INVITATION

I have the pleasure of
inviting you to attend
the public defence of
my dissertation entitled:

Retrieval of Soil Physical
Properties: Field Investigations,
Microwave Remote Sensing and
Data Assimilation

Which will take place on
Thursday, 17 June, 2021

De Waaier, room 4,
University of Twente,
Enschede

14.30 Layman's talk

14.45 PhD defence

Hong Zhao

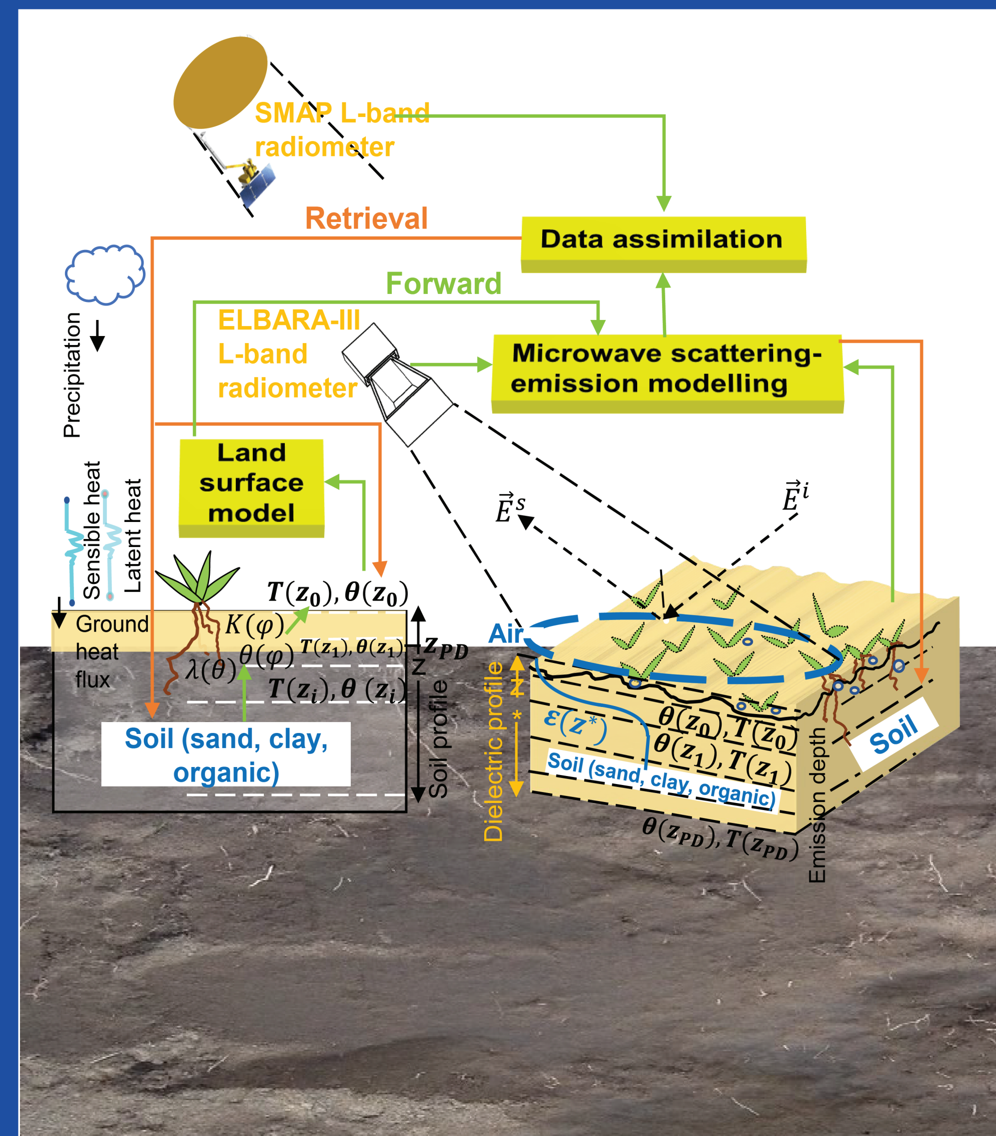
UNIVERSITY OF TWENTE. | ITC

ISBN: 978-90-365-5198-4
DOI: 10.3990/1.9789036551984
ITC dissertation number: 397

Retrieval of Soil Physical Properties: Field Investigations, Microwave Remote Sensing and Data Assimilation

Hong Zhao

Retrieval of Soil Physical Properties: Field Investigations, Microwave Remote Sensing and Data Assimilation



Hong Zhao

RETRIEVAL OF SOIL PHYSICAL PROPERTIES: FIELD INVESTIGATIONS, MICROWAVE REMOTE SENSING AND DATA ASSIMILATION

Hong Zhao

RETRIEVAL OF SOIL PHYSICAL PROPERTIES: FIELD INVESTIGATIONS, MICROWAVE REMOTE SENSING AND DATA ASSIMILATION

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
Prof. dr. ir. A. Veldkamp,
on account of the decision of the Doctorate Board,
to be publicly defended
on Thurday 17 June 2021 at 14.45 hrs

by

Hong Zhao
Born on 06 November 1988
in Lanzhou, China

This thesis has been approved by

Prof. dr. Z. Su, supervisor
Dr. Y.J. Zeng, co-supervisor

ITC dissertation number 397
ITC, P.O. Box 217, 7500 AE Enschede, The Netherlands

ISBN 978-90-365-5198-4
DOI 10.3990/1.9789036551984

Cover designed by Hong Zhao and Job Duim
Printed by CTRL-P, Hengelo

Graduation committee:

Chairman/Secretary

Prof. dr. F.D. van der Meer

Supervisor(s)

Prof. dr. Z. Su	University of Twente

Co-supervisor(s)

Dr. Y.J. Zeng	University of Twente

Members

Prof.dr. D. van der Wal	University of Twente
Dr.ir. C. van der Tol	University of Twente
Prof. dr. A. Verhoef	University of Reading
Prof. dr. S.C. Steele-Dunne	Delft University of Technology
Prof. dr. X.J. Han	Southwest University

Learning without thought is puzzled; thought without learning is perilous.

Confucius

This thesis is dedicated to my grandmother (Shihua Song),

my father (Suping Zhao),

my mother (Jihua Zhou),

and my beloved grandfather (Quanxing Zhao).

Acknowledgments

Five years ago, a thought enters my mind: why not delve into the hydrology/land-atmosphere interaction field to observe the implementation of dedicated remote sensing tools. I hoped for an exciting learning experience and answers to questions on what/how/why many stuff occur as such. I followed this inspiration and wrote an application letter to Prof. Bob Su, whose research interests include water cycle and remote sensing. I was very lucky to receive a positive response and the recommendation to approach Dr. Yijian Zeng—a young pioneer for assistance in writing a proposal. With their great help, I obtained a scholarship from the China Scholarship Council (CSC) and started my Ph.D. adventure. This journey has been filled with ignorance, exploration, joy, frustration, courage, surprise, happiness and faith. Along the way, I encountered many brilliant people. It is my pleasure to acknowledge all who have contributed to this thesis and supported me throughout this amazing and memorable period.

First and foremost, I would like to express my sincere and deep gratitude to my promoter, Prof. Bob Su, for his insightful ideas, valuable suggestions and unwavering support. His fruitful discussions prompted my research progress and helped me move forward, especially when I got stuck. His constructive comments delivered through meetings/emails extended my horizon and deepened my understanding of microwave remote sensing. The more I digested the provided information, the more my critical thinking and research skills improved. His cheerful and inspiring talks in every workshop and department gathering greatly encouraged me. He is always enthusiastic about the story behind formulas from an engineer's point of view and explains physical processes from a scientific perspective. His scientific creativity and openness continue to inspire me. With his high-quality guidance throughout these five years, my previous stance on science was altered, and a new mindset is established. Thanks to his inspiration, I have gained knowledge and persevered in this difficult field, which will benefit me throughout my life. I am also very grateful for the provided open atmosphere. This reduced any mental tension and allowed me to spend time exploring my own mind. My great respect and thanks are owed to his sharing of professional and nonprofessional books. They were the most precious New Year gifts ever

received. Furthermore, I am indebted to his assistance in nominating me to obtain international fellowships to support my Ph.D. extension.

