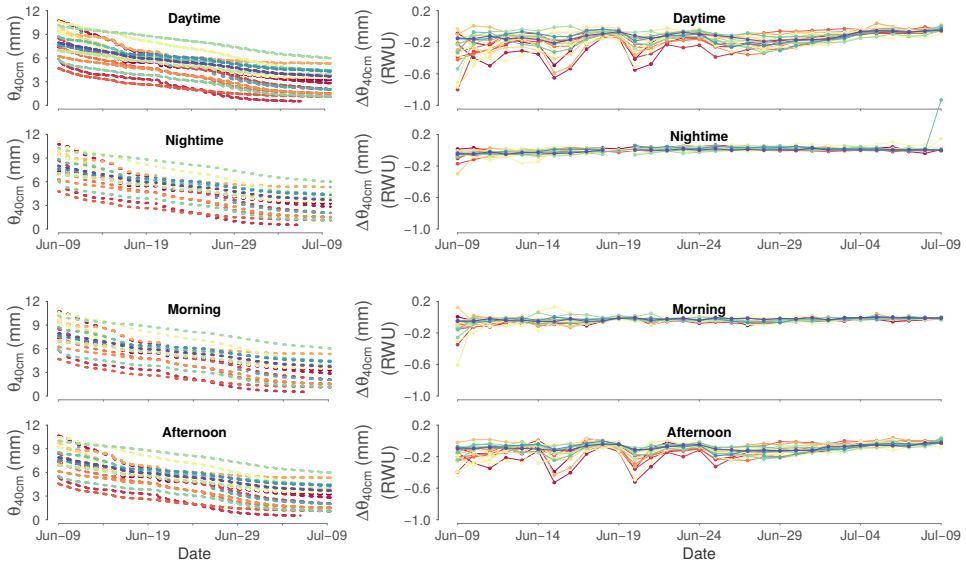
In the following winter period, SDRD values were smaller in comparison with the previous growing season. Since the ground was bare in winter, it is inferred that only soil heterogeneity contributes to SDRD, unlike the combined effects of both soil and vegetation during the

previous growing season. Variations in SDRD among the stations are small ( $\sim < 0.1$ ), implying that soil moisture is relatively stable during the winter period. Surprisingly, the MRD values calculated for winter 2017 are intermediate between those encountered during the two growing seasons (2017 and 2018). This somewhat deviates from the common expectation of wet or saturated soils during winter because of much lower evaporation demand (e.g. Western et al. 1998). However, Schwen et al. (2011a) suggested that hydraulically effective pores may decrease after post-harvest tillage in winter in response to rainfall. In addition, frozen surface conditions were recorded at the study site in early February based on surface soil measurements at the nearby RM07 station, although temperatures measured at 40 cm depth were still slightly above freezing. Frozen surface conditions may have further contributed to low infiltration rates at shallower soil depths that helped maintain lower subsurface soil moisture contents.

The values of MRD and SDRD obtained for the 2018 growing season are comparable to the 2017 growing season. In both cases, SDRD and MRD are higher compared to those obtained during the winter period. However in 2018, distinct trends observed in MRD and SDRD are not observed, which may be due to the presence of a single crop type. Since the TSA was calculated for each of the three measurement campaign periods, MRD and SDRD for the 2017 growing season may appear to be larger than the 2018 growing season because of the contrasting trends observed per crop type.

Similarities in MRD rankings are observed for points which were kept in the same location over the three seasons (stations 03 to 18, fig.3.1 and 3.4). For a few points (03, 16, 17), slight changes in their rankings suggest that they are time stable sites where subsurface variability did not dramatically change. However, most of the stations showed large shifts in the ranking over the seasons, indicating that they are non time-stable sites. These findings further demonstrate that time varying subsurface soil moisture trends maybe the rule rather than the exception for every location within a cultivated field. This is dependent on the type of crops grown per season and the sequence of crop rotation chosen by a farmer. Identification of systematic long-term changes in subsurface soil moisture dynamics may prove challenging since the choice of rotation of crops is still dependent on a farmer and is an information that may only become available in the short term. However, quantification of temporal changes over a whole growing season for commonly encountered crops could provide insights on the expected magnitude and/or distribution of soil moisture changes that may be useful for uncertainty estimation of, for instance, soil hydraulic parameters.

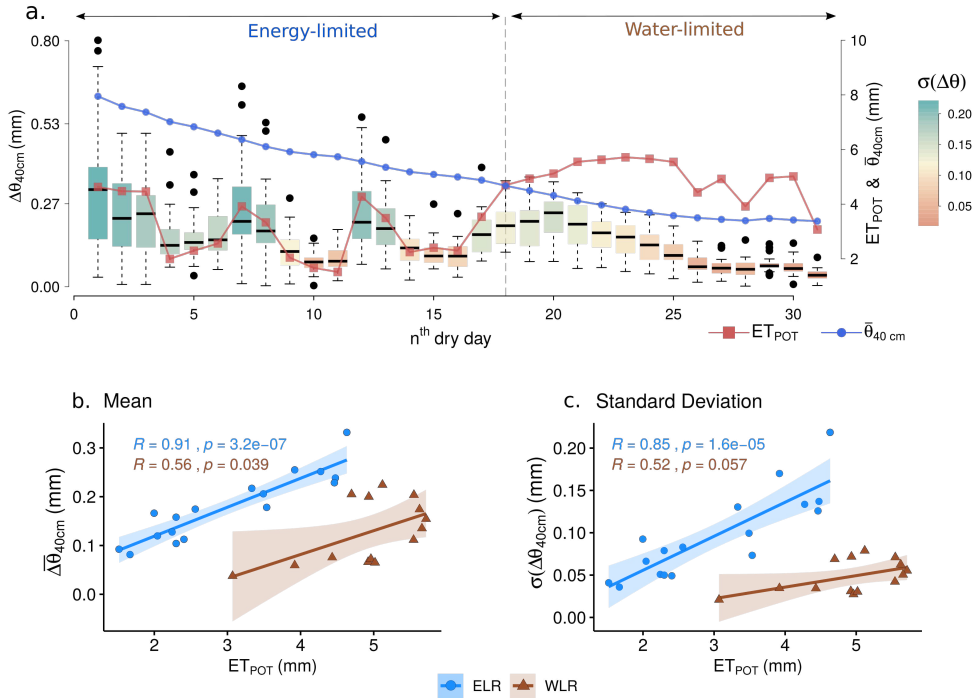
Overall, results from the TS analysis for the growing seasons demonstrate the combined influence of soil heterogeneity and vegetation activity on subsurface soil moisture variability. The approximate contributions of each component to the observed subsurface soil moisture variability were inferred from the changes in SDRD between 2017 winter and 2018 growing season, since the measurement points were all kept in the same locations. Focus was given to changes in SDRD since MRD from all stations were comparable for both seasons. For winter 2017 when the soil was bare, the SDRD calculated ( $0.021 - 0.125 \text{ m}^3 \text{ m}^{-3}$ ) is considered only to be a function of soil heterogeneity. For the 2018 growing season, SDRD ( $0.004 - 0.329 \text{ m}^3 \text{ m}^{-3}$ ) resulted from the combined influence of soil heterogeneity and vegetation activity. From win-



**Figure 3.5:** Soil moisture dynamics in drying phase from 9 June to 9 July 2018. *Left:* Time series plots of soil moisture for specific time periods in a day. *Right:* Diurnal changes in soil moisture for time specific periods.

ter 2017 to the 2018 growing season, there is a general increase in SDRD, since lower values in the latter season were only observed at three out of the 25 stations (21, 29, 31). The largest change in SDRD was calculated in station 25 (fig.3.4), where close to a 9-fold increase (8.86 times higher) was calculated (from  $0.023$  to  $0.203 \text{ m}^3 \text{ m}^{-3}$ ). If it is assumed that the same tillage activities were carried out post-harvest in 2017 and pre-planting in 2018, then the same magnitude of SDRD for soil heterogeneity in winter 2017 would be expected for 2018 growing season. Therefore, much higher SDRD calculated for 2018 growing season suggests that vegetation activity has a much larger contribution to soil moisture variability than soil heterogeneity. Based on higher SDRD measured using the temporal stability analysis in 2018, vegetation is interpreted to increase soil moisture variability at 40 cm depth from 1.26 to 8.86 times.

Studies on vegetation water uptake have shown that root mucilage also contribute to variable hydraulic properties in the rhizosphere. (Carminati et al., 2011; Kodešová et al., 2006; Lazarovitch et al., 2018). Carminati et al. (2010) showed that differences in hydraulic properties between the rhizosphere and bulk soil is actually beneficial for vegetation water uptake during water stressed periods. Since the sensors are installed in the active root zone of both corn and potato in this study, some, if not all of the measurements are potentially influenced by both rhizosphere and bulk soil hydraulic conditions. Currently, it is difficult to determine to what extent these different soil hydrological regions affect the measurements from sensors installed.



**Figure 3.6:** Relation between subsurface soil moisture dynamics at 40 cm depth and atmospheric demands during the drying phase. a) Boxplots showing the subsurface soil moisture changes ( $\Delta\theta_{40\text{cm}}$ ) from all 25 stations during the daytime period (between sunrise and sunset). The colors of each boxplot signifies the standard deviation in  $\Delta\theta_{40\text{cm}}$  ( $\sigma(\Delta\theta_{40\text{cm}})$ ). The blue line represents mean soil moisture ( $\bar{\theta}_{40\text{cm}}$ ) over the daytime period while the red line represents the daily potential evapotranspiration ( $ET_{\text{pot}}$ ). The drying period is subdivided into an Energy-limited regime (ELR) and Water-limited regime (WLR) based on the amount  $\bar{\theta}_{40\text{cm}}$  relative to  $ET_{\text{pot}}$ . On the 18<sup>th</sup> drying day, a transition from ENR to WLR occurred when  $\bar{\theta}_{40\text{cm}}$  started to become consistently below  $ET_{\text{pot}}$ . b and c) Scatter plots with trend lines showing the correlation between the mean and standard deviation of  $\Delta\theta_{40\text{cm}}$  with  $ET_{\text{pot}}$ . Pearson's correlation coefficient ( $R$ ) and P-values are given.

### 3.3.2 Diurnal soil moisture dynamics during a drought

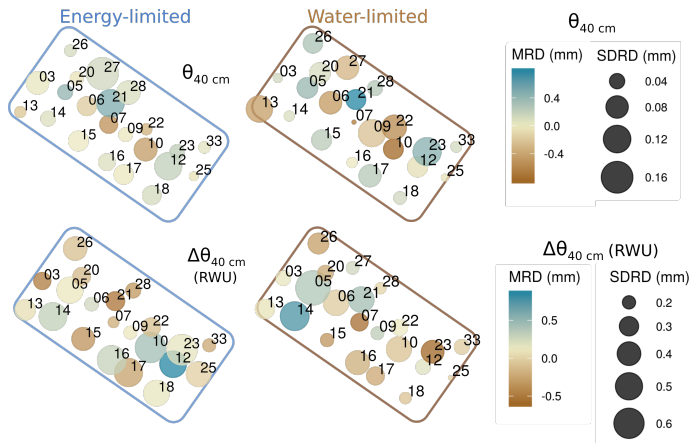
Diurnal changes in soil moisture at 40 cm depth over the month-long drying phase during the 2018 summer drought is plotted in fig.3.5. At the start of the drought on June 9 2018, diurnal soil moisture changes were not constantly decreasing. Larger diurnal changes smaller diurnal changes. Around June 24 2018, diurnal soil moisture changes appear to linearly decrease and approach zero until the end of the analysis period. Nighttime changes in soil moisture are constant and very minimal compared to daytime changes. Further comparison of morning and afternoon values reveal that latter accounts for most of the daytime changes. The observed

changes between night and day closely follow known trends in transpiration and confirms the assumption that they are indeed due to root water uptake (RWU).

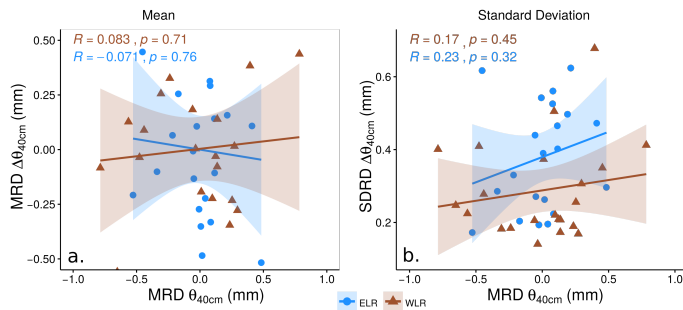
At the start of the drying phase, daily  $\theta_{40\text{cm}}$  were initially able to meet daily atmospheric demands, represented by  $ET_{pot}$  (fig.3.6a). On the 18<sup>th</sup> dry day, daily soil moisture values started to become lower than the daily the  $ET_{pot}$ . Soil moisture from the beginning of the drying phase up to day 18 is considered as the energy-limited regime (ELR) while succeeding period is the water-limited regime (WLR). In the former, soil moisture at 40 cm were, on average, higher and therefore sufficient to meet  $ET_{pot}$ . Continued depletion of soil moisture at 40 cm depth resulted in lower soil moisture compared to  $ET_{pot}$  from the 18<sup>th</sup> dry day until the last dry day investigated. Comparison of two transpiration regimes reveal that RWU trends only partly corresponds to  $ET_{pot}$ . The median of the daily RWU, indicated by the boxplot bars in fig.3.6a, was observed to change proportionally with daily  $ET_{pot}$  within the ELR. However, RWU begins to deviate from the  $ET_{pot}$  during the WLR. For this period, RWU is observed to be consistently smaller and no longer change proportionally with daily  $ET_{pot}$ . Aside from the median of boxplots, fig.3.6a further show that the variability in RWU differ between the two regimes. Within ELR, the variability in RWU based on the length of the box and whiskers, appears proportional with  $ET_{pot}$  across the stations, with lower  $ET_{pot}$  resulting in more a uniform RWU while higher  $ET_{pot}$  driving more variable RWU. Within WLR, the variability in RWU is consistently smaller than ENR and did not always change proportionally in response to changes in  $ET_{pot}$ . Furthermore, both the mean and standard deviation of RWU have lower correlations with  $ET_{pot}$  for WLR, as shown in fig.3.6b and c. The change in dynamics upon transition to a WLR fits the adaptive RWU framework wherein roots seek water from other (deeper) soil layers once the available supply at shallower depths are depleted and can no longer meet vegetation or atmospheric demands (Jarvis, 1989). Since  $\theta_{40\text{cm}}$  on day 18 started to be inadequate for meeting the daily  $ET_{pot}$ , the uptake of water at this depth gradually diminished. It is further hypothesized that from the 18<sup>th</sup> dry day, RWU was more aggressive at deeper layers where soil moisture not as dry as the 40 cm depth.

The critical soil moisture ( $\theta_{crit}$ ) signifying the transition between energy- and water-limited regimes may be affected by the measurement depth, as observed by Buitink et al. (2020) using all the soil moisture stations in the Raam and Twente (Dente et al., 2011) networks. In the Raam, they estimated a  $\theta_{crit}$  for the root zone up to 40 cm to be approximately  $\theta_{crit} = 0.20 \text{ m}^3 \text{ m}^{-3}$ . However, based on the measurements at 40 cm depth in this study, the  $\theta_{crit}$  identified on day 18 was 4.67 mm or  $0.089 \text{ m}^3 \text{ m}^{-3}$  (fig.3.6a). This discrepancy may be due to the depths analysed in both studies. Buitink et al. (2020) considered the depth-average values as opposed to a singular depth considered in the study. In addition, other stations in the Raam network may have higher soil moisture content compared to the field in RM07 and thus contributed to higher  $\theta_{crit}$  estimation by Buitink et al. (2020). Nevertheless, the day and soil moisture at which a transition to water-limited regime occurred is expected to differ when other depths are investigated.

The results obtained on RWU variability was found to be similar to the findings of Srayeddin and Doussan (2009) using Electrical Resistivity Tomography (ERT) in maize and sorghum. Aside from demonstrating the potential of the said geophysical technique to map subsurface



**Figure 3.7:** Spatial plots of temporal stability analysis for soil moisture and the changes in soil moisture at 40 cm ( $\theta_{40\text{cm}}$  and  $\Delta\theta_{40\text{cm}}$ ). The points are separated into the water conditions encountered and are represented by the colors of the boxes (blue for energy-limited and brown for water-limited regimes). The colors of the points reflect the mean relative difference (MRD) while the sizes reflect the standard deviation relative difference (SDRD).



**Figure 3.8:** scatter plots showing the relation of soil moisture ( $\theta_{40\text{cm}}$ ) and change in soil moisture ( $\Delta\theta_{40\text{cm}}$ ) in fig.3.7, for energy- (brown) and water-limited regimes (blue). MRD (mean relative difference) representing variability in  $\theta_{40\text{cm}}$  is correlated with MRD and SDRD (standard deviation relative difference) of  $\Delta\theta_{40\text{cm}}$ . Pearson's correlation coefficient ( $R$ ) and  $p$ -values are given.

soil moisture heterogeneity, they found variability in RWU over different soil water conditions. The 2D images of ERT obtained exhibited clear horizontal and vertical variation of soil electrical resistivity which qualitatively depict the pattern of water uptake in both crops. More importantly, they found the highest RWU heterogeneity under moderate water-stress conditions. In contrast, lower variability was observed for treatments with the highest water-stress. However, this could also potentially be due to some limitations of ERT for very low soil moisture contents. They primarily attributed the high RWU variability during moderate water-stress conditions to spatial variations in rooting pattern and/or root efficiency. Sim-

ilarly, it was shown in fig.3.6 that higher RWU variability was found during ELR. However, atmospheric demands may be the driving force for the observed variability in addition to root distribution.

Site-specific variations in plant water requirements is further hypothesized to affect RWU variability since a clear trend with  $\theta_{40\text{cm}}$  was not found, for both ELR and WLR, based on the TSA (fig.3.7). Weak correlations were obtained between MRD for RWU with the MRD and SDRD for  $\theta_{40\text{cm}}$  (fig.3.8). Based on MRD, the results show that locations with (lower) higher  $\theta_{40\text{cm}}$  does not necessarily result in a (smaller) larger RWU over the drying period. Furthermore, (lower) higher  $\theta_{40\text{cm}}$  also did not result in (less) more variable RWU. It is inferred that antecedent soil moisture state do not exert a significant control on RWU variability based on the in situ measurements.

### 3.3.3 Impacts on soil hydraulic properties

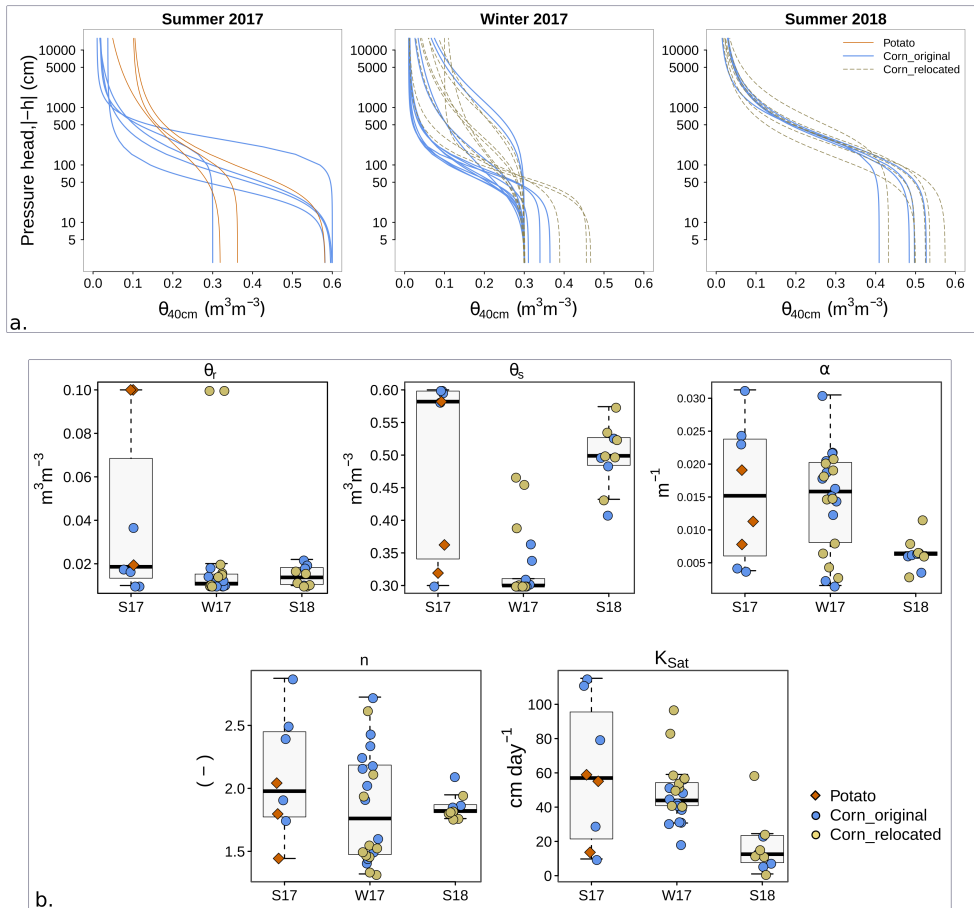
Inverse modeling resulted in varying accuracy of the estimated parameters, especially for the growing season. For some of the locations, lower  $R^2$  implies that the simulations did not suitably represent the conditions on the ground. A dual-porosity regime was tested but did not improve the accuracy or did not achieve model convergence. Nevertheless, for a little over half of the total dataset (40 of 75, 25 points per season), an  $R^2 \geq 0.5$  implies that a pore-flow model could reasonably describe the hydraulic properties at the corresponding stations. Most of the stations for winter 2017 could be reasonably represented by a pore flow model since 22 out of the 25 points resulted in  $R^2 > 0.5$  (table 3.1). However, for the two growing seasons, less than half of the 25 stations produced reasonable accuracy. This implies that a pore flow model is generally applicable for the bare winter period but not for the vegetated period investigated. Unsuccessful attempts using other flow regimes may be due to unsuitable initial hydraulic parameters used for inverse modeling optimization. Limited first-hand information on the macropore characteristics at the site or the lack of measurements for laboratory estimation of hydraulic parameters needed for, say a dual porosity flow regime, contributed to the outcome of the modelling attempts. Nevertheless, application of a non pore-flow regime merits future research that requires additional datasets other than those obtained in this study.

**Table 3.1:** Number of stations with  $R^2 \geq 0.5$ .

| Season      | $R^2 \geq 0.5$ | $R^2 < 0.5$ |
|-------------|----------------|-------------|
| summer 2017 | 8              | 17          |
| winter 2017 | 22             | 3           |
| summer 2018 | 10             | 15          |

For locations with  $R^2 \geq 0.5$ , water retention curves and hydraulic parameters per season are shown in fig.3.9. The remaining stations with  $R^2 < 0.5$  low were excluded and no longer further analyzed. Significant differences among the three groups were obtained for saturated water content ( $\theta_s$ ), alpha ( $\alpha$ ), and saturated conductivity ( $K_{sat}$ ) (Table 3.2) using all the points. However, when only points kept at the same locations (fig.3.9, blue lines and points), significant differences were only observed for  $\theta_s$  and  $K_{sat}$ . Further, between pairs of seasons,





**Figure 3.9:** Water retention curve and soil hydraulic parameters per season obtained using inverse modeling. a) Lines of the water retention curve colored based on the changes in their locations and crops at the site. Stations that were kept in the sample location are in blue while the orange lines were reinstalled at new locations after 2017 harvest, then represented by yellow lines. b) Box plots of the soil hydraulic parameters per season. The points are overlain to the box plots and colored per crop. Blue and yellow points represent corn while orange represent potato.

significant differences in  $\theta_s$  were found between 2017 winter and the two summers, considering all points and only those kept at their original locations (Table 3.3 and 3.5). However,  $\theta_s$  between the two summers were not significantly different, also for both cases. These results could be easily related to variability observed from TS analysis. Lower  $\theta_s$ 's estimated for winter 2017 is inferred to be due to the decrease in hydraulically effective pores after post-harvest tillage (see section 3.3.1) while higher  $\theta_s$  for the two growing seasons is in line with the current knowledge that roots or root growth improves soil hydraulic properties (e.g. Scanlan, 2009;

Scholl et al., 2014). Unlike  $\theta_s$ , p-values obtained for  $\alpha$  and  $K_{sat}$  parameters were not consistent when using all the points or only those at their original locations. For instance, the p-values for  $K_{sat}$  were conflicting in both cases tested, although it has been previously shown that structural changes in the soil greatly impacts  $K_{sat}$  (Ahuja et al., 1984; Fatichi et al., 2020). Results for  $\alpha$  and  $K_{sat}$  are therefore inconclusive compared to  $\theta_s$ . It is highly likely that the small number of samples contributed to the results even though non-parametric measures of significance were applied.

**Table 3.2:** Kruskal-Wallis test for significant differences among the soil hydraulic parameters for all stations with  $R^2 \geq 0.5$ .  $p$  – values  $< 0.05$  are marked with \*.

|          | $\theta_r$ | $\theta_s$ | $\alpha$  | n        | $K_{sat}$ |
|----------|------------|------------|-----------|----------|-----------|
| $\chi^2$ | 2.618      | 22.551     | 7.593     | 1.2091   | 13.23     |
| P-value  | 2.7e-01    | 1.277e-05* | 2.24e-02* | 5.46e-01 | 1.34e-03* |

**Table 3.3:** Pairwise Wilcoxon test to detect significant differences between pairs of seasons. Only parameters with  $p$  – value  $< 0.05$  in Table 3.2 are presented. S17 - summer 2017, W17 - winter 2017, and S18 - summer 2018.  $p$  – value  $< 0.05$  are marked with \*.

|           | $\theta_s$ | $\alpha$ | $K_{sat}$  |
|-----------|------------|----------|------------|
| S17 - W17 | 2.24e05*   | 1.61e02* | 7 2.87e04* |
| W17 - S18 | 2.91e03*   | 7.30e01  | 5.65e01    |
| S17 - S18 | 3.98e01    | 1.35e01  | 3.11e02*   |

**Table 3.4:** Kruskal-Wallis test for significant differences among the soil hydraulic parameters for stations kept in the same location and with  $R^2 \geq 0.5$ .  $p$  – value  $< 0.05$  are marked with \*.

|          | $\theta_r$ | $\theta_s$ | $\alpha$ | n     | $K_{sat}$ |
|----------|------------|------------|----------|-------|-----------|
| $\chi^2$ | 4.293      | 11.674     | 3.484    | 2.170 | 7.963     |
| P-value  | 0.117      | 0.003*     | 0.175    | 0.338 | 0.0187*   |

**Table 3.5:** Pairwise Wilcoxon test to detect significant differences between pairs of seasons. Only parameters with  $p$  – values  $< 0.05$  in Table 3.4 are presented. S17- summer 2017, W17 - winter 2017, and S18 - summer 2018.  $p$  – value  $< 0.05$  are marked with \*.

|           | $\theta_s$ | $K_{sat}$ |
|-----------|------------|-----------|
| S17 - W17 | 0.021*     | 0.506     |
| W17 - S18 | 0.007*     | 0.002*    |
| S17 - S18 | 0.176      | 0.063     |

## 3.4 Conclusion

In a cultivated fields, soil moisture in the active root zone of 40 cm depth shows considerable variability in space and time due to tillage and root-induced changes to soil structure. Based on in situ soil moisture measurements, root-induced subsurface soil moisture variability was much larger than tillage-induced variability. Furthermore, the variability in root water uptake is higher for periods with sufficient subsurface soil moisture to accommodate transpiration demands. Assessing the impacts of the soil moisture variability on the hydraulic properties proved to be challenging using the commonly-used pore-flow model. Among the soil hydraulic parameters estimated with reasonable accuracy, the changes in the saturated water content for different seasons were found to compliment the subsurface soil moisture variability observed from the temporal stability analysis. Changes for the remaining estimated parameters were not consistent and could not be easily related with the results from the temporal stability analysis. Perhaps the inverse modeling simulations could be improved with a different flow regime (e.g. dual-porosity). However, this was hampered by the availability of hydraulic parameters that could serve as initial values to ensure realistic optimized parameters.



"One person's trash is another person's treasure."

English proverb

# 4

## Can radar vegetation backscatter reflect root zone soil moisture?



Photo by: C. Nolet

This chapter is an adapted version of:

Carranza, C., van Emmerik, T., Benninga, H.-J., van der Velde, R., and van der Ploeg, M. J. (2020b). Exploring the sensitivity of vegetation backscatter to root zone soil moisture, *Submitted to Biogeosciences*

## Abstract

Monitoring root zone soil moisture dynamics is crucial for improving hydrological models and drought monitoring. This study explores the use of radar backscatter to estimate root zone soil moisture dynamics over agricultural areas. Sentinel-1 (C-band VV-polarization) backscatter was decomposed into the soil and vegetation contributions using the Water Cloud model for the growing seasons of 2016 to 2018, allowing us to specifically investigate the impact of increased vegetation water stress during the 2018 European drought. Soil backscatter strongly correlate with root zone soil moisture during non water-limited conditions. Surface and root zone soil moisture are de-coupled during water-limited conditions. However, vegetation backscatter was found to strongly correlate with soil moisture in the active root zone during the said conditions. These results suggest that radar remote sensing may provide a new source to estimate root zone moisture dynamics, especially during water-limited conditions.

## 4.1 Introduction

Root zone soil moisture is an essential component of the hydrologic cycle and is useful for drought monitoring, irrigation scheduling, and carbon cycle modeling (Bolten et al., 2009; Entekhabi et al., 1996; Falloon et al., 2011). In recent years, radar satellites have been a source of spatio-temporal root zone soil moisture information. Although radars only measure the upper surface soil layer, estimation of root zone soil moisture from surface soil moisture has been achieved through process-based models via data assimilation (e.g. Reichle et al., 2004; King et al., 2005; Draper et al., 2012) or data-driven approaches (e.g. Albergel et al., 2008; Kornelsen and Coulibaly, 2014; Gao et al., 2019). Estimation of root zone soil moisture conditions directly from surface values has been hampered by the variability along the soil profile due to non-linearities in the rate of processes controlling soil moisture state. For instance, varying rates of infiltration, evapotranspiration, root water uptake and the heterogeneity in soil physical properties contribute to vertical variability in soil moisture. Surface soil moisture may be de-coupled from deeper layers, especially during dry periods (e.g. Wilson et al., 2003; Hirschi et al., 2014; Carranza et al., 2018).

Over agricultural areas, radar backscatter contains information on both soil and vegetation water content since the incoming microwave signal is influenced by the geometric and dielectric properties of the scatterers (Ulaby et al., 1986). Radar specification equally influences the magnitude of scattering processes. The primary scatterers are elements (leaves or stalks) with sizes similar to or larger than the microwave wavelength or with an orientation similar to that of the incoming signal polarization. Elements smaller than the microwave wavelength contribute little to the backscatter but attenuate the signal. Radars operating at shorter wavelengths (X-band) do not usually penetrate into agricultural canopies in contrast to longer wavelengths (L- or C-band). Over bare soils or sparsely vegetated fields, the total backscatter is dictated (mostly) by soil characteristics. During vegetated periods, both the overlying vegetation canopy and the underlying surface soil layer influence the total backscatter (Attema and Ulaby, 1978; Imhoff, 1995). Vegetation results in volume scattering and attenuation of the soil signal. However, the soil signal is not totally diminished even under the mature agricultural canopy as it is carried within the soil-vegetation interaction signal. Estimation of the soil and vegetation contributions, for instance using the Water Cloud model (Attema and Ulaby, 1978), further allows for the estimation of each respective water contents.

The soil component of the radar backscatter has been widely studied in relation to estimating surface soil moisture for bare soils. When the influence of the overlying vegetation canopy on the total backscatter is significant, the vegetation component should be quantified to improve soil moisture estimation (e.g. Bindlish and Barros, 2001). The vegetation backscatter from (agricultural) canopies has been investigated to gain insights on the phenology (Bakar et al., 1997; Vreugdenhil et al., 2018; Asilo et al., 2019), presence of water stress (McNairn et al., 2002; van Emmerik et al., 2017), and productivity (Li et al., 2003; Setiyono et al., 2019). Vegetation water content (VWC) has been estimated from the vegetation backscatter (Notarnicola and Posa, 2007; Srivastava et al., 2015) and radar vegetation index (RVI, Kim and van Zyl, 2004; Kim et al., 2011, 2013; Huang et al., 2015). Diurnal variations in backscatter

have also used as an indicator of crop water stress (Steele-Dunne et al., 2012; van Emmerik et al., 2015; Van Emmerik et al., 2016). From agronomy, leaf and soil water potential are known to be in equilibrium at pre-dawn (e.g. Sellin, 1999). This indicates the potential of VWC to reflect root zone soil moisture conditions. Although soil water storage have been inferred from vegetation dynamics using optical remote sensing products at global Wang-Erlandsson et al. (2016) and catchment scales Sriwongsitanon et al. (2016), the utility of vegetation backscatter has not yet been explored despite the dependence of VWC to root zone soil moisture.

Both soil and vegetation components of the backscatter have the potential to indicate root zone soil moisture conditions. The former measures the uppermost soil layer which greatly influences the amount of water in the root zone, while the latter is connected via its roots, and whose only source of water is the soil. This study explores the potential of radar backscatter and its components, simulated using the Water Cloud Model (WCM), to estimate root zone soil moisture over corn and grass fields. We utilized measurements from two soil moisture networks (Raam and Twente) over three growing seasons (2016 to 2018) which allowed comparison of different meteorological conditions. More specifically, comparison is made between 2018 growing season, which coincided with the European summer drought, and the two previous years. With this paper, we explored the impact of different soil moisture conditions on the ability of radar backscatter components to reflect root zone conditions, that may be used for applications such as water stress detection, irrigation planning and fire risk assessment.

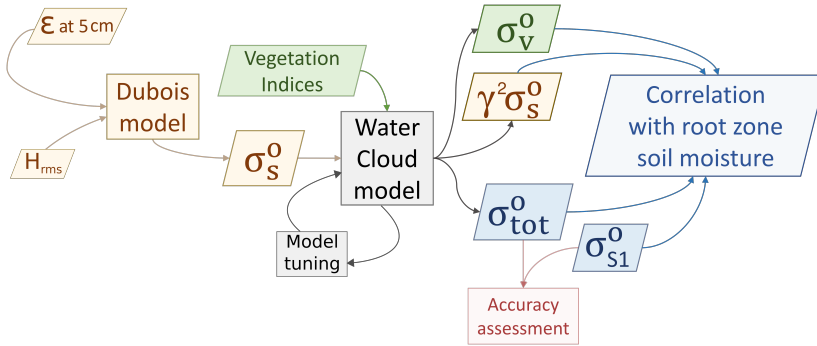
## 4.2 Materials and Methods

Correlations between radar backscatter and its components with root zone soil moisture were calculated to determine if the former is able to reflect the latter in agricultural fields. First, time series plots of radar backscatter from Sentinel-1 (S1) (VV- and VH- polarization) and soil moisture (surface and root zone) were investigated to show the overall trends over the growing seasons of 2016 to 2018. The soil and vegetation contributions to the total backscatter were simulated using the Water Cloud Model (WCM) and calibrated against S1, as summarized in fig.4.1. Only VV-polarization was investigated since its has been shown to have higher sensitivity to soil moisture (e.g. El Hajj et al., 2017; Amazirh et al., 2018; Benninga et al., 2019). Vegetation indices and Dubois et al. (1995)'s model for the soil backscatter were used to derive the vegetation and soil parameters of the WCM, respectively. We focused on the months of June - September to ensure that crops have had considerable growth in order to investigate the influence of both radar backscatter components.

### 4.2.1 Soil moisture network locations

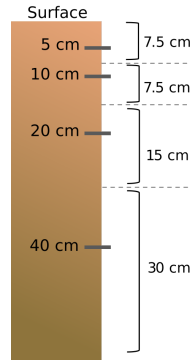
Agricultural fields within Twente (Dente et al., 2011) and Raam (Benninga et al., 2018) soil moisture networks were investigated over the growing seasons of 2016 - 2018 (fig.4.3a). Out of 35 stations, 25 were investigated (14 in Raam and 11 in Twente) based on availability of data in time and at depth. Furthermore, these stations hold sandy soils so the influence of soil texture on the correlations between radar backscatter and soil moisture can be assumed





**Figure 4.1:** Flowchart showing the sequence of analysis applied. The symbols used are:  $\sigma_s^0$  for soil backscatter,  $\gamma^2\sigma_s^0$  for attenuated soil backscatter,  $\sigma_v^0$  for vegetation backscatter,  $\sigma_{tot}^0$  for the total backscatter,  $\sigma_{S1}^0$  for Sentinel-1 VV polarization backscatter,  $\epsilon$  for dielectric constant. Box colors refer to different model components (green for vegetation, brown for soil, and blue for the total backscatter).

to be negligible. We focused on the two most common crops encountered: grass and corn (49 instances for grass and 18 for corn, table 4.1). At each station in the Raam network, Decagon ECH2O (EC-TM or 5TM) sensors were installed at 5, 10, 20, 40, and 80 cm depths, and records data every 15 minutes. In Twente, installation of sensors were not uniform at depth, so only stations which have at least up to 40 cm depths were utilized. The selected stations also have no or limited data gaps over the three-year period.



**Figure 4.2:** Diagram showing the installation setup for each station. Soil moisture sensors were installed at 5, 10, 20, and 40 cm depths for each station. For each measurement depth, an associated soil thickness is determined based on the midway distance between two adjacent measurement points. Root zone soil moisture is represented by the zone-weighted depth-average values that are aggregated based on both the measurement and its associated soil thickness (see eq.4.1).

The 15-minute data were further transformed into zone-weighted depth-averaged root zone values ( $\theta_{rz}$ ). This was calculated using measurements down to 40 cm for corn and 20 cm for

**Table 4.1:** Crop at the study sites for 2016 - 2018. Stations starting with RM belong to Raam network and those starting in TW are from Twente network

| Station | 2016         | 2017         | 2018         | Area (ha.) | % sand |
|---------|--------------|--------------|--------------|------------|--------|
| RM01    | grass        | grass        | grass        | 0.59       | 91.3   |
| RM02    | sugar beet   | fennel       | corn         | 2.92       | 90.4   |
| RM03    | grass        | grass        | grass        | 5.65       | 93.3   |
| RM04    | grass        | grass        | grass        | 2.04       | 90.0   |
| RM05    | onion        | lettuce      | lettuce      | 6.24       | 93.1   |
| RM06    | grass        | grass        | grass        | 1.62       | 83.7   |
| RM07    | corn         | potato       | grass        | 3.24       | 82.1   |
| RM08    | sugar beet   | winter wheat | beans        | 2.51       | 92.8   |
| RM09    | sugar beet   | potato       | corn         | 4.83       | 95.4   |
| RM11    | corn         | corn         | winter wheat | 3.49       | 94.8   |
| RM12    | grass        | grass        | grass        | 13.16      | 92     |
| RM13    | corn         | corn         | corn         | 8.48       | 96.7   |
| RM14    | grass        | grass        | grass        | 9.56       | 90     |
| RM15    | grass        | grass        | grass        | 2.05       | 88.6   |
| TW01    | grass        | grass        | grass        | 2.46       | 66.3   |
| TW02    | grass        | grass        | grass        | 4.76       | 80     |
| TW07    | winter wheat | corn         | corn         | 4.55       | 66.3   |
| TW08    | corn         | corn         | corn         | 6.14       | 88.3   |
| TW09    | corn         | corn         | grass        | 2.04       | 87     |
| TW10    | corn         | corn         | winter wheat | 8.88       | 88.3   |
| TW13    | grass        | grass        | grass        | 4.15       | 88.3   |
| TW14    | grass        | grass        | grass        | 3.18       | 78.67  |
| TW15    | grass        | grass        | grass        | 1.63       | 66.3   |
| TW17    | grass        | grass        | grass        | 2.03       | 88.33  |
| TW18    | corn         | corn         | corn         | 2.00       | 78.67  |

grass (fig.4.2). Measurements at 80 cm were absent for some stations and were, therefore not used in order to have uniform root zone across the study sites. Root zone soil moisture  $\theta_{rz}$  is calculated using:

$$\theta_{rz} = \frac{\sum_{j=1}^n \theta_j \Delta z_j}{z} \quad (4.1)$$

where  $\theta_j$  (in  $\text{m}^3 \text{m}^{-3}$ ) is the volumetric water content for measurement depth  $j$  (cm),  $\Delta z_j$  (cm) is the thickness of soil associated with the measurement depth, and  $z$  (cm) is the total averaging depth. The mean values from 0630 - 0700 in the morning were calculated from the 15-min.  $\theta_{rz}$  data to match timing of S1 descending pass over the study areas. This time of day is close to pre-dawn conditions when leaf and soil water potential are known to be in equilibrium.

#### 4.2.2 Sentinel-1 datasets

Radar backscatter observations were obtained from Sentinel-1A and -1B ( $\sigma_{S1}^0$ ), which operate in the C-band with a center frequency of 5.405 GHz. Over land, the Interferometric Wide Swath (IW) mode provides images at VV and VH polarization, with a pixel spacing of 10 m

$\times 10$  m and radiometric accuracy of 1 dB ( $3\sigma$ ) Torres et al. (2012). Datasets were acquired between June 2016 and September 2018. We used Level-1 Ground Range Detected (GRD) products, derived from focused synthetic aperture radar (SAR) images that were multi-looked and projected to a ground-range, using an Earth Ellipsoid model Filipponi (2019).

#### *Incidence angle normalization*

The incidence angle at which Sentinel-1 observes the earth surface depends on the location and orbit in which an image is acquired. For each scene in the IW mode, the incidence angle can vary between  $29.1^\circ$  and  $46^\circ$ . We normalized the observations to a common incidence angle of  $37.5^\circ$  by a cosine correction:

$$\sigma_{ref}^o = \sigma^o \frac{\cos^n(\varphi_{ref})}{\cos^n(\varphi_{inc})} \quad (4.2)$$

where  $\sigma_{ref}^o$  (in  $\text{m}^2 \text{m}^{-2}$ ) is the backscatter observation normalized to a reference angle  $\varphi_{ref}$  of  $37.5^\circ$ ,  $\sigma^o$  is the backscatter observation (in  $\text{m}^2 \text{m}^{-2}$ ) and  $\varphi_{inc}$  is the local projected incidence angle (in degrees). The coefficient  $n$  is equal to 1 if re-radiation of the satellite signal from the earth surface is isotropic (volume scattering) and  $n$  is equal to 2 if re-radiation follows Lambert's cosine law (Ulaby et al., 1982). Here we assume  $n = 2$ , because previous studies obtained good results with this assumption (Mladenova et al., 2013; Van der Velde and Su, 2009).

The values of S1 pixels were averaged over each field to reduce radiometric uncertainty. Benninga et al. (2019) showed that spatially averaging  $\sigma_{S1}^0$  over areas approximating the acreage of agricultural fields improves radiometric uncertainty by up to 0.85 dB for the VV polarization and up to 0.89 dB for the VH polarization. However, we still encountered considerable variability in the spatially averaged S1 that resulted in lower accuracy of the Water Cloud Model (WCM) simulations. Therefore, we investigated the mean  $\sigma_{S1}^0$  from all the fields for each S1 overpass per crop type to focus on the general trends over the growing season.

**Table 4.2:** Fitting parameters for Water Cloud Model (WCM) for each growing season and crop type

| Year | Corn   |        | Grass  |        |
|------|--------|--------|--------|--------|
|      | A      | B      | A      | B      |
| 2016 | 0.1713 | 0.0782 | 0.0382 | 0.4194 |
| 2017 | 0.0953 | 0.1502 | 0.0279 | 0.2941 |
| 2018 | 0.1625 | 0.1848 | 0.0288 | 0.2662 |

#### 4.2.3 Modeled radar backscatter components

At the C-band, the total backscatter for agricultural fields represents the integrated signal of the vegetation canopy and underlying surface soil layer. Using the Water Cloud model Attema

and Ulaby (1978) (WCM), components of the total backscatter ( $\sigma_{tot}^o$ ) for a dual-polarized signal  $pq$  ( $p$  and  $q$  is  $V$  or  $H$ ) were estimated using:

$$\sigma_{tot,pq}^o = \sigma_{veg,pq}^o + \gamma_{pq}^2 \sigma_{soil,pq}^o \quad (4.3)$$

$$\gamma_{pq}^2 = e^{-2B_{pq}V_1 \sec\theta} \quad (4.4)$$

$$\sigma_{veg,pq}^o = AV_2 \cos\theta (1 - \tau^2) \quad (4.5)$$

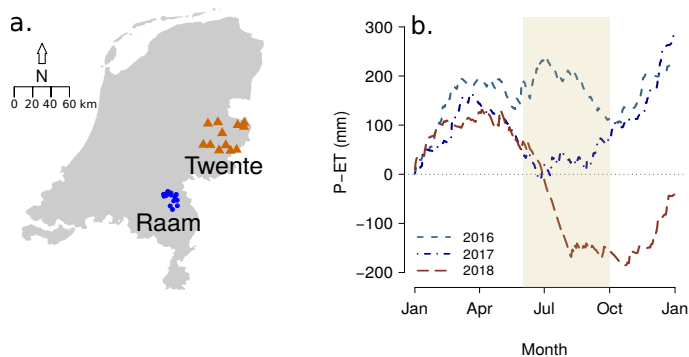
where  $\sigma_{veg}^o$  is the vegetation component,  $\sigma_{soil}^o$  is the soil component, and  $\gamma^2$  is the two-way attenuation.