My deep appreciation goes to my daily supervisor, Dr. Yijian Zeng, who offered professional guidance, enormous help, continuous encouragement and ample trust during my Ph.D. period. His full support helped me to step from my comfort zone and find my pace at the beginning of my Ph.D. journey. He spent much time on discussions of my research proposal, field/laboratory experimental design/implementation and result analyses that inspired me and enhanced my grasp of soil physics. He raised questions in a friendly manner and encouraged me to overcome difficulties. His enthusiasm, rigorous attitude and dedication to scientific research pushed me forward. With his great efforts in field campaigns on the Tibetan Plateau, I deeply investigated soil physics and drafted my first paper, with numerous draft versions involving major revisions and comments owing to his great patience and hard work, the quality of the paper is guaranteed. This training process was a good start for me, which made me more confident and promoted my independence. When my work moved to surface roughness parameterization based on a discrete scattering-emission model and soil property retrieval, he provided help to overcome obstacles. He carefully commented on each drafted file (e.g., reports, manuscript, press articles, etc.), helped collect the data I needed, and shared programming experience and sources/individuals to help me solve computer issues. Without his help and scientific literacy, I could not have achieved my goals in enhancing my critical thinking and modeling, scientific writing and communication skills. When promoting the research progress, two main inquiries came to mind. One involved ‘its explicate explanation and modeling in terms of its physical meaning’. The second was ‘to conduct research systematically in a well-organized manner’. I learned critical thinking and unifying efficient and careful actions from him, which will benefit me throughout the rest of my life.

My special thanks go to Prof. Xujun Han of Southwest University, China, for his achievement in developing the DasPy data assimilation framework. I obtained many benefits from his excellent work. I am very grateful for his very kind help, patience and guidance during my Ph.D. journey. He provided many constructive comments and useful solutions via WeChat to solve scientific and technical problems that I encountered in performing soil property retrieval. When I was in Chongqing in December 2019, he organized a fruitful discussion and a wonderful tour of the campus. The hillside scenery was so pleasing. Many thanks, Prof. Han.

I would like to acknowledge Prof. Wen Jun and his group at the Key Laboratory of Land Surface Process and Climate Change in Cold and Arid Regions, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, for their support in my field campaign and their wonderful establishment and maintenance of the Maqu network. Working with Dr. Xin Wang, Mr. Zuoliang Wang, Dr. Hui Tian, Dr. Jinlei Chen and Dr. Dongyu Jia in the field, I appreciate their kind help and humorous conversations. Special thanks go to Youqi Su at the same institute for his remote sharing of expertise on CLM operation.

My keen appreciation goes to my formal supervisor Prof. Yaowen Xie and Prof. Jizong Jiao of Lanzhou University (LZU), China, for their great efforts in assisting me in conducting soil analysis in the laboratory at LZU. Their guidance and unconditional supported my academic path. I thank academician Fahu Chen and Prof. Wei Huang for their patience and guidance to have the manuscript on the LZU climatic dataset published. The long process of paper revision and review taught me to not give up. I thank Dr. Lei Lu, who introduced me to the ITC and helped write a recommendation letter. Her continuous encouragement and support fostered a strong friendship. I thank Dr. Wei Liu, who shared his workplace to let me perform hydraulic experiments and helped process some of the samples. I thank Dr. Yuanyuan Hao and Dr. Yindan Zhang for their friendship and sweet comfort when I was back at LZU. I also thank Dr. Liqiang Zhao, Dr. Ping Liu, Hui Hua, Hui Lu and Zheng Huang for their help in my laboratory experiments.

I acknowledge the financial support received from the China Scholarship Council (CSC). Without Chinese government funding, my dream of studying overseas would have never been realized. I acknowledge the ITC Foundation Fellowship Program and the CEOP-AEGIS (Coordinated Asia-European long-term Observing system of Qinghai-Tibet Plateau hydro-meteorological processes and the Asian-monsoon system with Ground satellite Image data and numerical Simulations) project (https://www.futurewater.eu/projects/ceop-aegis-2/) for providing funding to support my Ph.D. extension. I also want to thank all the anonymous reviewers and editors for their constructive comments on my publications and manuscripts.

I am thankful to the ITC and UT for providing me with a nice working environment and facilities. I would like to thank Anke, Tina, Lindy and Murat for always being so kind and helpful throughout my Ph.D. study. My sincere thanks go to Loes and Theresa for their friendly assistance during my stay at the ITC. Big thanks to Annelies for her great help in my defense stuff, to Benno and Job in the poster design, and to Caroline for the provided lab experimental assistance. I thank Chris for his help on my Ph.D. extension. I thank Bas, Wim and Ali for their assistance in solving computer problems. I would also like to thank Daphne, Maciek, Christiaan and Rogier for their interesting presentations. Many thanks to the Yoga group at the ITC and group lessons at UT that enriched my time and energized me.

I thank all fellow Ph.D. students at the ITC. My special thanks go to Yifang and Chengliang for their company and support from the first day of my Ph.D. journey. I especially thank Yifang for her friendship and comfort. I am grateful to Linlin, who is like my sister, providing lots of care and convenience when I was new in the Netherlands. A debt is also due to Yijian, Shaoning and Lianyu for their great work on sampling soils on the Tibetan Plateau. I extend my gratitude to my roommates Peipei, Mengna, Qiang, Bagher, Behnaz, Binbin, Lianyu and Ming. Their smiles and hard work made our office a much better place. I thank Jan, Xiaojing, Donghai, Harm-Jan, Qiang and Laura for discussions in modeling topics. Big thanks go to Yijian & Rong & Gaoqi, Xuelong, Junping, Xiaolong & Xu, Ruosha, Peiqi & Jing, Xi, Fangyuan, Zhihui, Lichun, Yasser, Cesar, Tebogo, Egor, Georgios, Sammy, Kingsley, Novi, Mijun, Haili & Xin, Zhenchao and Yuhang for their friendly support.

I would like to thank my friends in China who have always remained supportive. I am deeply indebted to my close friend Hailan, Qinyu, Xingzhou and Ting for their trust and spiritual comfort, especially when I encountered difficulties. Thanks to Yuan, Xiulin, Rongrong and Bin who give me the best support for many years. I am also grateful to angel Huan. Her kindness and passion greatly affected me. I will never forget our wonderful tours throughout the Netherlands.

Finally, my deepest gratitude goes to my dearest parents, grandmother and my beloved grandfather. Thank you for all of your unconditional love and caring, as well as your wisdom, respect and tremendous support. My deepest gratitude is due to my younger sister Xing Zhao, brother-in-law Decai Wang and cute nephew Yuetong Wang for their company provided to my parents. Special thanks to my little aunt Xiulian Zhao and uncles Qixiang Tao and Chunping Hu who pictured a better world of knowledge and help broaden my views, and all members of the Zhao unified family for their love, continuous encouragement and assistance. I thank my in-laws for their love and support. Finally, I would like to thank my husband Wen Bai, for giving me the utmost support, patience and love. I am very fortunate and grateful to have enjoyed him and his company in the Netherlands,

especially during the strenuous coronavirus period in 2020 and 2021. Thank you for being my protective shield.