The  $\sigma_{soil}^o$  was estimated using Dubois et al. (1995)'s formulation for VV-polarization. The surface roughness parameter is represented by the height standard deviation ( $H_{rms}$ , in cm). A constant value of  $H_{rms} = 1$  was used over the study period. The selected  $H_{rms}$  was found to be well within the range of values measured by Sano et al. (1998) for grass (0.67 - 1.99 cm) and by Rakotoarivony et al. (1996) and Martinez-Agirre et al. (2017) for corn after sowing (0.8 - 1.5 cm, and  $\sim 0.8$  -  $\sim 1.8$  cm, respectively). In addition, Benninga et al. (2019) applied values very close to  $H_{rms} = 1$  for estimation of soil moisture in the Twente network.

The parameters  $V_1$  and  $V_2$  are vegetation descriptors which are used to express microwave scattering mechanisms related to canopy geometry and attenuation in the canopy layer related to the vegetation water content, respectively.  $A$  and  $B$  are fitting parameters that depend on the vegetation descriptor. We set Leaf Area Index (LAI) for  $V_1$  since biomass directly relates to canopy geometric properties. Furthermore, LAI has been commonly used as a vegetation descriptor in previous studies (e.g. Kumar et al., 2012; Dabrowska-Zielinska et al., 2007; Prevot et al., 1993). The Radar vegetation index (RVI) was used to represent  $V_2$  as it has been previously shown to relate with vegetation water content (e.g. Kim et al., 2011; Huang et al., 2015). LAI was obtained from the 8-day composite of MODIS. Missing data were linearly imputed to obtain values on days with available S1 and soil moisture data. RVI was estimated from S1 VH and VV observations using the formula:

$$RVI = \frac{4VH}{VH + VV} \quad (4.6)$$

This follows Charbonneau et al. (2005), who modified the original RVI formulation of Kim and van Zyl (2004) when only one type of cross polarization is available. However, Charbonneau et al. (2005) worked with HH+HV based on the assumption that HH $\sim$ VV. The mean value of all the fields for each crop was calculated, similar to the radar and soil moisture datasets. The WCM fitting parameters were constrained by calibrating the model per year using all the available S1 observations (Table 4.2). We minimized the RMSE as the objective function for calibration between model outputs and S1 using a conjugate gradient method (Fletcher and Reeves, 1964).



**Figure 4.3:** *a.* Location of study sites. *b.* Cumulative potential rainfall deficit relative to the start of each year to show water availability during the growing season (yellow region) of 2016 - 2018. Daily gridded  $P$  and  $ET_{pot}$  from the Royal Netherlands Meteorological Institute (KNMI), were used to calculate the mean potential water deficit ( $P - ET_{pot}$ ) from all the stations. Periods with potential rainfall deficits occur when  $ET_{pot}$  is higher than  $P$ , which translates to negative (cumulative) values.

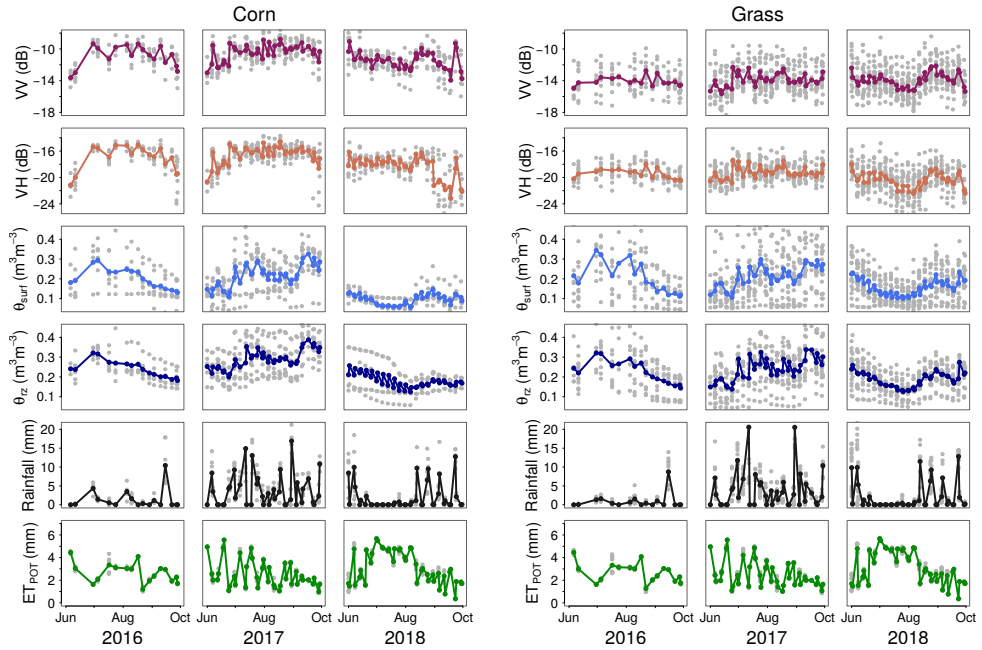
#### 4.2.4 Meteorological conditions during the 2016 - 2018 growing seasons

The growing season in the Netherlands for most (annual) crops falls between the months of May - October when the average temperature and sun hours increase. Corn seeds are sown annually in May, reach maturity in late July, and are harvested at the end of September or early October. Grasses are perennial crops which grow between April - October. Meteorological conditions during the growing seasons of 2016 - 2017 were generally considered to be energy - limited or non water - limited because rainfall events generally compensated the potential evapotranspiration demands and lead to minimal potential rainfall deficit over the growing seasons ( $P - ET_{pot}$ , fig.4.3b). In 2018, western Europe was hit by a drought during the summer months (Toreti et al., 2019; Vogel et al., 2019). In the Netherlands, the drought started at beginning of June up until early August. Compared to the two previous years, rainfall deficit was much higher during the summer of 2018. Furthermore, 2018 conditions saw water-limited conditions, particularly the the start of July when a large increase in water deficit occurred.

## 4.3 Results and Discussion

### 4.3.1 Radar backscatter and soil moisture time series trends

Figure 4.4 are the time series plots of S1 backscatter ( $\sigma_{S1}^0$ , VV- and VH polarization) and soil moisture ( $\theta_{surf}$  and  $\theta_{rz}$ ) over three growing seasons. In general, the trends for  $\sigma_{S1}^0$  in both channels were similar over the three year period for both crops. For 2016 and 2017, increasing  $\sigma_{S1}^0$  in VV and VH was observed as the growing season progressed for corn, albeit fewer observations were obtained for 2016 since Sentinel-1B was not yet operational. However,  $\sigma_{S1}^0$

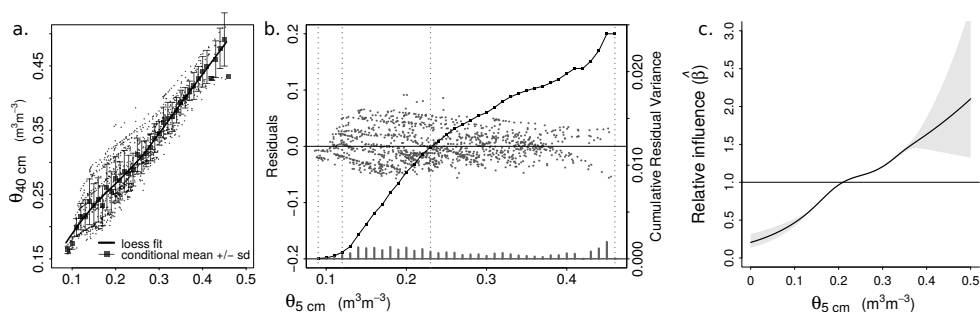


**Figure 4.4:** Time series plots of Sentinel-1 backscatter ( $\sigma_{S1}^0$ , VV and VH- polarization) and soil moisture at the study sites ( $\theta_{surf}$  and  $\theta_{rz}$ ).  $\theta_{surf}$  is represented by measurement at 5 cm while  $\theta_{rz}$  is calculated as the zone-weighted depth-averaged values (see section 4.2.1). Rainfall and potential evapotranspiration ( $ET_{pot}$ ) show the meteorological conditions over the growing seasons. Colored points represent the mean values while the grey dots are the complete observation points.

was relatively the same or had a flatter trend for grass fields. In 2018, however, a decreasing trend for both crops is visible in  $\sigma_{S1}^0$  until early August, coinciding with the 2018 summer drought. After which,  $\sigma_{S1}^0$  increased until the end of the growing season for both crops.

As expected, soil moisture trends ( $\theta$ ), both  $\theta_{surf}$  and  $\theta_{rz}$ , reflected meteorological conditions over the growing seasons. Soil moisture closely followed the amount and frequency of rainfall during the growing seasons (fig.4.4). Frequent rainfall events in June 2016 led to an increase in soil moisture. In 2017 soil moisture was slightly lower at the beginning of the growing season but constantly increased because of steady rainfall events. For 2018, the study area only received minimal rainfall at the beginning of the drought in June. Higher  $ET_{pot}$  combined with the lack of rainfall in July resulted in low soil moisture. Recovery in soil moisture started early to mid-August when steady amounts of rainfall resumed. The decrease in  $\theta_{surf}$  observed for both crops in 2018 was more evident for corn compared to the grass fields investigated. However, the decrease in  $\theta_{rz}$  for both crops appear similar.

Trends for  $\sigma_{S1}^0$  and  $\theta$  over the growing seasons in fig.4.4 appear to be more similar in 2018 compared to the two previous years. For instance in 2016, a decreasing trend in  $\theta$  at the two depths investigated is the opposite of the increasing  $\sigma_{S1}^0$  for corn or the static trend in grass. A



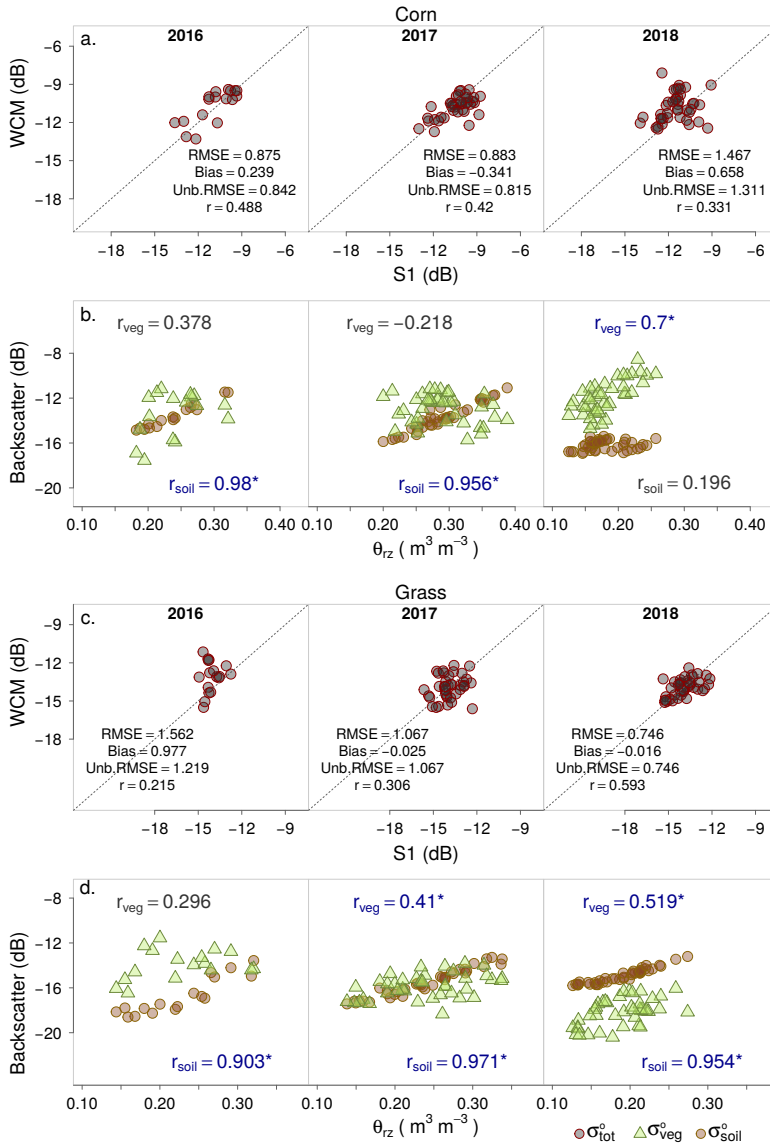
**Figure 4.5:** Results of coupling analysis between 5 cm and 40 cm soil moisture using average values from all the sites. The two statistical methods (residuals analysis and distributed lag non-linear model) applied following Carranza et al. (2018) showed complementary results. The scatterplot in a. shows the fit of a loess over the data points, from which the residuals in b. are derived. The plot of the residuals show higher variance for surface soil moisture  $< 0.23 \text{ m}^3 \text{ m}^{-3}$ . These results match well with the the range of values with  $\beta < 1$  in plot c. Lower  $\beta$  indicate weaker lag dependence between surface and subsurface values. Based on these two results, it is interpreted that decoupled conditions occur when surface soil moisture is approximately  $\theta_c < 0.2 \text{ m}^3 \text{ m}^{-3}$ .

strong correlation between  $\sigma_{S1}^0$  and  $\theta$  is therefore not expected for 2016 and 2017 based on the time series plots. However, dissimilarity in  $\sigma_{S1}^0$  and  $\theta$  confirms the non-negligible influence of vegetation on the former. Similar trends observed between  $\sigma_{S1}^0$  and  $\theta$  for 2018 may have been influenced by the summer drought that year. It further suggests that the total backscatter is roughly indicative root zone soil moisture trends over vegetated agricultural fields during droughts.

The 2018 summer drought resulted in decoupled soil moisture conditions, which means that the surface no longer reflected subsurface or root zone conditions (Carranza et al., 2018). Based on the soil moisture measurements at 5 cm and 40 cm depths from all the study sites, we found that coupling between surface and subsurface soil moisture is promoted for conditions above  $0.20 \text{ m}^3 \text{ m}^{-3}$ , referred to as  $\theta_c$  (fig.4.5). In this case, soil moisture measurements generally indicate that decoupling occurs during drier soil moisture conditions, which may, among others, be due to increased surface evaporation (e.g. Hirschi et al., 2014). In 2018,  $\theta_{surf}$  was generally lower than  $\theta_c$  which means that it is decoupled from root zone conditions. Based on fig.4.4,  $\theta_{rz}$  in 2018 were higher than  $\theta_{surf}$ . Furthermore, the difference between surface and root zone was observed to be larger for corn than for grass, which may be due to deeper rooting depth associated with the former. Decoupling of surface and subsurface soil moisture may have influenced the ability of the radar components to reflect root zone conditions, as discussed in the succeeding section.

### 4.3.2 Water Cloud model simulations

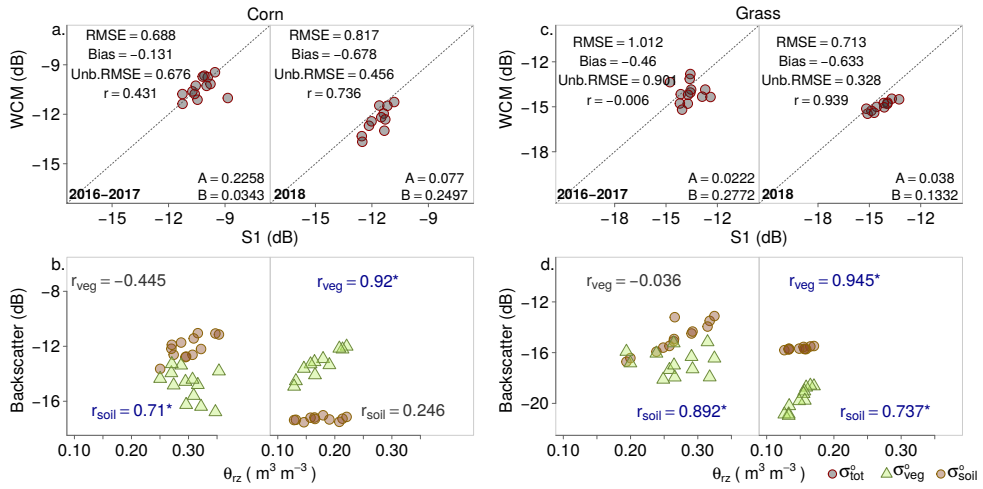
The results of the water cloud model (WCM) simulations are given in fig.4.6. The simulated values yielded RMSEs of 0.8 - 1.4 dB for corn and 0.75 - 1.5 dB for grass. The RMSEs obtained



**Figure 4.6:** Simulated backscatter in the VV-polarization for corn and grass using Water Cloud Model (WCM) calibrated using Sentinel-1 observation (a and c). Pearson's correlation coefficient ( $r$ ) between soil and vegetation backscatter ( $\sigma_{soil}^0$  and  $\sigma_{veg}^0$ ) with root zone soil moisture ( $\theta_{rz}$ ) are further calculated (b and d). Significant correlations ( $p$ -value < 0.05) are indicated by \*.

here were found to be similar to previous studies in C-band for both crops (e.g. Baghdadi et al., 2017; Aline and Pierre, 2013). Focusing on the estimated backscatter components,





**Figure 4.7:** Water Cloud Model simulation for July 2016 - 2017 and 2018 to compare the impact of water-limited conditions on the root zone - backscatter correlations ( $r$ ). Datasets for 2016 and 2017 were combined because of similar meteorological conditions. The fitting parameters used and accuracy of the Water Cloud Model (WCM) simulations are given in a and c. The correlations obtained for the backscatter components are given in b and d. Significant correlations ( $p$ -value < 0.05) are indicated by \*.

the behavior of  $\sigma_{soil}^0$  and  $\sigma_{veg}^0$  in relation with  $\theta_{rz}$  was observed to change over the growing seasons. For both corn and grass, the simulations for 2016 and 2017 yielded similar range of backscatter values for both  $\sigma_{soil}^0$  (-16 to -11 dB) and  $\sigma_{veg}^0$  (-17 to -9 dB). In corn,  $\sigma_{soil}^0$  showed strong correlations with  $\theta_{rz}$  for 2016 and 2017, but had much lower correlation in 2018. The correlations for  $\sigma_{veg}^0$  were the opposite of those obtained for  $\sigma_{soil}^0$ : higher in 2018 compared to the two previous years. In grass, strong correlations between  $\theta_{rz}$  and  $\sigma_{soil}^0$  were found over the three growing seasons. However, similar to corn, an increase in correlation between  $\theta_{rz}$  and  $\sigma_{veg}^0$  was also observed in 2018. The relative magnitude of the backscatter components differ for the two crops in 2018. Relatively higher  $\sigma_{veg}^0$  was simulated for corn compared to  $\sigma_{soil}^0$ . This apparent increase is due to lower  $\sigma_{soil}^0$  in 2018 compared to the two previous years. However, the opposite was obtained for grass, where the magnitude of  $\sigma_{veg}^0$  decreased relative to  $\sigma_{soil}^0$ .

Relating the estimated radar components with meteorological conditions, weaker correlations between  $\sigma_{soil}^0$  and  $\theta_{rz}$  coincided with decoupled soil moisture status in 2018. This is expected since  $\sigma_{soil}^0$  only measures the upper surface soil layer. In fig.4.4, the 2018 summer drought resulted in very dry  $\theta_{surf}$ , especially in July when the water deficit increased drastically (fig.4.3b). In contrast, the previous years promoted coupling between  $\theta_{surf}$  and  $\theta_{rz}$ . Hence,  $\sigma_{soil}^0$  was considered a good indicator of  $\theta_{rz}$  for 2016 and 2017.

For both crops, stronger correlations with  $\sigma_{veg}^0$  over the 2018 growing season suggest its potential as a direct measure of  $\theta_{rz}$  during dry periods. We hypothesize that during the 2018

summer drought, the root zone soil moisture - vegetation interaction was enhanced in  $\sigma_{veg}^0$  because very dry  $\theta_{surf}$  resulted in a static or relatively flat  $\sigma_{soil}^0$  signal, especially for corn. In fig.4.4, the decreasing trend in Sentinel-1 ( $\sigma_{S1}^0$ ) in 2018 must have reflected the decreasing vegetation water content in  $\sigma_{veg}^0$ .

To further assess the potential of  $\sigma_{veg}^0$  in reflecting  $\theta_{rz}$  for water-limited conditions, a comparison is made between the WCM simulations using only the measurements for the month of July. Based on fig.4.3b, water-limited conditions mostly occurred in July even though the 2018 summer drought started at the beginning of June. August already saw the start of recovery in soil moisture (fig.4.4) because of the resumption of rainfall events. WCM simulations for July 2016 - 2017 and 2018 are shown in fig.4.7. The fitting parameters (A and B) for the WCM were re-calibrated to the subsetted data. Observations for 2016 and 2017 were combined since both years have similar meteorological conditions. Plus, the former year had very few points (only 3 - 4) for model re-calibration. In fig.4.7, July only simulations for the backscatter components closely follow the overall trends for the growing season. However, the RMSEs obtained for the WCM improved, especially for the July 2018 simulations. Correlations obtained for  $\sigma_{soil}^0$  and  $\sigma_{veg}^0$  are also similar to those in fig.4.6. In both plots, strong significant correlations are found between  $\sigma_{veg}^0$  and  $\theta_{rz}$  for 2018. Even higher correlations for July further strengthens the hypothesis that  $\sigma_{veg}^0$  can potentially reflect  $\theta_{rz}$  for dry periods.

Fig.4.6 and 4.7 further demonstrate contrasting magnitudes of the simulated  $\sigma_{veg}^0$  for grass and corn during water-limited conditions. As shown in fig.4.6 and 4.7, drying and water-limited conditions resulted in higher  $\sigma_{veg}^0$  for corn, a broad-leaf crop, while a much lower  $\sigma_{veg}^0$  was estimated for grass which has narrow leaves. Previous studies have demonstrated the impact of vegetation geometry (i.e. "narrow" vs "broad" leaf crops) on the volumetric scattering of the radar signal. Macelloni et al. (2001) showed that vegetation with similar biomass but different geometries will result in different backscatter coefficients. Furthermore, in the broad-leaf crops, an increase in biomass results in an increase in backscatter because of the dominance of scattering process. For narrow-leaf plants, a flat or decreasing trend occurs with an increase in biomass because of the major contribution of absorption. Based on the WCM simulations for 2018 in fig.4.6 and 4.7, increased absorption of the incoming radar signal may have further decreased the magnitude of  $\sigma_{veg}^0$  in grass as compared to corn, in addition to the decreasing vegetation water content during the summer drought.

If  $\sigma_{veg}^0$  strongly correlate with  $\theta_{rz}$  during water-limited conditions, why then would they correlate poorly for average meteorological or wet conditions? The  $\sigma_{veg}^0 - \theta_{rz}$  interactions may have been highlighted during water-limited conditions because of the absence of factors that can obscure the said relation. The lack of rainfall during 2018 summer drought facilitated accounting of the decrease in  $\theta_{rz}$  solely to vegetation water uptake. Therefore, this easily allows for the changes in vegetation water content from  $\sigma_{veg}^0$  to be related with the depleting  $\theta_{rz}$ . For non-water limited conditions, there may be water surplus in the root zone that is not necessarily taken up by vegetation to meet transpiration demands. The subsurface fluxes from rainfall and drainage to deeper soil layers must first be accounted before a measure of the changes due to root water uptake can be quantified. Therefore, investigating  $\sigma_{veg}^0 - \theta_{rz}$  interactions appear to be less straightforward during non-water limited conditions. The signal from

$\sigma_{soil}^0$  may also dominate the total backscatter with higher  $\theta_{surf}$ . Furthermore, some studies found increased root water uptake variability for moderately stressed crops (e.g. Srayeddin and Doussan, 2009), which may contribute to uncertainties in assessing  $\sigma_{veg}^0 - \theta_{rz}$  relations during non-water limited conditions. Nevertheless, strong correlations between  $\sigma_{veg}^0$  and  $\theta_{rz}$  during the 2018 summer drought suggest that the limitations  $\sigma_{soil}^0$  could potentially be complemented by information contained in  $\sigma_{veg}^0$ . This may allow derivation of empirical relations using both components of  $\sigma_{S1}^0$  for the whole range of root zone soil moisture. The results obtained further suggest that vegetation-only scattering may potentially be used to directly monitor subsurface soil moisture changes during agricultural droughts or to assess wildfire hazards, especially for broad-leaf vegetation in agricultural and forested areas.

### 4.3.3 Implications for water management in vegetated regions

The sensitivity of the vegetation backscatter to root zone soil moisture highlighted during periods of drought is foreseen to be beneficial for managing vegetated areas (i.e. agricultural and forests) in order to maintain productivity and to prevent loss of biodiversity.

In agricultural areas, S1 vegetation backscatter will allow frequent root zone soil moisture information for monitoring and irrigation advisory that is crucial during (agricultural) droughts. Using the vegetation backscatter will allow for a simplified method to estimate root zone soil moisture that eliminates the uncertainties associated with soil moisture retrieval using the soil backscatter during vegetated periods. Figure 4.4 shows a decreasing trend in S1 total backscatter during the 2018 drought, suggesting that it can directly be applied for estimating root zone soil moisture without decomposing the signal into its components. The results found here can be further extended to forested areas and may provide information for assessing the probability of wildfire occurrence. Previous studies already demonstrated the importance of SAR for mapping and monitoring pre-fire conditions such as fuel type (e.g. Arroyo et al., 2008; Chuvieco et al., 2003) and live fuels content (e.g. Rao et al., 2020). Root zone soil moisture from SAR will provide complementary information on the water status in the soil that is valuable for mitigating wildfire occurrence.

## 4.4 Conclusions

The potential use of radar backscatter components to monitor root zone soil moisture dynamics in agricultural fields is demonstrated for varying meteorological conditions. At the regional or catchment scale, meteorological conditions dictate which component of the total backscatter can reflect root zone soil moisture better. During non water-limited conditions, soil backscatter shows strong and significant correlations with root zone soil moisture. However, water-limited conditions such as the drought in 2018, highlights the ability of vegetation backscatter to reflect root zone soil moisture. This further indicates the potential use of vegetation backscatter for water stress and drought monitoring or for fire risk-assessment of agricultural and forested areas.



"We sit in the mud, my friend, and reach for the stars."

Ivan Turgenev

# 5

## Monitoring agricultural field trafficability using Sentinel-1



Photo by: E. Niles

This chapter is based on:

Carranza, C., Benninga, H.-J., van der Velde, R., and van der Ploeg, M. J. (2019). Monitoring agricultural field trafficability using Sentinel-1, *Agricultural Water Management* 224, 105 698, doi: 10.1016/j.agwat.2019.105698

## Abstract

The use of heavy mobile machinery in agriculture for tillage and harvesting is now indispensable since it facilitates farming over large areas. However, one of the impacts of regular and prolonged use of heavy mobile machinery is soil compaction. To help minimize this harmful effect, trafficability of agricultural fields at any particular time needs to be determined. Soil moisture acts as one of the dominant controls for field trafficability. Satellites such as Sentinel-1, which is one source of spatio-temporal soil moisture information, could be useful in assessing trafficable conditions. One limitation of satellite-derived soil moisture is that only the surface layer of soil is mapped. In this study, we determined the feasibility of using Sentinel-1 surface soil moisture readings to monitor trafficability during 2016 and 2017. We first determined coupled conditions when surface soil moisture is a good indicator for values at the subsurface. We applied a probabilistic approach to determine trafficability using extensive in situ measurements of penetration resistance and surface soil moisture over a variety of crops. Trafficability is expressed as the probability that penetration resistance will exceed a threshold, for a given soil moisture value. Furthermore, we investigated the variability encountered from these measurements to gain insights into other temporal controls. Our results show coupled conditions for soil moisture  $\geq 0.19 \text{ cm}^3 \text{ cm}^{-3}$  and there is an almost 1:1 correspondence between surface and subsurface values. For decoupled conditions, values in the subsurface can be twice those at the surface. An increase in penetration resistance variability coincided with the maturity of crops for cultivated fields. Aside from soil moisture, root growth may also have a significant impact on the temporal variability of soil's penetration resistance. The status of trafficability can be monitored through the high temporal resolution of Sentinel-1. However, aggregation to coarser resolutions may be necessary as its original 10 m resolution appears to be suboptimal, based on validation against in situ measurements. This information can aid farmers in the timing of tillage activities or for water managers in deciding to adjust water levels to meet agricultural demands.

## 5.1 Introduction

Modern agriculture relies on heavy mobile machinery to carry out farming operations such as tillage and harvesting. Mechanization increases productivity and enables farming activities to be carried out over larger areas. This is driven both by the high demand for food production as well as economic factors to make agriculture profitable. However, one of the negative impacts of using heavy machinery is soil compaction (Lal, 1991; Raghavan et al., 1990; Hamza and Anderson, 2005; FAO, 2015). As a form of land degradation, soil compaction leads to structural damages to soil (Baumgartl and Horn, 1991; Eckelmann et al., 2006). Destruction of soil structure also leads to poor infiltration (Van Dijck and Van Asch, 2002) which in turn leads to water logging, run-off and erosion (Ekwue and Harrilal, 2010). Compaction can have negative consequences for crop root growth and soil ecosystems.

The use of heavy machinery in agriculture is indispensable so a balance in the timing of farming operations and susceptibility of soils to compaction must be achieved in order to minimize the harmful and long-term effects. A way to minimize the exposure to the negative impacts of using heavy machinery is to determine trafficable conditions for the soil. Assessment of agricultural field trafficability as well as the operating machinery during periods suitable for traffic can help slow down the rate at which soils are being compacted so that the use of heavy machinery can be sustainable in the long term. Trafficability is defined by Campbell and O'Sullivan (1991) as the ability of soil to (1) provide adequate traction for vehicles and (2) withstand traffic without excess compaction or structural damage. This is an extension to earlier studies for military purposes which were only concerned with vehicle mobility (Knight and Freitag, 1961). More recent studies on field trafficability focus on soil-vehicle interaction to quantify the amount of soil compaction with different vehicle specifications (Keller and Lamandé, 2010; Nawaz et al., 2013; Rücknagel et al., 2015). Two sets of factors determine whether soils are able to support the weight of overlying machinery without increased risk for compaction. On the one hand are soil physical characteristics that dictate the mechanical strength of the soil. These include texture and bulk density of the soil (Müller et al., 2011). For soil texture, the strength of aggregated soils increases as clay content increases. Texture wise, soils with higher bulk density at the onset of field traffic can withstand higher pressures before undergoing deformation. Both soil texture and bulk density do not change significantly over short time scales (weeks or months) but they are important controls over the spatial patterns of trafficability. On the other hand are external factors that affect the grain-to-grain contact of individual soil particles. Perhaps the most important of these factors is soil moisture. For any given texture or bulk density, soil strength decreases towards wetter soil moisture conditions. Therefore soils become more prone to compaction with increasing soil moisture. The amount of overlying pressure that soils can accommodate decreases with increasing soil moisture since the grain-to-grain contact disappears as water fills up the pore space. Spatio-temporal soil moisture variability is influenced by atmospheric conditions (Seneviratne et al., 2010), soil properties (Rawls et al., 1991), and vegetation (Hupet and Vanclooster, 2002). Soil moisture exerts significant temporal control over soil strength as it varies greatly over short time scales because of changes in the prevailing atmospheric conditions. Other factors that control compaction are the set of vehicle specifications that determine the impacts of

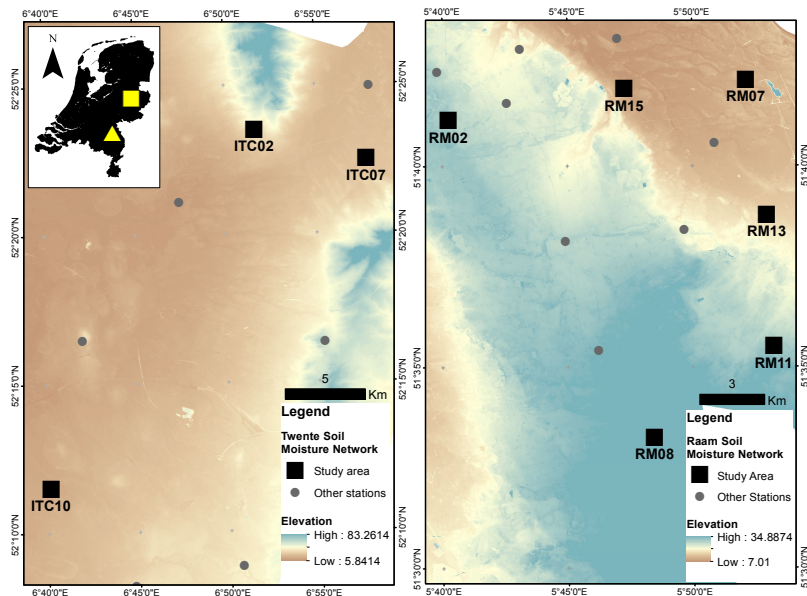
overlying machinery to the soil. Soil-vehicle interaction studies focus on the influence of tire inflation pressure, wheel and axle load on soil deformation (Müller et al., 2011).

The mechanical strength of the soil can be determined using cone penetrometers (Kuang et al., 2012; Upadhyaya, 2005). These are widely used instruments for determining soil's penetration resistance, which is equivalent to the force per unit base area required to push the cone penetrometer through a specified increment of soil depth (Bengough et al., 2000; ASAE EP542, 1999). They also have the advantage of providing relatively quick and easy measurements in the field. Cone penetrometer measurements are also referred to as cone index (CI). Existing field trafficability models have related soil moisture with CI using a decreasing exponential function (e.g. Henderson et al., 1988; Ayers and Perumpral, 1982; Sojka et al., 2001; Vaz et al., 2001), with an increasing CI trend towards drier soil moisture conditions. Several studies are geared towards monitoring and identifying conditions when the ground is less susceptible to compaction. For instance, Earl (1997) and Droogers et al. (1996) related trafficability and workability to soil hydraulic parameters to determine the number of workable days for a field. Other studies investigated the spatial variability of soil moisture and CI which can assist farmers in avoiding less trafficable areas within an agricultural field (Carrara et al., 2007; Ferrero et al., 2005). These studies have shown the impacts of soil moisture on field trafficability. Trafficability determined from soil moisture would be beneficial for agriculture, but this has been hampered in the past by the lack of continuous and available soil moisture data. Datasets from satellites are potential sources of regular and/or frequent soil moisture information that also cover a considerable spatial extent.

In the last few years, developments in mapping soil moisture using microwave remote sensing have been reported, with increasing spatial and temporal resolutions and accuracy (Kornelsen and Coulibaly, 2013; Vereecken et al., 2014). Techniques using microwave remote sensing are divided into active and passive methods. Passive microwave remote sensing measures the intensity of microwave emissions from the Earth's surface, expressed in terms of brightness temperatures. These measurements are performed with microwave radiometers. Active microwave remote sensing supply their own source of illumination. Active microwave sensors transmit signals towards a target and measures the portion scattered back. Synthetic aperture radar (SAR) is an active microwave sensing technique providing observations with a higher spatial resolution. SAR backscatter signals depend on the technical configuration of the sensor as well as the geometric and dielectric properties of objects on Earth. For soils, dielectric properties are highly influenced by its moisture content (Cihlar and Ulaby, 1974).

Sentinel-1 satellites, which carry a SAR instrument, are promising sources of soil moisture information that would be suitable for mapping and monitoring field trafficability at field scale. Its revisit time can be up to 2-4 days for certain areas in Europe (Torres et al., 2012), which makes it highly suitable for monitoring changes in soil moisture. The acquired images are also freely and operationally available. Sentinel-1 measurements are only sensitive to soil moisture in the upper surface layer ( $\sim 5$  cm). However, assessment of trafficability requires soil moisture values over the topsoil or critical layer (Droogers et al., 1996; Earl, 1997; Priddy and Willoughby, 2006; Reintam et al., 2016). Although values used in literature vary, this corresponds roughly to the upper 20-30 cm of the soil layer. This means that surface





**Figure 5.1:** Location of the study sites in Twente (yellow square) and the Raam (yellow triangle) soil moisture networks (inset). The stations utilized from Twente are plotted on the left while those from the Raam are plotted on the right.

soil moisture from Sentinel-1 needs to be translated into subsurface values before it can be used in certain applications. The vertical variability of soil moisture may lead to decoupling between surface and subsurface values (Capehart and Carlson, 1997; Carranza et al., 2018) wherein conditions in the former no longer represents those at the latter. This complicates the estimation of depth-average soil moisture needed for assessment of trafficability. Identifying coupled soil moisture conditions would be beneficial as it facilitates the use of satellite-derived surface soil moisture to assess field trafficability.

In this study, our main objective is to assess whether Sentinel-1-derived surface soil moisture can be used to monitor field trafficability. For Sentinel-1, several aggregation schemes were tested to determine the optimal pixel size. The accuracy of these aggregation schemes were validated against reference in situ measurements from two monitoring networks (Raam and Twente networks). Given the measurement depth limitations of Sentinel-1, we first aim to identify conditions when surface soil moisture values are good indicators of those at subsurface. In addition, the variability encountered from extensive penetration resistance measurements was analyzed in order to gain further insights on its temporal controls. A probabilistic approach is applied to express trafficability based on surface soil moisture and to incorporate variabilities encountered from in situ measurements. We demonstrate the results over a small site within the Raam catchment as an example to show the potential of Sentinel-1 for monitoring agricultural field trafficability.

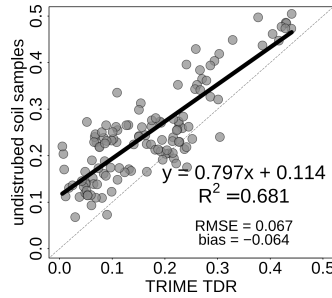
**Table 5.1:** Characteristics of the study sites. The description of soil types are taken from BODem-Fysische EenhedenKaart (BOFEK 2012, Wosten et al., 2013). For the fields within Raam, the percentage of mineral components and organic matter content are taken from Benninga et al. (2018) while those from Twente are also from BOFEK 2012

| Station | Soil Description   | %Silt | %Clay | %OM | Crop 2016    | Crop 2017    | Size (ha) |
|---------|--|-------|-------|-----|--------------|--------------|-----------|
| RM02    | Weakly loamy sandy soil on sub-soil of coarse sand (305)     | 3.7   | 2.1   | 3.8 | Sugar beets  | Winter Wheat | 2.96      |
| RM07    | Loamy sandy soil with thick man-made earth soil (317)        | 10.5  | 5.2   | 2.2 | Corn/Cichory | Corn/Potato  | 4.79      |
| RM08    | Weakly loamy podzol soil (304)                               | 1.6   | 1.4   | 4.1 | Sugar beets  | Winter Wheat | 2.51      |
| RM11    | Weakly loamy podzol soil (304)                               | 1.7   | 1.6   | 1.9 | Corn         | Corn         | 3.49      |
| RM13    | Weakly loamy soil partly on sub-soil of coarse sand (309)    | 1.1   | 0.8   | 1.4 | Corn         | Grass        | 8.48      |
| RM15    | Weakly loamy sandy soil with thick man-made earth soil (311) | 5.5   | 2.8   | 3.1 | Grass        | Grass        | 2.05      |
| TW02    | Loamy sandy soil with thick man-made earth soil (317)        | 21    | 4     | 5.2 | Grass        | Grass        | 4.76      |
| TW07    | loamy sandy soil with a clay deck (316)                      | 35    | 13    | 2.4 | Winter Wheat | Corn         | 4.55      |
| TW10    | Weakly loamy podzol soil (304)                               | 13    | 3     | 4.1 | Corn/Potato  | Corn         | 8.88      |

## 5.2 Material and methods

### 5.2.1 Raam and Twente soil moisture networks

We utilized several locations within the existing soil moisture monitoring network in Twente (Dente et al., 2011) and Raam catchment (Benninga et al., 2018) as our study sites. These two networks contain stations covering the eastern (Twente network) and southeastern (Raam network) parts of the Netherlands (fig.5.1). They were installed previously to serve as validation sites for satellite-derived data products. Each station contains sensors that continuously monitor soil moisture over the soil profile. Soil moisture and temperature sensors (ECH2O EC-TM or 5TM) were installed at discrete depths below the surface (5, 10, 20, 40, and 80 cm). Measurement loggers store values every 15 minutes. For this study, we utilized year-long measurements from 2016 - 2017. In both monitoring networks, stations were installed in agricultural areas, at the edge or corner of a field to allow continuous measurements. Installation in the middle of the field was not permitted because of tillage, harvesting or grazing of animals. From the Twente network, three out of the 20 stations were utilized; from the Raam network, six out of the 15 stations were utilized as study sites. These were chosen to capture the variability in the crops encountered at the study sites during the two-year study period. The nine agricultural fields for this study were either grass or cultivated fields. The most common cultivated crop encountered was corn, followed by potato, sugar beet, winter



**Figure 5.2:** Volumetric soil moisture measured using TRIME TDR vs. undisturbed soil samples (in  $\text{cm}^3 \text{cm}^{-3}$ ). Calibration of TRIME TDR measurements was performed by fitting a linear function of over the points. The fit is deemed reasonable based on an  $R^2$  value of  $\sim 0.7$ . Values for bias and RMSE are in  $\text{cm}^3 \text{cm}^{-3}$ .

wheat, and chicory (Table 5.1). The terrain in these fields was generally flat to gently sloping. Sandy soils were encountered in most of the fields, except for TW07 where the soil holds slightly higher loam content.

### 5.2.2 In situ surface soil moisture and cone index measurements

Aside from the soil moisture networks, we also collected in situ surface soil moisture and cone index measurements from the nine agricultural fields (fig.5.1). These were taken from two growing seasons, specifically from May 26, 2016 to October 09, 2017. For each growing season, a field was visited at least twice to collect measurements during different soil moisture conditions. The number of measurements ranged from 10 to 30 points per field per measurement day, depending on the size of the field. The points are 15 - 20 m apart and forms somewhat of a grid for each field. In addition, measurements were always taken at the same location for both surface soil moisture and penetration resistance. The measurement locations were also kept the same during succeeding measurement days for a growing season. However, this was not feasible for both years because changes in the crops planted made it difficult to keep the locations of existing points. In total, 840 actual measurement points were collected over all study sites.

Surface soil moisture was measured using a hand held time domain reflectometry (TDR) device (TRIME-IMKO) with 5 cm pins. The calibration of TDR for sandy soils at the study sites was performed against volumetric soil moisture values from undisturbed samples collected during several occasions within the whole field campaign (fig.5.2). A total of 127 sample points were used for calibration using linear regression. The calibrated volumetric soil moisture (VWC) is given by the function:

$$VWC = 0.797x + 0.114 \quad (5.1)$$

where  $x$  is the soil moisture measured using the TDR device. The linear function used adequately fitted the points based on  $R^2 = 0.681$ . In addition,  $RMSE = 0.067 \text{ cm}^3 \text{cm}^{-3}$  indicated

small differences between these two measurements and  $bias = -0.064 \text{ cm}^3 \text{ cm}^{-3}$  indicated that measurements using the TDR device gave slightly lower readings compared to those from undisturbed samples.

Penetration resistance measurements were obtained using a hand held penetrometer (Eijkelkamp Penetrologger) with a 1 cm cone diameter and a 60 degree angle. We referred to the penetration resistance measurements also as the cone index (CI). The measurements were taken from the soil surface until 20 cm for every 1 cm-depth interval.

### 5.2.3 Sentinel-1 imagery

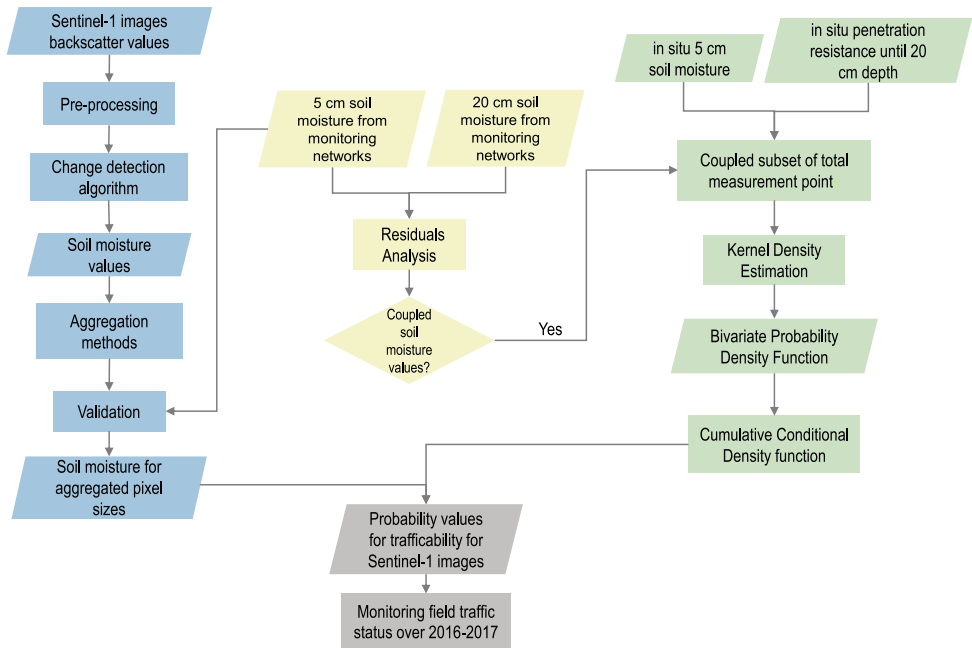
The SAR instrument on board the Sentinel-1 satellites operates in C-band (5.405 GHz), which in Interferometric Wide Swath (IW) mode provides over land images at VV and VH polarization, with pixel spacing of 10 m  $\times$  10 m and a reported radiometric accuracy of 1 dB ( $3\sigma$ ) (Torres et al., 2012). Sentinel-1A and Sentinel-1B provide images since 3 October 2014 and 26 September 2016, respectively. The combination of Sentinel-1A and Sentinel-1B results in a revisit time of 3 days over the Raam study area and a revisit time of 1.5 days over the Twente study area. Given the higher sensitivity to soil moisture of backscatter observations acquired in VV polarization than in VH polarization (e.g. Baghdadi et al., 2017; Bousbih et al., 2017; Hajj et al., 2017), we used the observations in VV polarization to retrieve soil moisture. The Sentinel-1 images are freely available via the Copernicus Open Access Hub (ESA, 2019).

We applied the following operations to convert raw pixel values into radar backscatter ( $\sigma$ ): (1) Range Doppler Terrain Correction (RDTC) using the tool in the Sentinel Application Platform (SNAP) software, which includes (a) radiometric calibration, (b) reprojection to correct for distortions due to topographical variations and tilt of the satellite sensor, and (c) radiometric normalization with projected local incidence angles, and (2) a 5  $\times$  5 median speckle filter to suppress speckle noise.

Projected local incidence angles of the Sentinel-1 observations vary between 36.8° and 40.7° for the Raam study area and 31.5° and 46.7° for the Twente study area. We normalized the backscatter observations to a reference angle of 40° using a cosine correction (Ulaby et al., 1986) :

$$\sigma_{ref}^o = \sigma^o \frac{\cos^n(\varphi_{ref})}{\cos^n(\varphi_{inc})} \quad (5.2)$$

where  $\sigma_{ref}^o$  (in  $\text{m}^2 \text{ m}^{-2}$ ) is the backscatter observation normalized to a reference angle  $\varphi_{ref}$  of 40°,  $\sigma^o$  is the backscatter observation (in  $\text{m}^2 \text{ m}^{-2}$ ) and  $\varphi_{inc}$  is the local projected incidence angle (in degrees). We assumed a value of  $n = 2$ , based on C-band SAR observations in previous studies (Lievens et al., 2011; Van der Velde and Su, 2009; Mladenova et al., 2013; van der Velde et al., 2015), which corresponds to the assumption that re-radiation from the soil surface follows Lambert's cosine law (Ulaby et al., 1986). Sentinel-1 observations that exceed the upper limit of -2 dB or the lower limit of -22 dB (the maximum noise equivalent to sigma zero) were taken out.



**Figure 5.3:** Flowchart summarizing the methods applied in this study. The analysis was carried out in three parts. One section focuses on translating Sentinel-1 backscatter to soil moisture values and validating the results with in situ time series measurements. Another section focuses on determining coupled soil moisture conditions. The third section describes the probabilistic framework applied to express trafficability using surface soil moisture and penetration resistance values.

### 5.3 Methods

The analysis carried out for monitoring field trafficability using Sentinel-1 was divided into three parts (fig.5.3). The first section describes the soil moisture retrieval from Sentinel-1 backscatter values. This section also includes the validation of soil moisture retrievals using time series field measurements. The second section describes the coupling of surface and topsoil soil moisture values by looking at the variability between the two. This was needed in order to identify which surface soil moisture values (e.g. from Sentinel-1) are representative for the total soil moisture content within the topsoil. Using these results, the third section describes the probabilistic framework applied to express field trafficability using in situ measurements of surface soil moisture and cone index.

### 5.3.1 Soil moisture retrieval - Change Detection Algorithm

The change detection algorithm, developed by Wagner et al. (1999), linearly relates backscatter (in dB) to a relative soil moisture index  $S_w$  of the surface layer over time:

$$S_w(t) = \frac{\sigma_{40^\circ}^\circ(t) - \sigma_{40^\circ,Min}^\circ}{\sigma_{40^\circ,Max}^\circ - \sigma_{40^\circ,Min}^\circ} \quad (5.3)$$

where  $\sigma_{40^\circ,Min}^\circ$  and  $\sigma_{40^\circ,Max}^\circ$  are the backscatter observations under dry and wet soil conditions (in  $\text{cm}^3 \text{cm}^{-3}$ ), respectively. The absolute minimum and maximum values of  $\sigma_{40^\circ}^\circ$  are probably outliers due to radiometric noise, speckle or exceptional surface conditions, such as wet snow or surface inundation (Pathe et al., 2009; Bauer-Marschallinger et al., 2018). To exclude these outliers, we estimated the  $\sigma_{40^\circ,Min}^\circ$  and  $\sigma_{40^\circ,Max}^\circ$  by the 2.5% and 97.5% percentile of the time series. For the calculation of  $\sigma_{40^\circ,Min}^\circ$  and  $\sigma_{40^\circ,Max}^\circ$  we used time series of two complete hydrological years, between March 1, 2016 and March 1, 2018, resulting in 395 images over the Twente region and 198 over the Raam catchment. For backscatter observations that exceed  $\sigma_{40^\circ,Min}^\circ$  or  $\sigma_{40^\circ,Max}^\circ$ ,  $S_w(t)$  is set equal to 0 or 1, respectively.

The main assumption of the change detection algorithm is that the parameters, other than soil moisture, are considered time-invariant, such as surface roughness and vegetation. The change detection algorithm is promising for operational applications of soil moisture retrievals from satellite observations, because the model is build only on a statistical analysis of backscatter time series and no detailed ground parameters are required (Wagner et al., 1999; Pathe et al., 2009; Hornacek et al., 2012; Bauer-Marschallinger et al., 2018). Although Wagner et al. (1999) developed the change detection algorithm for coarse satellite observations (European Remote Sensing scatterometer satellites, spatial resolution 50 km), several studies obtained acceptable correlations with in situ soil moisture measurements using SAR observations at C-band (Pathe et al., 2009; Hornacek et al., 2012; Bauer-Marschallinger et al., 2018).

The change detection algorithm results in a relative soil moisture estimate whereas absolute soil moisture, corresponding to the volumetric water content, is needed for mapping field trafficability. This problem is circumvented here by assuming that soil moisture varies between the wilting point ( $\theta_{wp}$ ) and saturated soil moisture content ( $\theta_{sat}$ ), which allowed us to linearly scale  $S_w$  to volumetric soil moisture  $\theta$  (in  $\text{cm}^3 \text{cm}^{-3}$ ), as follows:

$$\theta(t) = (\theta_{sat} - \theta_{wp}) * S_w(t) + \theta_{wp} \quad (5.4)$$

The BOdemFysische EenhedenKaart (BOFEK2012) provides the soil type classes and associated soil physical characteristics (including van Genuchten parameters) for the soil units in the Netherlands (Wösten et al., 2001; Wosten et al., 2013). Then, the equation proposed by van Genuchten (1980) gives the soil moisture content at wilting point and saturated conditions (table 5.2). Benninga et al. (2018) showed that the wilting point and saturated soil moisture content calculated from BOFEK2012 align with the minimum and maximum soil moisture measurements at 5 cm depth of the individual stations of the Raam network.

**Table 5.2:** Values of  $\theta_{sat}$  and  $\theta_{wp}$  derived from BOFEK 2012 (Wosten et al., 2013). These parameters were used to linearly scale relative  $S_W$  values and derive absolute soil moisture values (see Eq.5.4).

| Station | BOFEK code | $\theta_{sat}$ | $\theta_{wp}$ |
|---------|------------|----------------|---------------|
| RM02    | 305        | 0.43           | 0.03          |
| RM07    | 317        | 0.45           | 0.05          |
| RM08    | 304        | 0.43           | 0.03          |
| RM11    | 304        | 0.43           | 0.03          |
| RM13    | 309        | 0.43           | 0.03          |
| RM15    | 311        | 0.43           | 0.03          |
| TW02    | 317        | 0.45           | 0.05          |
| TW07    | 316        | 0.40           | 0.12          |
| TW10    | 313        | 0.45           | 0.05          |

#### *Validation of Sentinel-1 estimates using in situ measurements*

The accuracy of soil moisture retrievals were evaluated against reference soil moisture values at the study sites. We utilized the 5 cm soil moisture measurements from the nine stations as reference values to validate the soil moisture retrievals from Sentinel-1. A subset of the total Sentinel-1 dataset was obtained based on the dates in common with in situ measurements. Aside from the original 10 m resolution, we tested four aggregation methods in order to determine the optimal pixel resolution for Sentinel-1. The first method employed aggregation to coarser resolutions of all the pixels in the Sentinel-1 image after masking out those which were not agricultural lands or low vegetation nature areas. For the second method, we aggregated only the pixels within each field boundary after masking out pixels outside each field. For the third method, we aggregated only the surrounding pixels, with the station located at the center. A circular buffer around each station was created to mask out the pixels outside the buffer. The fourth method involved calculation of the field mean. We aggregated the pixels to 50 m, 100 m, and 150 m, for each of the first three aggregation methods. The choice of upper limit (150 m) for aggregation was based on the size of the smallest field encountered at the study sites (table 5.1) so that it is comparable to area of the largest pixel size used. The smallest field is at RM15 with an area of 2.05 ha while a  $150 \times 150$  m pixel size has an area of 2.25 ha.

To select the optimal resolution and aggregation method for Sentinel-1, we calculated several performance metrics to compare the results. We computed RMSE, bias, unbiased RMSE, and Spearman's rank correlation coefficient for the results of aggregation methods used. The optimal resolution and aggregation method was further utilized to derive the trafficability status at selected fields over 2016 to 2017.

### **5.3.2 Soil moisture and Cone index variability**

#### *Vertical soil moisture variability*

As the first step to relate satellite-derived surface soil moisture to field trafficability, we investigated when surface soil moisture is a good indicator of subsurface soil moisture conditions.

We utilized the time series measurements at 5 cm and 20 cm from the selected study sites within Twente and Raam monitoring networks. The 5 cm values represent surface soil moisture measurements as they approximate the depths at which most satellites are able to extract soil moisture information. The 20 cm values correspond to measurements at the topsoil depth which carries the weight of overlying machinery. Several studies have also referred to such depths as the critical layer (Knight and Freitag, 1961; Paul and de Vries, 1979; Priddy and Willoughby, 2006). We referred to vertical variability as the irregularity between soil moisture contents between 5 cm and 20 cm depths.

Using the time series soil moisture datasets from the Raam and Twente monitoring networks, we applied residuals analysis to look at (de)-coupling between surface and topsoil moisture conditions. Carranza et al. (2018) inferred that there is lower vertical variability during coupled soil moisture conditions. The analysis involved fitting a non-parametric loess function to relate surface and topsoil values. After which, residuals from the fitted loess function were used to calculate the residuals variance. This showed the vertical variability throughout the whole soil moisture range encountered at the study sites. To determine whether coupling or decoupling is present given any soil moisture value, the cumulative residuals variance line was plotted. This allowed us to observe changes in the variance of residuals as they were reflected as changes in the slope of the cumulative variance line. The soil moisture range with flatter slopes indicated lower variability or coupled range, and vice versa.

To get an impression of the similarity in soil moisture at 5 cm and 20 cm during coupled and decoupled conditions, the ratio between two was computed as a simple quantitative measure of correspondence. A Bayesian approach, developed and explained in detail by Kruschke (2013), was applied to determine the probabilistic mean of these ratios. The mean of 5 - 20 cm ratio is the parameter of interest that will be estimated. Briefly, the method involved Monte Carlo Markov Chain (MCMC) to generate a large representative sample (100,000 samples) to approximate the posterior distribution of the parameter of interest. From the posterior distribution, the mean as well as the high density interval (HDI), where 95% of the estimates lie was calculated. MCMC sampling was performed separately for the coupled and decoupled range.

#### *Cone index seasonal variability*

We also investigated if there is a trend in CI values over a growing season using the median and interquartile range (IQR) of CI over the topsoil (upper 20 cm values). This was performed for four fields where the most frequent measurements were made. These fields are RM07, RM08, TW02, and TW10. TW02 is a grass field while the other three are cultivated fields (see table 5.1). We accounted for differences in soil moisture by correcting CI values to a single soil moisture value based on the method of Busscher et al. (1997). Visual assessment of the median and IQR for different dates was performed to compare CI values and to identify if there are trends over the growing seasons of 2016 and 2017.



### 5.3.3 Probabilistic framework to model field trafficability

A probabilistic approach was applied in identifying trafficable conditions, This allowed for the variabilities encountered in the in situ measurements to be incorporated in the analysis. First, we determined the joint empirical probability density function (pdf) for cone index and surface soil moisture using kernel density estimation (KDE). The pdf's were then utilized to calculate the conditional probabilities to express field trafficability.

#### *Estimating joint probability density function (pdf) for CI and surface soil moisture*

We selected a subset of the in situ measurement points using the coupled range identified in section 5.3.2. At this range, surface soil moisture can be directly related to CI as it is a good indicator of values for the whole topsoil. In situ measurement points that belong to the decoupled range were discarded.

Initial visual inspection of the measurement values revealed that those collected from grass fields tend to have higher cone index values than those in cultivated fields. The probabilistic method in section 5.3.2 was applied and extended to test whether these measurements indeed form two separate groups, and merited separate analyses. This method is similar to a t-test with the main goal of determining whether the CI values are distinct from each group. The posterior distributions of grass and cultivated field were estimated using MCMC, also using 100,000 representative samples. To test the similarity between these two groups, the difference of means between every combination of representative values was obtained. Two groups were deemed similar if the difference between the means between their posterior distribution were close to or equal to zero.

Kernel Density Estimation (KDE) was applied afterwards to determine the empirical bivariate distribution of surface soil moisture and cone index. This is a non-parametric method to estimate the underlying probability density function (pdf) of a random variable (Parzen, 1962; Rosenblatt, 1956) using kernels. It is a suitable method for datasets with complicated distribution since no assumption is made on the shape of the underlying density function. Density estimation is carried out by centering a kernel at the location of each data point. The overall density estimate is obtained by summing all the densities estimated at each point. Points that plot closer to each other will have more kernels centered nearby which yields higher density estimates.

In the bivariate case, the data points were represented by two vectors  $x_1 = [x_{11}, x_{12}, x_{13}, \dots, x_{1n}]$  and  $x_2 = [x_{21}, x_{22}, x_{23}, \dots, x_{2n}]$  where  $x_i = (x_{1i}, x_{2i})$  was a sample from a bivariate distribution  $f$ . These two vectors represent the datasets collected for surface soil moisture and CI. The bivariate kernel density estimator  $f$  was given by:

$$\hat{f}(x, \mathbf{H}) = \frac{1}{n} \sum_{i=1}^n K_{\mathbf{H}}(x - x_i) \quad (5.5)$$

where  $K_H$  is a non-negative kernel function and  $H$  is the kernel bandwidth that controls the amount of smoothing.

To calculate the joint pdf of soil moisture and CI using KDE, we utilized a Gaussian kernel with bandwidths (represented by  $h$  for a univariate case) obtained using Scott's rule (Scott, 1992). For cultivated fields, we obtained a value of  $h_{SM} = 0.022 \text{ cm}^3 \text{ cm}^{-3}$  and  $h_{CI} = 0.138 \text{ MPa}$ ; and for grass fields, we obtained  $h_{SM} = 0.001 \text{ cm}^3 \text{ cm}^{-3}$  and  $h_{CI} = 0.184 \text{ MPa}$ . From the joint pdf of CI and soil moisture, we generated 500 random samples from a multivariate normal distribution using the bandwidth matrix derived (Gentle, 2009). The random samples were drawn with replacement along the length of  $x$  (total number of paired data points), and then adding random noise or perturbations to the sampled values using bandwidth matrix  $H$ . These random samples were used to calculate the conditional probabilities.

#### *Determining conditional probabilities for CI to express trafficability*

For a given soil moisture, we calculated the cumulative conditional probability that the corresponding CI will take a value less than or equal to a known threshold. Higher probability values indicated poor trafficability. We used a value from Droogers et al. (1996) who applied 0.70 MPa as the threshold for conditions suitable for agricultural traffic. Using the randomly generated points, trafficability is evaluated using the conditional cumulative distribution function (ccdf) given by  $F_{X|A}$ . This expressed the cumulative probability of a random variable  $X$  conditioned on the occurrence of an event  $A$ :

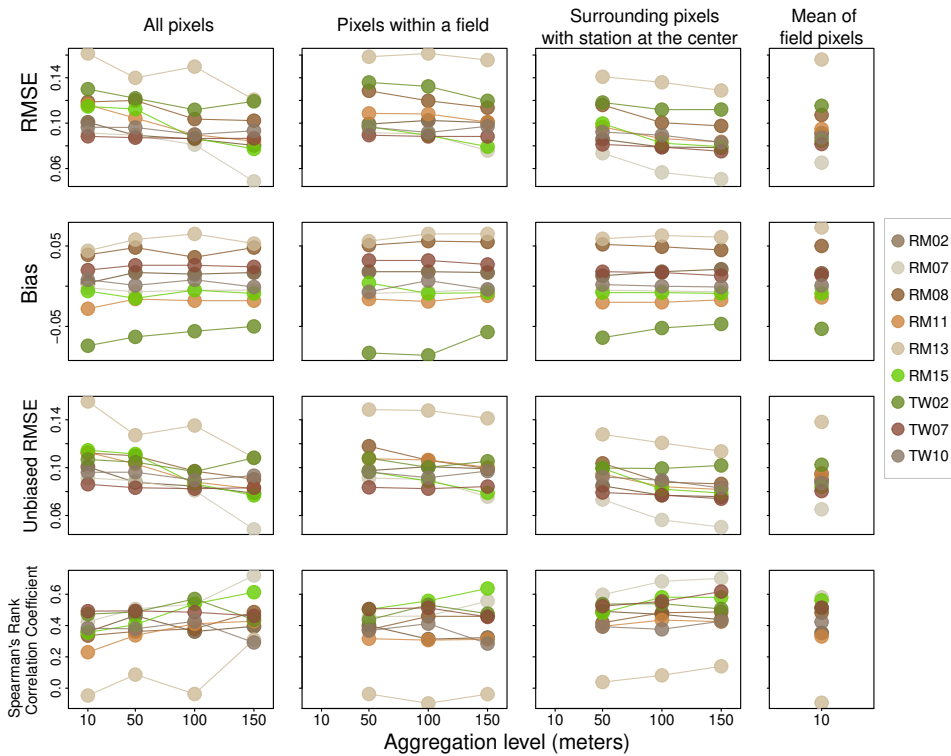
$$F_{X|A}(x) = P(X \leq x|A) \quad (5.6)$$

In this case, we determined the cumulative probability that a cone index value  $X$  will less than or equal to  $x = 0.7 \text{ MPa}$  given a soil moisture value  $A$ . Higher probability values indicate conditions not favorable for traffic, vice versa. Based on the results obtained in Section 5.4.1, soil moisture from Sentinel-1 images were transformed into its corresponding probability values for trafficability.

## 5.4 Results and Discussion

### 5.4.1 Surface soil moisture from Sentinel-1

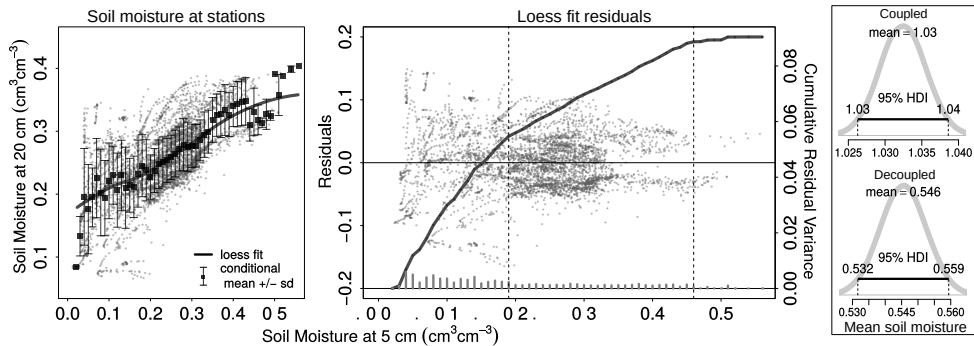
Figure 5.4 shows the calculated accuracy metrics using the four aggregation methods applied. The plot shows improvement in accuracy in all four metrics when larger pixel sizes are used compared to the original 10 m pixel size of Sentinel-1. Values for RMSE, bias, and unbiased RMSE are lower while Spearman's correlation is higher for larger pixel sizes compared to the original 10 m Sentinel-1 pixel size. From the four resolutions tested, using 150 m pixel resolution provides the best results. Furthermore, aggregation of all the pixels in a scene results in slightly higher accuracy values for all the study sites. We observed smaller variability for each metric calculated when aggregating all the pixels in a scene compared to the other methods. Radiometric uncertainties at the field scale, which affect the soil moisture retrieval accuracy, may be influenced by speckle effects (Ulaby et al., 1986). Aggregation of the original Sentinel-1 observations to larger pixels sizes implies a larger number of independent samples



**Figure 5.4:** Plots of four accuracy metrics used to compare the aggregation schemes and resolutions tested. RMSE, bias, unbiased RMSE, and Spearman's rank correlation coefficient were computed to compare the results. All metrics are in  $\text{cm}^3 \text{cm}^{-3}$ , except for Spearman's rank correlation coefficient which is unitless

that leads to the suppression of speckle effects; this is critical for improving  $\sigma^\circ$  uncertainties. Our results are in line with other studies who showed that aggregation of SAR pixels produced better results. For instance, Pierdicca et al. (2013) obtained higher soil moisture retrieval accuracy when aggregating pixels to field scale using synthetically generated  $\sigma^\circ$  representing Sentinel-1 observations of bare soil. Pathe et al. (2009) found that retrieval errors may be dominated by noise in SAR measurements, even when assuming high model parameter errors to account for the neglect of vegetation effects. They further suggested that several pixels should be averaged to decrease the noise level, even at the expense of the spatial resolution of the soil moisture maps.

For the different land cover types, we see no apparent trend in the results. For instance, points in fig.5.4 for grass (green points) and cultivated fields (brown points) do not reveal any clustering from the plots; which implies that the accuracy of soil moisture retrievals from Sentinel-1 is comparable for grass and cultivated fields.



**Figure 5.5:** *Left:* Scatterplot of 5 cm and 20 cm soil moisture values. 5 cm values represent surface soil moisture while 20 cm values represent the topsoil. A loess function is fitted for the overall trend while standard deviation bars indicate the variability across the whole soil moisture range. *Center:* Plot of the residuals from the fitted loess function over the surface soil moisture range. The variance from the residuals for every  $0.01 \text{ cm}^3 \text{ cm}^{-3}$  interval is given by the vertical bars at the bottom of the plot. The cumulative variance of residuals is plotted as a black line. A change in the slope of the line at  $0.19 \text{ cm}^3 \text{ cm}^{-3}$  separates the decoupled range ( $< 0.19 \text{ cm}^3 \text{ cm}^{-3}$ ) from the coupled range ( $\geq 0.19 \text{ cm}^3 \text{ cm}^{-3}$ ). *Right:* Posterior distributions of 5 to 20 cm soil moisture ratio has a mean of  $1.03 \text{ cm}^3 \text{ cm}^{-3}$  for coupled range compared to  $0.55 \text{ cm}^3 \text{ cm}^{-3}$  for decoupled range. The distributions also show the range corresponding to the 95% high density interval of the ratios calculated.

Aside from soil moisture, a component of the total radar backscatter is due to the influence of surface roughness and vegetation. In cultivated fields, tillage activities at the beginning of a growing season effect changes to surface roughness. Based on field measurements, Callens et al. (2006) showed that after the onset of the first rainfall events, surface roughness remains fairly stable. A sensitivity analysis performed by Joseph et al. (2010) on soil moisture retrieval across the corn growth cycle also came to a similar conclusion. In grass fields and meadows, surface roughness may be considered invariable over longer periods of time; an assumption frequently adopted in soil moisture retrieval methods (e.g. Álvarez-Mozos et al. (2006); van der Velde et al. (2012)). The magnitude of vegetation effects on the total backscatter depends on sensor specifications (e.g. frequency, incidence angle, polarization) and vegetation type. Effects of grass on  $\sigma^\circ$  is generally expected to be weak because the dimensions of scatterers (i.e. leaves and stems) are small in comparison to the SAR wavelength (e.g. Van der Velde and Su, 2009). However, the effects may be significant in cultivated fields. For example, Joseph et al. (2010) quantified the effect of corn across its growth cycle on C- and L-band backscattering at incidence angles of 15, 35 and 55°. They showed that throughout the corn growth cycle, both an attenuated soil return and vegetation scattering can dominate the measured  $\sigma^\circ$ . However, the measured  $\sigma^\circ$  still displayed some sensitivity to soil moisture even at peak biomass. Despite the known impacts of surface roughness and vegetation, previous studies have shown that change detection method can reasonably track changes in soil moisture using multi-temporal satellite imagery (e.g. Moran et al., 2000; Baghdadi et al., 2007; Pathe et al., 2009) since it

tries to overcome the difficulties in applying bare soil and vegetation backscatter models by only interpreting changes in backscatter rather than absolute backscatter levels.

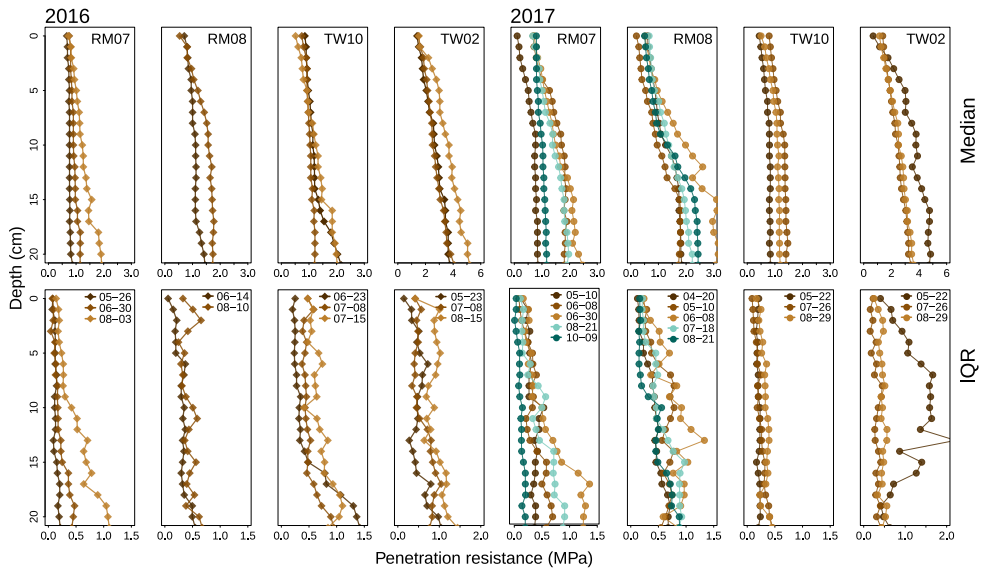
For the performance of linear transformation applied (see Eq.5.4), the accuracy might vary depending on the actual soil moisture distribution. At the study sites, the range of soil moisture encountered span from dry to wet conditions, which are close to the  $\theta_{wp}$  and  $\theta_{sat}$  parameter values applied. However, in regions where soil moisture distribution is skewed (e.g. very dry or very wet climates), the linear transformation method might not be the optimal choice if the range of soil moisture values encountered is limited, and  $\sigma_{Min}$  and  $\sigma_{Max}$  are not measured under conditions of  $\theta_{wp}$  and  $\theta_{sat}$ . In such cases, another transformation may be more appropriate. For instance, cumulative distribution function (cdf) matching (Reichle and Koster, 2004) can be applied as this method allows to incorporate the soil moisture distribution in scaling  $S_w$  values.

#### 5.4.2 Soil moisture and Cone index variability

The scatterplot in fig.5.5 (left) shows a general linear trend across the soil moisture range. The variability across the soil moisture range is not constant based on the length of the standard deviation bars. However, it is not very easy to visualize trends in variability from the scatterplot alone. Further examination of the variance of residuals from the fitted loess function (fig.5.5, center) reveals that for dry conditions, variance is consistently higher compared to those during wet conditions. A change from higher to lower variance occurs at  $0.19 \text{ cm}^3 \text{ cm}^{-3}$  based on the change in slope in the cumulative variance line. From this, we determine that values from  $< 0.19 \text{ cm}^3 \text{ cm}^{-3}$  corresponds to a decoupled soil moisture range while those  $\geq 0.19 \text{ cm}^3 \text{ cm}^{-3}$  corresponds to the coupled range (Carranza et al., 2018).

For assessment of trafficability, wetter soil moisture conditions are of primary interest since soils are generally more susceptible to compaction. During wet conditions, we found coupling between surface and topsoil soil moisture values, therefore the former is a good representative of the total soil moisture within the latter. This is further confirmed by the probabilistic mean calculated. A value equal to  $1.036 \text{ cm}^3 \text{ cm}^{-3}$  for the 5-20 cm ratio (fig.5.5, right) indicates an almost 1:1 correspondence between the two. This implies that surface soil moisture, for instance those derived from Sentinel-1, can already be used as a proxy for topsoil values without the need for a separate analysis to convert surface values to topsoil soil moisture content. For the decoupled range, the probabilistic mean is  $0.547 \text{ cm}^3 \text{ cm}^{-3}$ , which means that subsurface values can be twice as much as the surface.

CI plots over the topsoil for four fields show the variability over a growing season fig.5.6. CI values in cultivated fields increase in median and IQR towards mid-growing season and then decreases at the end or after harvest. This shows that CI becomes higher and more variable mid-growing season when the crops are mature. However, this trend was not observed in the grass field (TW02). Aside from soil moisture, we hypothesize the root growth affects the temporal variability of CI for cultivated fields. For the grass field, the lack of trend observed is still inconclusive since we only had one site. Several studies have already looked at importance of roots in increasing soil strength, but has been investigated mainly in relation to erosion

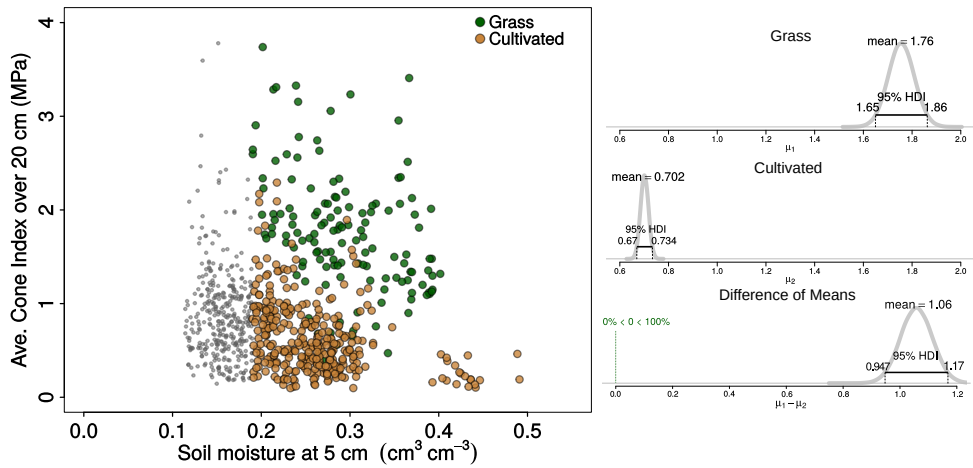


**Figure 5.6:** Profile CI values at four selected fields where most measurements were collected. The top panel shows the median CI per depth per measurement date (mm-dd) while the bottom panel shows the interquartile range (IQR). Dates when the measurements are collected are indicated in the bottom panel. For 2016 and 2017, a slight increase in both median and IQR of CI is observed up to mid-growing season (brown colors). These values decrease towards the end or after harvesting (blue colors). This pattern is observed for the three sites which are cultivated fields, with exception of TW02 which is a grass field.

and hazards (De Baets et al., 2008; Fan and Su, 2008; Mickovski et al., 2009). Studies investigating the impact of root growth on soil hydraulic properties and field trafficability is still rare in literature (e.g. Wieder and Shoop, 2017).

### 5.4.3 Probabilistic modeling of field trafficability

The calibrated values plotted in fig.5.7 (left) show that grass fields have higher CI values compared to those in cultivated fields. These two groups also show different probabilistic mean CI, which warrants separate analysis for trafficability (fig.5.7, right). Grass fields have a probabilistic mean CI of 1.74 MPa while cultivated fields have 0.709 MPa. The posterior distribution of the mean CI for cultivated and grass fields do not overlap and differ by at least 1 MPa. This difference can be attributed to the type of land management in these fields. For cultivated fields, yearly tillage loosens the soil so that CI values become lower at the beginning of the growing season. For grass fields, regular mowing and trampling of animals results in higher CI. Fields which remain as grass fields become more compacted over time as they are not being tilled. Faunal activities (e.g. from burrowing animals) may alleviate soil compaction to a certain degree but based on our measurements, these are not enough for grass fields to have similar CI values to cultivated fields.



**Figure 5.7:** *Left:* Scatterplot of in situ measurements of surface soil moisture and penetration resistance. Colored points are within the coupled soil moisture range while grey points are in the decoupled range. *Right:* Distribution of the probabilistic mean CI for grass and cultivated fields. The difference in mean CI for these two groups is given in the bottom panel. The probability that the difference between these two sets of mean CI is zero is also equal to zero.

Figure 5.8 shows the bivariate pdf's obtained and random samples generated for cultivated and grass fields. Aside from having generally higher values, grass fields also have higher variability as seen from the spread of points in the scatterplot (fig.5.7) and bivariate pdf's (fig.5.8). The variability observed in the pdf's implies that CI values does not depend on soil moisture alone. Aside from soil moisture, bulk density also controls CI values but was implicitly incorporated when we separated the analysis for cultivated and grass fields. The differences in CI for cultivated and grass fields already indicate their differences in bulk densities.

The distribution of random samples closely approximate the original measurement data distribution for CI but not for soil moisture (fig.5.8, bottom panel). However, the soil moisture values from the random samples generated still approximate the realistic range and distribution of soil moisture. This was confirmed by also plotting the 5 cm time series measurements from both Raam and Twente networks (plot not shown). In addition, our measurements might also be incomplete since we were not able to encounter all possible soil moisture values. From fig.5.7, we lack measurements beyond  $0.4 \text{ cm}^3 \text{ cm}^{-3}$  for grass and between  $\sim 0.35 - 0.4 \text{ cm}^3 \text{ cm}^{-3}$  for cultivated fields. Nevertheless, we were still able to generate random samples from the underlying distribution of soil moisture with the bandwidths used (see section 5.3.3)

The pdf's generated over the study two-year period only reflect average/normal weather conditions in the Netherlands, which is characterized by having moderate rainfall over the whole year. Similar to our measurements (fig.5.8), intermediate soil moisture values ( $20\text{-}30 \text{ cm}^3 \text{ cm}^{-3}$ ) are the most frequent while very dry and very wet conditions occur less frequently. Therefore, the pdf's obtained are more suitable for predicting trafficability during average weather years.

Since the shape of the pdf used is not universal, the results of the analysis may be suboptimal for periods when or areas where very wet or very dry weather conditions are prevalent. Modeling trafficability for such conditions can be done if the appropriate bivariate pdf's can be selected.

Trends in soil strength over the whole soil moisture range are consistent with previous studies (e.g. Henderson et al., 1988; Vaz et al., 2011) which demonstrate increasing CI with lower soil moisture. Similar findings found in this study despite using only surface soil moisture further strengthens the potential for using satellite-derived surface soil moisture for monitoring field trafficability.

#### 5.4.4 Monitoring trafficability with Sentinel-1

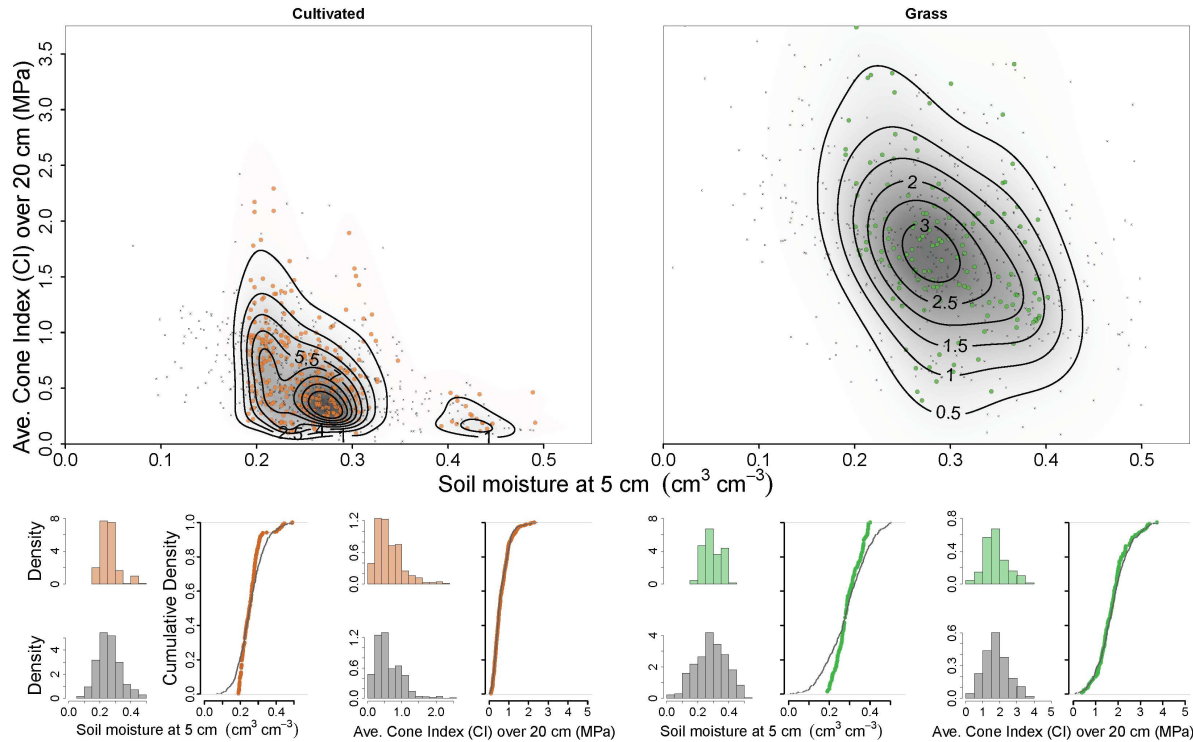
A small area around the vicinity of RM07 in the Raam catchment is selected as a test site to demonstrate monitoring of trafficability in grass and cultivated fields. Figure 5.9 (top panel) shows plots at four locations (2 in grass and 2 in cultivated fields). A 150 m pixel resolution was used based on the results in section 5.4.1. Time series plots in fig.5.9 (bottom) show the soil moisture trends over the fields. Monitoring of the trafficability status can be easily carried out using the high temporal resolution of Sentinel-1, although fewer images were available in 2016 as compared to 2017. Probability for exceeding the threshold are much higher for cultivated fields than grass fields even though similar range of soil moisture were encountered for both. Furthermore, soil moisture trends at the selected locations all show seasonality, but cultivated fields show larger inter-field variability than grass fields. We also observed the same trends in inter-field variability for other locations (both grass and cultivated fields) in this selected area.

From the time series plots in fig.5.9, periods with lower probability of exceeding the threshold translates to favorable trafficability conditions. For the two cultivated fields, this is observed in early spring (around April) of 2017, and coincides with the timing of tillage for most fields in the Netherlands. For the spring of 2016, however, the trafficability conditions between these two cultivated fields vary. For C1, soil moisture conditions indicate good trafficability, while for C2 it is the opposite case. Soil moisture contents in C2 did not decrease, so the trafficability conditions did not improve over the spring period. This difference may due to the types of crops grown as well as how these fields were managed in previous years.

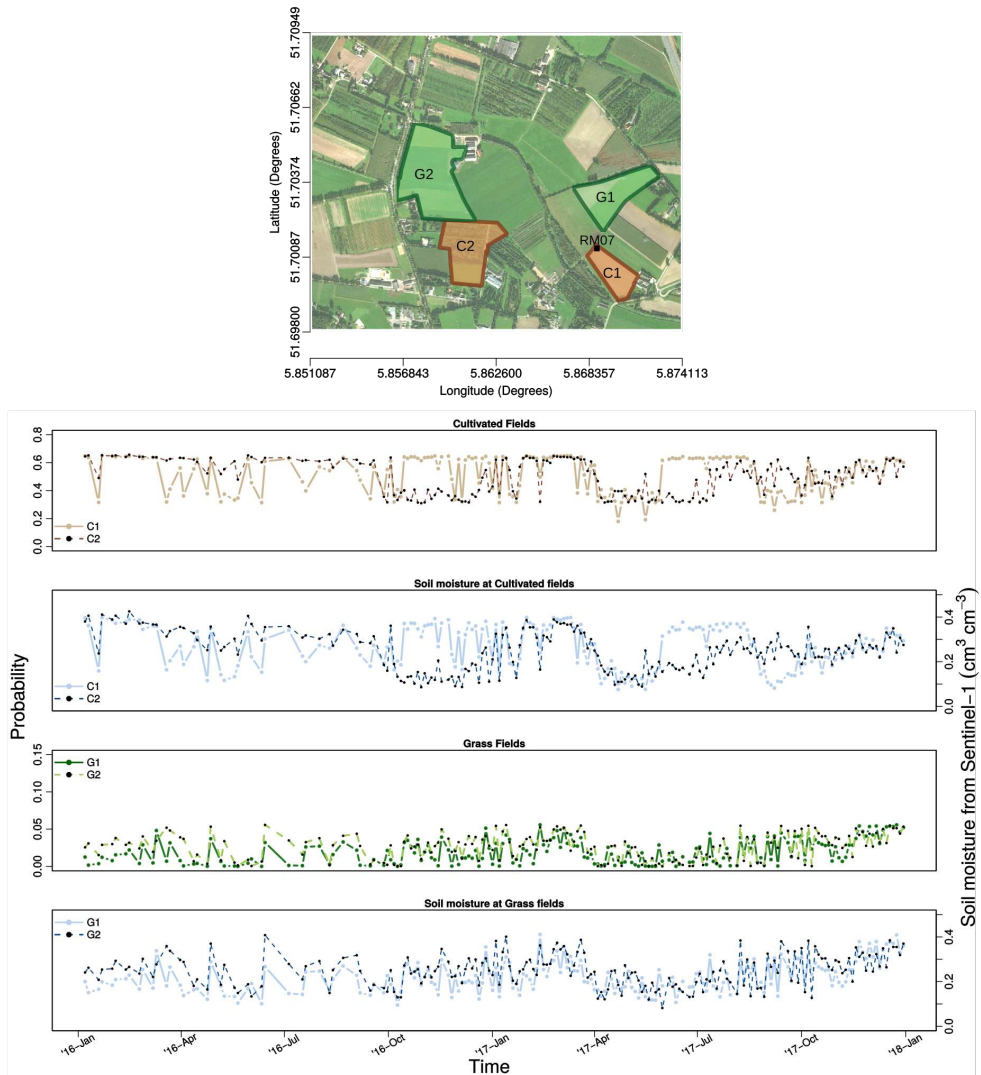
Grass fields always appear to be trafficable given the threshold used. This is not the case for cultivated fields where there are more pronounced fluctuations in trafficable conditions. However, the results for grass fields may not always hold true, especially for tractors with very heavy weights. In principle, these vehicles are legally allowed to reach up to 21 tons per axle (EU Regulation 167/2013). In contrast to our results for grass fields, soils may not be highly trafficable for very heavy vehicles during saturated conditions. The threshold value we used may be a more reasonable indicator of trafficability using average tractor weights ranging from 6-8 tons.

These results obtained could potentially be used by farmers and water managers alike. For farmers, the results can be used to identify periods suitable for traffic and to aid in deciding





**Figure 5.8:** Top: Plots of the joint pdf's of surface soil moisture and cone index generated by kernel density estimation for cultivated (left) and grass fields (right). Colored points are measurement data while grey points are random samples drawn using the joint pdf's. Contour lines indicate the density of points. Bottom: Histograms and cumulative distribution function (cdf's) of in situ measurements and random samples drawn are plotted for comparison.