Table of Contents

Chapter 1. General Introduction

1.1 Scientific Background

1.1.1 Soil physical properties

Soil is defined as the weathered and fragmented outer layer of the Earth's terrestrial surface that serves as a home to innumerable microscopic and macroscopic plant communities (Hillel, 2003). As soil constitutes a natural body, the soil is engaged in dynamic interactions with the atmosphere above and strata below and affects the Earth's climate and hydrological and carbon cycles (Fatichi et al., 2020; Paustian et al., 2016). Soil strata (horizons) comprise a soil profile, which reflects the character of soil as a whole. An example is shown in Figure 1.1, where the O horizon contains organic litter derived from plants and animals, which are usually present on the surface, the A mineral horizon occurs at the soil surface or below the O horizon, which contains organic matter mixed with mineral material, the B mineral horizon constitutes a subsoil layer below the A horizon, and the C mineral horizon underlying the B horizon comprises the soil parent material (Buol et al., 2011).

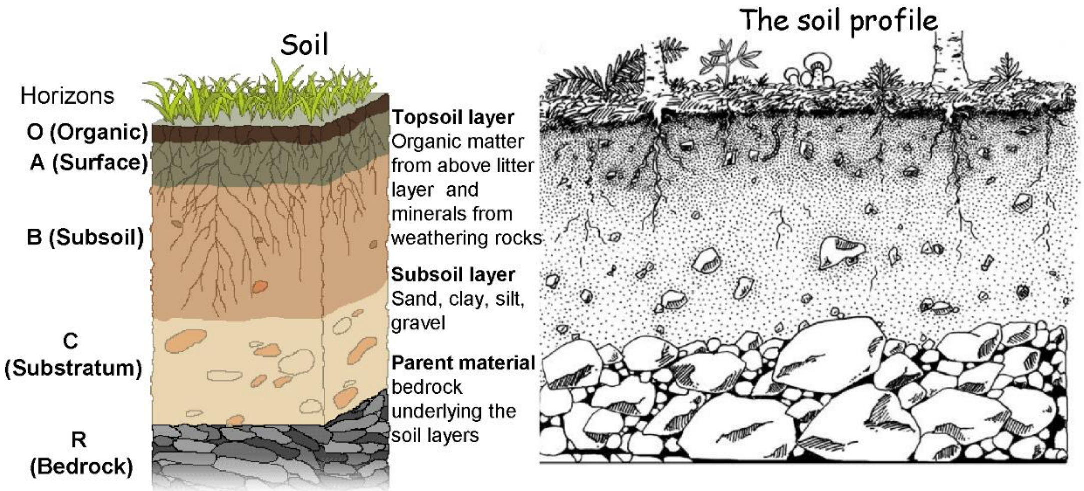


Figure 1.1 Schematic representation of a hypothetical soil profile (Ben-Dor, 2019) and the United States Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS).

Generally, soil is regarded as a three-phase (i.e., solids, air and water) disperse system (Hillel, 2003): 1) the soil solid phase comprises mineral particles, which have different sizes and shapes, as well as attached amorphous compounds, such as organic matter and hydrated chemicals (e.g., iron oxides). The soil matrix is the solid phase of soil, and the range of particle sizes in the soil is termed the soil texture. 2) Once soil particles (i.e., sand, clay and silt particles, as shown in Figure 1.2) are organized naturally, the pore space is formed, in which air and water are transmitted and retained. 3) The geometric characteristics of the pore space can be characterized through the soil dry bulk density (the ratio of the mass of solids to the total soil volume (the volume of both the solids and pores)), and the total volume fraction of pores is termed as the porosity.

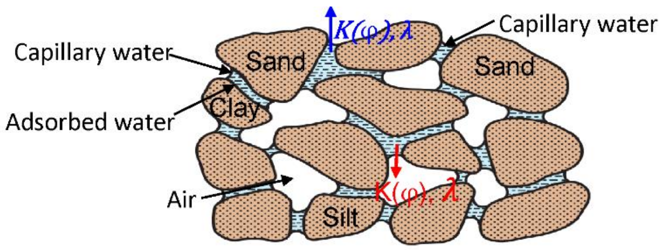


Figure 1.2 Schematic soil composition and soil physical process. $K(\varphi)$ and $\lambda(\theta)$ are the soil hydraulic and thermal conductivity, respectively. φ denotes the soil matric potential and θ is soil water content.

Soil water generally refers to the amount of water contained in the unsaturated soil zone (also referred to as the vadose zone) (Hillel, 1998). Soil water in the unsaturated zone involves two types, namely capillary water and adsorption water (as shown in Figure 1.2). Due to the adhesive intermolecular forces between water and the solid surfaces of pores exceeding the cohesive intermolecular forces among water molecules, capillary motion occurs, which is demonstrated by the upward/horizontal movement of water through pores against the force of gravity (Lu & Likos, 2004). The height to which water rises depends on the pore size. The smaller the soil pores are, the greater the capillary height. Compared to the capillary phenomenon, adsorption water envelopes the particle surface, especially that of clay particles. Clay particles typically carry a net negative electrostatic charge due to ion replacement occurring during crystallization and incomplete charge neutralization of terminal ions along lattice edges (Hillel, 2003). When hydrated, polar water molecules become attached to clay surfaces and form an electrostatic double layer, and the thusly attached water is referred to as adsorption water. Compared to capillary water, adsorbed water, which is denoted as bound water in soil dielectric modeling (Mironov et al., 2004) does not freely flow. The state of soil water is characterized by the potential, and the matric potential quantifies the tenacity with which soil water is retained by the soil matrix (Hillel, 2003) (as shown in Figure 1.2). When all pores are filled with water, the soil is saturated.

The soil water content and matric potential are functionally related. The graphical representation of this relationship is the soil moisture retention curve (SWRC) (Baver et al., 1956), which is affected by the direction and the rate of change of soil moisture. Moreover, the SWRC is indispensable in hydraulic transport modeling as it permits the definition of the hydraulic conductivity (transmitting property of the conducting medium) function (Buckingham, 1907). Parametric functions that fit a wide range of experimental data were proposed to describe the SWRC. Brooks and Corey (1964) proposed a two-parameter power function of the soil matric potential to represent the effective saturation, which was validated in several studies, such as Campbell (1974) and Clapp and Hornberger (1978). However, this model does not suitably depict soil water retention near saturation in finer-textured soils and undisturbed field soils (Clapp & Hornberger, 1978;

Milly, 1987; van Genuchten & Nielsen, 1985), which exhibit S-shaped retention characteristics. To remedy this condition, continuously S-shaped curves were proposed, such as the models of Brutsaert (1966), Laliberte (1969) and van Genuchten (1980). The van Genuchten (1980) model was found to provide a good fit to SWRC data for many soils, particularly considering near-saturation data (van Genuchten & Nielsen, 1985), although this model may yield poor results at low water contents (Nimmo, 1991; Ross & Smettem, 1993).

A series of studies has been carried out to mitigate the limitation of empirical equations applied within a certain suction range. For instance, Campbell and Shiozawa (1992) used an observed linear function between the capillary pressure in the dry region and liquid saturation (on a semilogarithmic plot) and added the van Genuchten (1980) function to obtain the SWRC over the range from complete saturation to zero-liquid residual saturation (approximating oven dryness). Fredlund and Xing (1994) developed a general SWRC equation over the entire suction range from 0 (full saturation) to 10^6 kPa (approximating oven dryness) (Ross et al., 1991) based on the soil pore-size distribution. Rossi and Nimmo (1994) modified Brooks-Corey equations and fitted two- and three-parameter models describing the SWRC at a zero residual saturation. Fayer and Simmons (1995) further modified Brooks–Corey and van Genuchten functions by replacing the residual water content with an adsorption equation developed by Campbell and Shiozawa (1992). Morel - Seytoux and Nimmo (1999) extended the Brooks–Corey model to oven dryness with the three-parameter junction model of Rossi and Nimmo (1994). The results in these studies revealed good data-model comparisons, but considering the ease of implementation, Webb (2000) extended the van Genuchten model with a dry-region expression according to the adsorption equation of Campbell and Shiozawa (1992).