**Figure 5.9:** Plot as an example for for monitoring field trafficability. *Top:* Map of the vicinity of RM07 station within the Raam soil moisture network where a few fields were selected. Two in grass (labeled G) and two in cultivated (labeled C) fields. One of the fields (C1) is also the location of RM07. *Bottom:* Time series plots for the probability values to express trafficability at the selected fields. Surface soil moisture derived from Sentinel-1 are also plotted (blue).

when to initiate tillage activities. Although in practice, this might still be difficult to apply since farmers still need to work on their fields despite conditions unfavorable for traffic. Nevertheless, the results can help in increasing awareness as to when soils are more susceptible to compaction.

For water managers, the results could aid in their decisions to adjust water levels in channels and/or canals in order to meet the water demands of farmers.

## 5.5 Conclusions

In this paper, we demonstrated how Sentinel-1 can be used to monitor trafficability status in agricultural fields. Our findings show that by identifying coupled conditions, satellite-derived surface soil moisture can be directly related with cone index values. Since we focused only on soil moisture as the dominant temporal control, expressing trafficability as probability values is advantageous because uncertainties are already incorporated. However, further attention should be given to the impacts of root growth as it may also act as a significant temporal control for cone index. Finally, we showed that the high temporal resolution of Sentinel-1 is suitable for tracking the changes in agricultural field trafficability. However, our results show that aggregation to coarser resolution may be necessary, which may also preclude identifying the spatial variability within a field.



"Two roads diverged in a wood, and I—  
I took the one less traveled by,  
And that has made all the difference."

Robert Frost

# 6

## Root zone soil moisture estimation using Random Forests



This chapter is an adapted version of:

Carranza, C., Nolet, C., Pezij, M., and van der Ploeg, M. (2020a). Root zone soil moisture estimation with Random Forest, *Journal of Hydrology* p. 125840, doi: 10.1016/j.jhydrol.2020.125840

## Abstract

Accurate estimates of root zone soil moisture (RZSM) at relevant spatio-temporal scales are essential for many agricultural and hydrological applications. In situ measurements have been increasing in the last decade, but are often of limited relevance due to their temporal and spatial specificity which does not provide long-term or spatially-distributed information. Quantitative models, which are either process-based or data-driven, will perhaps always be useful alternatives to direct soil moisture measurements. Data-driven machine learning techniques present unique opportunities to develop quantitative models that do not require assumptions about the processes operating within the system being investigated. In this study, Random Forest (RF), which is an ensemble learning algorithm, was applied with the main goal of identifying the utility of a data-driven model for estimating RZSM. Interpolation and extrapolation of RZSM on a daily timescale was carried out using RF over a small agricultural catchment from 2016 to 2018. In order to assess the performance and utility of RF, model results were compared to simulations from a process-based pore flow model combined with data assimilation. In addition, optimization steps were applied to increase model performance for both methods. In general, the accuracy obtained from both methods indicates that they are equally capable of accurately estimating RZSM. Both methods gave poorer estimates at extreme soil moisture conditions, which suggests that certain processes or covariates are poorly represented in the modeling efforts. However, in data-poor regions where soil hydraulic parameters are limited or lacking, RF may be more advantageous than a process-based model, especially when the primary goal is confined to the estimation of soil moisture states.

## 6.1 Introduction

Root zone soil moisture (RZSM) is an important environmental variable that impacts hydrological processes relevant for agriculture and climate-related studies. It is one of the main drivers for agricultural productivity (Rigden et al., 2020) and serves as an indicator for crop water stress which is valuable for drought monitoring (Bolten et al., 2009). Outside the hydrological cycle, RZSM dynamics play a role in quantifying soil carbon fluxes (e.g. Kurc and Small, 2007).

Accurate estimates of RZSM are necessary in order to have a better understanding of agricultural and environmental processes it controls. Direct RZSM measurements can be obtained from in situ sensors installed along the soil profile or at specific depths (Vereecken et al., 2008; Dobriyal et al., 2012). Achieving distributed spatial measurements of RZSM can be a challenge because installation of sensors at the subsurface can be a tedious task and are likely to disturb the soil properties. It has become relatively common to extract RZSM from surface soil moisture (SSM), which may be in situ or satellite-derived (Ulaby et al., 1996), since they are more easily obtained. Satellite-derived SSM has the advantage of providing spatially distributed soil moisture while in situ measurements offer higher temporal frequency (second or minutes) compared to satellites, which only provide snapshots at regular time intervals (days or weeks).

Analytical solutions are applied in cases when direct RZSM measurements are lacking or insufficient. These methods are based on theoretical or empirical relations between environmental variables controlling RZSM state. Arguably, the most common approach is to apply process-based hydrological models which are based on conceptual understanding of the system (e.g. Cordova and Bras, 1981; Porporato et al., 2004). These models employ numerical solutions of flow and transport equations in unsaturated porous media (Feddes et al., 1988). Information on soil hydraulic properties, either measured directly or from pedo-transfer functions (Schaap et al., 2001; Van Looy et al., 2017), are required to estimate water movement across a chosen flow domain. It may be necessary to optimize soil hydraulic parameters, for instance using inverse modeling (e.g. Ritter et al., 2003), in order to improve model simulation accuracy. The prevailing meteorological conditions, as well as variables that describe vegetation growth are necessary to determine the amount of water entering and exiting a given flow domain. In the last couple of decades, data assimilation methods have been applied to improve process-based model estimates (Houser et al., 1998; Peziz et al., 2019b), which may take advantage of satellite-derived SSM information.

Data driven methods to estimate RZSM include time series analysis (TSA) and machine learning (ML) techniques. These methods aim to extract knowledge by evaluating patterns or variability in that data and further stimulate actions that are dictated by the data. In the context of RZSM estimation, data-driven methods implicitly incorporate and evaluate all the interacting processes that produced a given RZSM state. TSA methods, such as the application of an exponential filter (Wagner et al., 1999; Albergel et al., 2008), a cumulative distribution function (cdf-matching, Gao et al., 2019; Zhuang et al., 2020), or transfer-functions (Peziz et al., 2020) primarily utilize surface soil moisture data to derive a functional relation with

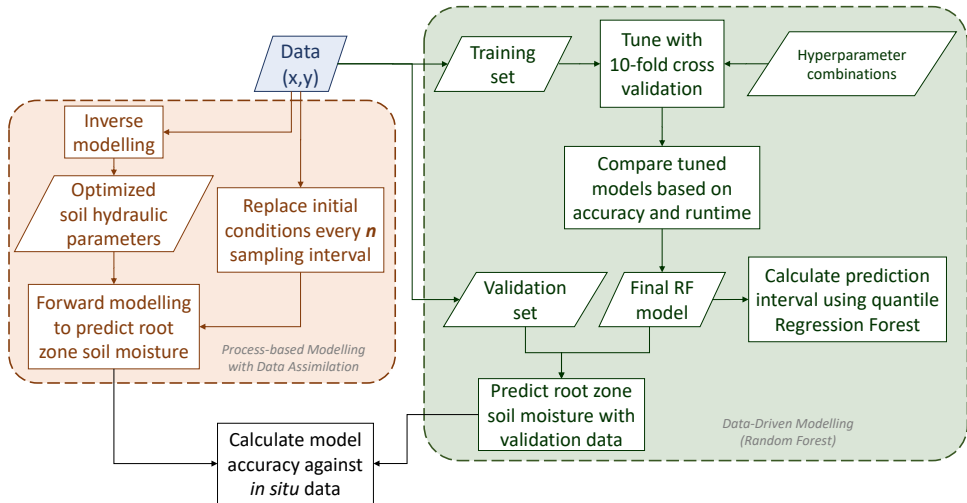
RZSM. However, calibration of functional parameters may be necessary each time it is applied to a different study area in order to obtain high accuracy. ML algorithms build mathematical models based on training sets and covariates to extract information from data. Furthermore, they are tuned to handle diverse and large volumes of data sets, which may be relevant for large scale studies or for operational (water) management. Application of ML methods in soil hydrology have started to gain attention in the last couple of decades. For instance, ML techniques have been applied to estimate model-derived RZSM (Kornelsen and Coulibaly, 2014) or SSM (Ahmad et al., 2010). Using satellite data, ML techniques allow up- or downscaling of soil moisture values (Srivastava et al., 2013; Zhang et al., 2017). Comparison of ML methods have been made for forecasting of soil moisture using values at discrete soil moisture depths (Prasad et al., 2018) or soil layers (Matei et al., 2017) at regional scales. Interestingly, SSM has also been estimated from in situ measurements of soil moisture at deeper layers using ML (Coopersmith et al., 2016). In a comparison study, Karandish and Šimnek (2016) showed that the ML techniques may provide a useful alternative to process-based models using limited input data.

We applied Random Forest (RF), an ensemble learning algorithm, to estimate RZSM based on in situ datasets within the Raam catchment in the Netherlands. Among the advantages outlined by Tyrallis et al. (2019) is that RF has been found to produce consistent predictions and it reduces the variance without increasing the bias of the predictions. So far, there are still limited studies applying RF in soil hydrology, particularly in estimations of RZSM. A data-driven method will ensure that all the processes operating in the system under study are incorporated in the predictive RF model developed. In addition, there has been proliferation of RZSM in situ measurements in the last couple of decades from various soil moisture monitoring networks worldwide (e.g. International Soil Moisture Network (ISMN, Dorigo et al., 2011a) which provides an excellent opportunity to capitalize on ML techniques. In this study, daily measurements in agricultural fields for almost two years were used for RF modeling in two ways: 1) interpolation at randomly selected points within the time series and 2) extrapolation of future RZSM state based on past values. A comparison is then made between the RF results and a process-based model in order to assess the capabilities of a data-driven method. A pore-flow model with data assimilation via direct insertion of in situ measurements was applied to simulate RZSM at the study sites.

## 6.2 Materials and Methods

As an overview, Random Forest (RF) was applied for interpolation and extrapolation of root zone soil moisture (RZSM) within a small catchment. For comparison of results, RZSM were also simulated using a process-based (PB) pore-flow model which was combined with data assimilation. For both methods, steps to optimize model (hyper) parameters were applied to improve model performance and to allow objective comparison of the results. Briefly, a selection of hyperparameters were tuned for RF while the soil hydraulic parameters were optimized via inverse modeling for PB. Data assimilation via direct insertion of in situ measurements was further applied to improve PB model results. The succeeding sections further describe in

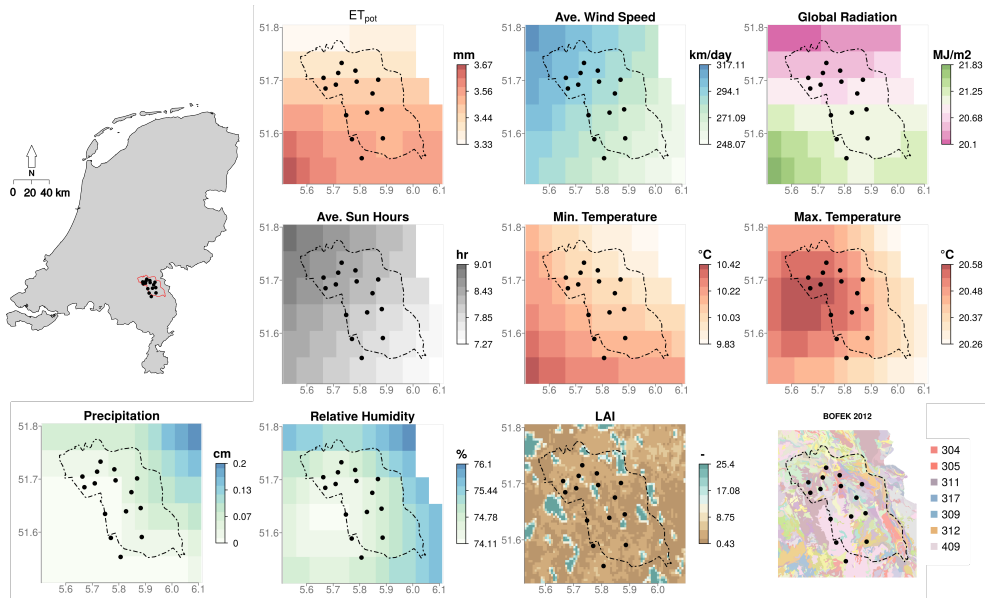




**Figure 6.1:** Flowchart to summarize the the analysis. A data-driven (Random Forest, RF) and a process-based model were applied to predict and forecast root zone soil moisture. For RF, hyperparameters were tuned to obtain the optimal model using different training sets. In addition to the RF estimates, prediction intervals were calculated using quantile regression forest to quantify the prediction uncertainties. For the process-based model, the soil hydraulic parameters were first optimized using inverse modeling. Forward modeling was then ran with data assimilation via direct insertion of *in situ* measurements into the initial conditions of the simulation runs for every 20-day sampling interval ( $n$ ).

**Table 6.1:** List of crops types encountered in the Raam network

| Station | 2016      | 2017    | 2018    |
|---------|-----------|---------|---------|
| RM01    | grass     | grass   | grass   |
| RM02    | sugarbeet | fennel  | corn    |
| RM03    | grass     | grass   | grass   |
| RM04    | grass     | grass   | grass   |
| RM05    | onion     | lettuce | lettuce |
| RM06    | grass     | grass   | grass   |
| RM07    | chicory   | potato  | corn    |
| RM08    | sugarbeet | wheat   | beans   |
| RM09    | sugarbeet | potato  | corn    |
| RM10    | grass     | grass   | grass   |
| RM11    | corn      | corn    | wheat   |
| RM12    | grass     | grass   | grass   |
| RM13    | corn      | corn    | corn    |
| RM14    | grass     | grass   | grass   |
| RM15    | grass     | grass   | grass   |



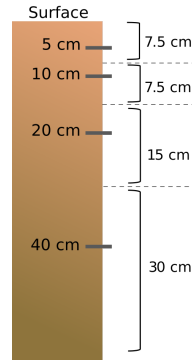
**Figure 6.2:** Site characteristics. The image on the top left shows the location of the Raam catchment (red area) in the southeastern portion of the Netherlands. Black dots indicate the 15 soil moisture stations used in the analysis. The remaining images show the covariates used for the Random Forest modeling, including meteorological, vegetation, and soil characteristics at the sites. Meteorological data are obtained from KNMI (Royal Dutch Meteorological Institute), Leaf Area Index (LAI) is from MODIS, and soil hydrological groups are from BOFEK2012 (Wosten et al., 2013). Except for BOFEK2012, the images shown are snapshots of the datasets on July 3, 2016.

detail the datasets used and methods applied while a summary is shown in fig.6.1.

### 6.2.1 Raam Soil Moisture Network

The Raam catchment is located in the southeastern portion of the Netherlands which holds mostly sandy soils. A total of 15 operational soil moisture stations are distributed across the whole catchment (fig.6.2). At each station, soil moisture and temperature sensors (Decagon EC-H20 5TM) were installed at 5, 10, 20, 40, and 80 cm depths and measurements were recorded every 15 minutes. The soil moisture stations were located in agricultural fields, which are the characteristic land cover type within the catchment area. The most common crop type at the stations is grass, followed by corn, potato, sugarbeet, and other vegetable crops (table 6.1). A more detailed description of the Raam soil moisture network is provided in Benninga et al. (2018).

Measurements down to 40 cm depth were integrated over a 60 cm averaging depth to calculate for root zone soil moisture (fig.6.3). This was chosen in order to have a uniform root zone across the study sites which have varying crop types. For grass fields, the active root zone



**Figure 6.3:** Diagram showing the installation setup for each station. Soil moisture sensors were installed at 5, 10, 20, and 40 cm depths for each station. For each measurement depth, an associated soil thickness is determined based on the midway distance between two adjacent measurement points. Root zone soil moisture is represented by the zone-weighted depth-average values that are aggregated based on both the measurements and its associated soil thickness (see eq.6.1).

may only be up to 20 cm because of its shallow rooting system while for crops such as corn or potato, the root zone can extend beyond 1 m. Nevertheless, the depth used for the analysis generally captures the active root zone for the crops at the study sites. Furthermore, the methods applied in this study could also be customized for other depths that would suitably represent the root zone depths. Root zone soil moisture  $\theta_{rz}$  is given by:

$$\theta_{rz} = \frac{\sum_{j=1}^n \theta_j \Delta z_j}{z} \quad (6.1)$$

where  $\theta_j$  (in  $\text{m}^3 \text{m}^{-3}$ ) is the volumetric water content for measurement depth  $j$  (cm),  $\Delta z_j$  (cm) is the thickness of soil associated with the measurement depth, and  $z$  (cm) is the total averaging depth. Measurements from all 15 stations starting from April 2016 up to December 2018 (33 months) were used for the analysis. The daily mean values from the 15-mins. data were calculated in order to match the meteorological datasets acquired.

### 6.2.2 Random Forest regression

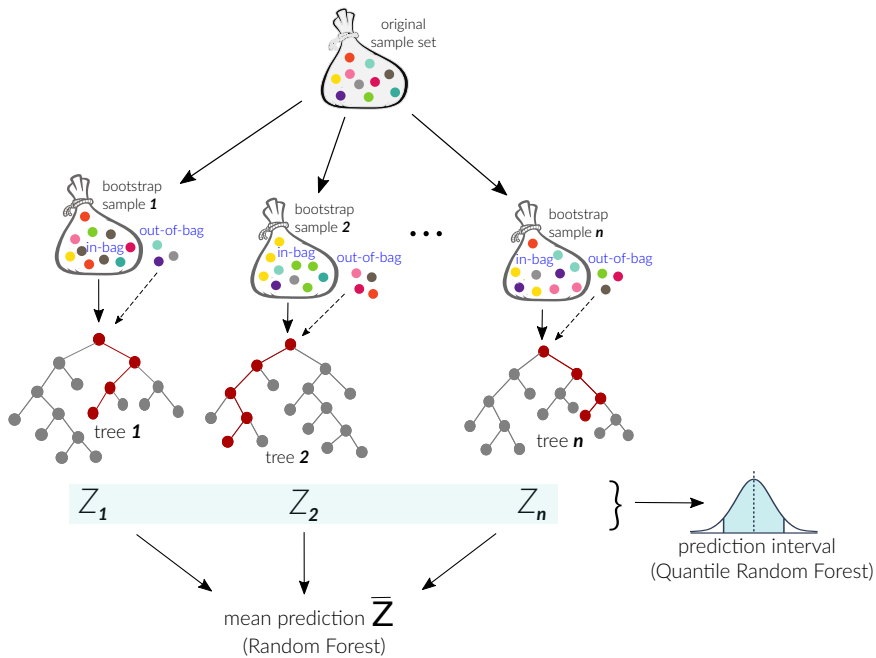
Random Forest (RF) is an ensemble-learning algorithm that combines the concepts of decision trees and bagging (fig.6.4, Breiman, 2001). Decision trees (DT), either for classification or regression, partition the variable space using a set of hierarchical rules such that the dataset are grouped recursively based on similar instances. A set of covariates (continuous or categorical variables) are used for recursively splitting the values of the variable of interest, which results in multiple parent and child nodes that resemble a tree-like structure. Splitting at each node involves random selection of candidate variables from the total number of covariates, referred to as the *mtry* parameter. The DT will evaluate each candidate to find the optimal split that maximizes the 'purity' or alternatively, that results in the largest decrease in the impurity at

each child node. In this case, the estimated response variance for regression trees was used as a measure for impurity (Wright and Ziegler, 2015). RF creates diverse DTs to avoid highly correlated predictors by growing them from different subsets of the training data through a procedure called bagging. Often a large number of trees are created, and is referred to as the *n<sub>tree</sub>* parameter. Bagging, an abbreviation for 'bootstrap aggregation', is a technique for generating multiple training data by resampling with replacement of the original training set. This means that some data may be used more than once in the training, while others might never be used. For each bootstrap sample, a regression DT generates multiple parent and child nodes until the stopping criterion is reached. In this case, when the value for the minimum node size (*min.nodesize*) parameter is achieved. After all the trees are grown, the RF regression predictor is the mean from all the predictions from each individual tree. More detailed description of RF methods and parameters are given in Breiman (2001) and Hastie et al. (2009).

Estimation of RZSM using RF was implemented in two ways: 1) interpolation of randomly selected points within the whole time series data, and 2) extrapolation of 'future' RZSM based on 'past' values. For each method, a single RF model was built based on the combined measurements from all 15 soil moisture stations in the Raam. For RF interpolation, random samples were obtained from the daily time series data at each station. For RF extrapolation, the length of the time series at each station was first split based on the sampling proportion used. The first part of the time series was selected for training and constitutes the "past" data, while the remaining was used for model validation and constitutes the "future" data. Proportions of 50% up to 80% (with increments of 10%) of the daily time series measurements at each station were used to generate the samples from each station. These were then combined into one training set for building each RF model. The value of the *n<sub>tree</sub>* parameter was made proportional to the samples in each training set, and was set to a tenth of the amount of each training set. This corresponded to 600, 700, 900 and 1000 trees for interpolation and 600, 800, 900, 1100 trees extrapolation for each training set (50%, 60%, 70%, 80% of total measurements). Optimization of RF models were carried out by tuning *mtry* and *min. node size* parameters for each training set proportion tested.

#### *Hyperparameter optimization*

The RF model was tuned in order to select the combination of the hyperparameters *mtry* and *min. node size* that would yield the highest accuracy. Hyperparameters are parameters that need to be set prior to training a model and somewhat defines the configuration of the regression trees. Their values directly control the behaviour of the learning algorithm and have a significant effect on the performance of the model being trained. Other RF model parameters that are not tuned will simply 'learn' on their own during model training. The values for *mtry* will dictate splitting of RZSM values at the nodes of the regression tree while the minimum number of elements per node (*min. node size*) will serve as a stopping criterion in building the regression trees. We tested values of *mtry* from 1 to 25 and *min. node size* of 5, 10, 20, and 30. A total of 100 combinations of hyperparameters (*mtry* and *min. node size*) were tested for each of the four training set proportions (50% to 80%) in the tuning phase. A



**Figure 6.4:** Schematic diagram for Random Forest. For each sample or training set, a large number of bootstrap samples ( $n_{tree}$ ) are obtained. A regression tree is built from each bootstrap sample which involves hierarchical partitioning of the dataset. This involves splitting of the datasets based on candidate variables that are randomly selected from the total number of covariates used ( $m_{try}$ ). The final RF estimate is the average of all estimated values from each regression tree. Based on the distribution of all the RSZM estimates obtained from all the trees, the prediction intervals at the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles were determined to quantify the uncertainties in the RF model.

10-fold cross-validation (CV) scheme was applied for each hyperparameter combination in the tuning phase. The mean root mean square error (RMSE) computed for each hyperparameter combination in the 10-fold CV scheme were compared to assess model performance, which in this case in the RMSE. Aside from having a separate validation set, a CV scheme is a preventive measure for model overfitting (Lever et al., 2016).

RMSE's were examined further to select the final RF model as the one with highest accuracy ('best model') might also be computationally expensive. Therefore, we compared the model with the best RMSE to another one that has a faster computation time but comparable RMSE as a 'tradeoff' to evaluate a simpler model without sacrificing accuracy. RMSE's were first ranked from lowest to highest and then a pairwise elimination process was applied by evaluating the improvement in RMSE. The final 'tradeoff' model was selected once a  $< 1\%$  improvement in RMSE was found.

*Random Forest covariates*

Covariates or the set of predictor variables used to build RF regression trees include information on meteorological conditions, soil properties, land cover and vegetation characteristics at each site (fig.6.2 and table 6.2). Daily meteorological data from 36 KNMI (Royal Dutch Meteorological Institute) stations distributed within the entire Netherlands were interpolated to produce a 5 x 5 km gridded image in order to extract daily values at the location of each soil moisture station. Spatial estimates of the values at each KNMI station were obtained using a Thin Plate Splines interpolation (Sluiter, 2012). Temperature, wind speed, relative humidity, sun hours, potential evapotranspiration, and radiation, were selected among the total meteorological datasets available as these were also the input variables in the process-based model applied in this study. They are, therefore, indicative of surface processes that influence the RZSM state. Gridded values of daily rainfall measurements with a 1 x 1 km pixel size were obtained directly from KNMI. Ordinary kriging was applied to around 300 measurement locations of rain gauges distributed across the Netherlands to produce the rainfall maps (Soenario et al., 2010). Leaf area index (LAI) from an 8-day MODIS composite with 500 m resolution was used to capture vegetation characteristics over the study sites. The values for days in between LAI measurements were linearly interpolated to obtain daily estimates.

Both crop type and soil hydro-physical groups were also included as categorical covariates. The former is based on field observations while the latter was obtained from Bodemfysische Eenhedenkaart (BOFEK2012, Wosten et al., 2013), which is a map of soil hydro-physical properties for the Netherlands. Figure 6.2 indicates the BOFEK2012 codes for the soils within the Raam network only. Further description of each code is given in Wosten et al. (2013). These two categorical variables are re-coded into dummy or indicator variables for the RF regression. The categorical variables are transformed into a dichotomous (1 or 0) representation of its presence or absence for each data point. For example, the categorical variable "Crop" with a type "Corn", a value of 1 is assigned for measurements having the said crop, and 0 for measurements with another crop type.

The current soil moisture state is inevitably affected by its past values and past meteorological conditions. The so-called soil moisture memory (or persistence) has been widely investigated because of its importance in climate-related studies (e.g. Koster and Suarez, 2001). Therefore, lagged values were also calculated for meteorological and soil moisture datasets in order to incorporate past information in the RF model. This may be useful especially for forecasting where only past information is available. For surface soil moisture and meteorological datasets, values with a lag of 1 day were obtained. Additional lagged surface soil moisture values of 3 and 40 days were also calculated based on findings of soil moisture memory studies at global (McColl et al., 2017) and continental (European, Orth and Seneviratne, 2012) scale, respectively. A total of 39 covariates were used for the RF models (table. 6.2).

*RF prediction intervals*

Uncertainties in RF estimates are defined based on the 95% prediction interval (PI) obtained using quantile regression forest (qRF, fig.6.4). The idea behind qRF is that instead of recording

**Table 6.2:** Covariates used in Random Forest

| Meteorological |                     | Vegetation     |                   | Soil      |                         |
|----------------|---------------------|----------------|-------------------|-----------|-------------------------|
| Symbol         | Description         | Symbol         | Description       | Symbol    | Description             |
| RG             | Ave. Wind speed     | LAI            | Leaf Area Index   | VWC5      | Soil moisture at 5cm    |
| Q              | Radiation           | LAI.lag        | 1-day lag         | VWC.lag   | 1-day lag               |
| rd             | Rainfall            | Crop.grass     | Crop type (dummy) | VWC.lag3  | 3-day lag               |
| SQ             | Sun Hours           | Crop.corn      |                   | VWC.lag40 | 40-day lag              |
| TN             | Min. Temp           | Crop.potato    |                   | BOFEK.305 | BOFEK2012 codes (dummy) |
| TX             | Max. Temp           | Crop.sugarbeet |                   | BOFEK.304 |                         |
| UG             | Relative Humidity   | Crop.wheat     |                   | BOFEK.311 |                         |
| EV24           | Evapo-transpiration | Crop.onion     |                   | BOFEK.409 |                         |
| RG.lag         | 1-day lag           | Crop.fennel    |                   | BOFEK.317 |                         |
| Q.lag          |                     | Crop.beans     |                   | BOFEK.309 |                         |
| rd.lag         |                     | Crop.lettuce   |                   | BOFEK.312 |                         |
| SQ.lag         |                     |                |                   |           |                         |
| TN.lag         |                     |                |                   |           |                         |
| TX.lag         |                     |                |                   |           |                         |
| UG.lag         |                     |                |                   |           |                         |
| E24.lag        |                     |                |                   |           |                         |
| DOY            |                     | Day of year    |                   |           |                         |

the mean value of response variables from the trees, all responses for each tree are recorded (Meinshausen, 2006). This allows not only for the estimation of the conditional mean but also a good approximation of the full conditional distribution. PIs were defined using quantile regression based on the chosen quantiles ( $\alpha$ 's). For a given random variable, the conditional distribution function  $F(y|X = x)$  is given by the probability that, for  $X = x$ ,  $Y$  is smaller than  $y$ . For a continuous distribution function, the  $\alpha$ -quantile ( $Q_\alpha(x)$ ) specifies a value such that the probability of  $x$  being smaller than  $Q_\alpha(x)$  is, for a given random variable  $X = x$ , exactly equal to  $\alpha$ . A 95% PI ( $I_{95}$ ) for the RZSM estimates is based on 2.5% and 97.5% quantiles ( $[Q_{0.025}(x), Q_{0.975}(x)]$ ).

### Variable Importance

Variable importance from the RF models determined using a permutation method (Wright and Ziegler, 2015). Rankings for covariates were based on the mean decrease in model accuracy after shuffling or randomly permuting the values of a predictor  $X_i$ , where  $i = 1 \dots n$  for each of the covariates used. By permuting the values of  $X_i$ , its association with the response variable  $Y$  (i.e. RZSM) is broken. Therefore, if the predictor  $X_i$  is associated with the response  $Y$ , a substantial decrease in accuracy is expected after prediction using the permuted and remaining non-permuted variables.

### 6.2.3 Process-based modelling with data assimilation

A soil water balance model was carried out to simulate one-dimensional daily RZSM at the study sites. In a water balance model, mass and energy fluxes over time and/or space are calculated to estimate soil moisture along the profile. We assumed that soil water movement would be restricted along the vertical dimension since the study sites are generally characterized by homogeneously textured soils and the terrain at the study sites is generally flat (e.g. De Laat, 1980, for modeling unsaturated flow in the Netherlands). The vertical water flow in unsaturated porous media is solved numerically using Richard's equation:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial Z} \left[ K(h) \left( \frac{\partial h}{\partial Z} + 1 \right) \right] - S \quad (6.2)$$

where  $t$  is the time (days),  $\theta$  is the volumetric water content ( $\text{cm}^3 \text{cm}^{-3}$ ),  $h$  is the soil water pressure head (cm),  $Z$  is the spatial coordinate (cm) defined as positive upward,  $K(h)$  is the unsaturated hydraulic conductivity function ( $\text{cm d}^{-1}$ ) and  $S$  is a sink term representing water uptake by plant roots ( $\text{cm d}^{-1}$ ).  $K(h)$  is derived from a water retention curve, given by van Genuchten (1980):

$$\theta(h) = \frac{\theta_s - \theta_r}{[1 + (\alpha h)^n]^m}, h \leq 0 \quad (6.3)$$

$$\begin{aligned} K(h) &= K_s S_e^l (1 - (1 - S_e^{l/m})^m)^2 \\ m &= 1 - \frac{1}{n} \\ S_e &= \frac{\theta - \theta_r}{\theta_s - \theta_r} \end{aligned} \quad (6.4)$$

where  $\theta_r$  and  $\theta_s$  denote residual and saturated volumetric water contents ( $\text{cm}^3 \text{cm}^{-3}$ ), respectively;  $\alpha$  ( $\text{cm}^{-1}$ ) and  $n$  (–) are fitting parameters of soil water characteristic curve;  $K_s$  is the saturated hydraulic conductivity ( $\text{cm d}^{-1}$ );  $l$  (–) is the pore connectivity parameter; and  $S_e$  (–) is the relative saturation.

#### *Inverse modeling for parameter optimization*

The soil water balance was carried in two parts using Hydrus-1D software (Simunek et al., 2005). The first part involved optimization of soil parameters describing the shape of the water retention curve ( $\theta_s$ ,  $\theta_r$ ,  $\alpha$ ,  $n$ ) and hydraulic conductivity curve ( $K_s$ ,  $l$ ) using inverse modeling. We initially carried out simulations using soil hydraulic parameters available from (BOFEK2012, Wosten et al., 2013), but found the results to be unsatisfactory. Optimization of soil hydraulic parameters was based on Marquardt-Levenberg parameter estimation method (Marquardt, 1963) as implemented in Hydrus-1D, using soil water content measurements. The soil domain considered is 1 m to cover depths similar to the measurements stations. A variable atmospheric condition, based on rainfall and evapotranspiration, was set as the

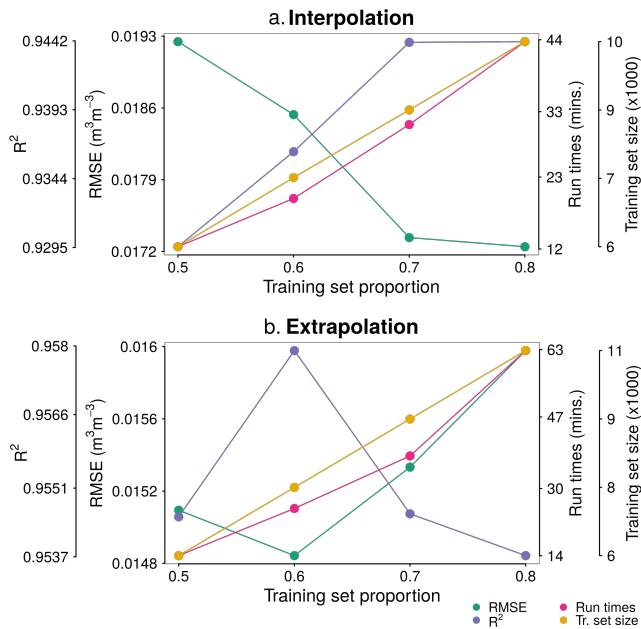


upper boundary conditions while a free drainage condition was set as the lower boundary conditions. Daily meteorological datasets from KNMI, as described in section 6.2.2, were used for the upper boundary conditions. Initial conditions for the inverse modeling were set to the pressure head at field capacity with the assumption that the soil is close to saturation the start of a year when the simulations commenced. In addition, simulations from January until just before the start of the in situ measurements in April were part of the spin-up period for the model. A single porosity van Genuchten - Mualem model without hysteresis was used for the simulation. The flow domain was subdivided based on the number of soil layers present in BOFEK2012 (see table 8.1, Appendix). For instance, station RM02 was subdivided into two layers while RM15 has only one layer in the flow domain. Simunek et al. (2005) provides more detailed information regarding the theory, methods and default parameters in H1D software. Subsequently, the second part was the forward modeling to estimate soil water content using the optimized set of soil hydraulic parameters. Observation points within the flow domain were selected at the same depths as the in situ measurement points. Furthermore, depth-averaged zone-weighted root zone soil moisture was calculated in a similar manner as the in situ values.

#### *Sequential Data Assimilation*

Data-assimilation is an often applied method to improve the accuracy of hydrological modelling using up-to-date measurements. The goal of data assimilation is to combine measurements and modelling efforts into an optimal state estimate of the variable of interest (Reichle, 2008). The difference between machine-learning and data-assimilation is that the latter depends on a dynamical model of the system, in this case H1D. To show the added-value of the machine-learning method, we show an application of data-assimilation with the modelling instrument used in this study. We recognize that data-assimilation should use information on uncertainties in both observations and modelling efforts, for which sequential methods such as the Ensemble Kalman Filter can be used (Evensen, 2009; Houtekamer and Mitchell, 1998; Peziz et al., 2019b). However, in this study we only focus on a simple data-assimilation method, which is relatively easy implemented. Houser et al. (1998) and Heathman et al. (2003) showed the value of direction insertion for soil moisture modelling. Therefore, we applied a direct insertion data-assimilation method to update the soil moisture state.

We applied direction insertion by replacing the model state by the in situ measurements for every 20-day interval over the whole simulation period which covered two years. Measurements along the entire soil profile were used. At the end of each 20-day period, the model state was replaced by the soil profile provided by the in situ measurements. The model was subsequently run for the next 20-day period. The said assimilation interval was tested as it approximated the revisit times of some microwave satellites (e.g. Radarsat-2 or ALOS PALSAR-2) which have been assimilated into process-based models in the past. The days when data were assimilated were excluded for model evaluation.



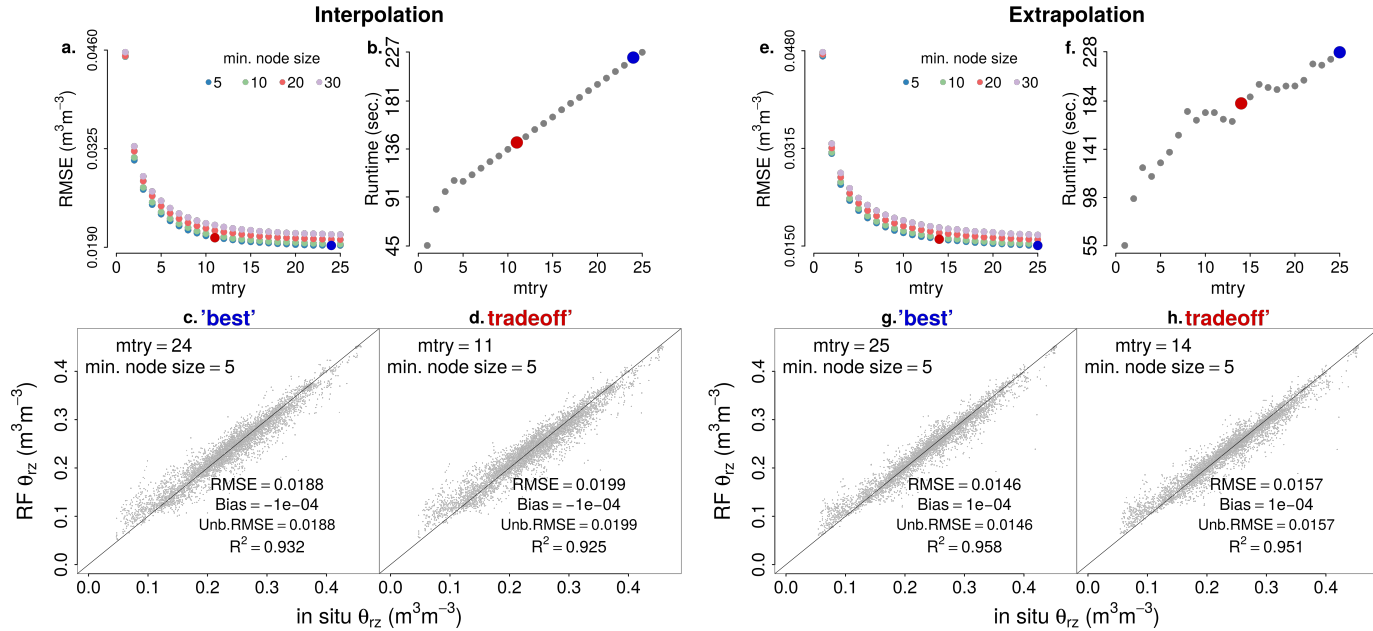
**Figure 6.5:** Accuracy metrics and RF model specifications for different training set sample proportions.

## 6.3 Results and Discussions

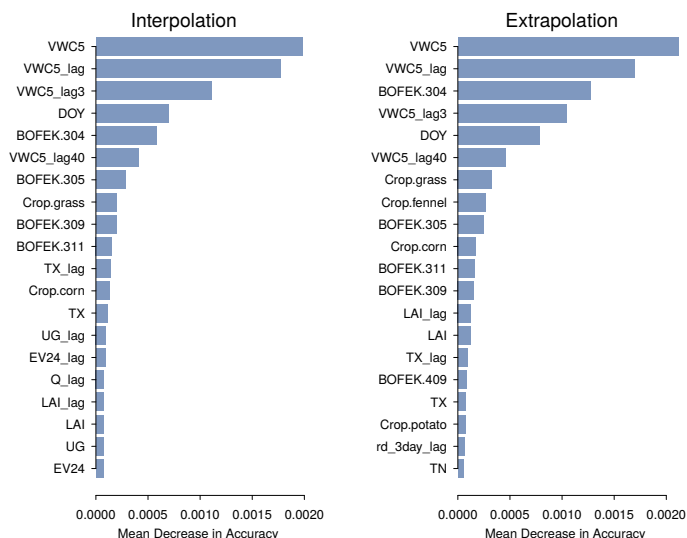
### 6.3.1 Random Forest model tuning

The RF models generated using different training sets indicate that the highest and lowest RMSEs are based on 50% and 80% for interpolation and 80% and 60% for extrapolation (fig.6.5a and b). However, a 50% training set performed relatively well in both cases based on in very minimal decrease in the RF model performance; RMSEs are only  $0.002 \text{ m}^3 \text{m}^{-3}$  and  $0.0003 \text{ m}^3 \text{m}^{-3}$  higher than those obtained using 80% and 60% training proportion for interpolation and extrapolation, respectively. Furthermore, the runtime is fastest with a 50% training set, decreasing the computation time of the best performing model by at least 44% (e.g. from 25 to 14 mins. for extrapolation). This aspect is of importance for machine learning techniques, especially as the volume of datasets become larger. Therefore, we selected the RF model from a 50% training set for further evaluation of the hyperparameter tuning results.

A comparison of the results obtained from the tuning process using 50% training set is given in fig.6.6. The RMSEs are observed to exponentially decrease with increasing *mtry* values, which is combined with a consistent increase in accuracy with smaller *min. node size* (fig.6.6a and e). For both interpolation and extrapolation, a large *mtry* and a small *min. node size* resulted in the best RF model based on RMSEs (fig.6.6c and g). This is somewhat expected



**Figure 6.6:** Random Forest model tuning results using 50% training proportion. RMSE's for all hyperparameter combinations are given in *a* and *c* while the runtimes are shown in *b* and *f*. The points for the 'best' (red dot) and 'tradeoff' (blue dot) models are highlighted. Scatterplots (*c,d,g,h*) with corresponding accuracy metrics show the differences between the 'best' and 'tradeoff' models.

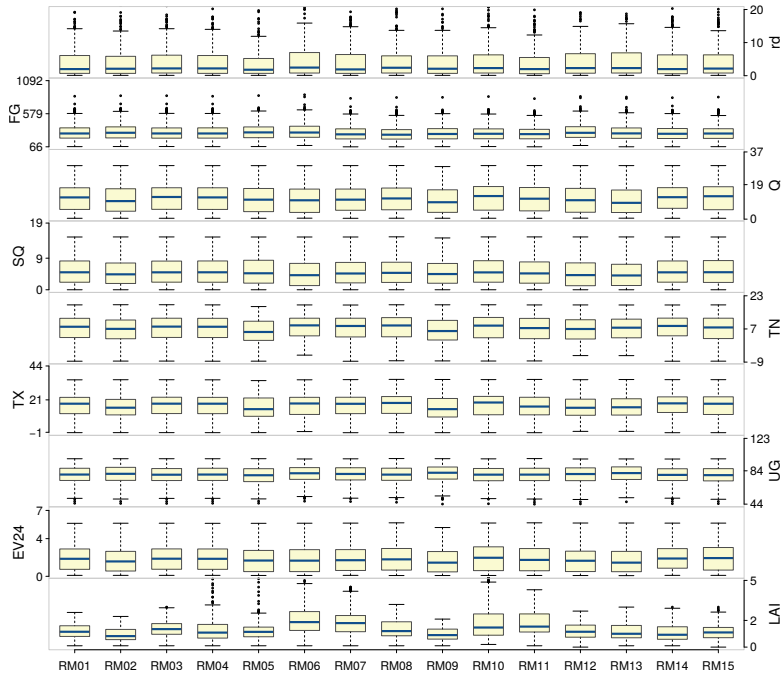


**Figure 6.7:** Bar graphs enumerating the important variables (VI) for the RF model, decreasing from top to bottom. Only the top 20 VI's are listed for presentation purposes. The description of the variables are provided in table 6.2.

because the homogeneity of elements at each node is higher when the *min. node size* value is kept smaller. In addition, the largest *mtry* costs the most computing time, as expected. In contrast, the 'tradeoff' model with a smaller *mtry* value halves the computing time (from 25 to 10 sec.) but only has a slightly lower RMSE fig.6.6 (bottom panel). To balance accuracy and computation time, the hyperparameters from the 'tradeoff' model were used further for model evaluation.

### Variable Importance

Based on fig.6.7, surface soil moisture (SSM), soil properties and land cover types have larger impacts on RF model accuracy compared to meteorological variables. Lagged soil moisture values appear higher on the list of important variables (VI) in the RF model. For both interpolation and extrapolation, SSM with lags of up to 40 days are still highly relevant in estimating RZSM. Temperature appears to have the most important effects on RF model performance among all the meteorological covariates included. Since soil moisture is directly influenced by precipitation, it is surprising that current and antecedent rainfall did not rank higher in the VI list. Rather, the impact of precipitation on the RF model may be included within the SSM variables (VWC5 and VWC5\_lag), which ranked highest in the VI list. Within the Raam catchment, meteorological conditions among the Raam stations were found to be similar due to its small areal coverage (fig.6.8). Since RF capitalizes differences or unique values in the covariates to separate RZSM into groups in building the regression trees, meteorological variables over the stations may have been less able to differentiate between RZSM among different stations



**Figure 6.8:** Comparison of values of temporally varying covariates among the different stations. The description of the variables are given in table 6.2.

based on the single RF model trained (i.e. one each for interpolation and extrapolation). We hypothesize that the influence of meteorological conditions may also be encapsulated within the DOY variable, which represents effects of seasonal changes on RZSM. The combination of DOY and temperature may have been adequate for the RF models to estimate RZSM in the Raam catchment. The results obtained, however, do not imply that meteorological variables are not important controls for RZSM. In this case, meteorological variability over the Raam catchment were secondary to variability in crop types and soil characteristics among the stations for estimating RZSM using the single RF model developed. Perhaps representing all the meteorological variables into one (or two) variables via dimensionality reduction methods (e.g. Principal Component) may potentially result in different VI ranking for a consequent RF model.

### 6.3.2 Root zone soil moisture estimation in the Raam catchment

The RF model performance for stations with the best and worst accuracy obtained from RZSM interpolation and extrapolation are given in fig.6.9 (See Appendix, fig.8.2 and fig.8.3 for the complete results). RF interpolations (fig.6.9a and 6.9b) for RZSM have high accuracy in comparison to RF extrapolations (fig.6.9 c and d). However, the soil moisture dynamics

**Table 6.3:** Accuracy metrics for RF and Hydrus-1 (H1D) for interpolation and extrapolation of root zone soil moisture. The range of values (min.-max.) encountered from all 15 stations in the Raam network are reported. Except for the unitless  $R^2$ , the metric are expressed in  $\text{m}^3 \text{m}^{-3}$ .

|           | Interpolation    |                  | Extrapolation    |                  |
|-----------|------------------|------------------|------------------|------------------|
|           | RF               | H1D              | RF               | H1D              |
| RMSE      | 0.0097 – 0.0313  | 0.0185 – 0.0507  | 0.0168 – 0.0621  | 0.0201 – 0.0544  |
| Bias      | -0.0128 – 0.0178 | -0.0204 – 0.0259 | -0.0235 – 0.0526 | -0.0232 – 0.0219 |
| Unb. RMSE | 0.0095 – 0.0263  | 0.0175 – 0.0506  | 0.0167 – 0.0422  | 0.0192 – 0.0542  |
| $R^2$     | 0.7985 – 0.9730  | 0.6829 – 0.8652  | 0.6821 – 0.9611  | 0.4030 – 0.8443  |

(i.e. an increase or decrease) are still captured in the extrapolated values, even though soil moisture state may be over- or underestimated. The accuracy of Hydrus-1D (H1D) simulations are generally lower than RF interpolations but are comparable with RF extrapolations (table 6.3). For instance, the values from H1D simulations are closer to in situ values at the station with the worst performing RF extrapolations (RM02).

The results from RF generally have high  $R^2$  ( $> 0.75$ ) and low RMSEs ( $> 0.06 \text{ m}^3 \text{m}^{-3}$ ), indicating the capability of a data-driven method to accurately estimate RZSM. They are comparable, or may even be better than those from H1D simulations, which further adds weight to the utility of the RF model applied. Differences in accuracy between RF interpolation and extrapolation could be related to the impact of the training samples used to build each respective RF model. Higher accuracy for RF interpolations have resulted from inclusion of most, if not all, of the possible RZSM conditions within the Raam catchment using the randomly selected training set. This may not be the case for the RF extrapolation model trained, which consequently contributed to lower accuracy in the validation set. The 'past' data used to build the RF extrapolation model may exclude some of the meteorological or soil moisture conditions possible in the Raam catchment. Therefore, 'future' soil moisture conditions that are not represented in the training set are 'unseen' or 'foreign' values to the RF model, and are more likely to be poorly estimated. Poor extrapolation of values outside the training set is a known drawback of RF and other similar ML techniques (e.g. Hengl et al., 2018). This can be resolved by inclusion of the full suite of soil moisture conditions and corresponding covariates in the training dataset. However, this may not be always possible from in situ measurements since not all soil moisture conditions are encountered in the field within a short time span ( $< 5$  years). Remote sensing is seen as an additional source of SSM or RZSM information provided that the spatio-temporal resolutions from satellite images matches the intended scale of study. Another potential complementary dataset are those simulated by process-based models, especially for extreme meteorological conditions that are not encountered during field measurements. However, process-based model outputs should also demonstrate acceptable to high accuracy levels in order to be used as inputs.

For the H1D simulations, data assimilation (DA) by direct insertion (DI) of in situ measurements improved the root zone soil moisture estimates. H1D simulations at the stations with the highest and lowest RMSE's (fig.6.10) both show improvement after DA (See Appendix,

fig.8.4 for the complete results). Simulations based on soil hydraulic parameters from BOFEK generally underestimated RZSM values for all the study sites. A substantial increase in model accuracy is obtained after optimizing the hydraulic parameters. Further improvement in the simulation accuracy were obtained after applying DA. Using the DI approach, model estimates are pushed towards the observations. However, for some stations, a large spike or drop in RZSM estimates are observed immediately after data was assimilated into the simulation. The model reverts back to the original state quite quickly, which could either imply suboptimal DA sampling interval or that the model physics and/or parameters may not be completely adequate in explaining the measurements. Such effects can be mitigated by applying other types of DA (e.g. Ensemble Kalman Filter), which allow continuous estimation of model uncertainties. However, despite some limitations of the DI method selected, the primary goal was to demonstrate the improvement in model accuracy, and therefore, other (more complex) DA methods were not further pursued.

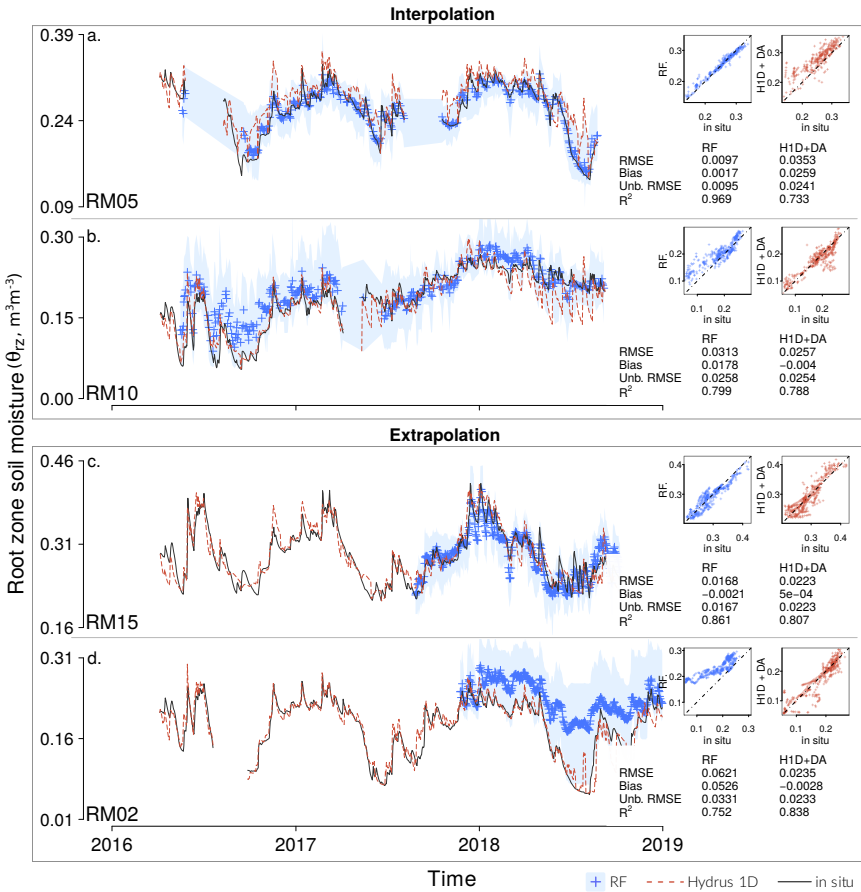
### 6.3.3 Model residuals for extreme soil moisture conditions

Residuals of the model estimates against in situ measurements from all stations in the Raam were further assessed to compare the model performance for extreme RZSM conditions. Accurate RZSM estimates of extremes conditions are vital in understanding the environmental impacts of climate or extreme climatology. In contrast to a single overall metric provided by RMSE or  $R^2$ , residuals allow investigation of specific RZSM values that are poorly estimated using the two methods applied. The results from the residual analysis generally reflect the accuracy obtained for RF (both inter- and extrapolation) and H1D, as described in the previous section. The range of residual values are smaller for RF interpolation and higher for RF extrapolation (fig.6.11c and h). However, based on the residuals, less accurate estimates are found towards drier and wetter soil moisture values from both RF and H1D.

The variability in the residuals for extreme conditions representing the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles (equivalent to  $\leq 0.08 \text{ m}^3 \text{ m}^{-3}$  and  $\geq 0.38 \text{ m}^3 \text{ m}^{-3}$ ) from the total dataset distribution are given as boxplots in fig.6.11 (*d, e, i, j*). For the two RF models, extreme dry conditions tend to be overestimated while extreme wet conditions tend to be underestimated, based on larger than zero residuals for the former and smaller than zero residuals in the latter. The degree of over- or underestimation is larger for RF extrapolation than RF interpolation. Furthermore, H1D simulations have smaller residuals than RF for extreme dry conditions but have worse estimates for extreme wet conditions.

Since extreme conditions represent only a small proportion of the total dataset, the probability of being excluded from the bootstrap samples used for building the regression trees is higher than other frequently encountered soil moisture values. This may have resulted to poor learning of the RF model, which is clearly demonstrated in the large residuals for RF extrapolations of extreme dry conditions. RF extrapolations, with a median of  $0.05 \text{ m}^3 \text{ m}^{-3}$ , mostly overestimated extremely dry conditions, and are worse than those from H1D simulations which had a median close to zero.

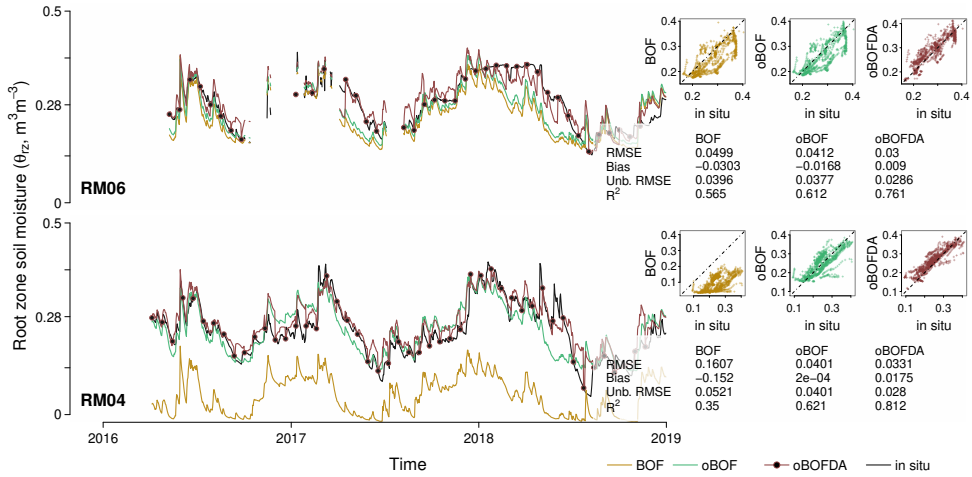
Aside from the impact of the frequency of extreme conditions to the bootstrap samples, large



**Figure 6.9:** Time series plots of RF estimates (blue +) for stations with the lowest (a and c) and highest (b and d) RMSEs. The prediction intervals (PI) within the 95<sup>th</sup> percentiles, shown as blue bands, are based on quantile regression forest. The results of H1D simulations with Data Assimilation are plotted as brown dotted lines. A scatterplot on the upper right compares RF vs in situ values. Accuracy metrics are given below each scatterplot.

residuals obtained for RF extrapolations and H1D (fig.6.11 d, e, i, j) may be related to the covariates used in the former and the type of flow model applied in the latter. Since only a pore-flow model was applied for simulation root zone soil moisture using H1D, the impact of preferential flow was excluded in the analysis. Preferential flow paths generated by biotic (plant roots or animals) activity are likely to be present at the study sites. However, additional model parameters for incorporating preferential flow are not readily available for the study sites and would require separate investigation. Arguably, pore-flow models are still the widely implemented for most applications of process-based models. Migration to a framework that routinely incorporates preferential flow might be necessary for modelling at spatial-temporal



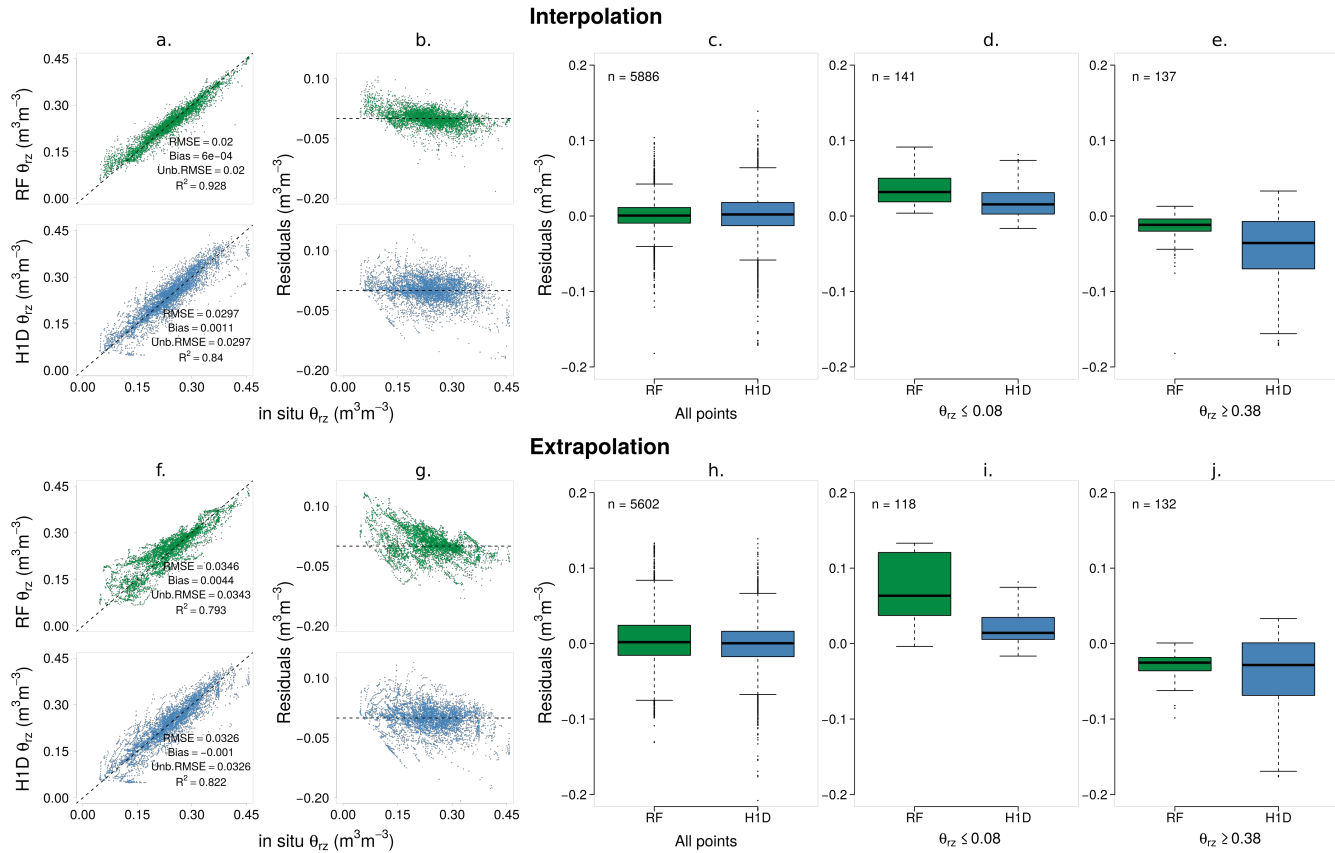


**Figure 6.10:** Hydrus-1D simulations at the stations with the lowest (top) and highest (bottom) RMSE’s using soil hydraulic parameters from BOFEK2012 (yellow). The results from the optimized soil hydraulic parameters (green) as well as those from data assimilation (DA) via direct insertion of in situ measurements (red) are also plotted for comparison. Black dots represent data assimilated at every 20-day sampling interval. Accuracy metrics for each are provided below the scatterplots (bottomright)

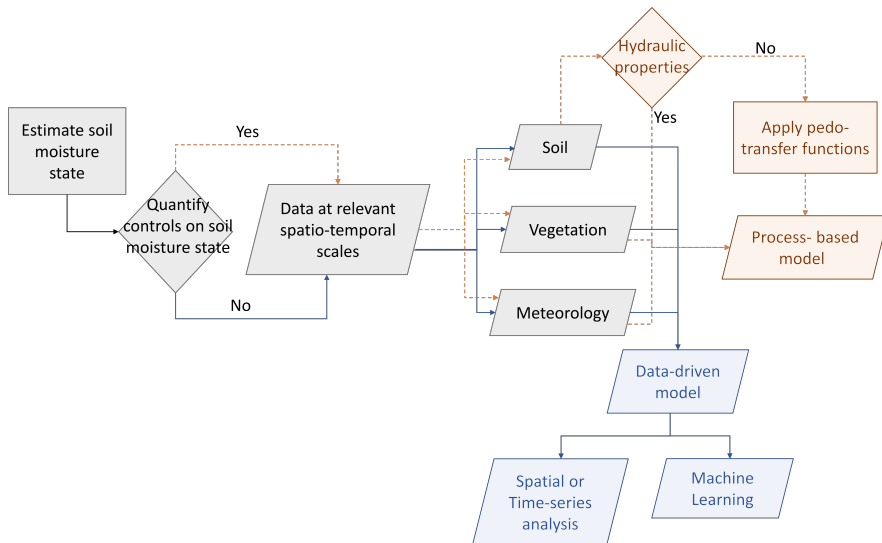
scales where its impact are substantial. Similarly, covariates used in RF also dominantly reflect processes that are important for simulating pore-flow. As mentioned in section 6.2.2, they were chosen based on the knowledge that they are inputs for the process-based model. Underestimation of the two RF models for extreme wet conditions indicates that the covariates selected may have been insufficient to achieve higher accuracy for the said conditions. Addition of covariates which directly represent or indicate the likelihood of occurrence of preferential flow paths in the soil could potentially be beneficial for the RF model. However, deriving such covariates with suitable spatio-temporal resolutions remained elusive.

### 6.3.4 Contextualizing utility of data-driven methods for RZSM estimation

Comparison for the results from RF and H1D show that both methods are equally able to accurately estimate RZSM state, although they operate differently. On the one hand, process-based models determine the rate of water movement along the soil profile which always require soil hydraulic properties. On the other hand, ML methods such as RF performs focus on patterns than allow hierarchical splitting of the dataset using suitable covariates. For both methods, techniques are available in order to optimize and improve naively implemented models that can elevate accuracy to acceptable levels. The question of utility for different scenarios or applications therefore arises. In other words, what are the advantages/disadvantages of one over the other, and how does this affect model selection for a certain application? For RF, one of the advantages of a data-driven method is its ability to create a single model that



**Figure 6.11:** Root zone soil moisture predictions from RF and H1D. *a&f* Scatterplots of predicted vs. in situ root zone soil moisture. *b&g* Residual scatter plots based on *a&f*. Boxplot showing the distribution of residuals for *c&h* all soil moisture conditions, *d&i* extreme dry, and *e&j* extreme wet conditions. Extreme conditions are based on the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles ( $\leq 0.12 \text{ m}^3 \text{ m}^{-3}$  and  $\geq 0.38 \text{ m}^3 \text{ m}^{-3}$ ) of the total dataset distribution, respectively



**Figure 6.12:** A flowchart for model selection between data-driven or process based methods.

will fit very large datasets without any assumption on the system dynamics. An RF approach may be attractive for areas with limited information on soil hydraulic properties because it can be applied using easily obtained meteorological and satellite-derived variables. Process-based models maybe applied over large areas in a spatially distributed manner but they need to explicitly account for heterogeneities in soil properties by modifying the hydraulic parameters and/or the type of flow mechanisms expected for different parts of a study area. One common supplementary analysis for process-based models is to apply pedo-transfer functions to estimate soil hydraulic parameters from commonly measured soil properties such as texture and organic matter content. RF, however, circumvents the need to carry out this intermediate and supplementary step by not requiring prior assumptions on the system dynamics, thus not anchoring its estimates on soil hydraulic properties. For this study, another difference between the two models applied is the use of SSM values in developing the RF model. Although theoretically, RF could be carried out excluding SSM, the results from the variable importance list show that SSM is relevant for achieving good model performance. Satellite-derived SSM is a good alternative for the RF model in cases where in situ measurements are insufficient. Lastly, high RF accuracy for prediction of soil moisture values further opens opportunities for filling data gaps in highly non-linear time-series datasets.

The increasing amount of available soil moisture measurements globally could be a resource for expanding the application of data-driven methods in soil hydrology. Similar to what is carried out in this study, creation of a single model from numerous soil moisture networks could potentially allow for operational RZSM prediction or forecasting at different spatio-temporal scales. In situations where the primary goal is to determine the soil moisture state, RF is seen as a good approach as it can be applied based on accessible surficial datasets.

However, it is not capable of determining the dominant processes that control the soil moisture state, although a hint may be provided in the important variables list identified by the RF model. The impact of certain processes on soil moisture state may be better analyzed with using a process-based model. The context in which each method may be the better option is summarized in fig.6.12

## 6.4 Conclusions

In this study, we demonstrated the capabilities of a data-driven method using Random Forest for estimating root zone soil moisture with high accuracy, similar to process-based models. It may be advantageous to apply a Random forest framework for areas with limited information on soil hydraulic properties, and may circumvent the need to apply pedo-transfer functions. Increasing availability of datasets from in situ measurements worldwide, and also from satellites provide opportunities in data-driven methods for large scale studies or operational (water) management. However, the results from the Random forest model does not explicitly elaborate on process controlling soil moisture state and may suffer from poor extrapolation results.

*"One never notices what has been done; one can only see what remains to be done."*

Marie Skłodowska Curie

# 7

## Synthesis



In this PhD project, the potential of radar satellites, particularly Sentinel-1, to support resource efficient agriculture and operational water management was explored. Sentinel-1 continuously collect information that is related to soil moisture, but it has not been clear how to directly exploit the information for agricultural applications. The findings within this PhD project have advanced understanding of the dynamics between surface and sub-surface moisture and, more to the point, provide new and novel insights into how to directly harness the strengths of radar satellite technology for effective water management and soil protection.

This chapter provides a synthesis of the results obtained separately from the previous research chapters. First, a summary of results or highlights from each of the research chapter is presented. These aim to answer the research questions posed in chapter 1. Afterwards, new scientific insights drawn from new methods and results obtained from the research chapters are described. The third section focuses on interpretations and implications of the results for relevant fields. Finally, the last section reflects on the limitations/challenges faced within the research as well as future outlook.

## **7.1 Summary of major findings**

### **7.1.1 How does soil moisture vary over depth and over which conditions are surface and subsurface soil moisture coupled?**

Investigation of vertical soil moisture variability in agricultural field resulted in identification of coupled and decoupled conditions in Chapter 2. The former indicates conditions when the surface moisture is a good representation of subsurface conditions, while the latter does not. Previous studies that investigated spatio-temporal soil moisture variability in similar landscapes observed decoupled conditions primarily during a drying phase, which was attributed to variable surface evaporation rates (e.g. Hirschi et al., 2014). However, a new insight from our research is that decoupled conditions are not necessarily unique to dry periods since they were also observed during wet periods as reported in Ch. 2. This has implications for understanding the controls on soil moisture dynamics. For instance, as inferred from the results in Ch. 2 soil heterogeneity and the presence of macropores at the sensor location contributed to decoupling during wet conditions. The statistical methods applied to analyze (de)coupling is considered robust as no prior assumptions are made regarding the relation between surface and subsurface soil moisture values. Furthermore, the distributed lag non-linear model (DLNM, Gasparrini et al., 2010) applied is capable of considering all lag effects in comparison to singular lags in correlograms. In addition, it is flexible enough to be extended and include other environmental variables that may have lagged relations with subsurface soil moisture. Knowledge of when (de)coupling between surface and subsurface occurs has implications for remote sensing and data assimilation methods, especially for uncertainties related to the use of surface soil moisture to obtain integrated soil moisture values. In addition, this knowledge can be used to explore surface and subsurface soil moisture variability to identify processes that may be excluded or not adequately represented in process-based hydrological models.

### 7.1.2 How does subsurface soil moisture in agricultural areas vary in space and time?

In Chapter 3, meteorological conditions over a cultivated field are considered uniform and their direct influence on subsurface soil moisture dynamics is much less than their influence on surface soil moisture. Therefore, in a cultivated field, the effects of soil heterogeneity and root water uptake on subsurface soil moisture at 40 cm depth were analyzed over different seasons. Although Chapter 3 focused on a single field, the results obtained have implications on the subsurface variability expected in other similar landscapes.

During winter when the soil was bare, variability in measured subsurface soil moisture was attributed to variability in soil structure alone since only a single soil textural composition is present in the field. Over vegetated periods, the variability in subsurface soil moisture was found to increase substantially due to vegetation activity. Comparison of the results under the bare and vegetated conditions presented show that subsurface soil moisture variability in the latter may be up to nine times greater compared to the former. Aside from taking up water, root growth and distribution can alter the structure of the soil over a growing season which has additional direct impacts on soil moisture variability.

Crop types and land management associated with each crop type can further manipulate the soil moisture dynamics. Within the same field planted with two crops, distinct subsurface soil moisture patterns were observed under each crop. In particular, soils under the potato crops had consistently higher subsurface soil moisture measurements compared to the relatively lower soil moisture measured under corn, leading us to the conclusion that potato cultivation results in increased or longer soil water retention than cultivation of corn.

During the 2018 drought, the variability in subsurface soil moisture due to root water uptake was proportional to atmospheric demands. In other words, there was higher variability in subsurface soil moisture when the potential evapotranspiration was higher. However, this is only true for conditions with sufficient subsurface soil moisture to accommodate atmospheric demands. Once subsurface soil moisture is depleted, the amount and variability in root water uptake decreased, even during days when the potential evapotranspiration remained high. Since measurements were taken only at a single depth, it was inferred that vegetation sought other (deeper) layers with sufficient water supply. These results contribute further measurement-based evidence for the adaptive root water uptake strategies of vegetation, especially during dry periods.

The soil hydraulic parameters estimated using inverse modeling demonstrated significant differences in saturated water contents. The results highlight that the pore-flow regime applied produced reasonably accurate parameters only for bare conditions. It was challenging to estimate soil hydraulic parameters during vegetated periods not only using a pore-flow model, but with other flow regimes as well. These results suggest that the current information on subsurface characteristics during vegetated periods or that current modeling frameworks are not sufficient to successfully apply inverse modeling for estimation of hydraulic parameters from in situ measurements.

### **7.1.3 During saturated water conditions, how can satellite-derived surface soil moisture be used for management of agricultural fields?**

In temperate regions, tillage activities at the start of the growing season typically coincide with saturated soil moisture conditions. Even though water supply is not a concern, field trafficability can be a serious problem, not only in relation to mobility of heavy machinery but also for aggravating the levels of subsurface soil compaction. Results presented in chapter 5 indicate that satellite-derived soil moisture information has great potential for use in monitoring the trafficability status in agricultural fields. This clearly demonstrates a direct application of Sentinel-1 surface soil moisture data since coupled conditions were found for the months of March, April and May when much is conducted.

The results further highlight the role of land management in different agricultural fields. For example, cultivated fields are likely to suffer more from problems associated with trafficability than grass fields because tillage activities in the former tend to 'reset' soil strength annually or seasonally. Different stages of vegetation growth were observed to impact the soil's penetration resistance measurements, especially in cultivated fields. The value and variability in penetration resistance increased with crop growth. This suggests the importance of vegetation in improving the ability of soils to support traffic, similar to the results of several studies as reviewed by Wieder and Shoop (2017)

### **7.1.4 For drier periods, does vegetation backscatter from Sentinel-1 have the potential to reflect root zone soil moisture?**

Results presented in chapter 4 focus on exploring the potential for vegetation backscatter to reflect root zone soil moisture over a three-year study period. Estimation of soil moisture from radar satellites has largely focused on the soil component of the total backscatter. However, one of the main limitations of this approach is that only the upper surface soil layer is measured. During vegetated periods, the impact of the vegetation component on the total backscatter is non-negligible especially for broad-leaf crops. This affects the retrieval of the soil component and, consequently soil moisture estimation. Interestingly, the vegetation component is also sensitive to vegetation water content much like the soil component is sensitive to soil moisture. What is notable about this is that the vegetation water content is highly dependent on the available water in the soil. Theoretically therefore, the vegetation backscatter could offer insights into the amount of root zone soil moisture since studies have shown that at pre-dawn soil and leaf-water potential are in equilibrium (e.g Sellin, 1999).

In chapter 4, vegetation backscatter showed stronger correlations with root zone soil moisture during a dry period in 2018 when compared to the soil backscatter. In contrast, stronger correlations with soil backscatter were observed for 2016 and 2017 which were characterized as years with average meteorological conditions. We hypothesize that the correlations obtained were influenced by surface-root zone soil moisture (de)coupling. Due to the dry conditions, surface conditions were generally decoupled from root zone soil moisture in 2018 but were coupled in 2017 and 2016. It is suggested that the vegetation backscatter has the potential to be another source of root zone soil moisture information during extremely dry or water-limited



conditions. It may overcome the limitations of radar-derived surface soil moisture which is unable to represent the root zone during such extreme conditions.

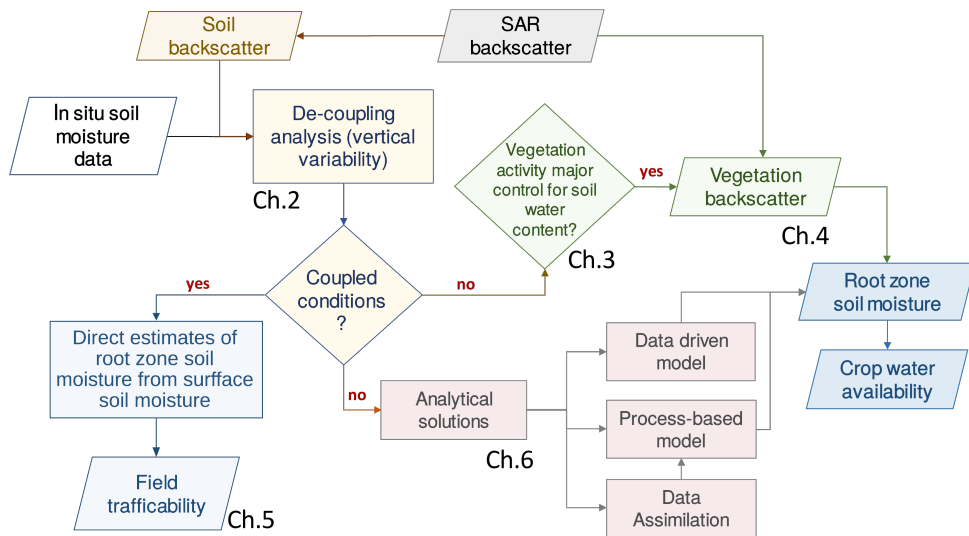
### **7.1.5 Can a data-driven approach based on surface conditions accurately estimate root zone soil moisture?**

Data-driven methods provide opportunities to develop quantitative models without prior assumptions about the dynamics of the system under investigation. Results contained in Chapter 6 show that a purely data-driven machine learning (ML) method, in this case using Random Forests (RF), can achieve a level of accuracy comparable to process-based models for estimating root zone soil moisture. Furthermore, RF may be advantageous for areas with limited information on soil hydraulic properties. Available remote sensing datasets for relevant covariates, such as surface soil moisture from Sentinel-1, would be valuable in developing ML-based models for root zone soil moisture. The proliferation of soil moisture networks also provides an opportunity to explore the capabilities of ML for estimating root zone soil moisture over much larger scales. While ML and process - based models operate on different fundamental frameworks, existing knowledge from both can complement future endeavors. For instance, meteorological variables used as inputs in physical models can be used as the covariates in data driven methods. Processes that affect the current soil moisture state should be represented in one or more of the covariates used for ML methods. As shown in chapter 6, lower accuracy was obtained for extreme dry and wet conditions, which was related to the choice of covariates in the RF model.

## **7.2 General conclusions**

The underlying hypothesis of this project was that radar satellites can be used for estimation of agricultural water requirements since their sensitivity to surface soil moisture can be exploited to gain further insights into soil moisture at deeper soil layers (e.g. topsoil or root zone).

The main goal was discovery of how satellite- derived surface soil moisture information can be used in agricultural fields and operational water management. The focus was on analysis of methodologies to maximize the use of radar satellites for estimation of root zone soil moisture conditions while also addressing the limitations. This novel approach considered how well radar satellites represent subsurface conditions for different periods of a growing season instead of focusing solely on existing approaches such as data assimilation into process-based hydrological models. Emphasis was also given to subsurface soil moisture dynamics in order to advance understanding of its variability and controls over agricultural fields as this is closely linked with the ability and effectiveness of radar satellites to represent subsurface conditions (Ch. 3). An important contribution of this thesis is the quantification of (de)coupled soil moisture values which is key to the applications for satellite-derived surface soil moisture information (Ch. 2). For cases when surface - subsurface coupling predominates (i.e. surface soil moisture reflects subsurface conditions), it was shown that satellite-derived surface soil moisture can be used directly as representative of subsurface/rootzone conditions, for instance in monitoring trafficability to reduce problems with heavy machinery and compaction (Ch.5). However,



**Figure 7.1:** Flowchart showing how the findings of different sections of this thesis could aid in operational water management agricultural fields. The chapter related to a section of the flowchart is indicated.

decoupled conditions require more complex or sequential analysis for root zone soil moisture estimation. In that regard, another important contribution of this thesis project is discovery or validation of two innovative approaches for addressing the challenges of decoupled conditions. One innovative approach to using satellite data for decoupled conditions discovered in this research is to utilize the vegetation component of the backscatter instead of focusing only on the soil backscatter, especially for water-limited periods (Ch.3 and 4). Another approach to estimating rootzone moisture during decoupled conditions that still uses the satellite surface soil information is to apply data-driven methods which do not require assumptions about system dynamics (Ch.6.) Both of these approaches have advantages, particularly for data poor regions.

With respect to the main hypothesis, this thesis research has verified that agricultural water requirements can indeed be estimated from radar satellites in multiple ways and has developed/identified several approaches for doing so that overcome limitations or questions that existed before this research was conducted. These results will allow incorporation of radar-derived surface soil moisture into an adaptive methodological framework for agriculture and operational water management that reflects close to real-time conditions, instead of relying on a single method (e.g. data assimilation into process-based model) that will have varying accuracy across the full range of soil moisture encountered.

## 7.3 Discussion

### 7.3.1 Scientific Advancements/Implications

Multiple methods have been explored and developed in this research project for estimating root zone soil moisture from Sentinel-1, primarily for agricultural areas. Prior to this research, satellite-derived surface soil moisture information has been commonly applied using data assimilation into process-based hydrological models (e.g. Houser et al., 1998; Draper et al., 2012; Peziz et al., 2019b). Other less commonly applied data-driven methods have been developed, primarily using time series analysis (e.g. Soil Water Index, Wagner et al., 1999). This thesis differs from previous studies in that the direct applicability of satellite-derived soil moisture was sought by investigating the spatio-temporal soil moisture dynamics. Instrumental to the conclusions made in this thesis was the identification and quantification of soil moisture coupling. Although the surface-subsurface coupling has been described in literature (e.g. Capehart and Carlson, 1997, also in relation to the utility of remote sensing products), this thesis further demonstrates its importance for gaining insights on factors controlling soil moisture or developing methods that extends the direct utility of soil moisture information derived from radar satellites for agricultural applications.

This thesis further highlights that the vegetation backscatter is an untapped resource for directly estimating root zone soil moisture during extremely dry conditions. It complements the lack of sensitivity of the soil backscatter to root zone soil moisture during such extreme conditions. Utilization of the vegetation backscatter can result in a more simplified way of estimating root zone soil moisture during for water-limited periods that require lesser data and fewer model parameters, in comparison to using data assimilation with process-based models. This can further eliminate the need to address the uncertainties in surface soil moisture retrieval during vegetated periods. The vegetation backscatter can be beneficial in assessing vegetation water-stress levels, crucial during droughts, that will relevant for giving out timely management or disaster-risk advisories over agricultural and forested areas. Currently, studies that focus on radar satellite-derived vegetation characteristics to determine root zone soil moisture dynamics are still limited. Future studies are needed to ascertain the utility of vegetation backscatter in providing direct and accurate root zone soil moisture, which could make radar satellites such as Sentinel-1 (S1) indispensable water management tools for different agro-hydrological conditions. In relation to data assimilation, perhaps there is potential for the vegetation backscatter to be assimilated into process-based models, similar to the recent investigation of Kumar et al. (2020) who assimilated the vegetation optical depth derived from passive microwave radiometry in land surface models for improved estimations of water and carbon fluxes or for joint assimilation with surface soil moisture.

### 7.3.2 Implications for satellite-aided operational water management in agricultural fields

One of the primary motivations for investigating soil moisture variability in this research was to assess the potential of using satellite-derived soil moisture data, particularly S1, for operational water management. This involves translating data from sensors into useful information or indi-

cators for advisory or decision-making. The findings of this research collectively demonstrate a framework for how S1 can aid operational water management, as summarized in fig.7.1.

The results presented in Ch.5 demonstrate that estimation of field trafficability could be easily achieved from S1 with minimal analysis. This would facilitate timely advisory for farmers of when to initiate tillage every growing season. Arguably, the onset of tillage is still dependent on the farmers knowledge and assessment of local conditions based mostly on prevailing rainfall events. However, their assessments alone have individual subjectivity that may risk exacerbating the existing subsoil compaction levels in a field. Nevertheless, their years of experience should also not be discounted since they are well accustomed to conditions in their fields they may also be considered experts. Therefore, field trafficability advisories based on S1 could supplement their initial assessment of the suitability of the soil for tillage. This is valuable in order to balance a farmer's productivity and with further mitigation of present subsoil compaction levels.

Consequences of poor farming practice during saturated conditions are, arguably, given less attention than those during dry periods. Perhaps this is due to the direct and more noticeable impacts on crop yield during dry periods as compared to the invisibility of subsoil compaction due to poor field traffic management. Although soil compaction has been extensively studied in the past, indicators for practical applications that support and facilitate mitigation of soil compaction are not extensively available. This is important since the ubiquity of compacted soils in agricultural fields have resulted in continued adverse effects in soil functions (Keller and Lamandé, 2010) and may have consequences not only for the estimation of root zone soil moisture, but for soil productivity as well. In the Netherlands, for instance, Brus and Van Den Akker (2018) estimates that 43% of the areas estimated to be at risk of subsoil compaction, excluding peat areas, are already overcompacted. In a review paper a few years ago, Nawaz et al. (2013) estimated that about 68 million ha. of the soils worldwide are affected by soil compaction from vehicular traffic. Although an outlook was not further elaborated in both studies, the demands of increasing global population will put further pressure on agricultural production. Therefore, these estimated values will increase in the near future.

Root zone soil moisture is a valuable information for farmers and water managers further into the growing season to determine irrigation demands. Most of the indicators needed by farmers to execute the necessary water management for their field (e.g. soil water storage, available crop water) are based on estimates of root zone soil moisture. Water managers use root zone soil moisture information in deciding the appropriate levels of weirs along canals to meet water demands in agricultural fields. The occurrence of extreme meteorological conditions, such as the 2018 European summer drought, may also coincide with the growing season or crop maturity. With climate change, such extreme conditions are predicted to be more common in the future (Seneviratne et al., 2012). Therefore, maximizing crop water resources will remain essential for farmers and water managers to meet crop demands now and in the future.

Results in this thesis, specifically in Ch.3-5, agree with previous studies that accurate estimations of root zone soil moisture from S1 is not trivial because of both the increased subsurface soil moisture variability and influence of vegetation backscatter for vegetated periods (e.g. Bindlish and Barros, 2001; Gherboudj et al., 2011; Kornelsen and Coulibaly, 2013). However,

instead of focusing on improved parameter estimation or optimization of currently applied data assimilation and process-based modelling, alternative methods that circumvent the limitations in these existing methods are presented in this thesis. By assessing the direct utility of S1 for root zone soil moisture estimation based on coupling between the surface and the subsurface, additional options are provided to soil hydrologists and water managers in cases when the traditional process-based models prove to be inadequate. This may occur because currently, process-based modelling frameworks do not systematically incorporate and embed the dynamic nature of the subsurface soil conditions, which is especially true for agricultural areas. Furthermore, having multiple methods for root zone soil moisture estimation is beneficial for areas with data availability concerns, and not only for areas with sufficient data such as the Netherlands. In such areas, operational water management could still be carried out by capitalizing on data-driven methods using freely available satellite or meteorological datasets.

Integrating the methods developed in this thesis with new technology that facilitate end-user services could definitely aid farmers in optimizing their activities in the absence of direct field measurements. For instance, the probabilistic measure of trafficability estimated in Ch.5 or root zone soil moisture estimation in Ch.4 could be incorporated in cloud-based services that allow near-real time assessment monitoring.

### **7.3.3 Implications for measuring and modeling spatio-temporal soil moisture dynamics**

The effects of land management and biological activity on soil moisture dynamics in agricultural areas have been extensively examined using process-based models in literature. Although the field of hydrological modeling is mature, the lack of consistent modeling results in Ch.3 highlight aspects that need further investigation in order to realistically represent spatio-temporal changes in (agricultural) field conditions within computer simulations. This is important since the results of process-based models are increasingly being relied on for operational management, as well as policy and decision-making.

The results in this thesis show that vegetation act as agents of change that make agricultural areas dynamic environments even at short time scales spanning months or seasons. However, the exact role of roots and root development on the physical properties of the rhizosphere is still a matter of debate. For instance, some studies suggest that roots may compact soils around them (e.g. Dexter, 1987; Koebernick et al., 2019), similar to the increased penetration resistance with crop growth in Ch.5. Others found that roots and plant growth improved porosity and hydraulic conductivity in the rhizosphere (e.g. Helliwell et al., 2017; Whalley et al., 2005). Similarly, improved saturated water contents from bare to vegetated periods were estimated in Ch.3, implying improved porosity. Therefore, more research is needed to ascertain the role roots play in modifying soil physical properties in order to reconcile contrasting results in literature, such as those presented above. One example is the recent study of Lucas et al. (2019) who, through microscopic investigations, demonstrated that root interactions with existing soil pore structure determine the subsequent effect of root growth on soil structure. In other words, they showed that roots can compact the surrounding soil

depending on the structures they encounter. Furthermore, by demonstrating that existing soil conditions influence plant-soil interactions, Lucas et al. (2019) hint at the importance of feedback loops in their results. In relation to the modeling results in this thesis, dynamic soil hydraulic properties estimated for the cultivated field investigated in Ch.3 are also likely influenced by feedback mechanisms. Current modeling frameworks to simulate water content in the unsaturated zone over longer time periods (i.e. years) therefore need to include provisions to systematically integrate known feedback mechanisms. This is important since some of the changes in soil hydraulic function due to biological feedbacks could be irreversible, as demonstrated by Robinson et al. (2016, 2019). Robinson et al. (2019) further stressed that incorporating all the major feedbacks between soil properties and soil-biological-atmospheric processes, from land use and climate drivers of change, is crucial to better model ecosystem dynamics.

Currently, mathematical models for different types of flow regimes are available in modeling platforms despite lacking feedback mechanisms. They can potentially allow for dynamic soil hydraulic properties in a stop-and-go framework, but may not always be easy and straightforward to implement. The findings in Ch.3 support previous studies which demonstrate improved modeling results when dynamic soil hydraulic parameters are applied (e.g. Alletto and Coquet, 2009; Schwen et al., 2011b). Because of seasonal changes in agricultural fields, dynamic soil properties is, arguably, the rule rather than the exception in these types of landscapes. It is clear that time-invariant soil hydraulic properties should be substituted with more dynamic soil functions in soil hydrologic model structures and should not be seen as a static occurrence.

At present, pore-flow regimes are still commonly used for most practical applications in agricultural water management. The systematic use of other types of flow regimes is, perhaps, hindered by limited knowledge of when to transition from a pore-flow dominated or a preferential-flow dominated system. For instance, it was demonstrated in Ch.2 that decoupling could be related to the presence of hydraulically effective macropores at the measurement location. However, the extent at which hydraulically effective macropores persist in the soil over a growing season, whether they dominate or are only sparsely distributed in the subsurface, is difficult to ascertain. This information is essential as it can assist modelers in determining when or during which conditions would a particular type of flow regime more appropriately describe a system under investigation. Not only does this strengthen the need to incorporate feedback mechanisms that are deemed precursors of dynamic soil properties, it also highlights the limited availability of tools to estimate effective hydraulic parameters suitable for preferential flow models.