The SWRC is based on the soil pore structure. Assuming that the shape of the SWRC depends on the soil pore-radius (size) distribution, the SWRC is uniquely determined if the pore-size distribution of a given soil is obtained or predicted (Fredlund & Xing, 1994). For instance, the Brooks-Corey power function is a unique case when the pore-size distribution inversely varies with the power of the radius, while the van Genuchten model follows a different pore-size distribution (Fredlund & Xing, 1994). Given that the distribution of the particle size of many soils is approximately lognormal, Kosugi (1994) applied three-parameter lognormal distribution to the pore-radius distribution and pore capillary pressure distribution function to obtain a new SWRC model. To represent soil water retention under all matric potentials, Khlosi et al. (2006) modified the Kosugi (1994) function by replacing residual water content with the adsorption equation of Campbell and Shiozawa (1992). Compared to other models, the Khlosi et al. (2006) model was found to be the most consistent among different soils, and they attained a significant correlation between the model parameters and basic soil properties (Khlosi et al., 2008).

Based on the assumption of an ideal capillary medium characterized by a certain pore-size distribution model, conductivity models such as the Burdine (1953) and Mualem (1976) models were proposed. These two models were incorporated to derive closed-form expressions to quantify soil the unsaturated hydraulic conductivity based on SWRC models. Leij et al. (1997) and Assouline and Or (2013) provide a comprehensive review of popular closed-form expressions describing soil water retention and hydraulic conductivity. Among land surface models (LSMs), the Clapp and Hornberger (1978) model and van Genuchten (1980)-Mualem (1976) model represent the most widely applied schemes for the parameterization of soil hydraulic properties (SHPs) (i.e., SWRC and hydraulic conductivity). For instance, the community Noah land surface model (Chen & Dudhia, 2001) and community land model (CLM) (Oleson et al., 2008) reply on the former (Chen & Dudhia, 2001; Oleson et al., 2008), whereas the Hydrology-Tiled European Center for Medium-Range Weather Forecasts (ECMWF) Scheme for Surface Exchanges over Land (H-TESSEL) (Balsamo et al., 2009) employs the latter.

Soil thermal properties (STPs) involve the soil heat capacity and thermal conductivity. Soil volumetric heat capacity is defined as the change in the heat content of a unit bulk volume of soil per unit change in temperature (De Vries, 1963). Soil thermal conductivity is defined as the rate at which heat energy flows across a unit area of soil due to a unit temperature gradient (De Vries, 1963). STPs depend on the composition of the solid phase (mineral and organic constituents), volumetric water content, porosity, dry density, and temperature (Farouki, 1986). Moreover, STPs were found to be sensitive to the sizes, shapes and spatial arrangement of soil particles (De Vries, 1963). A number of studies has been conducted to estimate STPs, such as the empirical Johansen (1975) model, which has been adopted in LSMs, including Noah, H-TESSEL and CLM. There are other semi-empirical STP parameterizations, such as Kersten (1949), De Vries (1963), Farouki (1981), Campbell (1985), Côté and Konrad (2005), Balland and Arp (2005), Lu et al. (2007), Tarnawski and Leong (2012) and the simplified De Vries method (Tian et al., 2016). The performance of these different soil thermal conductivity schemes has been evaluated for application in land surface modeling (Dai et al., 2019).

With both SHPs and STPs quantified, LSMs model soil water (vertical) flow and heat transport with a one-dimensional Richards (1931) equation and the Fourier diffusion law (Kreith & Black, 1980) respectively.

1.1.2 Soil moisture estimation

Soil moisture is an essential climate variable (ECV), as designated by the Global Climate Observing System (GCOS, 2006), and included in the European Space

Agency (ESA) Climate Change Initiative (CCI) project (Hollmann et al., 2013). An ECV is defined as a physical, chemical or biological variable that critically contributes to the characterization of Earth's climate (GCOS, 2006). Soil moisture is a well-known main variable to specify the lower boundary condition of the atmosphere, control the partition of incoming energy into latent and sensible heat fluxes, adjust surface runoff and soil drainage, and regulate canopy transpiration and photosynthesis (Koster et al., 2004; Seneviratne et al., 2010). Through its role in land and atmosphere interactions, soil moisture is important in weather/climate predictions. In numerical weather prediction models, a realistic initialization of soil moisture is a necessity given the inferred impacts of this variable on temperature and precipitation (Chen & Avissar, 1994; Koster et al., 2004).

There are a variety of methods to estimate soil moisture, such as ground observations, remote sensing and land surface model simulations (Njoku et al., 2002; Owe et al., 2008; Seneviratne et al., 2010; Su et al., 2011; Su et al., 2020b; Yang et al., 2013; Zeng et al., 2016). Since ground observations are very limited in space and remote sensing estimates suffer limitations regarding their temporal and/or spatial coverage, as well as the accuracy and exact link with root-zone soil moisture, LSMs synergized with *in situ* and remote sensing observations within a data assimilation framework have become a useful alternative to produce spatiotemporally consistent (profile) soil moisture information. Therefore, enhancement of understanding of the physics of soil water flow and energy exchange in the vadose zone is imperative to obtain soil moisture estimates.

As SHPs & STPs described in section 1.1.1 govern water and heat transport processes and the partitioning of soil moisture between infiltration and evaporation fluxes, they are mandatory physical parameters for the estimation of soil moisture profiles. In certain cases, soil moisture simulations with LSMs were found to be much more dependent upon the specification of SHPs & STPs than on the specification of atmospheric forcing or surface conditions (Gutmann & Small, 2005; Pitman, 2003; Santanello et al., 2001). Soils with different textures have varied SHPs & STPs. In contrast, as the soil is repeatedly wetted by rain, drained by gravity and dried by evaporation and root extraction, SHPs & STPs also vary in time and space. Moreover, most SHP & STP models adopted are pore-scale derivations. The appropriate representation remains problematic when these models are implemented in LSMs, which are characterized by large-scale heterogeneous processes. Nevertheless, the appropriate characterization of SHPs & STPs under different spatial (e.g., in the field approximately at ~m scale and regionally at the ~km scale) conditions and at the appropriate process scale is critical to ensure the success of land surface modeling for spatiotemporally soil moisture estimation (Mohanty, 2013).

Additionally, these parameters are required in agronomy, as they can be considered to schedule management practices, especially irrigation and fertilization, and in contaminant hydrology and geochemistry regarding pollutant evaluation and treatment. In engineering applications such as site selection for railway construction, especially in alpine regions, knowledge of SHPs & STPs is also necessary.

1.1.3 Estimates of soil physical properties

Field and laboratory measurements to quantify SHPs & STPs require a substantial investment in time, manpower and resources (Angulo-Jaramillo et al., 2000). As basic soil properties such as soil texture, organic matter content, dry bulk density and porosity (please refer to section 1.1.1) play a fundamental role in determining SHPs & STPs, pedotransfer functions (PTFs) (Bouma, 1989) were proposed to relate the parameters in SHP & STP schemes to these readily available data on basic soil properties. PTFs can be categorized as class and continuous PTFs (Wösten et al. 1990). Class PTFs predict average hydraulic characteristics based on soil texture classes, and continuous PTFs consider basic soil properties to estimate SHPs & STPs, such as through statistical regression equations (Cosby et al., 1984; Saxton et al., 1986) or artificial neural networks (Schaap et al., 1998; Zhang & Schaap, 2017).