Despite the growing consensus about its importance to soil functions and hydraulic properties (Rabot et al., 2018), in situ characterization of the temporal changes in soil structures using sensors and measuring devices at relevant modeling scales are difficult to obtain. Furthermore, available datasets of hydraulic parameters of macroporous soils based field or laboratory techniques are far less common than for matrix or pore flow. At present, hydraulic properties of macroporous soils may be estimated through inverse modeling solutions based on measured soil moisture data (Haws et al., 2005). Characterization of 3D soil structure from soil sam-

ples have been carried out using SEM or X-ray CT at the microscopic scale (e.g. Luo et al., 2008; Sammartino et al., 2012), fluorescence imaging at millimeter scales (e.g. Haas et al., 2020), and using dye traces in the field (e.g. Alaoui and Goetz, 2008; Cey and Rudolph, 2009). Future efforts are needed to translate these observations into effective hydraulic parameters to facilitate more practical applications of preferential flow modelling studies. In addition, Chandrasekhar et al. (2018) noted that applications of temporally variable soil hydraulic parameters in models is not limited by the lack of theoretical framework but because of the lack of adequate and relevant soil structural and hydrologic data.

Overall, soil moisture estimation in agricultural fields requires dynamic soil functions and hydraulic properties in order to accurately represent subsurface processes. The cyclical changes in the soil due to vegetation growth and land management adds additional dimensions that make subsurface soil moisture equally as variable as surface soil moisture. Mapping and monitoring of subsurface soil moisture variability may, however, prove to be more challenging because of the complexity of data collection or sampling, and quantifying plant responses and effects to soil physical properties. The growing need to provide accurate soil moisture information at local scales and higher temporal frequencies in agricultural landscapes may provide the necessary push to adapt a dynamic modeling framework in future scientific efforts.

One avenue that could be explored to improve the current structure of process-based models is the vast, and continuously increasing amount of global in situ soil moisture datasets that have been collected over the years, and some made available through data hosting platforms such as the International Soil Moisture Network (ISMN, Dorigo et al., 2011b). Chapter 6 demonstrates that machine learning (ML) can achieve high accuracy for estimating soil moisture states using available satellite- and measurement-based datasets. The applications of ML, and also deep learning (DL), to hydrology is expanding further with recent advances in the fusion of data-driven and process-based models through a physics-guided machine learning framework (e.g. Karpatne et al., 2017; Read et al., 2019; Zhao et al., 2019). However, the potential of ML or DL methods goes beyond spatio-temporal forecasting and prediction of the environmental variables or parameters. Reichstein et al. (2019) presented several opportunities of DL approaches to hydrological and earth sciences. Among them were in recognition of the particularities of the data, plausibility and interpretability of inferences, uncertainty estimation, and testing against complex physical models. They did not however, explicitly mention the identification or assessment of feedback loops in models, but may have implicitly included it in recognition of particularities of the data. Using available soil moisture datasets, ML and DL are potentially useful tools for exploring consistent patterns and relationships (Steinbach et al., 2001) and well as causality between variables (Pearl, 2009; Runge et al., 2019) that may shed light to important processes and feedback loops that poorly represented in current hydrological modeling frameworks.

#### 7.3.4 Research challenges/limitations and outlook

- The study areas were located in dominantly sandy soils that are homogeneous over the depths investigated. The surface-subsurface coupling behaviour for other soil textures or for structured soil were not investigated. Furthermore, it is challenging to project

and extend the results found in this thesis for other soil textures or soil configurations because of the non-linear behaviour of soil hydraulic properties. Therefore, future research is needed to determine the (de)coupling behaviour for other soil textures;

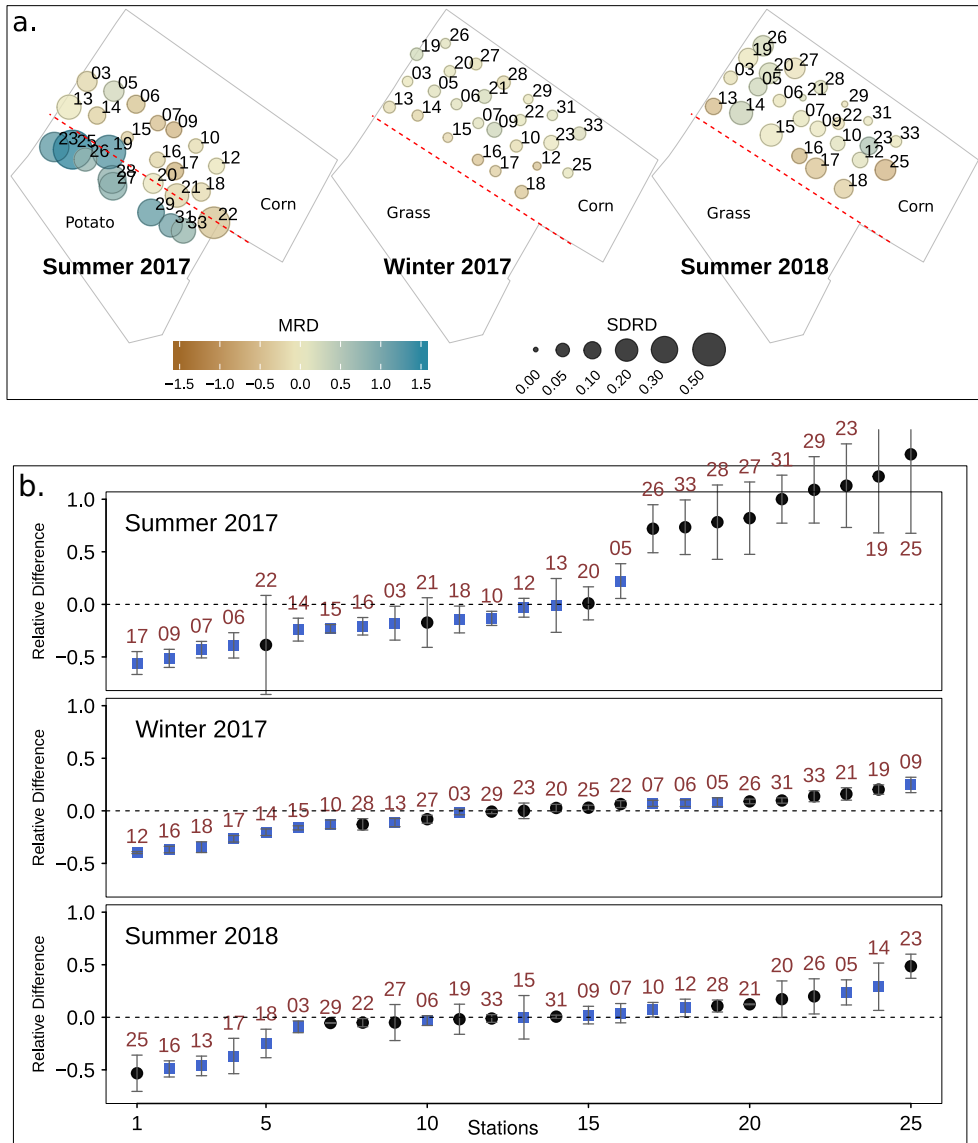
- The study sites are in a temperate region with generally flat terrain. This has consequences for the processes included in the models applied. Agricultural areas in other climatic regions and in mountainous terrains may be further influenced by lateral flows and surface run-off, which are not considered applicable for the areas investigated;
- The areas investigated for vertical soil moisture (de)coupling had low water tables and, because of their occurrence in sandy soils, were not significantly affected by groundwater fluctuations or capillary rise. However, local topography, groundwater levels and texture is foreseen to affect the estimation of (de)coupled soil moisture states. For areas where groundwater is foreseen to be impactful, relevant datasets should be included for the distributed lag nonlinear model applied.
- The results obtained focused on water management in agricultural areas but may potentially find parallel applications for forested regions. For instance, the potential of vegetation backscatter for monitoring root zone in forested areas and their potential use for wildfire assessment. Furthermore, arid areas may be interesting regions to further explore the utility of the vegetation backscatter in representing root zone soil moisture because water-limited conditions are commonly encountered;
- The field measurements gathered in this research focused on the soil but with limited characterization of vegetation traits. Future research should aim to have simultaneous measurements of vegetation and soil characteristics to reduce the assumptions made in investigating soil-plant interactions;
- The assessment of trafficability focused only on changes to soil strength with soil moisture. However, the impact of vehicular characteristics (e.g. tire inflation pressure) on trafficability and vehicle-soil interactions is not explored. The assumption is that the threshold applied for estimation of trafficability is applicable to average tractor weights of 6-8 tons. However, incorporating soil mechanics with soil moisture impacts on soil strength could improve assessment of field trafficability.



# 8

## Appendix

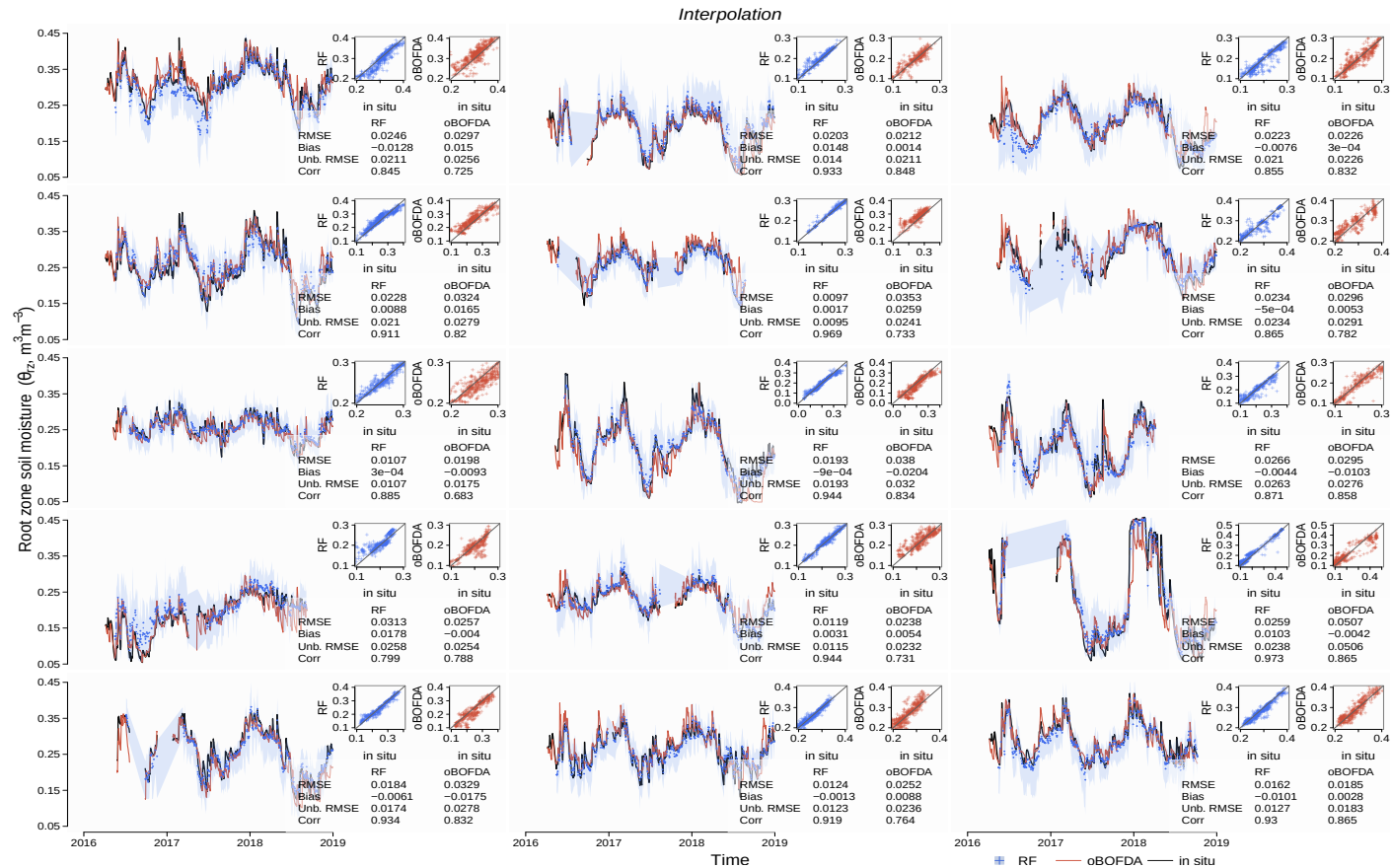




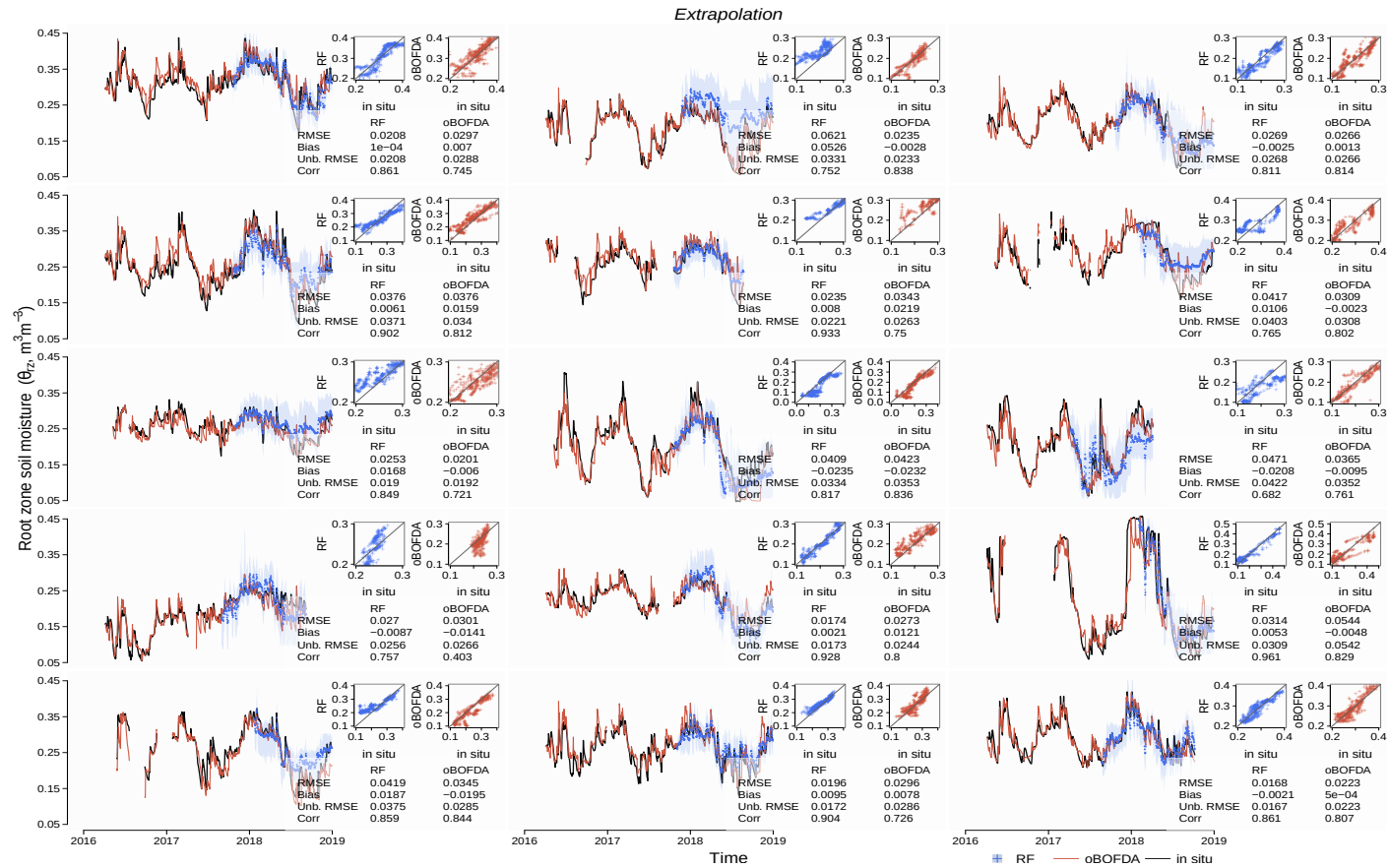
**Figure 8.1:** Plots of relative differences (RD) over different seasons based on temporal stability analysis. a) Spatial plots of median and standard deviations in RD (MRD and SDRD). b) The stations are ranks based on increasing MRD for the different seasons. The dots indicate MRD while the bars indicate SDRD for each station. Blue square dots are stations kept in the same locations while the black circles were changed after 2017 harvest. MRD and SDRD are in  $m^3 m^{-3}$  units.

**Table 8.1:** Soil classes and hydraulic properties based on BOFEK2012 (Wosten et al., 2013). The optimized parameters obtained from inverse modelling with Hydrus-1D are given on the last few columns.

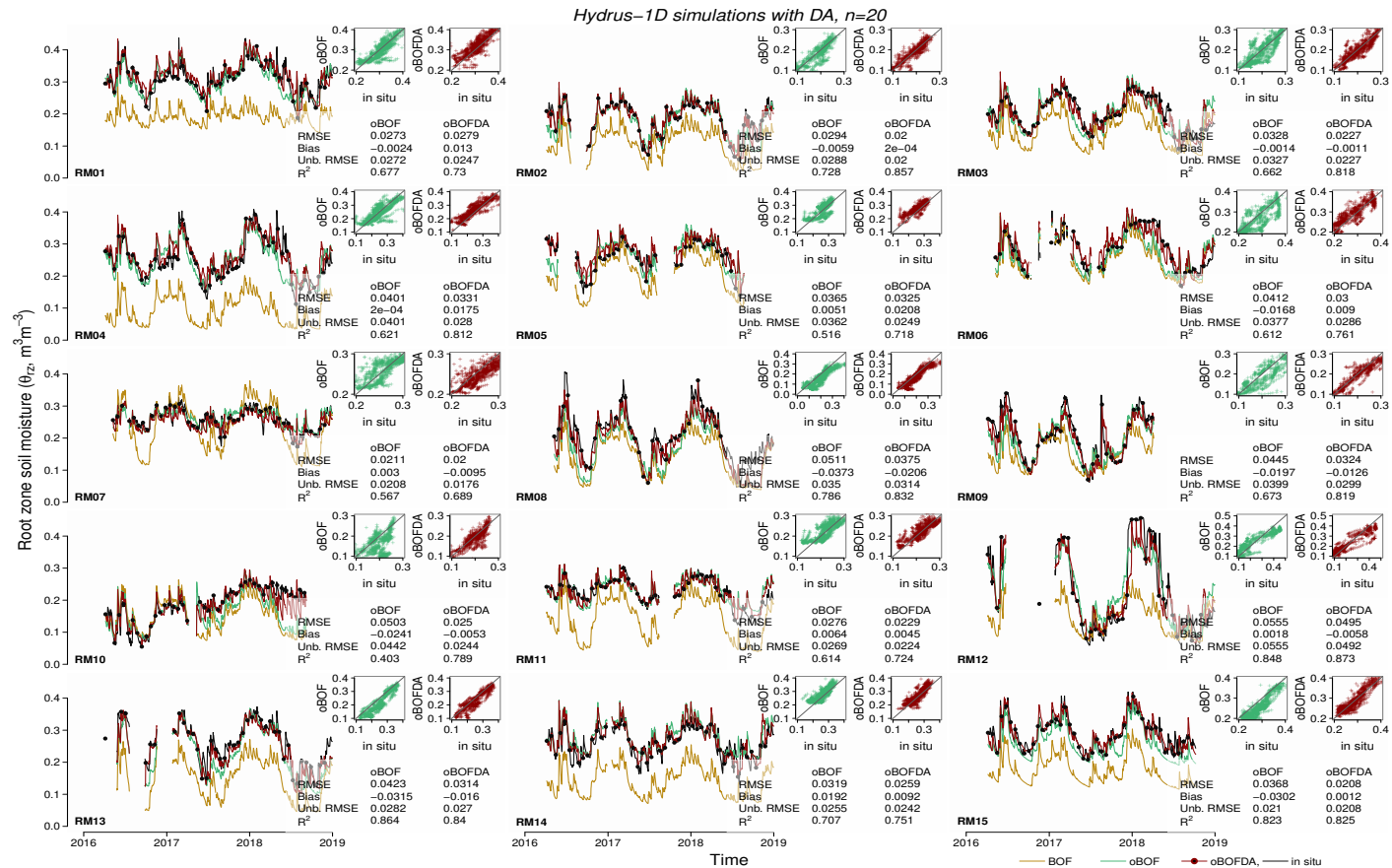
| Name | BOFEK2012 unit code | PAWN code | Upper Depth | Lower Depth | Startingreeks code | BOFEK 2012 |            |          |       |       |        | HYDRUS-1D optimized |            |          |       |         |        |
|------|---------------------|-----------|-------------|-------------|--------------------|------------|------------|----------|-------|-------|--------|---------------------|------------|----------|-------|---------|--------|
|      |                     |           |             |             |                    | $\theta_r$ | $\theta_s$ | $\alpha$ | n     | $K_s$ | l      | $\theta_r$          | $\theta_s$ | $\alpha$ | n     | $K_s$   | l      |
| RM01 | 305                 | 10        | 0           | 25          | B1                 | 0.00       | 0.37       | 0.0208   | 1.646 | 33.34 | 0.571  | 0.077               | 0.545      | 0.009    | 1.459 | 57.463  | 0.163  |
|      |                     |           | 25          | 90          | O1                 | 0.00       | 0.35       | 0.0220   | 2.186 | 99.70 | 0.796  | 0.078               | 0.445      | 0.003    | 1.594 | 2.703   | 0.204  |
| RM02 | 305                 | 10        | 0           | 25          | B1                 | 0.00       | 0.37       | 0.0208   | 1.646 | 33.34 | 0.571  | 0.090               | 0.361      | 0.011    | 1.895 | 114.810 | 0.580  |
|      |                     |           | 25          | 90          | O1                 | 0.00       | 0.35       | 0.0220   | 2.186 | 99.70 | 0.796  | 0.041               | 0.313      | 0.020    | 1.454 | 38.716  | 0.785  |
| RM03 | 304                 | 9         | 0           | 25          | B2                 | 0.00       | 0.43       | 0.0224   | 1.436 | 32.21 | -0.304 | 0.000               | 0.402      | 0.022    | 1.516 | 63.779  | -0.304 |
|      |                     |           | 25          | 60          | O2                 | 0.00       | 0.38       | 0.0182   | 1.870 | 63.90 | 0.911  | 0.055               | 0.500      | 0.016    | 1.690 | 33.119  | 0.911  |
|      |                     |           | 60          | 120         | O1                 | 0.00       | 0.35       | 0.0220   | 2.186 | 99.70 | 0.796  | 0.062               | 0.628      | 0.021    | 1.718 | 22.317  | 0.796  |
| RM04 | 305                 | 10        | 0           | 25          | B1                 | 0.00       | 0.37       | 0.0208   | 1.646 | 33.34 | 0.571  | 0.012               | 0.484      | 0.012    | 1.403 | 99.219  | 0.164  |
|      |                     |           | 25          | 90          | O1                 | 0.00       | 0.35       | 0.0220   | 2.186 | 99.70 | 0.796  | 0.013               | 0.635      | 0.014    | 1.619 | 29.852  | 0.002  |
| RM05 | 311                 | 12        | 0           | 90          | B2                 | 0.00       | 0.43       | 0.0224   | 1.436 | 32.21 | -0.304 | 0.000               | 0.434      | 0.015    | 1.287 | 49.987  | -0.503 |
| RM06 | 409                 | 20        | 0           | 20          | B8                 | 0.00       | 0.40       | 0.0313   | 1.200 | 22.90 | -3.578 | 0.000               | 0.480      | 0.043    | 1.231 | 26.590  | -3.578 |
|      |                     |           | 20          | 40          | O10                | 0.00       | 0.44       | 0.0231   | 1.212 | 25.60 | -2.220 | 0.000               | 0.476      | 0.027    | 1.253 | 55.259  | -2.220 |
|      |                     |           | 40          | 65          | O9                 | 0.00       | 0.41       | 0.0280   | 1.283 | 24.00 | -1.559 | 0.019               | 0.490      | 0.026    | 1.306 | 33.161  | -1.559 |
| RM07 | 317                 | 12        | 0           | 90          | B3                 | 0.00       | 0.45       | 0.0152   | 1.412 | 17.81 | -0.213 | 0.095               | 0.419      | 0.014    | 1.402 | 26.936  | -0.056 |
| RM08 | 304                 | 9         | 0           | 25          | B2                 | 0.00       | 0.43       | 0.0224   | 1.436 | 32.21 | -0.304 | 0.000               | 0.538      | 0.017    | 1.690 | 12.896  | -0.304 |
|      |                     |           | 25          | 60          | O2                 | 0.00       | 0.38       | 0.0182   | 1.870 | 63.90 | 0.911  | 0.117               | 0.452      | 0.011    | 2.299 | 45.388  | 0.911  |
|      |                     |           | 60          | 120         | O1                 | 0.00       | 0.35       | 0.0220   | 2.186 | 99.70 | 0.796  | 0.017               | 0.559      | 0.020    | 1.343 | 77.552  | 0.796  |
| RM09 | 304                 | 9         | 0           | 25          | B2                 | 0.00       | 0.43       | 0.0224   | 1.436 | 32.21 | -0.304 | 0.010               | 0.413      | 0.023    | 1.654 | 88.567  | -0.304 |
|      |                     |           | 25          | 60          | O2                 | 0.00       | 0.38       | 0.0182   | 1.870 | 63.90 | 0.911  | 0.000               | 0.434      | 0.008    | 1.822 | 13.537  | 0.911  |
|      |                     |           | 60          | 120         | O1                 | 0.00       | 0.35       | 0.0220   | 2.186 | 99.70 | 0.796  | 0.065               | 0.499      | 0.022    | 1.472 | 25.717  | 0.796  |
| RM10 | 304                 | 9         | 0           | 25          | B2                 | 0.00       | 0.43       | 0.0224   | 1.436 | 32.21 | -0.304 | 0.047               | 0.453      | 0.024    | 1.455 | 59.433  | -0.304 |
|      |                     |           | 25          | 60          | O2                 | 0.00       | 0.38       | 0.0182   | 1.870 | 63.90 | 0.911  | 0.060               | 0.498      | 0.018    | 1.866 | 41.270  | 0.911  |
|      |                     |           | 60          | 120         | O1                 | 0.00       | 0.35       | 0.0220   | 2.186 | 99.70 | 0.796  | 0.083               | 0.486      | 0.017    | 1.783 | 61.426  | 0.796  |
| RM11 | 304                 | 9         | 0           | 25          | B2                 | 0.00       | 0.43       | 0.0224   | 1.436 | 32.21 | -0.304 | 0.085               | 0.337      | 0.024    | 1.243 | 71.797  | -0.304 |
|      |                     |           | 25          | 60          | O2                 | 0.00       | 0.38       | 0.0182   | 1.870 | 63.90 | 0.911  | 0.051               | 0.300      | 0.016    | 1.216 | 21.441  | 0.911  |
|      |                     |           | 60          | 120         | O1                 | 0.00       | 0.35       | 0.0220   | 2.186 | 99.70 | 0.796  | 0.134               | 0.599      | 0.030    | 1.527 | 127.540 | 0.796  |
| RM12 | 304                 | 9         | 0           | 25          | B2                 | 0.00       | 0.43       | 0.0224   | 1.436 | 32.21 | -0.304 | 0.000               | 0.478      | 0.016    | 1.615 | 80.487  | -0.304 |
|      |                     |           | 25          | 60          | O2                 | 0.00       | 0.38       | 0.0182   | 1.870 | 63.90 | 0.911  | 0.004               | 0.502      | 0.014    | 1.725 | 49.370  | 0.911  |
|      |                     |           | 60          | 120         | O1                 | 0.00       | 0.35       | 0.0220   | 2.186 | 99.70 | 0.796  | 0.069               | 0.567      | 0.030    | 1.588 | 5.319   | 0.796  |
| RM13 | 309                 | 13        | 0           | 30          | B2                 | 0.00       | 0.43       | 0.0224   | 1.436 | 32.21 | -0.304 | 0.003               | 0.483      | 0.003    | 2.527 | 163.940 | -0.304 |
|      |                     |           | 30          | 60          | O2                 | 0.00       | 0.38       | 0.0182   | 1.870 | 63.90 | 0.911  | 0.116               | 0.409      | 0.003    | 1.614 | 154.260 | 0.911  |
|      |                     |           | 60          | 120         | O1                 | 0.00       | 0.35       | 0.0220   | 2.186 | 99.70 | 0.796  | 0.001               | 0.574      | 0.014    | 1.330 | 230.460 | 0.796  |
| RM14 | 312                 | 13        | 0           | 20          | B3                 | 0.00       | 0.45       | 0.0152   | 1.412 | 17.81 | -0.213 | 0.053               | 0.372      | 0.013    | 1.414 | 15.818  | -0.213 |
|      |                     |           | 20          | 55          | O3                 | 0.00       | 0.34       | 0.0265   | 1.543 | 44.60 | -0.333 | 0.084               | 0.426      | 0.028    | 1.277 | 14.439  | -0.333 |
|      |                     |           | 55          | 120         | O2                 | 0.00       | 0.38       | 0.0182   | 1.870 | 63.90 | 0.911  | 0.047               | 0.584      | 0.016    | 1.945 | 10.991  | 0.911  |
| RM15 | 311                 | 12        | 0           | 90          | B2                 | 0.00       | 0.43       | 0.0224   | 1.436 | 32.21 | -0.304 | 0.056               | 0.570      | 0.016    | 1.430 | 35.873  | -0.258 |



**Figure 8.2:** Random forest (RF) interpolation for each station in the Raam network (blue points) from a model trained using 50% the total dataset. The blue bands represent the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentile prediction interval. Hydrus-1D (H1D) simulations are drawn as brown lines. Scatterplot of RF vs in situ values show the accuracy of the predictions. The accuracy metrics are given in  $\text{m}^3\text{m}^{-3}$ , except for the correlation coefficient, which is dimensionless.



**Figure 8.3:** Random forest (RF) extrapolation for each station in the Raam (blue points) from a model trained using 50% the total dataset. The bands represent the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentile prediction interval. Hydrus-1D (H1D) simulations are drawn as brown lines. Scatterplot of RF vs in situ values show the accuracy of the predictions. The accuracy metrics are given in  $\text{m}^3 \text{m}^{-3}$ , except for the correlation coefficient, which is dimensionless.



**Figure 8.4:** Root zone soil moisture estimates from Hydrus-1D simulations. Results using soil hydraulic parameters (shp's) obtained from BOFEK2012 (Wosten et al., 2013) are plotted in grey. The green lines show the results based on the set of optimized shp's while the red lines are results after data was assimilated via direct insertion of in situ measurements at every 20-day sampling interval ( $n$ ), as represented by the black dots. The accuracy metrics are given in  $\text{m}^3 \text{m}^{-3}$ , except for the correlation coefficient, which is dimensionless.

# Bibliography

- Ahmad, S., Kalra, A., and Stephen, H. (2010). Estimating soil moisture using remote sensing data: A machine learning approach, *Advances in Water Resources* 33, 69–80, doi: 10.1016/j.advwatres.2009.10.008.
- Ahuja, L., Naney, J., Green, R., and Nielsen, D. (1984). Macroporosity to characterize spatial variability of hydraulic conductivity and effects of land management, *Soil Science Society of America Journal* 48, 699–702, doi: 10.2136/sssaj1984.03615995004800040001x.
- Alaoui, A. and Goetz, B. (2008). Dye tracer and infiltration experiments to investigate macropore flow, *Geoderma* 144, 279–286, doi: 10.1016/j.geoderma.2007.11.020.
- Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Pignatelli, B., and Martin, E. (2008). From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations, *Hydrology and Earth System Sciences* 12, 1323–1337, doi: 10.5194/hess-12-1323-2008.
- Albertson, J. D. and Montaldo, N. (2003). Temporal dynamics of soil moisture variability: 1. Theoretical basis, *Water Resources Research* 39, doi: 10.1029/2002WR001616.
- Aline, L. and Pierre, D. (2013). Complementarity of linear polarizations in C-band SAR imagery to estimate leaf area index for maize and winter wheat, in: *ESA Living Planet Symposium*, in ESA SP-722, edited by Ouwehand, L.
- Alippi, C., Camplani, R., Galperti, C., and Roveri, M. (2010). A robust, adaptive, solar-powered WSN framework for aquatic environmental monitoring, *IEEE Sensors Journal* 11, 45–55, doi: 10.1109/JSEN.2010.2051539.
- Allen, R. G., Pereira, L. S., Smith, M., Raes, D., and Wright, J. L. (2005). FAO-56 dual crop coefficient method for estimating evaporation from soil and application extensions, *Journal of irrigation and drainage engineering* 131, 2–13, doi: 10.1061/(ASCE)0733-9437(2005)131:1(2).
- Alletto, L. and Coquet, Y. (2009). Temporal and spatial variability of soil bulk density and near-saturated hydraulic conductivity under two contrasted tillage management systems, *Geoderma* 152, 85–94.
- Almon, S. (1965). The distributed lag between capital appropriations and expenditures, *Econometrica: Journal of the Econometric Society* pp. 178–196.
- Álvarez-Mozos, J., Casali, J., González-Audicana, M., and Verhoest, N. E. (2006). Assessment of the operational applicability of RADARSAT-1 data for surface soil moisture estimation, *IEEE Transactions on Geoscience and Remote Sensing* 44, 913–924, doi: 10.1109/TGRS.2005.862248.
- Amazirh, A., Merlin, O., Er-Raki, S., Gao, Q., Rivalland, V., Malbetreau, Y., Khabba, S., and Escorihuela, M. J. (2018). Retrieving surface soil moisture at high spatio-temporal resolution from a synergy between Sentinel-1 radar and Landsat thermal data: A study case over bare soil, *Remote sensing of environment* 211, 321–337, doi: 10.1016/j.advwatres.2009.10.008.
- Angers, Denis A. and Caron, J.: (1998). *Plant-induced changes in soil structure: Processes and feedbacks*, pp. 55–72, Springer Netherlands, Dordrecht, doi: 10.1007/978-94-017-2691-7.3.
- Arroyo, L. A., Pascual, C., and Manzanera, J. A. (2008). Fire models and methods to map fuel types: the role of remote sensing, *Forest ecology and management* 256, 1239–1252, doi: 10.1016/j.foreco.2008.06.048.

- ASAE EP542, F. (1999). Procedures for Using and Reporting Data Obtained with the Soil Cone Penetrometer, Tech. rep., American Society of Agricultural and Biological Engineers.
- Asilo, S., Nelson, A., de Bie, K., Skidmore, A., Laborte, A., Maunahan, A., and Quilang, E. J. P. (2019). Relating X-band SAR Backscattering to Leaf Area Index of Rice in Different Phenological Phases, *Remote Sensing* 11, 1462, doi: 10.3390/rs11121462.
- Attema, E. P. W. and Ulaby, F. T. (1978). Vegetation modeled as a water cloud, *Radio Science* 13, 357–364, doi: 10.1029/RS013i002p00357.
- Ayers, P. and Perumpral, J. (1982). Moisture and density effect on cone index, *Transactions of the ASAE* 25, 1169–1172, doi: doi:10.13031/2013.33691.
- Baghdadi, N., Aubert, M., Cerdan, O., Franchistéguy, L., Viel, C., Eric, M., Zribi, M., and Desprats, J. (2007). Operational mapping of soil moisture using synthetic aperture radar data: application to the Touch basin (France), *Sensors* 7, 2458–2483, doi: 10.3390/s7102458.
- Baghdadi, N., El Hajj, M., Zribi, M., and Bousbih, S. (2017). Calibration of the water cloud model at C-Band for winter crop fields and grasslands, *Remote Sensing* 9, 969, doi: 10.3390/rs9090969.
- Bakar, S., Shaari, A., Chuah, H., and Ewe, H. (1997). A preliminary study of phenological growth stages of wetland rice using ERS1/2 SAR data, in: IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing-A Scientific Vision for Sustainable Development, vol. 2, pp. 1069–1071, IEEE, doi: 10.1109/IGARSS.1997.615345.
- Baroni, G., Ortuani, B., Facchi, A., and Gandolfi, C. (2013). The role of vegetation and soil properties on the spatio-temporal variability of the surface soil moisture in a maize-cropped field, *Journal of Hydrology* 489, 148–159, doi: 10.1016/j.jhydrol.2013.03.007.
- Bauer-Marschallinger, B., Paulik, C., Hochstöger, S., Mistelbauer, T., Modanesi, S., Ciabatta, L., Massari, C., Brocca, L., and Wagner, W. (2018). Soil Moisture from Fusion of Scatterometer and SAR: Closing the Scale Gap with Temporal Filtering, *Remote Sensing* 10, doi: 10.3390/rs10071030.
- Baumgartl, T. and Horn, R. (1991). Effect of aggregate stability on soil compaction, *Soil and Tillage Research* 19, 203–213, doi: 10.1016/0167-1987(91)90088-F.
- Belmans, C., Wesseling, J., and Feddes, R. A. (1983). Simulation model of the water balance of a cropped soil: SWATRE, *Journal of hydrology* 63, 271–286, doi: 10.1016/0022-1694(83)90045-8.
- Bengough, A. G., Campbell, D. J., and O'Sullivan, M. F. (2000). Penetrometer Techniques in Relation to Soil Compaction and Root Growth, in: Soil and environmental analysis: physical methods, revised, and expanded, edited by Smith, K. A. and Mullins, C. E., pp. 377–404, CRC Press.
- Benninga, H.-J. F., Carranza, C. D., Pezij, M., van Santen, P., van der Ploeg, M. J., Augustijn, D. C., and van der Velde, R. (2018). The Raam regional soil moisture monitoring network in the Netherlands, *Earth System Science Data* 10, 61, doi: 10.5194/essd-10-61-2018.
- Benninga, H.-J. F., van der Velde, R., and Su, Z. (2019). Impacts of Radiometric Uncertainty and Weather-Related Surface Conditions on Soil Moisture Retrievals with Sentinel-1, *Remote sensing* 11, 2025, doi: 10.3390/rs11172025.
- Beven, K. and Germann, P. (2013). Macropores and water flow in soils revisited, *Water Resources Research* 49, 3071–3092, doi: 10.1002/wrcr.20156.
- Bhaskaran, K., Gasparrini, A., Hajat, S., Smeeth, L., and Armstrong, B. (2013). Time series regression studies in environmental epidemiology, *International journal of epidemiology* 42, 1187–1195, doi: 10.1093/ije/dyt092.
- Bindlish, R. and Barros, A. P. (2001). Parameterization of vegetation backscatter in radar-based, soil moisture estimation, *Remote sensing of environment* 76, 130–137, doi: 10.1016/S0034-4257(00)00200-5.
- Bolten, J. D., Crow, W. T., Zhan, X., Jackson, T. J., and Reynolds, C. A. (2009). Evaluating the utility of remotely sensed soil moisture retrievals for operational agricultural drought monitoring, *IEEE Journal of*



- Selected Topics in Applied Earth Observations and Remote Sensing* 3, 57–66, doi: 10.1109/JSTARS.2009.2037163.
- Bousbih, S., Zribi, M., Lili-Chabaane, Z., Baghdadi, N., El Hajj, M., Gao, Q., and Mougenot, B. (2017). Potential of Sentinel-1 Radar Data for the Assessment of Soil and Cereal Cover Parameters, *Sensors* 17, 2617, doi: 10.3390/s17112617.
- Breiman, L. (2001). Random forests, *Machine learning* 45, 5–32.
- Brocca, L., Melone, F., Moramarco, T., and Morbidelli, R. (2009). Soil moisture temporal stability over experimental areas in Central Italy, *Geoderma* 148, 364–374, doi: 10.1016/j.geoderma.2008.11.004.
- Brocca, L., Ciabatta, L., Massari, C., Camici, S., and Tarpanelli, A. (2017). Soil Moisture for Hydrological Applications: Open Questions and New Opportunities, *Water* 9, 140, doi: 10.3390/w9020140.
- Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A., and Fernández-Prieto, D. (2018). How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products, *International journal of applied earth observation and geoinformation* 73, 752–766, doi: 10.1016/j.jag.2018.08.023.
- Brus, D. J. and Heuvelink, G. B. (2007). Optimization of sample patterns for universal kriging of environmental variables, *Geoderma* 138, 86–95, doi: 10.1016/j.geoderma.2006.10.016.
- Brus, D. J. and Van Den Akker, J. J. (2018). How serious a problem is subsoil compaction in the Netherlands? A survey based on probability sampling, *Soil* 4, 37–45, doi: 10.5194/soil-4-37-2018.
- Buitink, J., Swank, A. M., van der Ploeg, M., Smith, N. E., Benninga, H.-J. F., van der Bolt, F., Carranza, C. D. U., Koren, G., van der Velde, R., and Teuling, A. J. (2020). Anatomy of the 2018 agricultural drought in The Netherlands using in situ soil moisture and satellite vegetation indices, *Hydrology and Earth System Sciences Discussions* 2020, 1–17, doi: 10.5194/hess-2020-358.
- Busscher, W., Bauer, P., Camp, C., and Sojka, R. (1997). Correction of cone index for soil water content differences in a coastal plain soil, *Soil and Tillage Research* 43, 205 – 217, doi: 10.1016/S0167-1987(97)00015-9.
- Callens, M., Verhoest, N. E., and Davidson, M. W. (2006). Parameterization of tillage-induced single-scale soil roughness from 4-m profiles, *IEEE Transactions on Geoscience and Remote Sensing* 44, 878–888, doi: 10.1109/TGRS.2005.860488.
- Campbell, D. and O'Sullivan, M. (1991). The cone penetrometer in relation to trafficability, compaction and tillage, in: *Soil Analysis: Physical Methods*, edited by K.A., S. and C.E., M., pp. 399–429, Marcel Dekker Inc., New York, doi: 10.1016/S0167-1987(96)01072-0.
- Capehart, W. J. and Carlson, T. N. (1997). Decoupling of surface and near-surface soil water content: A remote sensing perspective, *Water Resources Research* 33, 1383–1395, doi: 10.1029/97WR00617.
- Carminati, A., Moradi, A. B., Vetterlein, D., Vontobel, P., Lehmann, E., Weller, U., Vogel, H.-J., and Oswald, S. E. (2010). Dynamics of soil water content in the rhizosphere, *Plant and soil* 332, 163–176, doi: 10.1007/s11104-010-0283-8.
- Carminati, A., Schneider, C. L., Moradi, A. B., Zarebanadkouki, M., Vetterlein, D., Vogel, H.-J., Hildebrandt, A., Weller, U., Schüler, L., and Oswald, S. E. (2011). How the rhizosphere may favor water availability to roots, *Vadose Zone Journal* 10, 988–998, doi: 10.2136/vzj2010.0113.
- Carranza, C. and van der Ploeg, M. J. (2020). Characterizing root zone soil moisture dynamics in a cultivated field, *in prep.*
- Carranza, C., Van der Ploeg, M. J., and Torfs, P. J. (2018). Using lagged dependence to identify (de)coupled surface and subsurface soil moisture values, *Hydrology and Earth System Sciences* 22, 2255–2267, doi: 10.5194/hess-22-2255-2018.
- Carranza, C., Benninga, H.-J., van der Velde, R., and van der Ploeg, M. J. (2019). Monitoring agricultural field trafficability using Sentinel-1, *Agricultural Water Management* 224, 105698, doi: 10.1016/j.agwat.