As global and regional soil property maps (e.g., FAO-UNESCO Soil Map of the World (2007) and the Harmonized World Soil Database (FAO/IIASA/ISRIC/ISSCAS/JR, 2012)) have become available, most LSMs including CLM, Noah and H-TESSEL, utilize PTFs to obtain SHPs & STPs. As published soil property maps only show a single soil texture type on a spatial grid, a single set of SHPs & STPs is generated. However, these SHPs & STPs may not represent all spatial extents due to their high heterogeneity in space (Vereecken et al., 2007). Additionally, in a certain area, there might be varied soil types according to different soil map sources, which might cause discrepancies in LSM simulations. Moreover, the uncertainties in these existing soil property maps are hard to quantify (Dai et al., 2013; Shangguan et al., 2012). Inaccurate soil property maps might lead to an inaccurate specification of SHPs & STPs through PTFs and further introduce biases in the results of LSM simulations. For instance, Su et al. (2013) found that due to the unrepresentative soil property map of the Tibetan Plateau adopted in H-TESSEL, the estimated saturated hydraulic conductivity was two orders of magnitude lower than that based on field measurements, resulting in soil moisture analyses at the ECMWF significantly overestimating the regional soil moisture level in the cold-semiarid Naqu area during the monsoon season. Therefore, accurately and objectively obtaining soil physical properties across a wide spatial scale range and quantifying their uncertainties are imperative to improve estimates of land surface states/fluxes. For this purpose, remote sensing that provides fast and repetitive coverage of large areas for many applications (e.g., ranging from weather forecasts to reports on natural disasters) deserves to be investigated.

Remote sensing (RS) is the process of inferring surface parameters from distant measurements of the upwelling emitted or reflected electromagnetic radiation of the land surface in both passive and/or active modes (Ben-Dor, 2019). As electromagnetic radiation interacts with soil materials, this provides information regarding a range of physical parameters of the soil matrix (mostly at the surface). Compared to optical measurements, the microwave spectral region (e.g., the C-band at 4.75 GHz and L-band at 1.4 GHz) is transparent in regard to the atmosphere. As such, this approach is more advantageous to obtain microwave thermal radiation (i.e., brightness temperature) observations of the Earth's surface from space under all meteorological conditions. The propagation of electromagnetic waves in the medium is determined by the complex dielectric constant (Ulaby et al., 2014). Due to the molecular structure of water, liquid water has a higher dielectric constant ($\approx$ 80 at a low frequency, i.e., $<$ 4 GHz) in the microwave band than that of dry soil ($\approx$ 5 at the same frequency), air ($\approx$ 1) and ice ($\approx$ 3). Owing to these characteristics, different types of soil with varied water contents can be detected via microwave remote sensing.

To obtain quantitative information on soil physical properties across a wide spatial scale range, a retrieval approach based on remotely sensed data in conjunction with a coupled soil physics and microwave radiative transfer model has been proposed. The retrieval method has a physical foundation and does not require intensive laboratory SHP & STP measurements. With the use of the SHP & STP retrieved at the appropriate process scale, land state variables (i.e., soil moisture and temperature) and surface fluxes (i.e., land heat and sensible fluxes) can also be calculated. The conceptual framework for SHP & STP retrieval via microwave remote sensing is shown in Figure 1.3.

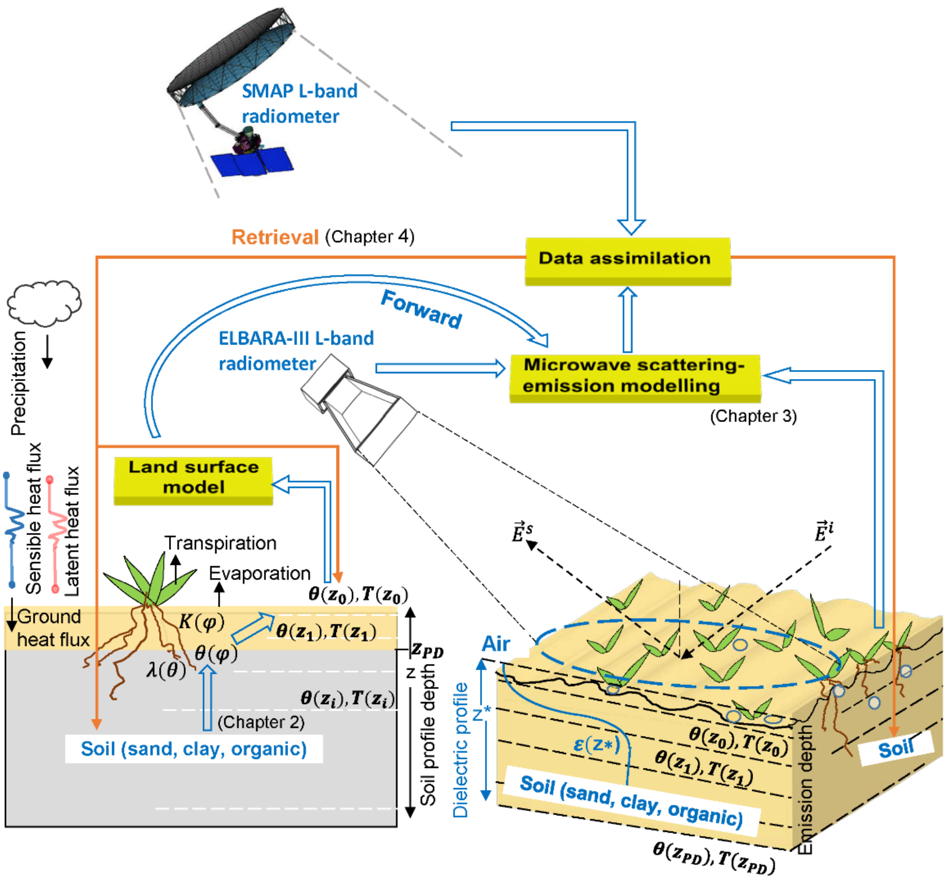


Figure 1.3 Conceptual framework for SHP & STP retrieval via passive microwave remote sensing. The yellow rectangle represents the hypothetical zone of emission measured by the microwave radiometer. $K(\varphi)$ and $\lambda(\theta)$ are soil hydraulic and thermal conductivity, respectively. φ denotes soil matric potential, θ denotes soil water content, T is soil temperature, z is soil depth and $0/\ 1/i$ denotes the soil layer. z_{PD} represents soil penetration/emission depth. T_B^p is brightness temperature (with p = H or V polarization). z^* denotes the dielectric profile ($\varepsilon(z^*)$) from air to soil, and $\vec{E}^i$ and $\vec{E}^s$ denote incident and scattering electric field, respectively.

Camillo et al. (1986) were among the early researchers envisioning that SHPs could be estimated by calibrating energy and moisture balance models with microwave brightness temperature data. Following the studies of Feddes et al.

(1993), Burke et al. (1997), etc. verified the success of retrieval modeling based on soil moisture information to estimate SHPs at a large spatial scale. With the availability of soil moisture data derived from proximal-, air- or satellite-based observations, soil moisture became the main state variable used for SHP & STP retrieval combined with LSMs (Bandara et al., 2014; Bandara et al., 2015; Montzka et al., 2011; Peters-Lidard et al., 2008; Qin et al., 2009; Santanello et al., 2007). In addition to soil moisture, the land surface temperature (LST), leaf area index (LAI), and evapotranspiration (ET) are applied in SHP & STP retrieval (Charoenhirunyingyos et al., 2011; Corbari et al., 2015; Dong et al., 2016; Gutmann & Small, 2010; Lu et al., 2016). It should be noted that LST, LAI and ET, as state variables, are derived from spectral remotes sensing (e.g., MODIS, Landsat), which is not directly linked to the signal penetrating into the soil and is often limited by cloud cover signal contamination. The use of these state variables to estimate SHPs & STPs may be associated with a high probability of contamination and unexpected uncertainties. In contrast, soil moisture derived from microwave remote sensing is not affected by weather conditions, although it only represents the soil moisture content in the upper centimeters (Jackson, 1993; Njoku & Kong, 1977). The application of soil moisture information derived from microwave remote sensing in SHP & STP estimates has become a preferred approach.

Since passive microwave L-band remote sensing has become the most promising technique to measure near-surface soil moisture due to its high penetrating capability, the soil moisture derived from L-band brightness temperature observations has been applied for SHP & STP retrieval purposes. Two innovative passive microwave missions are currently operating in the L-band to monitor surface soil moisture over continental surfaces. One mission is the Soil Moisture and Ocean Salinity (SMOS) mission launched in November 2009 dedicated to global soil moisture mapping (Kerr et al., 2001; Kerr et al., 2010). The other mission is the Soil Moisture Active Passive (SMAP) mission launched in January 2015 (Entekhabi et al., 2010), which incorporates a radiometer (continues to operate as planned) and radar (failed in July 2015) to measure and map the global soil moisture and freeze/thaw state. Researchers have seized this unique opportunity to estimate SHPs at a large spatial scale using SMOS-retrieved soil moisture products, such as Lee et al. (2014) and Bandara et al. (2015). However, any uncertainties in the adopted remotely sensed data at the calibration and retrieval stages can propagate to the retrieved soil physical properties at the pixel-scale (Ines & Mohanty, 2009) (as shown in Figure 1.3).

LSMs driven by atmospheric forcing provide profile soil moisture and soil temperature data, and these data can be fed into a microwave emission model for brightness temperature (T_B^p, with p = H, V polarization) simulations. With the use of coupled LSMs and microwave emission models (i.e., a forward observation simulator) to assimilate T_B^p observations, SHPs & STPs can be retrieved based on the optimization between simulated and observed T_B^p values (please refer to Figure 1.3). Compared to the direct retrieval technique using soil moisture, this approach represents a more consistent physical method. Cutting-edge studies were conducted by Han et al. (2014a), Dimitrov et al. (2014; 2015) and Yang et al. (2016).

Finally, SHPs & STPs can be retrieved either directly by bypassing PTFs (e.g., Gutmann et al. 2010; Bandara et al. 2014; Lee 2014; Dimitrov et al. 2014) or indirectly by using PTFs (e.g., Han et al. 2014; Santanello et al. 2007; Yang et al. 2016). The direct retrieval method is common, and the results are regarded as effective SHPs & STPs while the scale- and model-dependent. Moreover, the characteristics of different SHP & STP schemes can only be depicted in a statistical manner but cannot be analyzed from a physical perspective. The PTF-based method considering inherent soil properties can facilitate soil property mapping, and its accuracy can be evaluated with *in situ* and existing soil property datasets. The latter method guarantees soil physical consistency in land-atmosphere processes and has been advised as a more suitable way to retrieve spatially aggregated SHPs & STPs (Cooper et al., 2020; Santanello et al., 2007; Soet & Stricker, 2003).

1.2 Problem Description

The Tibetan Plateau observatory for soil moisture and soil temperature (Tibet-Obs) was built in 2016 and has been maintained onwards (Su et al., 2013; Su et al., 2011), providing comprehensive observations (including atmospheric forcing data and soil moisture and temperature profiles) for land surface modeling and validation of SM retrieval from satellite microwave remote sensing and reanalysis SM datasets (Dente et al., 2012; Yu et al., 2018; Zeng et al., 2016; Zheng et al., 2015a; Zheng et al., 2018a; Zhuang et al., 2020). In 2016, a ground-based ELBARA-III L-band (1.4 GHz) radiometer was mounted at the Tibet-Obs Maqu site (33.91°N, 102.16°E), providing T_B^p observations for L-band microwave radiometry investigation (Su et al., 2020a; Zheng et al., 2019; Zheng et al., 2018a; Zheng et al., 2018b; Zheng et al., 2017). Considering its comprehensive *in situ* observations, the Maqu site is selected as the study area in this thesis to investigate soil physical property retrieval. However, soil properties and SHPs & STPs are measured only for the topsoil layer by the Tibet-Obs (Zheng et al., 2015a). Moreover, these data are not available in the area where the ELBARA-III radiometer is operated. Soil physical property data obtained from available global and regional soil maps can be used as an alternative for land surface modeling and retrieved result validation. However, it is important to obtain *in situ* measurements because there might exist uncertainties in soil property data extracted from these datasets due to very few accessible *in situ* soil profiles in spatial data interpolation on the Tibetan Plateau (TP) (Su et al., 2013). Therefore, the first research question this thesis tries to address is as follows:

(1) What is the accuracy of existing soil datasets and their derived SHPs & STPs across the TP?

To answer the above question, there is a need to carry out an intensive field campaign to collect soil samples across the three climate zones (i.e., from east to west, sub-humid, semi-arid, and arid zones occur) on the TP. Nevertheless, as mentioned in section 1.1.3, *in situ* samplings experience limitations in generating continuous spatial soil property maps of the TP. To derive a physically consistent soil property map, it is necessary to combine different sources of information retrieved from *in situ*, satellite, and model data. Such kind of "information blending" can be realized via a coupled LSM with a microwave emission model within a data assimilation framework to retrieve basic soil properties and associated SHPs & STPs, through the assimilation of T_B^p observations (as shown in Figure 1.3). On the other hand, because the uncertainties in the observation operator propagate into soil physical property retrieval, the second focus of this thesis is to investigate the physical process that may affect L-band radiometry modeling.

It is to note that all observation operators applied in current retrieval approaches (Han et al., 2014a; Yang et al., 2016) are zeroth-order radiative transfer models (RTMs). Specifically, regarding the vegetation part, vegetation regarded as scatters is assumed to exhibit equal dimension, shape and permittivity characteristics. As such, vegetation emission can be characterized by two global parameters (Wigneron et al., 2007)—the single scattering albedo ω and the optical thickness τ. ω is assumed to remain constant for different vegetation types, such as the zero value considered in the SMOS soil moisture retrieval algorithm (Kerr et al., 2012). τ is estimated via empirical relationships using vegetation parameters such as the LAI (e.g., used in SMOS) and normalized difference vegetation index (NDVI, e.g., used in SMAP). Regarding the soil part, the surface reflectivity is calculated with planar Fresnel equations combined with a surface roughness correction model such as the Q/H model (Choudhury et al., 1979; Wang & Choudhury, 1981).

However, as mentioned in section 1.1.1, natural surfaces are complex, and soil characteristics (e.g., structure, moisture content) are highly heterogeneous. Topsoil structures and inhomogeneous moisture distribution in a given soil volume may induce volume scattering and affect the L-band microwave surface emission (please refer to Figure 1.3). However, the surface roughness model adopted in the zeroth-order RTM is a site-specific empirical model and cannot be used to investigate relevant physical processes. Therefore, the second and third research questions are as follows:

(2) How do the topsoil moisture distribution and geometric structure affect the dielectric roughness and surface roughness? How can we model these roughness effects in the observation operator to gain an improved understanding of the L-band radiometry? (as shown in Figure 1.3)

(3) Based on the observation operator improved in (2), does the T_B^p observation assimilation enable improvements in soil property retrieval? Consequently, does it improve estimates of (profile) soil moisture and temperature and land surface fluxes?