- 2019.105698.
- Carranza, C., Nolet, C., Peziz, M., and van der Ploeg, M. (2020a). Root zone soil moisture estimation with Random Forest, *Journal of Hydrology* p. 125840, doi: 10.1016/j.jhydrol.2020.125840.
- Carranza, C., van Emmerik, T., Benninga, H.-J., van der Velde, R., and van der Ploeg, M. J. (2020b). Exploring the sensitivity of vegetation backscatter to root zone soil moisture, *Submitted to Biogeosciences*.
- Carrara, M., Castrignanò, A., Comparetti, A., Febo, P., and Orlando, S. (2007). Mapping of penetrometer resistance in relation to tractor traffic using multivariate geostatistics, *Geoderma* 142, 294–307, doi: 10.1016/j.geoderma.2007.08.020.
- Çerçioğlu, M., Anderson, S. H., Udawatta, R. P., and Alagele, S. (2019). Effect of cover crop management on soil hydraulic properties, *Geoderma* 343, 247–253, doi: 10.1016/j.geoderma.2019.02.027.
- Cey, E. E. and Rudolph, D. L. (2009). Field study of macropore flow processes using tension infiltration of a dye tracer in partially saturated soils, *Hydrological Processes: An International Journal* 23, 1768–1779, doi: 10.1002/hyp.7302.
- Chandrasekhar, P., Kreiselmeyer, J., Schwen, A., Weninger, T., Julich, S., Feger, K.-H., and Schwärzel, K. (2018). Why we should include soil structural dynamics of agricultural soils in hydrological models, *Water* 10, 1862, doi: 10.3390/w10121862.
- Charbonneau, F., Trudel, M., and Fernandes, R. (2005). Use of Dual-Polarization and Multi-Incidence SAR for soil permeability mapping, in: *Proceedings of the 2005 Advanced Synthetic Aperture Radar (ASAR) Workshop*, St-Hubert, QC, Canada.
- Chuvieco, E., Riaño, D., Van Wagtenok, J., and Morsdorf, F. (2003). Fuel loads and fuel type mapping, in: *Wildland fire danger estimation and mapping: The role of remote sensing data*, pp. 119–142, World Scientific.
- Cihlar, J. and Ulaby, F. T. (1974). Dielectric properties of soils as a function of moisture content, Tech. Rep. 177-4, The University of Kansas Center for Research, Inc., Lawrence, Kansas.
- Cleveland, R. B., Cleveland, W. S., and Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess, *Journal of Official Statistics* 6, 3.
- Cleveland, W. S. and Devlin, S. J. (1988). Locally weighted regression: an approach to regression analysis by local fitting, *Journal of the American statistical association* 83, 596–610, doi: 10.1080/01621459.1988.10478639.
- Cobos, D. and Campbell, C. (2007). Correcting temperature sensitivity of ECH2O soil moisture sensors, *Appl. Note. Decagon Devices, Pullman, WA*.
- Coopersmith, E. J., Cosh, M. H., Bell, J. E., and Boyles, R. (2016). Using machine learning to produce near surface soil moisture estimates from deeper in situ records at US Climate Reference Network (USCRN) locations: Analysis and applications to AMSR-E satellite validation, *Advances in Water Resources* 98, 122–131, doi: 10.1016/j.advwatres.2016.10.007.
- Cordova, J. R. and Bras, R. L. (1981). Physically based probabilistic models of infiltration, soil moisture, and actual evapotranspiration, *Water Resources Research* 17, 93–106, doi: 10.1029/WR017i001p00093.
- Dabrowska-Zielinska, K., Inoue, Y., Kowalik, W., and Gruszczynska, M. (2007). Inferring the effect of plant and soil variables on C- and L-band SAR backscatter over agricultural fields, based on model analysis, *Advances in Space Research* 39, 139–148, doi: 10.1016/j.asr.2006.02.032.
- Das, N., Mohanty, B., Cosh, M., and Jackson, T. (2008). Modeling and assimilation of root zone soil moisture using remote sensing observations in Walnut Gulch Watershed during SMEX04, *Remote Sensing of Environment* 112, 415–429, doi: 10.1016/j.rse.2006.10.027.
- Das, N. N. and Mohanty, B. P. (2006). Root zone soil moisture assessment using remote sensing and vadose zone modeling, *Vadose Zone Journal* 5, 296–307, doi: 10.2136/vzj2005.0033.
- De Baets, S., Poesen, J., Reubens, B., Wemans, K., De Baerdemaeker, J., and Muys, B. (2008). Root tensile

- strength and root distribution of typical Mediterranean plant species and their contribution to soil shear strength, *Plant and soil* 305, 207–226, doi: 10.1007/s11104-008-9553-0.
- De Laet, P.: (1980). *Model for unsaturated flow above a shallow water-table applied to a regional sub-surface flow problem*, Pudoc, Centre for Agricultural Publishing and Documentation.
- De Lange, W. J., Prinsen, G. F., Hoogewoud, J. C., Veldhuizen, A. A., Verkaik, J., Essink, G. H. O., Van Walsum, P. E., Delsman, J. R., Hunink, J. C., Massop, H. T. L., et al. (2014). An operational, multi-scale, multi-model system for consensus-based, integrated water management and policy analysis: The Netherlands Hydrological Instrument, *Environmental Modelling & Software* 59, 98–108, doi: 10.1016/j.envsoft.2014.05.009.
- De Lannoy, G. J., Verhoest, N. E., Houser, P. R., Gish, T. J., and Van Meirvenne, M. (2006). Spatial and temporal characteristics of soil moisture in an intensively monitored agricultural field (OPE 3), *Journal of Hydrology* 331, 719–730, doi: 10.1016/j.jhydrol.2006.06.016.
- Défosse, P. and Richard, G. (2002). Models of soil compaction due to traffic and their evaluation, *Soil and Tillage Research* 67, 41–64, doi: 10.1016/S0167-1987(02)00030-2.
- Dente, L., Verkedy, Z., Su, Z., and Ucer, M. (2011). Twente Soil Moisture and Soil Temperature Monitoring Network, Tech. rep., University of Twente—ITC: Enschede, The Netherlands, URL [http://www.itc.nl/library/papers\\_2011/scie/dente\\_twe.pdf](http://www.itc.nl/library/papers_2011/scie/dente_twe.pdf).
- Dexter, A. (1987). Compression of soil around roots, *Plant and Soil* 97, 401–406, doi: 10.1007/BF02383230.
- Dobriyal, P., Qureshi, A., Badola, R., and Hussain, S. A. (2012). A review of the methods available for estimating soil moisture and its implications for water resource management, *Journal of Hydrology* 458, 110–117, doi: 10.1016/j.jhydrol.2012.06.021.
- Dorigo, W., Van Oevelen, P., Wagner, W., Drusch, M., Mecklenburg, S., Robock, A., and Jackson, T. (2011a). A new international network for in situ soil moisture data, *Eos, Transactions American Geophysical Union* 92, 141–142, doi: 10.1029/2011EO170001.
- Dorigo, W., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Oevelen, P. v., et al. (2011b). The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, *Hydrology and Earth System Sciences* 15, 1675–1698, doi: 10.5194/hess-15-1675-2011.
- Draper, C., Reichle, R., De Lannoy, G., and Liu, Q. (2012). Assimilation of passive and active microwave soil moisture retrievals, *Geophysical Research Letters* 39, doi: 10.1029/2011GL050655.
- Droogers, P., Fermont, A., and Bouma, J. (1996). Effects of ecological soil management on workability and trafficability of a loamy soil in the Netherlands, *Geoderma* 73, 131–145, doi: 10.1016/0016-7061(96)00053-5.
- Dubois, P. C., Van Zyl, J., and Engman, T. (1995). Measuring soil moisture with imaging radars, *IEEE Transactions on Geoscience and Remote Sensing* 33, 915–926, doi: 10.1109/36.406677.
- D'odorico, P., Laio, F., Porporato, A., and Rodriguez-Iturbe, I. (2003). Hydrologic controls on soil carbon and nitrogen cycles. II. A case study, *Advances in Water Resources* 26, 59–70, doi: 10.1016/S0309-1708(02)00095-7.
- Earl, R. (1997). Prediction of trafficability and workability from soil moisture deficit, *Soil and tillage research* 40, 155–168, doi: 10.1016/S0167-1987(96)01072-0.
- Eckelmann, W., Baritz, R., Bialousz, S., Bielek, P., Carré, F., Hrušková, B., Jones, R. J., Kibblewhite, M., Kozak, J., Le Bas, C., et al.: (2006). *Common criteria for risk area identification according to soil threats*, Office for Official Publications of the European Communities.
- Ekwue, E. and Harrilal, A. (2010). Effect of soil type, peat, slope, compaction effort and their interactions on infiltration, runoff and raindrop erosion of some Trinidadian soils, *Biosystems Engineering* 105, 112–118, doi: 10.1016/j.biosystemseng.2009.10.001.

- El Hajj, M., Baghdadi, N., Zribi, M., and Bazzi, H. (2017). Synergic use of Sentinel-1 and Sentinel-2 images for operational soil moisture mapping at high spatial resolution over agricultural areas, *Remote Sensing* 9, 1292, doi: 10.3390/rs9121292.
- Entekhabi, D., Rodriguez-Iturbe, I., and Castelli, F. (1996). Mutual interaction of soil moisture state and atmospheric processes, *Journal of Hydrology* 184, 3–17, doi: 10.1016/0022-1694(95)02965-6.
- ESA (2019). Copernicus Open Access Hub, URL <https://scihub.copernicus.eu/>.
- Evensen, G.: (2009). *Data assimilation: the ensemble Kalman filter*, Springer Science & Business Media.
- Fabrizzi, K., Garcia, F., Costa, J., and Picone, L. (2005). Soil water dynamics, physical properties and corn and wheat responses to minimum and no-tillage systems in the southern Pampas of Argentina, *Soil and Tillage Research* 81, 57–69, doi: 10.1016/j.still.2004.05.001.
- Falloon, P., Jones, C. D., Ades, M., and Paul, K. (2011). Direct soil moisture controls of future global soil carbon changes: An important source of uncertainty, *Global Biogeochemical Cycles* 25, doi: 10.1029/2010GB003938.
- Famiglietti, J. S., Rudnicki, J. W., and Rodell, M. (1998). Variability in surface moisture content along a hill-slope transect: Rattlesnake Hill, Texas, *Journal of hydrology* 210, 259–281, doi: 10.1016/S0022-1694(98)00187-5.
- Famiglietti, J. S., Ryu, D., Berg, A. A., Rodell, M., and Jackson, T. J. (2008). Field observations of soil moisture variability across scales, *Water Resources Research* 44, doi: 10.1029/2006WR005804.
- Fan, C.-C. and Su, C.-F. (2008). Role of roots in the shear strength of root-reinforced soils with high moisture content, *Ecological Engineering* 33, 157–166, doi: 10.1016/j.ecoleng.2008.02.013.
- FAO, I. (2015). Status of the World's Soil Resources (SWSR)—Main Report, *Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils, Rome, Italy* 650.
- Fatichi, S., Or, D., Walko, R., Vereecken, H., Young, M. H., Ghezzehei, T. A., Hengl, T., Kollet, S., Agam, N., and Avissar, R. (2020). Soil structure is an important omission in Earth System Models, *Nature communications* 11, 1–11, doi: 10.1038/s41467-020-14411-z.
- Feddes, R., Kabat, P., Van Bakel, P., Bronswijk, J., and Halbertsma, J. (1988). Modelling soil water dynamics in the unsaturated zone—state of the art, *Journal of Hydrology* 100, 69–111, doi: 10.1016/0022-1694(88)90182-5.
- Ferrero, A., Usowicz, B., and Lipiec, J. (2005). Effects of tractor traffic on spatial variability of soil strength and water content in grass covered and cultivated sloping vineyard, *Soil and Tillage Research* 84, 127–138, doi: 10.1016/j.still.2004.10.003.
- Filipponi, F. (2019). Sentinel-1 GRD Preprocessing Workflow, in: Proceedings of 3rd International Electronic Conference on Remote Sensing, vol. 18, p. 11, doi: 10.3390/ecrs-3-06201.
- Fletcher, R. and Reeves, C. M. (1964). Function minimization by conjugate gradients, *The computer journal* 7, 149–154, doi: 10.1093/comjnl/7.2.149.
- Ford, T., Harris, E., and Quiring, S. (2014). Estimating root zone soil moisture using near-surface observations from SMOS, *Hydrology and Earth System Sciences* 18, 139–154, doi: 10.5194/hess-18-139-2014.
- Fry, J. E. and Guber, A. K. (2020). Temporal stability of field-scale patterns in soil water content across topographically diverse agricultural landscapes, *Journal of Hydrology* 580, 124260, doi: 10.1016/j.jhydrol.2019.124260.
- Fung, A. K.: (1994). *Microwave scattering and emission models and their applications*, Artech House.
- Fung, A. K., Li, Z., and Chen, K.-S. (1992). Backscattering from a randomly rough dielectric surface, *IEEE Transactions on Geoscience and remote sensing* 30, 356–369, doi: 10.1109/36.134085.
- Gao, X., Wu, P., Zhao, X., Wang, J., and Shi, Y. (2014). Effects of land use on soil moisture variations in a semi-arid catchment: implications for land and agricultural water management, *Land Degradation & Development* 25, 163–172, doi: 10.1002/ldr.1156.

- Gao, X., Zhao, X., Brocca, L., Pan, D., and Wu, P. (2019). Testing of observation operators designed to estimate profile soil moisture from surface measurements, *Hydrological processes* 33, 575–584, doi: 10.1002/hyp.13344.
- Gasparrini, A. (2011). Distributed lag linear and non-linear models in R: the package dlnm, *Journal of statistical software* 43, 1.
- Gasparrini, A., Armstrong, B., and Kenward, M. (2010). Distributed lag non-linear models, *Statistics in medicine* 29, 2224–2234, doi: 10.1002/sim.3940.
- Gasparrini, A., Scheipl, F., Armstrong, B., and Kenward, M. (2017). A penalized framework for distributed lag non-linear models, *Biometrics* doi: 10.1111/biom.12645.
- Gebbers, R. and Adamchuk, V. I. (2010). Precision agriculture and food security, *Science* 327, 828–831, doi: 10.1126/science.1183899.
- Geerts, S. and Raes, D. (2009). Deficit irrigation as an on-farm strategy to maximize crop water productivity in dry areas, *Agricultural water management* 96, 1275–1284, doi: 10.1016/j.agwat.2009.04.009.
- Gentle, J. E.: (2009). *Computational Statistics*, chap. Generation of Random Numbers, pp. 305–331, Springer New York, New York, NY, doi: 10.1007/978-0-387-98144-4.
- Ghafoor, A., Koestel, J., Larsbo, M., Moeys, J., and Jarvis, N. (2013). Soil properties and susceptibility to preferential solute transport in tilled topsoil at the catchment scale, *Journal of Hydrology* 492, 190–199, doi: 10.1016/j.jhydrol.2013.03.046.
- Gherboudj, I., Magagi, R., Berg, A. A., and Toth, B. (2011). Soil moisture retrieval over agricultural fields from multi-polarized and multi-angular RADARSAT-2 SAR data, *Remote sensing of environment* 115, 33–43, doi: 10.1016/j.rse.2010.07.011.
- Gonzalez-Sosa, E., Braud, I., Dehotin, J., Lassabatère, L., Angulo-Jaramillo, R., Lagouy, M., Branger, F., Jacqueminet, C., Kermadi, S., and Michel, K. (2010). Impact of land use on the hydraulic properties of the topsoil in a small French catchment, *Hydrological processes* 24, 2382–2399, doi: 10.1002/hyp.7640.
- González-Zamora, Á., Sánchez, N., Martínez-Fernández, J., and Wagner, W. (2016). Root-zone plant available water estimation using the SMOS-derived soil water index, *Advances in Water Resources* 96, 339–353, doi: 10.1016/j.advwatres.2016.08.001.
- Govaerts, B., Fuentes, M., Mezzalama, M., Nicol, J. M., Deckers, J., Etchevers, J. D., Figueroa-Sandoval, B., and Sayre, K. D. (2007). Infiltration, soil moisture, root rot and nematode populations after 12 years of different tillage, residue and crop rotation managements, *Soil and Tillage Research* 94, 209–219, doi: 10.1016/j.still.2006.07.013.
- Grayson, R. B., Western, A. W., Chiew, F. H., and Blöschl, G. (1997). Preferred states in spatial soil moisture patterns: Local and nonlocal controls, *Water resources research* 33, 2897–2908, doi: 10.1029/97WR02174.
- Guber, A., Gish, T., Pachepsky, Y., van Genuchten, M. T., Daughtry, C., Nicholson, T., and Cady, R. (2008). Temporal stability in soil water content patterns across agricultural fields, *Catena* 73, 125–133, doi: 10.1016/j.catena.2007.09.010.
- Haas, C., Horn, R., Ellerbrock, R., and Gerke, H. (2020). Fluorescence imaging for mm-scale observation of macropore-matrix mass transfer: Calibration experiments, *Geoderma* 360, 114002, doi: 10.1016/j.geoderma.2019.114002.
- Hajj, M. E., Baghdadi, N., Zribi, M., and Bazzi, H. (2017). Synergic use of Sentinel-1 and Sentinel-2 images for operational soil moisture mapping at high spatial resolution over agricultural areas, *Remote Sensing* 9, 1292, doi: 10.3390/rs9121292.
- Hamza, M. and Anderson, W. (2005). Soil compaction in cropping systems: a review of the nature, causes and possible solutions, *Soil and tillage research* 82, 121–145, doi: 10.1016/j.still.2004.08.009.
- Hartge, K. H., Horn, R., Horton, R., Bachmann, J., and Peth, S.: (2016). *Essential soil physics*, Schweizerbart'sche Verlagsbuchhandlung.

- Hastie, T., Tibshirani, R., and Friedman, J.: (2009). *The elements of statistical learning: data mining, inference, and prediction*, Springer Science & Business Media.
- Haws, N. W., Rao, P. S. C., Simunek, J., and Poyer, I. C. (2005). Single-porosity and dual-porosity modeling of water flow and solute transport in subsurface-drained fields using effective field-scale parameters, *Journal of Hydrology* 313, 257–273, doi: 10.1016/j.jhydrol.2005.03.035.
- Heathman, G. C., Starks, P. J., Ahuja, L. R., and Jackson, T. J. (2003). Assimilation of surface soil moisture to estimate profile soil water content, *Journal of Hydrology* 279, 1–17, doi: 10.1016/S0022-1694(03)00088-X.
- Helliwell, J., Sturrock, C. J., Mairhofer, S., Craigon, J., Ashton, R., Miller, A., Whalley, W., and Mooney, S. J. (2017). The emergent rhizosphere: imaging the development of the porous architecture at the root-soil interface, *Scientific reports* 7, 1–10, doi: 10.1038/s41598-017-14904-w.
- Henderson, C., Levett, A., and Lisle, D. (1988). The effects of soil water content and bulk density on the compactibility and soil penetration resistance of some Western Australian sandy soils, *Soil Research* 26, 391–400, doi: 10.1071/SR9880391.
- Henderson, F. M. and Lewis, A. J.: (1998). *Principles and applications of imaging radar. Manual of remote sensing: Volume 2*, John Wiley and Sons, Inc., Somerset, NJ (United States).
- Hengl, T., Nussbaum, M., Wright, M. N., Heuvelink, G. B., and Gräler, B. (2018). Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables, *PeerJ* 6, e5518, doi: 10.7287/peerj.preprints.26693v3.
- Hillel, D.: (1998). *Environmental soil physics: Fundamentals, applications, and environmental considerations*, Elsevier.
- Hirschi, M., Mueller, B., Dorigo, W., and Seneviratne, S. (2014). Using remotely sensed soil moisture for land–atmosphere coupling diagnostics: The role of surface vs. root-zone soil moisture variability, *Remote sensing of environment* 154, 246–252, doi: 10.1016/j.rse.2014.08.030.
- Hornacek, M., Wagner, W., Sabel, D., Truong, H.-L., Snoeij, P., Hahmann, T., Diedrich, E., and Doubková, M. (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, 1303–1311, doi: 10.1109/JSTARS.2012.2190136.
- Houser, P. R., Shuttleworth, W. J., Famiglietti, J. S., Gupta, H. V., Syed, K. H., and Goodrich, D. C. (1998). Integration of soil moisture remote sensing and hydrologic modeling using data assimilation, *Water Resources Research* 34, 3405–3420, doi: 10.1029/1998WR900001.
- Houtekamer, P. L. and Mitchell, H. L. (1998). Data assimilation using an ensemble Kalman filter technique, *Monthly Weather Review* 126, 796–811, doi: 10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.
- Huang, Y., Walker, J. P., Gao, Y., Wu, X., and Moneris, A. (2015). Estimation of vegetation water content from the radar vegetation index at L-band, *IEEE Transactions on Geoscience and Remote Sensing* 54, 981–989, doi: 10.1109/TGRS.2015.2471803.
- Hupet, F. and Vanclooster, M. (2002). Intraseasonal dynamics of soil moisture variability within a small agricultural maize cropped field, *Journal of Hydrology* 261, 86–101, doi: 10.1016/S0022-1694(02)00016-1.
- Imhoff, M. L. (1995). A theoretical analysis of the effect of forest structure on synthetic aperture radar backscatter and the remote sensing of biomass, *IEEE Transactions on Geoscience and remote sensing* 33, 341–351, doi: 10.1109/TGRS.1995.8746015.
- Jackson, T. J. (1993). III. Measuring surface soil moisture using passive microwave remote sensing, *Hydrological processes* 7, 139–152, doi: 10.1002/hyp.3360070205.
- Jarvis, N. (1989). A simple empirical model of root water uptake, *Journal of Hydrology* 107, 57–72, doi: 10.1016/0022-1694(89)90050-4.
- Joseph, A., van der Velde, R., O’neill, P., Lang, R., and Gish, T. (2010). Effects of corn on C-and L-band radar backscatter: A correction method for soil moisture retrieval, *Remote Sensing of Environment* 114,

- 2417–2430, doi: 0.1016/j.rse.2010.05.017.
- Joshi, C. and Mohanty, B. P. (2010). Physical controls of near-surface soil moisture across varying spatial scales in an agricultural landscape during SMEX02, *Water Resources Research* 46, doi: 10.1029/2010WR009152.
- Karandish, F. and Šimnek, J. (2016). A comparison of numerical and machine-learning modeling of soil water content with limited input data, *Journal of Hydrology* 543, 892–909, doi: 10.1016/j.jhydrol.2016.11.007.
- Karpatne, A., Watkins, W., Read, J., and Kumar, V. (2017). Physics-guided neural networks (pgnn): An application in lake temperature modeling, *arXiv preprint arXiv:1710.11431* .
- Keesstra, S., Pereira, P., Novara, A., Brevik, E. C., Azorin-Molina, C., Parras-Alcántara, L., Jordán, A., and Cerdà, A. (2016). Effects of soil management techniques on soil water erosion in apricot orchards, *Science of the Total Environment* 551, 357–366, doi: 10.1016/j.scitotenv.2016.01.182.
- Keller, T. and Lamandé, M. (2010). Challenges in the development of analytical soil compaction models, *Soil and Tillage Research* 111, 54–64, doi: 10.1016/j.still.2010.08.004.
- Kim, G. and Barros, A. P. (2002). Space–time characterization of soil moisture from passive microwave remotely sensed imagery and ancillary data, *Remote sensing of environment* 81, 393–403, doi: 10.1016/S0034-4257(02)00014-7.
- Kim, Y. and van Zyl, J. (2004). Vegetation effects on soil moisture estimation, in: IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium, vol. 2, pp. 800–802, IEEE, doi: 10.1109/IGARSS.2004.1368525.
- Kim, Y., Jackson, T., Bindlish, R., Lee, H., and Hong, S. (2011). Radar vegetation index for estimating the vegetation water content of rice and soybean, *IEEE Geoscience and Remote Sensing Letters* 9, 564–568, doi: 10.1109/LGRS.2011.2174772.
- Kim, Y., Jackson, T., Bindlish, R., Hong, S., Jung, G., and Lee, K. (2013). Retrieval of wheat growth parameters with radar vegetation indices, *IEEE Geoscience and Remote Sensing Letters* 11, 808–812, doi: 10.1109/LGRS.2013.2279255.
- King, C., Baghdadi, N., Lecomte, V., and Cerdan, O. (2005). The application of remote-sensing data to monitoring and modelling of soil erosion, *Catena* 62, 79–93, doi: 10.3997/2214-4609.2020geo131.
- Knight, S. and Freitag, D. (1961). Measuring Soil Trafficability Characteristics, Tech. rep., Army Engineer Waterways Experiment Station Vicksburg MS.
- Kodešová, R., Kodeš, V., Žigová, A., and Šimnek, J. (2006). Impact of plant roots and soil organisms on soil micromorphology and hydraulic properties, *Biologia* 61, S339–S343, doi: 10.2478/s11756-006-0185-7.
- Koebnick, N., Daly, K. R., Keyes, S. D., Bengough, A. G., Brown, L. K., Cooper, L. J., George, T. S., Hallett, P. D., Naveed, M., Raffan, A., et al. (2019). Imaging microstructure of the barley rhizosphere: particle packing and root hair influences, *New Phytologist* 221, 1878–1889, doi: 10.1111/nph.15516.
- Kornelsen, K. C. and Coulibaly, P. (2013). Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications, *Journal of Hydrology* 476, 460–489, doi: 10.1016/j.jhydrol.2012.10.044.
- Kornelsen, K. C. and Coulibaly, P. (2014). Root-zone soil moisture estimation using data-driven methods, *Water Resources Research* 50, 2946–2962, doi: 10.1002/2013WR014127.
- Koster, R. D. and Suarez, M. J. (2001). Soil moisture memory in climate models, *Journal of hydrometeorology* 2, 558–570, doi: 10.1175/1525-7541(2001)002<0558:SMMICM>2.0.CO;2.
- Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C., Kanae, S., Kowalczyk, E., Lawrence, D., et al. (2004). Regions of strong coupling between soil moisture and precipitation, *Science* 305, 1138–1140, doi: 10.1126/science.1100217.
- Kováč, K., Macak, M., and Švančárková, M. (2005). The effect of soil conservation tillage on soil moisture dynamics under single cropping and crop rotation, *Plant Soil Environ* 3, 124–130.
- Kruschke, J. K. (2013). Bayesian estimation supersedes the t test, *Journal of Experimental Psychology: General* 142, 573, doi: 10.1037/a0029146.

- Kuang, B., Mahmood, H. S., Quraishi, M. Z., Hoogmoed, W. B., Mouazen, A. M., and van Henten, E. J. (2012). Chapter four - Sensing Soil Properties in the Laboratory, In Situ, and On-Line: A Review, in: *Advances in Agronomy*, edited by Sparks, D. L., vol. 114 of *Advances in Agronomy*, pp. 155 – 223, Academic Press, doi: 10.1016/B978-0-12-394275-3.00003-1.
- Kumar, K., Hari Prasad, K., and Arora, M. (2012). Estimation of water cloud model vegetation parameters using a genetic algorithm, *Hydrological Sciences Journal* 57, 776–789, doi: 10.1080/02626667.2012.678583.
- Kumar, S. V., Reichle, R. H., Koster, R. D., Crow, W. T., and Peters-Lidard, C. D. (2009). Role of subsurface physics in the assimilation of surface soil moisture observations, *Journal of hydrometeorology* 10, 1534–1547, doi: 10.1175/2009JHM1134.1.
- Kumar, S. V., Holmes, T. R., Bindlish, R., Jeu, R. d., and Peters-Lidard, C. (2020). Assimilation of vegetation optical depth retrievals from passive microwave radiometry, *Hydrology and Earth System Sciences* 24, 3431–3450, doi: 10.5194/hess-24-3431-2020.
- Kurc, S. A. and Small, E. E. (2007). Soil moisture variations and ecosystem-scale fluxes of water and carbon in semiarid grassland and shrubland, *Water Resources Research* 43, doi: 10.1029/2006WR005011.
- Lal, R. (1991). Tillage and agricultural sustainability, *Soil and Tillage Research* 20, 133–146, doi: 10.1016/0167-1987(91)90036-W.
- Lazarovitch, N., Vanderborght, J., Jin, Y., and van Genuchten, M. T. (2018). The root zone: soil physics and beyond, *Vadose Zone Journal* 17, 1–6, doi: 10.2136/vzj2018.01.0002.
- Lever, J., Krzywinski, M., and Altman, N. (2016). Points of significance: model selection and overfitting, doi: 10.1038/nmeth.3968.
- Li, Y., Liao, Q., Li, X., Liao, S., Chi, G., and Peng, S. (2003). Towards an operational system for regional-scale rice yield estimation using a time-series of Radarsat ScanSAR images, *International Journal of Remote Sensing* 24, 4207–4220, doi: 10.1080/0143116031000095970.
- Li, Y., Ye, W., Wang, M., and Yan, X. (2009). Climate change and drought: a risk assessment of crop-yield impacts, *Climate research* 39, 31–46, doi: 10.3354/cr00797.
- Liao, L., Zhang, L., and Bengtsson, L. (2008). Soil moisture variation and water consumption of spring wheat and their effects on crop yield under drip irrigation, *Irrigation and drainage systems* 22, 253–270, doi: 10.1007/s10795-008-9055-5.
- Lievens, H., Verhoest, N. E., Keyser, E. D., Vernieuwe, H., Matgen, P., Álvarez-Mozos, J., and Baets, B. D. (2011). Effective roughness modelling as a tool for soil moisture retrieval from C- and L-band SAR, *Hydrology and Earth System Sciences* 15, 151–162, doi: 10.5194/hess-15-151-2011.
- LoRaWAN (2021). LoRaWAN Specification, LoRa Alliance 2021, <https://lora-alliance.org/>.
- Lu, J., Zhang, Q., Werner, A. D., Li, Y., Jiang, S., and Tan, Z. (2020). Root-induced changes of soil hydraulic properties—A review, *Journal of Hydrology* p. 125203, doi: 10.1016/j.jhydrol.2020.125203.
- Lucas, M., Schlüter, S., Vogel, H.-J., and Vetterlein, D. (2019). Roots compact the surrounding soil depending on the structures they encounter, *Scientific reports* 9, 1–13, doi: 10.1038/s41598-019-52665-w.
- Luo, L., Lin, H., and Halleck, P. (2008). Quantifying soil structure and preferential flow in intact soil using X-ray computed tomography, *Soil Science Society of America Journal* 72, 1058–1069, doi: 10.2136/sssaj2007.0179.
- Macelloni, G., Paloscia, S., Pampaloni, P., Marliani, F., and Gai, M. (2001). The relationship between the backscattering coefficient and the biomass of narrow and broad leaf crops, *IEEE Transactions on Geoscience and Remote Sensing* 39, 873–884, doi: 10.1109/36.917914.
- Mahmood, R., Littell, A., Hubbard, K. G., and You, J. (2012). Observed data-based assessment of relationships among soil moisture at various depths, precipitation, and temperature, *Applied Geography* 34, 255–264, doi: 10.1016/j.apgeog.2011.11.009.
- Mallory, J., Mohtar, R., Heathman, G., Schulze, D., and Braudeau, E. (2011). Evaluating the effect of tillage



- on soil structural properties using the pedostructure concept, *Geoderma* 163, 141–149, doi: 10.1016/j.geoderma.2011.01.018.
- Malone, R., Logsdon, S., Shipitalo, M., Weatherington-Rice, J., Ahuja, L., and Ma, L. (2003). Tillage effect on macroporosity and herbicide transport in percolate, *Geoderma* 116, 191–215, doi: 10.1016/S0016-7061(03)00101-0.
- Mapa, R., Green, R., and Santo, L. (1986). Temporal variability of soil hydraulic properties with wetting and drying subsequent to tillage, *Soil Science Society of America Journal* 50, 1133–1138, doi: 10.2136/sssaj1986.036159950050000500008x.
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters, *Journal of the society for Industrial and Applied Mathematics* 11, 431–441, doi: 10.1137/0111030.
- Martinez, C., Hancock, G., Kalma, J., and Wells, T. (2008). Spatio-temporal distribution of near-surface and root zone soil moisture at the catchment scale, *Hydrological Processes* 22, 2699–2714, doi: 10.1002/hyp.6869.
- Martinez, G., Pachepsky, Y. A., Vereecken, H., Hardelauf, H., Herbst, M., and Vanderlinden, K. (2013). Modeling local control effects on the temporal stability of soil water content, *Journal of Hydrology* 481, 106–118, doi: 10.1016/j.jhydrol.2012.12.024.
- Martinez-Agirre, A., Álvarez-Mozos, J., Lievens, H., and Verhoest, N. E. (2017). Influence of surface roughness measurement scale on radar backscattering in different agricultural soils, *IEEE Transactions on Geoscience and Remote Sensing* 55, 5925–5936, doi: 10.1109/TGRS.2017.2717043.
- Matei, O., Rusu, T., Petrovan, A., and Mihaș, G. (2017). A data mining system for real time soil moisture prediction, *Procedia Engineering* 181, 837–844, doi: 10.1016/j.proeng.2017.02.475.
- McColl, K. A., Alemohammad, S. H., Akbar, R., Konings, A. G., Yueh, S., and Entekhabi, D. (2017). The global distribution and dynamics of surface soil moisture, *Nature Geoscience* 10, 100–104, doi: 10.1038/ngeo2868.
- McNairn, H., Decker, V., and Murnaghan, K. (2002). The sensitivity of C-band polarimetric SAR to crop condition, in: *IEEE International Geoscience and Remote Sensing Symposium*, vol. 3, pp. 1471–1473, IEEE, doi: 10.1109/IGARSS.2002.1026152.
- Meinshausen, N. (2006). Quantile regression forests, *Journal of Machine Learning Research* 7, 983–999.
- Melles, S., Heuvelink, G. B., Twenhöfel, C. J., van Dijk, A., Hiemstra, P. H., Baume, O., and Stöhlker, U. (2011). Optimizing the spatial pattern of networks for monitoring radioactive releases, *Computers & Geosciences* 37, 280–288, doi: 10.1016/j.cageo.2010.04.007.
- Mickovski, S. B., Hallett, P. D., Bransby, M. F., Davies, M. C., Sonnenberg, R., and Bengough, A. G. (2009). Mechanical reinforcement of soil by willow roots: impacts of root properties and root failure mechanism, *Soil Science Society of America Journal* 73, 1276–1285, doi: 10.2136/sssaj2008.0172.
- Mladenova, I. E., Jackson, T. J., Bindlish, R., and Hensley, S. (2013). Incidence angle normalization of radar backscatter data, *IEEE Transactions on Geoscience and Remote Sensing* 51, 1791–1804, doi: 10.1109/TGRS.2012.2205264.
- Mohanty, B. and Skaggs, T. (2001). Spatio-temporal evolution and time-stable characteristics of soil moisture within remote sensing footprints with varying soil, slope, and vegetation, *Advances in water resources* 24, 1051–1067.
- Mohanty, B. P., Cosh, M. H., Lakshmi, V., and Montzka, C. (2017). Soil Moisture Remote Sensing: State-of-the-Science, *Vadose Zone Journal* 16, doi: 10.2136/vzj2016.10.0105.
- Moran, M. S., Hymer, D. C., Qi, J., and Sano, E. E. (2000). Soil moisture evaluation using multi-temporal synthetic aperture radar (SAR) in semiarid rangeland, *Agricultural and Forest meteorology* 105, 69–80, doi: 10.1016/S0168-1923(00)00189-1.
- Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., and Papathanassiou, K. P. (2013). A tutorial

- on synthetic aperture radar, *IEEE Geoscience and remote sensing magazine* 1, 6–43, doi: 10.1109/MGRS.2013.2248301.
- Moret, D. and Arrúe, J. (2007). Dynamics of soil hydraulic properties during fallow as affected by tillage, *Soil and Tillage Research* 96, 103–113, doi: 10.1016/j.still.2007.04.003.
- Mosaddeghi, M., Hajabbasi, M., Hemmat, A., and Afyuni, M. (2000). Soil compactibility as affected by soil moisture content and farmyard manure in central Iran, *Soil and Tillage Research* 55, 87–97, doi: 10.1016/S0167-1987(00)00102-1.
- Mudgal, A., Anderson, S., Baffaut, C., Kitchen, N., and Sadler, E. (2010). Effects of long-term soil and crop management on soil hydraulic properties for claypan soils, *Journal of soil and water conservation* 65, 393–403, doi: 10.2489/jswc.65.6.393.
- Müller, L., Lipiec, J., Kornecki, T. S., and Gebhardt, S. (2011). Trafficability and workability of soils, in: *Encyclopedia of agrophysics*, edited by Gliński, J., Horabik, J., and Lipiec, J., pp. 912–924, Springer Netherlands, doi: 10.1007/978-90-481-3585-1.
- Nawaz, M. F., Bourrie, G., and Trolard, F. (2013). Soil compaction impact and modelling. A review, *Agronomy for Sustainable Development* 33, 291–309, doi: 10.1007/s13593-011-0071-8.
- Ndiaye, B., Molénat, J., Hallaire, V., Gascuel, C., and Hamon, Y. (2007). Effects of agricultural practices on hydraulic properties and water movement in soils in Brittany (France), *Soil and Tillage Research* 93, 251–263, doi: 10.1016/j.still.2006.04.005.
- Njoku, E. G., Jackson, T. J., Lakshmi, V., Chan, T. K., and Nghiem, S. V. (2003). Soil moisture retrieval from AMSR-E, *IEEE transactions on Geoscience and remote sensing* 41, 215–229, doi: 10.1109/TGRS.2002.808243.
- Notarnicola, C. and Posa, F. (2007). Inferring vegetation water content from C-and L-band SAR images, *IEEE Transactions on Geoscience and Remote Sensing* 45, 3165–3171, doi: 10.1109/TGRS.2007.903698.
- Notarnicola, C., Angiulli, M., and Posa, F. (2006). Use of radar and optical remotely sensed data for soil moisture retrieval over vegetated areas, *IEEE Transactions on Geoscience and Remote Sensing* 44, 925–935, doi: 10.1109/TGRS.2006.872287.
- Obermeier, V., Scheipl, F., Heumann, C., Wassermann, J., and Küchenhoff, H. (2015). Flexible distributed lags for modelling earthquake data, *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 64, 395–412, doi: 10.1111/rssc.12077.
- Oh, Y. and Kay, Y. C. (1998). Condition for precise measurement of soil surface roughness, *IEEE transactions on geoscience and remote sensing* 36, 691–695, doi: 10.1109/36.662751.
- Orth, R. and Seneviratne, S. I. (2012). Analysis of soil moisture memory from observations in Europe, *Journal of Geophysical Research: Atmospheres* 117, doi: 10.1029/2011JD017366.
- Osunbitan, J., Oyedele, D., and Adekalu, K. (2005). Tillage effects on bulk density, hydraulic conductivity and strength of a loamy sand soil in southwestern Nigeria, *Soil and Tillage Research* 82, 57–64, doi: 10.1016/j.still.2004.05.007.
- Pagliai, M., Vignozzi, N., and Pellegrini, S. (2004). Soil structure and the effect of management practices, *Soil and Tillage Research* 79, 131–143, doi: 10.1016/j.still.2004.07.002.
- Parzen, E. (1962). On estimation of a probability density function and mode, *The annals of mathematical statistics* 33, 1065–1076, doi: 10.1214/aoms/1177704472.
- Passioura, J. (1994). The yield of crops in relation to drought, *Physiology and determination of crop yield* pp. 343–359, doi: 10.1016/j.scitotenv.2018.10.434.
- Pathe, C., Wagner, W., Sabel, D., Doubkova, M., and Basara, J. B. (2009). Using ENVISAT ASAR global mode data for surface soil moisture retrieval over Oklahoma, USA, *IEEE Transactions on Geoscience and Remote Sensing* 47, 468–480, doi: 10.1109/TGRS.2008.2004711.
- Paul, B. K., Vanlauwe, B., Ayuke, F., Gassner, A., Hoogmoed, M., Hurisso, T., Koala, S., Lelei, D., Nd-

- abamenye, T., Six, J., et al. (2013). Medium-term impact of tillage and residue management on soil aggregate stability, soil carbon and crop productivity, *Agriculture, ecosystems & environment* 164, 14–22, doi: 10.1016/j.agee.2012.10.003.
- Paul, C. and de Vries, J. (1979). Prediction of soil strength from hydrologic and mechanical properties, *Canadian Journal of Soil Science* 59, 301–311, doi: 10.4141/cjss79-034.
- Pearl, J.: (2009). *Causality*, Cambridge university press, doi: 10.1017/CBO9780511803161.
- Penna, D., Brocca, L., Borga, M., and Dalla Fontana, G. (2013). Soil moisture temporal stability at different depths on two alpine hillslopes during wet and dry periods, *Journal of Hydrology* 477, 55–71, doi: 10.1016/j.jhydrol.2012.10.052.
- Petropoulos, G. P., Srivastava, P. K., Piles, M., and Pearson, S. (2018). Earth observation-based operational estimation of soil moisture and evapotranspiration for agricultural crops in support of sustainable water management, *Sustainability* 10, 181, doi: 10.3390/su10010181.
- Pezij, M., Benninga, H., Carranza, C., Augustijn, D. C., van der Velde, R., and van der Ploeg, M. (2016). Bodemvocht uit satellietdata voor optimalisatie waterbeheer, *Land+ water* 56, 26–27.
- Pezij, M., Augustijn, D. C., Hendriks, D. M., and Hulscher, S. J. (2019a). The role of evidence-based information in regional operational water management in the Netherlands, *Environmental science & policy* 93, 75–82, doi: 10.1016/j.envsci.2018.12.025.
- Pezij, M., Augustijn, D. C., Hendriks, D. M., Weerts, A. H., Hummel, S., van der Velde, R., and Hulscher, S. J. (2019b). State updating of root zone soil moisture estimates of an unsaturated zone metamodel for operational water resources management, *Journal of Hydrology X* 4, 100 040, doi: 10.1016/j.hydroa.2019.100040.
- Pezij, M., Augustijn, D. C., Hendriks, D. M., and Hulscher, S. J. (2020). Applying transfer function-noise modelling to characterize soil moisture dynamics: A data-driven approach using remote sensing data, *Environmental Modelling & Software* p. 104756, doi: 10.1016/j.envsoft.2020.104756.
- Pierdicca, N., Pulvirenti, L., and Pace, G. (2013). A prototype software package to retrieve soil moisture from Sentinel-1 data by using a bayesian multitemporal algorithm, *IEEE Journal of selected topics in applied earth observations and remote sensing* 7, 153–166, doi: 10.1109/JSTARS.2013.2257698.
- Porporato, A., Laio, F., Ridolfi, L., Caylor, K. K., and Rodriguez-Iturbe, I. (2003). Soil moisture and plant stress dynamics along the Kalahari precipitation gradient, *Journal of Geophysical Research: Atmospheres* 108, doi: 10.1029/2002JD002448.
- Porporato, A., Daly, E., and Rodriguez-Iturbe, I. (2004). Soil water balance and ecosystem response to climate change, *The American Naturalist* 164, 625–632, doi: 10.1086/424970.
- Powlson, D. S., Gregory, P. J., Whalley, W. R., Quinton, J. N., Hopkins, D. W., Whitmore, A. P., Hirsch, P. R., and Goulding, K. W. (2011). Soil management in relation to sustainable agriculture and ecosystem services, *Food policy* 36, S72–S87, doi: 10.1016/j.foodpol.2010.11.025.
- Prasad, R., Deo, R. C., Li, Y., and Maraseni, T. (2018). Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition, *Geoderma* 330, 136–161, doi: 10.1016/j.geoderma.2018.05.035.
- Prevot, L., Dechambre, M., Taconet, O., Vidal-Madjar, D., Normand, M., and Gallej, S. (1993). Estimating the characteristics of vegetation canopies with airborne radar measurements, *International Journal of Remote Sensing* 14, 2803–2818, doi: 10.1080/01431169308904310.
- Priddy, J. D. and Willoughby, W. E. (2006). Clarification of vehicle cone index with reference to mean maximum pressure, *Journal of terramechanics* 43, 85–96, doi: 10.1016/j.jterra.2004.10.001.
- Qiu, J., Crow, W. T., Nearing, G. S., Mo, X., and Liu, S. (2014). The impact of vertical measurement depth on the information content of soil moisture times series data, *Geophysical Research Letters* 41, 4997–5004, doi: 10.1002/2014GL060017.

- Qiu, Y., Fu, B., Wang, J., and Chen, L. (2001). Soil moisture variation in relation to topography and land use in a hillslope catchment of the Loess Plateau, China, *Journal of Hydrology* 240, 243–263, doi: 10.1016/S0022-1694(00)00362-0.
- Rabot, E., Wiesmeier, M., Schlüter, S., and Vogel, H.-J. (2018). Soil structure as an indicator of soil functions: a review, *Geoderma* 314, 122–137, doi: 10.1016/j.geoderma.2017.11.009.
- Raghavan, G., Alvo, P., and McKyes, E. (1990). Soil Compaction in Agriculture: A View Toward Managing the Problem, in: *Advances in Soil Science: Soil Degradation*, edited by Lal, R. and Stewart, B., pp. 1–36, Springer New York, New York, NY, doi: 10.1007/978-1-4612-3322-0.
- Rakotoarivony, L., Taconet, O., Vidal-Madjar, D., Bellemain, P., and Benallegue, M. (1996). Radar backscattering over agricultural bare soils, *Journal of Electromagnetic Waves and Applications* 10, 187–209, doi: 10.1163/156939396X00964.
- Rao, K., Williams, A. P., Flefil, J. F., and Konings, A. G. (2020). SAR-enhanced mapping of live fuel moisture content, *Remote Sensing of Environment* 245, 111 797, doi: 10.1016/j.rse.2020.111797.
- Raper, R. (2005). Agricultural traffic impacts on soil, *Journal of Terramechanics* 42, 259–280, doi: 10.1016/j.jterra.2004.10.010.
- Rawls, W., Gish, T., and Brakensiek, D. (1991). Estimating soil water retention from soil physical properties and characteristics, in: *Advances in soil science*, pp. 213–234, Springer, doi: 10.1007/978-1-4612-3144-8.5.
- Read, J. S., Jia, X., Willard, J., Appling, A. P., Zwart, J. A., Oliver, S. K., Karpatne, A., Hansen, G. J., Hanson, P. C., Watkins, W., et al. (2019). Process-guided deep learning predictions of lake water temperature, *Water Resources Research* 55, 9173–9190, doi: 10.1029/2019WR024922.
- Reichle, R. H. (2008). Data assimilation methods in the Earth sciences, *Advances in water resources* 31, 1411–1418, doi: 10.1016/j.advwatres.2008.01.001.
- Reichle, R. H. and Koster, R. D. (2004). Bias reduction in short records of satellite soil moisture, *Geophysical Research Letters* 31, doi: 10.1029/2004GL020938.
- Reichle, R. H., Koster, R. D., Dong, J., and Berg, A. A. (2004). Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation, *Journal of Hydrometeorology* 5, 430–442, doi: 10.1029/2002JD003334.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., et al. (2019). Deep learning and process understanding for data-driven Earth system science, *Nature* 566, 195–204, doi: 10.1038/s41586-019-0912-1.
- Reintam, E., Vennik, K., Kukk, L., Kade, S., Krebstein, K., Are, M., and Astover, A. (2016). Measuring and predicting soil moisture conditions for trafficability, *Acta Agriculturae Scandinavica* 66, 698–705, doi: 10.1080/09064710.2016.1210210.
- Renzullo, L. J., Van Dijk, A., Perraud, J.-M., Collins, D., Henderson, B., Jin, H., Smith, A., and McJannet, D. (2014). Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment, *Journal of Hydrology* 519, 2747–2762, doi: 10.1016/j.jhydrol.2014.08.008.
- Rigden, A. J., Mueller, N. D., Holbrook, N. M., Pillai, N., and Huybers, P. (2020). Combined influence of soil moisture and atmospheric evaporative demand is important for accurately predicting US maize yields, *Nature Food* 1, 127–133, doi: 10.1038/s43016-020-0028-7.
- Ritter, A., Hupet, F., Muñoz-Carpena, R., Lambot, S., and Vanclooster, M. (2003). Using inverse methods for estimating soil hydraulic properties from field data as an alternative to direct methods, *Agricultural Water Management* 59, 77–96, doi: 10.1016/S0378-3774(02)00160-9.
- Rivers, M., Coles, N., Zia, H., Harris, N. R., and Yates, R. (2015). How could sensor networks help with agricultural water management issues? Optimizing irrigation scheduling through networked soil-moisture sensors, in: *2015 IEEE sensors applications symposium (SAS)*, pp. 1–6, IEEE, doi: 10.1109/SAS.2015.7133593.