1.3 Structure of the Thesis

Correspondingly, this thesis is organized as follows:

Chapter 1 introduces the background of this research. The definition of soil physical properties is briefly reviewed in section 1.1.1, and estimates of soil moisture and soil physical properties are provided in sections 1.1.2 and 1.1.3, respectively. The research questions to be addressed in this thesis are introduced in section 1.2.

Chapter 2 describes field and laboratory experiments in terms of soil samples collected across the TP, to determine soil physical properties (as shown in Figure 1.3). Based on the compiled dataset, this chapter specifically addresses research question (1), thereby analyzing uncertainties in existing soil datasets via a comparison of basic soil properties, SHPs & STPs among the three climate regimes across the TP.

Chapter 3 develops an enhanced air-to-soil transition (ATS) model (please refer to Figure 1.3) that incorporates a new roughness parameterization to account for the effect of surface roughness on T_B^p modeling.

Chapter 4 retrieves soil physical properties with a forward observation simulator, namely, CLM 4.5 is coupled with a physically-based discrete scattering-emission model that integrates the developed ATS model in (3), within a data assimilation framework (as shown in Figure 1.3). Via comparison to *in situ* observations, this chapter addresses research question (3).

Chapter 5 synthesizes this study and presents an outlook for the future.

Chapter 2. Analysis of Soil Hydraulic and Thermal Properties for Land Surface Modeling on the Tibetan Plateau

This chapter is based on:

Zhao, H., Zeng Y., Lv S., and Su Z. "Analysis of Soil Hydraulic and Thermal Properties for Land Surface Modeling over the Tibetan Plateau." *Earth System Science Data* 10, no. 2 (2018): 1031-61.

Abstract: Soil information (e.g., soil texture and porosity) extracted from existing soil datasets over the Tibetan Plateau (TP) is claimed to be inadequate and even inaccurate for determining soil hydraulic properties(SHPs) and soil thermal properties (STPs), thus hampering the understanding of land surface processes over the TP. As soil varies across the three dominant climate zones (i.e., arid, semi-arid, and sub-humid) on the TP, the associated SHPs & STPs are expected to vary correspondingly. To obtain explicit insights into soil hydro-thermal properties on the TP, *in situ* and laboratory measurements of over 30 soil property profiles were obtained across the above climate zones. The results demonstrate that the porosity and SHPs & STPs differ across these climate zones and strongly depend on the soil texture. In particular, it is proposed that the gravel impact on the porosity and SHPs & STPs should be considered in both the arid zone and deep layers of the semi-arid zone. Parameterization schemes for the porosity, SHPs & STPs were investigated and compared to measurements. To determine SHPs, including the soil water retention curve and hydraulic conductivity, the pedotransfer functions (PTFs) developed by Cosby et al. (1984) (with the Clapp-Hornberger model) and the continuous Wösten et al. (1999) (with the van Genuchten-Mualem model) are recommended. The STP parameterization scheme proposed by Farouki (1981) based on the model of De Vries (1963) performed better across the TP than did other schemes. With the use of the parameterization schemes mentioned above, the uncertainties in five existing global and regional soil datasets and their derived SHPs & STPs on the TP are quantified through a comparison to *in situ* and laboratory measurements. This study suggests the SoilGrids1km dataset for the use in the arid and sub-humid zones, while the combination of the FAO-UNESCO Soil Map of the World dataset in shallow layers and the Harmonized World Soil Database (HWSD) dataset in deeper layers is recommended in the semi-arid zone on the TP. The

measured soil physical property dataset is available at https://data.4tu.nl/repository/uuid:c712717c-6ac0-47ff-9d58-97f88082ddc0.

Keywords: Soil hydraulic and thermal properties; Tibetan Plateau; Pedotransfer functions; Soil maps; Land surface model.

2.1 Introduction

As the highest plateau in the world, the Tibetan Plateau (TP) exerts a significant influence on the Earth's climate system and plays a prominent role in the evolution of the Asian monsoon system (Kang et al., 2010; Ma et al., 2017; Qiu, 2008; Yao et al., 2012). Studying this influence can advance our understanding of climate change (Ma et al., 2017). Soil moisture (hereafter referred to as SM), one of the lower boundary conditions of the atmosphere, is a crucial land surface state (Koster et al., 2004) and therefore of great interest to investigate land-atmosphere interactions, thereby reflecting the trend and variability in the feedback between the water cycle and climate on the TP (Su et al., 2013; Su et al., 2011). Accurate SM information is a necessity to improve precipitation and hydrology forecasts (Dirmeyer, 2000; Drusch, 2007; Robinson et al., 2008), especially on the TP, which experiences evident climate change (Douville et al., 2001; Ma et al., 2017; Yang et al., 2014; Yang et al., 2011). Consistent spatiotemporal SM data can be obtained by using land surface models (LSMs) assimilating *in situ* and satellite observations. In these models, the specification of soil hydraulic properties (SHPs) (i.e., soil water retention curve, hydraulic conductivity) and soil thermal properties (STPs) (i.e., thermal conductivity and heat capacity) is more decisive in SM simulation than is the specification of atmospheric forcing and land surface characteristics (Gutmann & Small, 2005; Kishné et al., 2017; Livneh et al., 2015; Shellito et al., 2016) because SHPs govern the partitioning of SM between infiltration and evaporation fluxes and STPs regulate water and heat transport processes (Garcia Gonzalez et al., 2012; Zeng et al., 2009a; Zeng et al., 2009b).

In situ measurements of basic soil properties and SHPs & STPs are crucial for soil moisture and heat flux simulations with LSMs. LSMs frequently adopt the Clapp and Hornberger (1978) model and the van Genuchten (1980)-Mualem

(1976) model formulated with SHPs, and the Farouki (1981) and Johansen (1975) schemes formulated with STPs. Since direct measurements of SHPs & STPs are always highly time-consuming, labor-intensive and costly, pedotransfer functions (PTFs) (Bouma, 1989; Van Looy et al., 2017) based on basic soil property information have been developed to estimate parameters of the above SHP & STP schemes. Examples include the Cosby et al. (1984) PTF (e.g., based on the sand fraction) for CH scheme estimation in the Noah and community land model (Chen & Dudhia, 2001; Oleson et al., 2008), and the soil class PTF (e.g., based on the soil texture type) for the VG scheme (Balsamo et al., 2009) in the Hydrology-Tiled European Center for Medium-Range Weather Forecasts (ECMWF) Scheme for Surface Exchanges over Land (H-TESSEL). However, these PTFs do not consistently predict SHPs &STPs well, especially when the soil contains organic matter or gravel (particle diameter ≥ 2 mm), because gravel and organic matter possess different hydraulic and thermal properties than those of other fine mineral soil, which suggests the necessity to obtain comprehensive soil property information (e.g., not only soil texture and porosity information but also soil organic content and gravel fraction information).

Furthermore, studies using the information on state variables (e.g., near-surface soil moisture or brightness temperature) can retrieve the effective SHPs & STPs directly or indirectly through PTFs within LSMs (Dimitrov et al., 2015; Dimitrov et al., 2014; Han et al., 2014a; Ines & Mohanty, 2008a; Yang et al., 2016). Nevertheless, most of these retrieval studies only focused on basic surface soil properties and SHPs, under the assumption of a homogenous soil column. If the system is highly heterogeneous (e.g., along the vertical profile), retrieval may be problematic (Ines & Mohanty, 2008b) and *in situ* measurements of soil property profiles may shed light on the retrieval of soil property along the vertical profile.