- Robertson, G. P. and Swinton, S. M. (2005). Reconciling agricultural productivity and environmental integrity: a grand challenge for agriculture, *Frontiers in Ecology and the Environment* 3, 38–46, doi: 10.1890/1540-9295(2005)003[0038:RAPAE]2.0.CO;2.
- Robinson, D. A., Jones, S. B., Lebron, I., Reinsch, S., Domínguez, M. T., Smith, A. R., Jones, D. L., Marshall, M. R., and Emmett, B. A. (2016). Experimental evidence for drought induced alternative stable states of soil moisture, *Scientific reports* 6, 1–6, doi: 10.1038/srep20018.
- Robinson, D. A., Hopmans, J. W., Filipovic, V., van der Ploeg, M., Lebron, I., Jones, S. B., Reinsch, S., Jarvis, N., and Tuller, M. (2019). Global environmental changes impact soil hydraulic functions through biophysical feedbacks, *Global change biology* 25, 1895–1904, doi: 10.1111/gcb.14626.
- Rodriguez-Iturbe, I., Porporato, A., Ridolfi, L., Isham, V., and Cox, D. (1999). Probabilistic modelling of water balance at a point: the role of climate, soil and vegetation, *Proceedings of the Royal Society of London Series A* 455, 3789, doi: 10.1098/rspa.1999.0477.
- Roger-Estrade, J., Anger, C., Bertrand, M., and Richard, G. (2010). Tillage and soil ecology: partners for sustainable agriculture, *Soil and Tillage Research* 111, 33–40, doi: 10.1016/j.still.2010.08.010.
- Roseberg, R. and McCoy, E. (1992). Tillage-and traffic-induced changes in macroporosity and macropore continuity: Air permeability assessment, *Soil Science Society of America Journal* 56, 1261–1267, doi: 10.2136/sssaj1992.03615995005600040042x.
- Rosenbaum, U., Bogena, H. R., Herbst, M., Huisman, J., Peterson, T., Weuthen, A., Western, A., and Vereecken, H. (2012). Seasonal and event dynamics of spatial soil moisture patterns at the small catchment scale, *Water Resources Research* 48, doi: 10.1029/2011WR011518.
- Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function, *The Annals of Mathematical Statistics* pp. 832–837.
- Rücknagel, J., Hofmann, B., Deumelandt, P., Reinicke, F., Bauhardt, J., Hülsbergen, K.-J., and Christen, O. (2015). Indicator based assessment of the soil compaction risk at arable sites using the model REPRO, *Ecological indicators* 52, 341–352, doi: 10.1016/j.ecolind.2014.12.022.
- Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M. D., Muñoz-Marí, J., et al. (2019). Inferring causation from time series in Earth system sciences, *Nature communications* 10, 1–13, doi: 10.1038/s41467-019-10105-3.
- Sammartino, S., Michel, E., and Capowicz, Y. (2012). A novel method to visualize and characterize preferential flow in undisturbed soil cores by using multislice helical CT, *Vadose Zone Journal* 11, doi: 10.2136/vzj2011.0100.
- Sano, E. E., Huete, A. R., Troufleau, D., Moran, M. S., and Vidal, A. (1998). Relation between ERS-1 synthetic aperture radar data and measurements of surface roughness and moisture content of rocky soils in a semiarid rangeland, *Water Resources Research* 34, 1491–1498, doi: 10.1029/98WR00032.
- Scanlan, C. A. (2009). *Processes and effects of root-induced changes to soil hydraulic properties*, Ph.D. thesis.
- Schaap, M. G., Leij, F. J., and Van Genuchten, M. T. (2001). Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions, *Journal of hydrology* 251, 163–176, doi: 10.1016/S0022-1694(01)00466-8.
- Schneider, K., Huisman, J., Breuer, L., Zhao, Y., and Frede, H.-G. (2008). Temporal stability of soil moisture in various semi-arid steppe ecosystems and its application in remote sensing, *Journal of Hydrology* 359, 16–29, doi: 10.1016/j.jhydrol.2008.06.016.
- Scholl, P., Leitner, D., Kammerer, G., Loiskandl, W., Kaul, H.-P., and Bodner, G. (2014). Root induced changes of effective 1D hydraulic properties in a soil column, *Plant and soil* 381, 193–213, doi: 10.1007/s11104-014-2121-x.
- Schwen, A., Bodner, G., and Loiskandl, W. (2011a). Time-variable soil hydraulic properties in near-surface soil water simulations for different tillage methods, *Agricultural Water Management* 99, 42–50, doi: 10.1016/j.agwat.2011.07.020.

- Schwen, A., Bodner, G., Scholl, P., Buchan, G. D., and Loiskandl, W. (2011b). Temporal dynamics of soil hydraulic properties and the water-conducting porosity under different tillage, *Soil and Tillage Research* 113, 89–98, doi: 10.1016/j.still.2011.02.005.
- Scott, D.: (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization*, A Wiley-interscience publication, Wiley, doi: 10.1002/9780470316849.
- Sellin, A. (1999). Does pre-dawn water potential reflect conditions of equilibrium in plant and soil water status?, *Acta oecologica* 20, 51–59, doi: 10.1016/S1146-609X(99)80015-0.
- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate: A review, *Earth-Science Reviews* 99, 125–161, doi: 10.1016/j.earscirev.2010.02.004.
- Seneviratne, S. I., Nicholls, N., Easterling, D., Goodess, C. M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., et al. (2012). Changes in climate extremes and their impacts on the natural physical environment, in: *Managing the risks of extreme events and disasters to advance climate change adaptation: Special Report of the Intergovernmental Panel on Climate Change*, pp. 109–230, Cambridge University Press.
- Setiyono, T., Quicho, E., Holecz, F., Khan, N., Romuga, G., Maunahan, A., Garcia, C., Rala, A., Raviz, J., Collivignarelli, F., et al. (2019). Rice yield estimation using synthetic aperture radar (SAR) and the ORYZA crop growth model: development and application of the system in South and South-east Asian countries, *International Journal of Remote Sensing* 40, 8093–8124, doi: 10.1080/01431161.2018.1547457.
- Shipitalo, M., Dick, W., and Edwards, W. (2000). Conservation tillage and macropore factors that affect water movement and the fate of chemicals, *Soil and tillage research* 53, 167–183, doi: 10.1016/S0167-1987(99)00104-X.
- Shumway, R. H. and Stoffer, D. S.: (2010). *Time series analysis and its applications: with R examples*, Springer Science & Business Media.
- Simunek, J., Van Genuchten, M. T., and Sejna, M. (2005). The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media, Tech. rep.
- Sluiter, R. (2012). Interpolation methods for the climate atlas, World Bank Policy Research Working Paper TR-335, Koninkrijk Nederlands Meteorologisch Instituut (KNMI), De Bilt, the Netherlands.
- Soenario, I., Plieger, M., and Sluiter, R. (2010). Optimization of Rainfall Interpolation, Tech. rep., Koninkrijk Nederlands Meteorologisch Instituut (KNMI), URL <http://bibliotheek.knmi.nl/knmipubIR/IR2010-01.pdf>.
- Sojka, R., Busscher, W., and Lehrs, G. (2001). In situ strength, bulk density, and water content relationships of a Durinodic Xeric Haplocalcid soil, *Soil science* 166, 520–529.
- Srayeddin, I. and Doussan, C. (2009). Estimation of the spatial variability of root water uptake of maize and sorghum at the field scale by electrical resistivity tomography, *Plant and soil* 319, 185–207, doi: 10.1007/s11104-008-9860-5.
- Srivastava, P. K., Han, D., Ramirez, M. R., and Islam, T. (2013). Machine learning techniques for downscaling SMOS satellite soil moisture using MODIS land surface temperature for hydrological application, *Water resources management* 27, 3127–3144, doi: 10.1007/s11269-013-0337-9.
- Srivastava, P. K., O'Neill, P., Cosh, M., Lang, R., and Joseph, A. (2015). Evaluation of radar vegetation indices for vegetation water content estimation using data from a ground-based SMAP simulator, in: 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 1296–1299, doi: 10.1109/IGARSS.2015.7326012.
- Sriwongsitanon, N., Gao, H., Savenije, H. H., Maekan, E., Saengsawang, S., and Thianpopirug, S. (2016). Comparing the Normalized Difference Infrared Index (NDII) with root zone storage in a lumped conceptual model, *Hydrology and Earth System Sciences* 20, 3361, doi: 10.5194/hess-20-3361-2016.
- Steele-Dunne, S. C., Friesen, J., and van de Giesen, N. (2012). Using diurnal variation in backscatter to

- detect vegetation water stress, *IEEE Transactions on Geoscience and Remote Sensing* 50, 2618–2629, doi: 10.1109/TGRS.2012.2194156.
- Steele-Dunne, S. C., McNairn, H., Monsivais-Huertero, A., Judge, J., Liu, P.-W., and Papathanassiou, K. (2017). Radar remote sensing of agricultural canopies: A review, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10, 2249–2273, doi: 10.1109/JSTARS.2016.2639043.
- Steinbach, M., Tan, P.-N., Kumar, V., Potter, C., Klooster, S., and Torregrosa, A. (2001). Clustering earth science data: Goals, issues and results, in: Proc. of the Fourth KDD Workshop on Mining Scientific Datasets.
- Stedle, E. (2000). Water uptake by plant roots: an integration of views, *Plant and soil* 226, 45–56, doi: 10.1023/A:1026439226716.
- Strudley, M. W., Green, T. R., and Ascough II, J. C. (2008). Tillage effects on soil hydraulic properties in space and time: State of the science, *Soil and Tillage Research* 99, 4–48, doi: 10.1016/j.still.2008.01.007.
- Tardieu, F., Draye, X., and Javaux, M. (2017). Root water uptake and ideotypes of the root system: Whole-plant controls matter, *Vadose Zone Journal* 16, 1–10, doi: 10.2136/vzj2017.05.0107.
- Toreti, A., Belward, A., Perez-Dominguez, I., Naumann, G., Luterbacher, J., Cronie, O., Seguini, L., Manfron, G., Lopez-Lozano, R., Baruth, B., et al. (2019). The exceptional 2018 European water seesaw calls for action on adaptation, *Earth's Future* 7, 652–663, doi: 10.1029/2019EF001170.
- Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., et al. (2012). GMES Sentinel-1 mission, *Remote Sensing of Environment* 120, 9–24, doi: 10.1016/j.rse.2011.05.028.
- Tyralis, H., Papacharalampous, G., and Langousis, A. (2019). A brief review of random forests for water scientists and practitioners and their recent history in water resources, *Water* 11, 910, doi: 10.3390/w11050910.
- Ulaby, F., Moore, R., and Fung, A.: (1982). *Microwave remote sensing: Active and passive. Volume 2-Radar remote sensing and surface scattering and emission theory*, Addison-Wesley Publishers, Reading, Mass.
- Ulaby, F. T., Moore, R. K., and Fung, A. K.: (1986). *Microwave Remote Sensing Active and Passive-Volume II: Radar Remote Sensing and Surface Scattering and Emission Theory*, Addison-Wesley Publishing Company/Advanced Book Program/World Science Division.
- Ulaby, F. T., Dubois, P. C., and Van Zyl, J. (1996). Radar mapping of surface soil moisture, *Journal of Hydrology* 184, 57–84, doi: 10.1016/0022-1694(95)02968-0.
- Upadhyaya, S. (2005). Stress-Strain and Soil Strength, in: Encyclopedia of Soils in the Environment, edited by Hillel, D., pp. 38 – 54, Elsevier, Oxford, doi: 10.1016/B0-12-348530-4/00507-5.
- Vachaud, G., Passerat de Silans, A., Balabanis, P., and Vauclin, M. (1985). Temporal stability of spatially measured soil water probability density function, *Soil Science Society of America Journal* 49, 822–828, doi: 10.2136/sssaj1985.03615995004900040006x.
- Van der Velde, R. and Su, Z. (2009). Dynamics in land-surface conditions on the Tibetan Plateau observed by Advanced Synthetic Aperture Radar (ASAR), *Hydrological sciences journal* 54, 1079–1093, doi: 10.1623/hysj.54.6.1079.
- van der Velde, R., Su, Z., van Oevelen, P., Wen, J., Ma, Y., and Salama, M. S. (2012). Soil moisture mapping over the central part of the Tibetan Plateau using a series of ASAR WS images, *Remote sensing of environment* 120, 175–187, doi: 10.1016/j.rse.2011.05.029.
- van der Velde, R., Salama, M. S., Eweys, O. A., Wen, J., and Wang, Q. (2015). Soil moisture mapping using combined active/passive microwave observations over the East of the Netherlands, *IEEE Journal of selected topics in applied earth observations and remote sensing* 8, 4355–4372, doi: 10.1109/JSTARS.2014.2353692.
- van der Velde, R., Benninga, H.-J., Pezij, M., Augustijn, D., Carranza, C., van der Ploeg, M., and Hendriks, D. (2018). Bodemvocht uit satellietdata: wat kan de Nederlandse waterbeheerder ermee?, *H2O online*

- 2018.
- Van Dijck, S. and Van Asch, T. W. (2002). Compaction of loamy soils due to tractor traffic in vineyards and orchards and its effect on infiltration in southern France, *Soil and Tillage Research* 63, 141–153, doi: 10.1016/S0167-1987(01)00237-9.
- van Emmerik, T., Steele-Dunne, S. C., Judge, J., and van de Giesen, N. (2015). Impact of diurnal variation in vegetation water content on radar backscatter from maize during water stress, *IEEE Transactions on Geoscience and Remote Sensing* 53, 3855–3869, doi: 10.1109/TGRS.2014.2386142.
- Van Emmerik, T., Steele-Dunne, S. C., Judge, J., and Van De Giesen, N. (2016). Dielectric response of corn leaves to water stress, *IEEE Geoscience and Remote Sensing Letters* 14, 8–12, doi: 10.1109/LGRS.2016.2606662.
- van Emmerik, T., Steele-Dunne, S., Paget, A., Oliveira, R. S., Bittencourt, P. R., Barros, F. d. V., and van de Giesen, N. (2017). Water stress detection in the Amazon using radar, *Geophysical Research Letters* 44, 6841–6849, doi: 10.1002/2017GL073747.
- van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil science society of America journal* 44, 892–898, doi: 10.2136/sssaj1980.03615995004400050002x.
- Van Groenigen, J. and Stein, A. (1998). Constrained optimization of spatial sampling using continuous simulated annealing, *Journal of Environmental Quality* 27, 1078–1086, doi: 10.2134/jeq1998.00472425002700050013x.
- Van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., Montzka, C., Nemes, A., Pachepsky, Y. A., Padarian, J., et al. (2017). Pedotransfer functions in Earth system science: challenges and perspectives, *Reviews of Geophysics* 55, 1199–1256, doi: 10.1002/2017RG000581.
- Vanderlinden, K., Vereecken, H., Hardelauf, H., Herbst, M., Martínez, G., Cosh, M. H., and Pachepsky, Y. A. (2012). Temporal stability of soil water contents: A review of data and analyses, *Vadose Zone Journal* 11, vzj2011–0178, doi: 10.2136/vzj2011.0178.
- Vaz, C. M., Bassoi, L. H., and Hopmans, J. W. (2001). Contribution of water content and bulk density to field soil penetration resistance as measured by a combined cone penetrometer–TDR probe, *Soil and Tillage Research* 60, 35–42, doi: 10.1016/S0167-1987(01)00173-8.
- Vaz, C. M., Manieri, J. M., De Maria, I. C., and Tuller, M. (2011). Modeling and correction of soil penetration resistance for varying soil water content, *Geoderma* 166, 92–101, doi: 10.1016/j.geoderma.2011.07.016.
- Vaz, C. M., Jones, S., Meding, M., and Tuller, M. (2013). Evaluation of standard calibration functions for eight electromagnetic soil moisture sensors, *Vadose Zone Journal* 12, 1–16, doi: 10.2136/vzj2012.0160.
- Vellidis, G., Tucker, M., Perry, C., Kvien, C., and Bednarz, C. (2008). A real-time wireless smart sensor array for scheduling irrigation, *Computers and electronics in agriculture* 61, 44–50, doi: 10.1016/j.compag.2007.05.009.
- Vereecken, H., Huisman, J., Bogena, H., Vanderborght, J., Vrugt, J., and Hopmans, J. (2008). On the value of soil moisture measurements in vadose zone hydrology: A review, *Water resources research* 44, doi: 10.1029/2008WR006829.
- Vereecken, H., Huisman, J., Pachepsky, Y., Montzka, C., Van Der Kruk, J., Bogena, H., Weihermüller, L., Herbst, M., Martinez, G., and Vanderborght, J. (2014). On the spatio-temporal dynamics of soil moisture at the field scale, *Journal of Hydrology* 516, 76–96, doi: 10.1016/j.jhydrol.2013.11.061.
- Vogel, M. M., Zscheischler, J., Wartenburger, R., Dee, D., and Seneviratne, S. I. (2019). Concurrent 2018 hot extremes across Northern Hemisphere due to human-induced climate change, *Earth's future* 7, 692–703, doi: 10.1029/2019EF001189.
- Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., and Strauss, P. (2018). Sensitivity of Sentinel-1 backscatter to vegetation dynamics: An Austrian case study, *Remote Sensing* 10, 1396, doi: 10.3390/rs10091396.



- 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, 191–207, doi: 0.1016/S0034-4257(99)00036-X.
- Wagner, W., Blöschl, G., Pampaloni, P., Calvet, J.-C., Bizzarri, B., Wigneron, J.-P., and Kerr, Y. (2007). Operational readiness of microwave remote sensing of soil moisture for hydrologic applications, *Hydrology Research* 38, 1–20, doi: 10.2166/nh.2007.029.
- Wagner, W., Sabel, D., Doubkova, M., Bartsch, A., and Pathe, C. (2009). The potential of Sentinel-1 for monitoring soil moisture with a high spatial resolution at global scale, in: Symposium of Earth Observation and Water Cycle Science.
- Wang, L., Manzoni, S., Ravi, S., Riveros-Iregui, D., and Caylor, K. (2015). Dynamic interactions of eco-hydrological and biogeochemical processes in water-limited systems, *Ecosphere* 6, 1–27, doi: 10.1890/ES15-00122.1.
- Wang, Y., Shao, M., Liu, Z., and Warrington, D. N. (2012). Regional spatial pattern of deep soil water content and its influencing factors, *Hydrological Sciences Journal* 57, 265–281, doi: 10.1080/02626667.2011.644243.
- Wang-Erlandsson, L., Bastiaanssen, W. G., Gao, H., Jägermeyr, J., Senay, G. B., van Dijk, A. I., Guerschman, J. P., Keys, P. W., Gordon, L. J., and Savenije, H. H. (2016). Global root zone storage capacity from satellite-based evaporation, *Hydrol. Earth Syst. Sci* 20, 1459–1481, doi: 10.5194/hess-20-1459-2016.
- Western, A. W., Blöschl, G., and Grayson, R. B. (1998). Geostatistical characterisation of soil moisture patterns in the Tarrawarra catchment, *Journal of Hydrology* 205, 20–37, doi: 10.1016/S0022-1694(97)00142-X.
- Whalley, W. R., Riseley, B., Leeds-Harrison, P. B., Bird, N. R., Leech, P. K., and Adderley, W. P. (2005). Structural differences between bulk and rhizosphere soil, *European Journal of Soil Science* 56, 353–360, doi: 10.1111/j.1365-2389.2004.00670.x.
- Wieder, W. L. and Shoop, S. A. (2017). Vegetation Impact on Soil Strength: A State of the Knowledge Review, Tech. rep., U.S. Army Engineer Research and Development Center (ERDC), Cold Regions Research and Engineering Laboratory, doi: 10.1016/j.jterra.2018.03.006.
- Wilson, D. J., Western, A. W., Grayson, R. B., Berg, A. A., Lear, M. S., Rodell, M., Famiglietti, J. S., Woods, R. A., and McMahon, T. A. (2003). Spatial distribution of soil moisture over 6 and 30cm depth, Mahurangi river catchment, New Zealand, *Journal of Hydrology* 276, 254–274, doi: 10.1016/S0022-1694(03)00060-X.
- Wilson, D. J., Western, A. W., and Grayson, R. B. (2004). Identifying and quantifying sources of variability in temporal and spatial soil moisture observations, *Water Resources Research* 40, doi: 10.1029/2003WR002306.
- Wood, S.: (2006a). *Generalized Additive Models: An Introduction with R*, CRC Press.
- Wood, S. N. (2006b). Low-Rank Scale-Invariant Tensor Product Smooths for Generalized Additive Mixed Models, *Biometrics* 62, 1025–1036, doi: 10.1111/j.1541-0420.2006.00574.x.
- Wösten, J., Veerman, G., de Groot, W., and Stolte, J. (2001). Waterretentie-en doorlatendheidskarakteristieken van boven-en ondergronden in Nederland: de Staringreeks; vernieuwde uitgave 2001, Tech. rep., Alterra.
- Wosten, J., de Vries, F., Hoogland, T., Massop, H., Veldhuizen, A., Vroon, H., Wesseling, J., Heijkers, J., and Bolman, A. (2013). BOFEK2012, de nieuwe bodemfysische schematisatie van Nederland, Tech. rep., Alterra.
- Wright, M. N. and Ziegler, A. (2015). ranger: A fast implementation of random forests for high dimensional data in C++ and R, *arXiv preprint arXiv:1508.04409*.
- Wu, W., Xiao, Y., Li, G., Zeng, W., Lin, H., Rutherford, S., Xu, Y., Luo, Y., Xu, X., Chu, C., et al. (2013). Temperature–mortality relationship in four subtropical Chinese cities: a time-series study using a distributed lag non-linear model, *Science of the Total Environment* 449, 355–362, doi: 10.1016/j.scitotenv.2013.01.090.
- Xu, Y.-P. and Tung, Y.-K. (2008). Decision-making in water management under uncertainty, *Water resources*

- management* 22, 535–550, doi: 10.1007/s11269-007-9176-x.
- Yu, G.-R., Zhuang, J., Nakayama, K., and Jin, Y. (2007). Root water uptake and profile soil water as affected by vertical root distribution, *Plant Ecology* 189, 15–30, doi: 10.1007/s11258-006-9163-y.
- Zanobetti, A., Schwartz, J., Samoli, E., Gryparis, A., Touloumi, G., Atkinson, R., Le Tertre, A., Bobros, J., Celko, M., Goren, A., et al. (2002). The temporal pattern of mortality responses to air pollution: a multicity assessment of mortality displacement, *Epidemiology* 13, 87–93, doi: 10.1097/00001648-200201000-00014.
- Zhang, D., Zhang, W., Huang, W., Hong, Z., and Meng, L. (2017). Upscaling of surface soil moisture using a deep learning model with VIIRS RDR, *ISPRS International Journal of Geo-Information* 6, 130.
- Zhang, H. and Oweis, T. (1999). Water–yield relations and optimal irrigation scheduling of wheat in the Mediterranean region, *Agricultural Water Management* 38, 195–211, doi: 10.1016/S0378-3774(98)00069-9.
- Zhao, W. L., Gentine, P., Reichstein, M., Zhang, Y., Zhou, S., Wen, Y., Lin, C., Li, X., and Qiu, G. Y. (2019). Physics-constrained machine learning of evapotranspiration, *Geophysical Research Letters* 46, 14 496–14 507, doi: 10.1029/2019GL085291.
- Zhao, Y., Peth, S., Wang, X., Lin, H., and Horn, R. (2010). Controls of surface soil moisture spatial patterns and their temporal stability in a semi-arid steppe, *Hydrological Processes* 24, 2507–2519, doi: 10.1002/hyp.7665.
- Zheng, H., Liu, W., Zheng, J., Luo, Y., Li, R., Wang, H., and Qi, H. (2018). Effect of long-term tillage on soil aggregates and aggregate-associated carbon in black soil of Northeast China, *PLoS One* 13, e0199523, doi: 10.1371/journal.pone.0199523.
- Zhu, Q. and Lin, H. (2011). Influences of soil, terrain, and crop growth on soil moisture variation from transect to farm scales, *Geoderma* 163, 45–54, doi: 10.1016/j.geoderma.2011.03.015.
- Zhuang, R., Zeng, Y., Manfreda, S., and Su, Z. (2020). Quantifying Long-Term Land Surface and Root Zone Soil Moisture over Tibetan Plateau, *Remote Sensing* 12, doi: 10.3390/rs12030509.
- Zotarelli, L., Dukes, M., Scholberg, J., Femminella, K., and Munoz-Carpena, R. (2011). Irrigation scheduling for green bell peppers using capacitance soil moisture sensors, *Journal of irrigation and drainage engineering* 137, 73–81, doi: 10.1061/(ASCE)IR.1943-4774.0000281.

# Data Availability

**Chapter 2:** Data from the Twente soil moisture network can be made available upon request from the Water Resources Department, Faculty of Geo-information and Earth Observation (ITC), Twente, University

**Chapter 3:** Carranza, Coleen; Gertsen, Henny; Peters, Piet; van der Ploeg, M.J. (2020): Spatio-temporal subsurface soil moisture (40 cm depth) within an agricultural field in the Netherlands. 4TU.ResearchData. Dataset. <https://doi.org/10.4121/13376960>

**Chapter 4:** Underlying datasets can be requested from Coleen Carranza with exception of the soil moisture measurements from Twente network, which can be made available upon request from the Water Resources Department, Faculty of Geo-information and Earth Observation (ITC), Twente, University

**Chapter 5:** Carranza, C.D. (2019): Cone index and surface soil moisture measurements in selected agricultural fields within the Raam and Twente soil moisture monitoring networks. 4TU.ResearchData. Dataset. <https://doi.org/10.4121/uuid:a451a8b6-b9f1-4cb5-b390-92778e7466f5>

**Chapter 6:** Carranza, Coleen; Nolet, Corjan; Pezij, Michiel; van der Ploeg, M.J. (2020): Data for the article: "Root zone soil moisture estimation with Random Forests". 4TU.ResearchData. Dataset. <https://doi.org/10.4121/13148231>

## **Complete dataset in the Raam soil moisture network:**

Benninga, H.F.; Carranza, C.D.; van der Ploeg, M.J.; Augustijn, D.C.M.; van der Velde, R.; Pezij, Michiel: Regional soil moisture monitoring network in the Raam catchment in the Netherlands. 4TU.ResearchData. Collection. <https://doi.org/10.4121/uuid:2411bbb8-2161-4f31-985f-7b65b8448bc9>

Benninga, H.F.; Carranza, C.D.; Pezij, Michiel; van der Ploeg, M.J.; Augustijn, D.C.M.; van der Velde, R. (2020): Regional soil moisture monitoring network in the Raam catchment in the Netherlands - 2018-04 / 2019-04. 4TU.ResearchData. Dataset. <https://doi.org/10.4121/uuid:b68e3971-c73e-4d7f-b52f-9ef7d7fe1ed2>



# Summary

Since modern day agriculture is commonly carried out over large areas, efficient land and water management strategies are essential in order to meet increasing crop productivity and resource demands. Soil moisture plays a central role in determining efficient agricultural management strategies because it provides information crucial for the selection of optimal tillage practices and for determining irrigation schedules. Radar satellites have been a source of soil moisture information at different spatio-temporal scales. Although they provide information for almost all weather and environmental conditions, they only measure soil moisture at the upper surface layer. This can be inadequate since most agricultural applications require soil moisture information over deeper layers (e.g. over rooting depth). This thesis therefore investigates the utility of radar satellites, particularly Sentinel-1, for agricultural water management by investigating surface and subsurface soil moisture dynamics. The main hypothesis in this research project is that the known sensitivity of radar satellites such as Sentinel-1 to surface soil moisture can be further exploited to gain insights and estimates of subsurface conditions, which can be further transformed into indicators relevant for agricultural water management.

The vertical variability between surface and subsurface soil moisture was investigated in Chapter 2. Two statistical methods are applied in order to examine soil moisture dynamics in agricultural areas using time series datasets. These enable in the identification of (de)coupled conditions that determine whether surface soil moisture can directly represent subsurface conditions. The results further highlight that the occurrence of decoupled conditions is not confined to dry periods, as commonly encountered in literature. Subsurface soil moisture dynamics, particularly the occurrence of preferential flow paths at measurement locations were inferred to have influenced decoupling during wetter surface conditions. The statistical methods applied to identify (de)coupled range are considered robust as no assumptions are made on the underlying relation between surface and subsurface soil moisture.

Subsurface soil moisture variability in a cultivated field was investigated in Chapter 3 to determine the impact of soil structure and vegetation activity on soil moisture dynamics, based on almost a year of situ soil moisture measurements. The results of the temporal stability analysis demonstrate that the contribution of vegetation to soil moisture variability is larger compared to soil structure. Furthermore, the impact of land management practices for different crop types can result in contrasting subsurface soil moisture contents within the same field. Temporally dynamic soil hydraulic properties were estimated from inverse modeling using in situ measurements. However, the changes in soil hydraulic properties may be challenging to determine since using a single type of flow regime has been observed to be insufficient.

Chapter 4 explores the potential of the vegetation backscatter to reflect root zone soil moisture conditions, especially during water-limited periods. Based on the the time series of Sentinel-1 backscatter for three growing seasons, a clear decrease in the total backscatter was observed during the 2018 European summer drought. Further investigation of the backscatter components using the Water Cloud model reveals that the soil backscatter correlates well with root zone soil moisture for non water-limited conditions. However, water-limited conditions highlights the ability of the vegetation backscatter to reflect root zone conditions based on good correlations obtained between the two. Unlike the soil backscatter which has been extensively studied in the past, the vegetation backscatter is deemed as an untapped source that can potentially allow direct estimation of root zone soil moisture from radar satellites.

During saturated or near-saturated conditions, assessment of field trafficability is important to ensure good vehicular mobility and to mitigate the present soil compaction rates in agricultural fields. Chapter 5 demonstrates a method to directly apply Sentinel-1-derived surface soil moisture for monitoring trafficability in agricultural fields. Using the statistical methods applied in Chapter 2, coupling between surface and topsoil layer was found to coincide with tillage periods, and facilitates the direct use of Sentinel-1 for estimating field trafficability. The field measurements of soil strength were related with surface soil moisture to generate a probabilistic measure of trafficability. Conditions favorable for traffic were found in the beginning of spring, and the changes in trafficability can be monitored further because of the high temporal frequency of Sentinel-1.

The dynamic and non-linear behaviour of soil moisture over time complicates the estimation of root zone soil moisture directly from surface soil moisture or using simple empirical relations during decoupled conditions. Chapter 6 demonstrates a data-driven machine learning approach using Random Forest to estimate daily root zone soil moisture based on in situ measurements. Similar to a process-based hydrological model with data assimilation, Random Forest achieves high, or even in slightly higher, accuracies for prediction and comparable accuracies with for forecasting of root zone soil moisture. Based on the model residuals, however, model estimates for extreme dry and wet conditions are more inaccurate than commonly encountered soil moisture states. It is inferred that poor learning of Random Forest for such infrequently encountered extreme conditions and inadequacy of the pore-flow model applied for the process-based model contributed to the said findings. One of the advantages of RF is it does not make assumptions on the system dynamics, implying that information on soil hydraulic properties are not required for prediction or forecasting.

The findings from the research chapters are discussed and tied together in Chapter 7. Overall, the results in this PhD research provide novel and innovative methods for applying Sentinel-1 in agricultural water management. The results can collectively be incorporated into an adaptive framework that supports operational water management. For saturated conditions at the beginning of the growing season when soil compaction is a concern, field trafficability can be directly estimated from Sentinel-1 derived surface soil moisture, which can assist farmers in determining the onset of tillage activities. Further into the growing season when adequate soil water supply is essential to ensure sufficient crop productivity, information on root zone soil moisture could potentially be derived from the vegetation backscatter especially for droughts

or by integrating Sentinel-1 surface soil moisture into data-driven methods. These methods may further be beneficial for data-poor regions which may not have accurate information on soil hydraulic properties. The results from this PhD further highlight that soil properties in agricultural areas are constantly changing and require dynamic soil functions and model structures in order for improved understanding of subsurface processes that are essential for accurate soil moisture estimation.





*♪ Coz I've got one hand in my pocket and the other one is giving a high-five ♪*

Alanis Morissette

## Acknowledgements

Every story has a backstory, and some of them quite literally. Behind each chapter are memorable interactions and experiences, both good and bad, that has made my thesis story whole. Before this booklet comes to a close, I would like to thank those who have been directly or least indirectly part of my research. And so my backstory began in Bangkok one day on September 2015, after coming from a short course and work holiday. On that day, I anxiously waited for it to be 16:00 local time when it would turn 10:00 in the Netherlands. I must have drank several cups of coffee and bubble teas by 16:00 when my interview for a PhD position began that I thought I hurriedly and nervously responded to the questions being asked. Of all of them, the only memorable one was "What about the Netherlands did you least like when you were here?". Compared to the technical and work related questions, this is the one I found hardest to answer. Surely I was thinking something like Dutch techno-sounding music but all I could manage to say was "Uuummm, the weather?". I wasn't particularly confident coming out of the interview that when I was informed of another succeeding meeting I thought: "Well, they must be really polite as they personally want to tell me through Skype that I didn't get the position". Surprisingly, the call turned out to be the opposite and has led to the next five challenging and exciting years.

To my promotor Coen and co-promotor Martine, the two people who interviewed me, I am grateful for giving me the opportunity to be part of SLM and to continue with my graduate degree. As the chairholder in SLM, Coen provided all the relevant resources that allowed me to carry out my field and lab measurements. Thanks for the freedom to explore different methods and to develop my own expertise.

To Martine, I would like to express my gratitude for mentoring me from the very beginning of my proposal up until the completion of my thesis manuscript. It was a pleasure not only to learn theories and techniques in soil physics from you but also to see what it's like and what it takes to be a scientist in our time. Among the many things you have shared with your students, it's your encouragement of their ideas and what they want to pursue, but at the same time being realistic of the potential outcomes, that helps keep their research afoot especially during those days when one constantly questions the necessity of doing a PhD. I appreciate the many advice you've given me for my work and career, for instance in building my scientific network. But yeah, it's still is a work-in-progress.

To the researchers I have collaborated with throughout my PhD, the scientific discussions have been inspiring for me to expand my knowledge but also my skills in communicating science.

The discussions can become very theoretical and philosophical that at times I lose sight of the original topic of our dialogues. Among those who I have collaborated with were members of the OWASIS project, in which my research was part of. Aside from the unique experience of being part of a Dutch-funded project, because of it I was also able to travel and briefly visit other places in NL outside the common tourist destinations. Although most of the times I literally could not follow the meetings in Dutch, the project gave me perspectives on what it is like to work overseas. To Michiel and Harm-jan, and their supervisors Denie and Rogier, being in the project allowed me to learn more about your expertise in water management and radar remote sensing. Apart from them, I would like to thank Paul Torfs for insightful statistical discussions for my first manuscript and Tim van Emmerik for guiding me through my radar manuscript.

I had to collect most of my datasets from the field and it would have been difficult without the assistance of Henny, but also Piet and Harm. Henny, thanks driving with me to Noord-Brabant and Twente regularly for around two years to install and collect measurements. Getting ready to leave WUR before 07:30, especially on very cold days (brrrr), was not particularly easy but our field works were always fun!

I've also learned a lot about agriculture in NL from the hospitable farmers during my field work. Most notable of them were Mr. and Mrs. Huup Lemmers, whose fields I frequented the most. Thank you for assisting during my numerous measurements and for allowing me to install my sensors for a few seasons in your fields. Heel hartelijk bedankt!

To the PhD students and SLM staff who were part of our lab meeting, I learned a great deal about your interesting science and our writing failures. Though there was this constant pressure of handing in a piece of our writing, our Friday morning meetings were nice breaks away from our screens and they provided refreshing interactions that always sent signals to my brain to stop working 'coz it's already weekend. I also enjoyed lunch conversations with the PhDs in SLM, although I was not a regular face among the group. It was a lot of fun being in a team with some of them for the Worm Charming challenge and the Soil health challenge. Funny how as a PhD, one's idea of enjoyment still has some serious elements to it. Seems very far from having regular night outs and getting headaches the mornings after. Perhaps it's part of what they call...adulting?

Thanks to Ricardo and Meng for being the paranymphs for my defence. I hope it was not too much pressure to ask you to start a work week sitting still for almost an hour. :)

Relocating halfway around the world to do my PhD meant having bouts of homesickness, which were certainly proportional to the drop in temperature in NL. So whenever I am able to go back home, I usually have a list of food that I miss while in NL. To my family in the Philippines, especially my mom and sister, thanks for feeding me (almost) all of my (fast)food wishes whenever I am home. They always make me gain weight but also regain energy and happiness that is very much needed when I return to do some serious PhD business.

To Corjan, thank you for being a big support throughout my PhD. Sharing my PhD with you was both good and bad. Good that we can have scientific discussions (or arguments?) and collaboration but bad that we continue with work-related stuff even when we come home.

Nevertheless, it was an experience worth having. Thanks for sharing the highs and the lows, for showing calmness to my volatility, for literally pulling me (and the WUR car haha) out of the mud, and most of all, for giving me a family. :)



## About the author



Coleen was born on July 20, 1987 in Daet, Camarines Norte, Philippines but she lived most of her childhood in Legazpi City, Philippines. She pursued Geology for her Bachelor's Degree in the University of the Philippines - Diliman from 2004-2008. After her studies, she successfully passed the licensure examination in order to work professionally as a Geologist in the Philippines in 2008, and was among the top 10 (6<sup>th</sup>) ranking examinees nationwide. She worked briefly as an exploration geologist in Cebu and afterwards as a lecturer in Geology in the University of the Philippines - Diliman from 2009 to 2011.

A scholarship grant from the Erasmus Mundus Mobility-Asia fund allowed her to take up an MSc degree from ITC - Twente University between 2011 to 2013, where she specialized in remote sensing and GIS methods for earth resources exploration (cum laude). From 2013 to 2014, she worked as a GIS specialist and geologist for landslide and flood hazards mapping program in several provinces in the Philippines. Then she was a researcher for a year (2014-2015) within the PhilLiDAR 2 project for mapping agricultural resources.

By the end of 2015, she started her PhD in Wageningen University. Her research was carried out within the OWASIS project where she focused on data-driven methods and applying geospatial analysis in investigating soil moisture dynamics. She was involved in setting up a soil moisture network in the southeastern part of the Netherlands, which was one of her fieldwork sites.

## Peer-reviewed publications related to this thesis

Benninga, H.-J. F., Carranza, C. D., Pezij, M., van Santen, P., van der Ploeg, M. J., Augustijn, D. C., and van der Velde, R. (2018). The Raam regional soil moisture monitoring network in the Netherlands, *Earth System Science Data* 10, 61, doi: 10.5194/essd-10-61-2018

Carranza, C., Van der Ploeg, M. J., and Torfs, P. J. (2018). Using lagged dependence to identify (de)coupled surface and subsurface soil moisture values, *Hydrology and Earth System Sciences* 22, 2255–2267, doi: 10.5194/hess-22-2255-2018

Carranza, C., Benninga, H.-J., van der Velde, R., and van der Ploeg, M. J. (2019). Monitoring agricultural field trafficability using Sentinel-1, *Agricultural Water Management* 224, 105 698, doi: 10.1016/j.agwat.2019.105698

Buitink, J., Swank, A. M., van der Ploeg, M., Smith, N. E., Benninga, H.-J. F., van der Bolt, F., Carranza, C. D. U., Koren, G., van der Velde, R., and Teuling, A. J. (2020). Anatomy of the 2018 agricultural drought in The Netherlands using in situ soil moisture and satellite vegetation indices, *Hydrology and Earth System Sciences Discussions* 2020, 1–17, doi: 10.5194/hess-2020-358

Carranza, C., Nolet, C., Pezij, M., and van der Ploeg, M. (2020a). Root zone soil moisture estimation with Random Forest, *Journal of Hydrology* p. 125840, doi: 10.1016/j.jhydrol.2020.125840

Carranza, C., van Emmerik, T., Benninga, H.-J., van der Velde, R., and van der Ploeg, M. J. (2020b). Exploring the sensitivity of vegetation backscatter to root zone soil moisture, *Submitted to Biogeosciences*

## Professional publications

Pezij, M., Benninga, H., Carranza, C., Augustijn, D. C., van der Velde, R., and van der Ploeg, M. (2016). Bodemvocht uit satellietdata voor optimalisatie waterbeheer, *Land+ water* 56, 26–27

van der Velde, R., Benninga, H.-J., Pezij, M., Augustijn, D., Carranza, C., van der Ploeg, M., and Hendriks, D. (2018). Bodemvocht uit satellietdata: wat kan de Nederlandse waterbeheerder ermee?, *H2O online* 2018





*Netherlands Research School for the  
Socio-Economic and Natural Sciences of the Environment*

# D I P L O M A

*for specialised PhD training*

The Netherlands research school for the  
Socio-Economic and Natural Sciences of the Environment  
(SENSE) declares that

***Coleen Dorothy  
Urbano Carranza***

born on 20 July 1987 in Daet, Camarines Norte, Philippines

has successfully fulfilled all requirements of the  
educational PhD programme of SENSE.

Wageningen, 25 January 2020

Chair of the SENSE board

Prof. dr. Martin Wassen

The SENSE Director

Prof. Philipp Pattberg

*The SENSE Research School has been accredited by the Royal Netherlands Academy of Arts and Sciences (KNAW)*



K O N I N K L I J K E N E D E R L A N D S E  
A K A D E M I E V A N W E T E N S C H A P P E N





The SENSE Research School declares that **Coleen Dorothy Urbano Carranza** has successfully fulfilled all requirements of the educational PhD programme of SENSE with a work load of 44.5 EC, including the following activities:

#### SENSE PhD Courses

- o Environmental research in context (2016)
- o Research in context activity: 'Soil Compaction episode for Soil4Life: Sustainable Soil MOOC' (2016)

#### Other PhD and Advanced MSc Courses

- o Summer School for Flow and Transport in Terrestrial Systems, TU Clausthal, Germany (2016)
- o Reviewing Scientific Manuscripts, Wageningen Graduate Schools (2016)
- o GEOSTAT Albacete 2016 Summer School, University of Castilla-La Mancha (2016)
- o Hydrus Course and Workshop, Czech University of Life science (2016)
- o Bayesian Statistics and Modeling, Wageningen University (2017)
- o Statistical Uncertainty Analysis of Dynamic Models, Wageningen University (2017)
- o R and Big data, Wageningen University (2019)

#### Management and Didactic Skills Training

- o Member of OWAS1S user committee meetings (2015-2019)
- o Camera training for MOOC's, Wageningen University (2016)
- o PhD peer consultation – a powerful tool to tackle challenges, Wageningen Graduate Schools (2016)
- o Organizing committee member, International Soil Modeling Consortium (ISMC) conference, Wageningen University (2018)
- o Assisting practicals of the BSc course 'Hydrogeology' (2016-2017)
- o Assisting practicals of the MSc course 'Environmental Data Collection and Analysis' (2019)

#### Oral Presentations

- o *Using lagged dependence to identify decoupled range*, Boussinesq Lecture, 26 October 2017, Delft, The Netherlands
- o *Using lagged dependence for assessing vertical soil moisture (de) coupling*, European Geosciences Union General Assembly, 8-13 April 2018, Vienna, Austria

SENSE coordinator PhD education

Dr. ir. Peter Vermeulen

This research was funded by the project entitled "Operational Water Management using Sentinel-1 Satellites" (OWAS1S) with project number 13871. The project is funded by Toegepaste and Technische Wetenschappen (TTW) which is part of the Netherlands Organization for Scientific Research (NWO).