Many global and local efforts have been made to compile and develop soil databases, but uncertainties in these soil datasets might also cause a certain bias in SHP & STP predictions, and hence introduce uncertainties in the representation of the land surface state by LSMs. It has been reported that the overestimations of ECMWF SM analyses in the central TP region could be partly attributed to the unrepresentative soil information extracted from the FAO Digital Soil Map (2003) as used in H-TESSEL (Su et al., 2013). Currently, there is only soil texture information and few soil organic content profiles available for the TP when relying on published globally *in situ* soil profiles (Batjes et al., 2017). The profiles of other vital soil properties, such as the dry bulk density (BD) and porosity, are commonly not provided (e.g., no *in situ* BD or porosity profiles are available). Moreover, there are no comprehensive *in situ* measurements of basic soil properties, SHPs and STPs suitable for land surface modeling on the TP.

In this study, we implemented *in situ* and laboratory measurements of soil physical property profiles across the three climate zones on the TP and compiled a Tibet-Obs soil property dataset. Based on the compiled dataset, variations in basic soil properties and SHPs & STPs across the three climate zones were investigated. Applications of the Tibet-Obs dataset were demonstrated in two cases: 1) appropriate parameterization schemes of the porosity and SHPs & STPs were examined in regard to their applicability in land surface modeling on the TP; 2) the uncertainties in five existing global and regional soil datasets and their correspondingly derived SHPs & STPs were evaluated on the TP. In section 2.2 of this chapter, the field campaign and laboratory experiments are described, as well as the parameterization schemes for the porosity and SHPs & STPs estimates. Specification of the Tibet-Obs dataset with the data availability is documented in section 2.3. The results of the application of this dataset are presented in section 2.4. Conclusions are presented in section 2.5. This chapter is expected to contribute to land surface modeling and hydro-climatology studies of the Third Pole environment and to soil research in terms of filling geographic gaps in existing published global soil databases.

2.2 Materials and Methods

2.2.1 Field experiments

On the TP, soils exhibit spatial variation due to varying soil formation factors (e.g., climate and parent material). The TP can be categorized into three main climatic zones, namely, an arid zone (0.03 < aridity index (AI) < 0.2), a semi-arid zone (0.2 < AI < 0.5) and a sub-humid zone (0.5 < AI < 1.0), according to the Food and Agriculture Organization (FAO) aridity index map (Figure 2.1a) (Zeng et al., 2016). The Tibetan Plateau observatory of plateau-scale soil moisture (SM) and soil temperature (ST) (Tibet-Obs) (Su et al., 2011) is distributed throughout these climatic zones, including: 1) the Ngari network in the arid zone, located in the western part of the TP with the elevation varying between 4200 and 6300 m above mean sea level (a.s.l), where the annual mean temperature is 1.01 °C, the annual mean precipitation amount is 66.4 mm, the land cover is a typical desert environment dominated by bare soil surrounding desert shrub, and soils are prevailed by sandy soils mixed with gravel (Figure 2.1b); 2) the Naqu network in the semi-arid zone, located in a flat terrain with rolling hills at an average elevation of 4500 m a.s.l, where the annual mean temperature is -0.6 °C, the annual mean precipitation amount is 482 mm, the land cover is characterized as grasslands consisting of prairie grasses and mosses, and soils are dominated by loamy sand with organic matter and gravel (Figure 2.1c); and 3) the Maqu network in the sub-humid zone, located at the northeastern edge of the TP at elevations varying between 3430 m and 3750 m, where the annual mean temperature is 1.8 °C and the annual precipitation is 600 mm with more than 70% occurring during the monsoon season (from June till September). The land cover is dominated by short grasslands, and soils are dominated by fine minerals with high silt proportions (Figure 2.1d). Of these networks, the Naqu network is collocated with the multiscale SMST monitoring network in the central Tibetan Plateau (CTP-SMTMN) area (Yang et al., 2013).

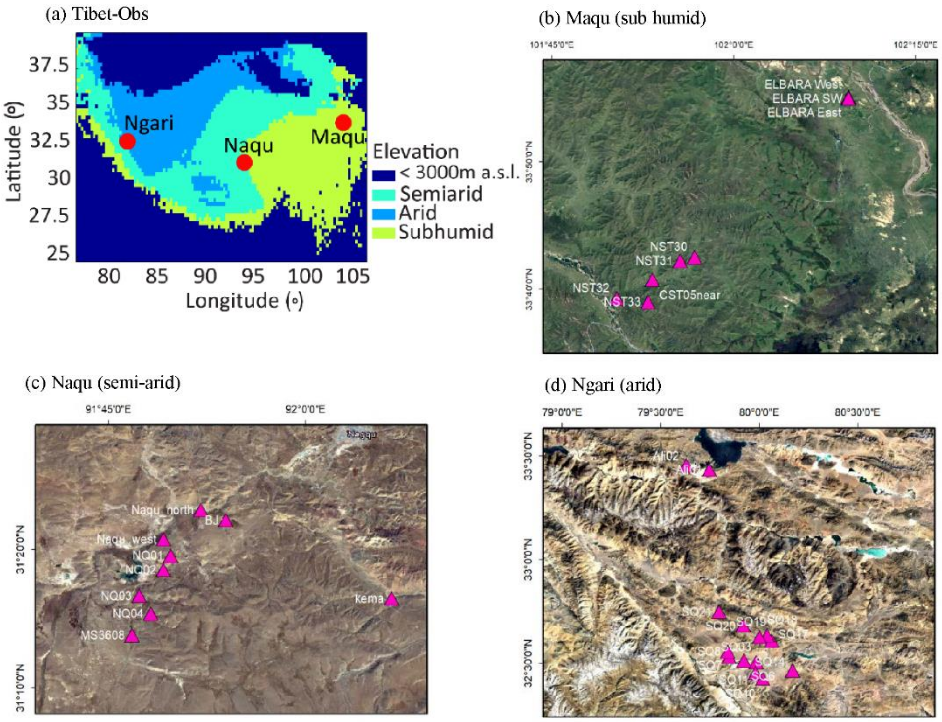


Figure 2.1 Location of Tibet-Obs and the spatial distribution of soil sampling sites across the three climate zones. (a) Tibet-Obs networks are distributed in the three climatic zones where the zones are classified based on the FAO aridity index map. The dark blue color represents the area around the Tibetan Plateau, with an elevation lower than 3000 m above sea level (a.s.l.) (Zeng et al. 2016). (b), (c) and (d) show sampling distributions of the Maqu network in the sub-humid zone, Naqu network in the semi-arid zone and Ngari network in the arid zone, respectively, in a kml image extracted from Google Earth. It should be noted that the image acquisition-times are August, February and December. The triangle in pink represents each sampling site.

A field experiment was carried out across the TP in August 2016, and soil core samples were collected and the field saturated hydraulic conductivity (K_s) was measured at various soil depths (Table 2.1 and Figure 2.1). Soils were vertically sampled using sample rings and augers (Eijkelkamp Soil & Water Company) in the vicinity of existing Tibet-Obs SMST stations (Su et al. 2011). Table 2.1 lists the specific sampling approach: 1) soil was sampled (c.a. 200 g) with a plastic bag used to measure the gravel content, soil texture and soil organic content (SOC); 2) soil was sampled with standard sample rings (5 cm in height, 100 cm^3 in volume) to determine the dry bulk density (BD), porosity and thermal conductivity (λ); 3) to generate soil water retention curves (SWRCs), a dedicated small-sample ring (1 cm in height, 20 cm^3 in volume) was used; 4) K_s was measured *in situ* with an Aardvark permeameter (2840 operating instructions - Eijkelkamp), which is a fully automated constant-head borehole permeameter. The Reynolds and Elrick solution aided with soil texture-structure category information (Elrick, 1989) was chosen to calculate K_s.

Table 2.1 Sampling approach for the basic soil properties, SHPs and STPs based on Tibet-Obs

Sampling Approach	Soil Texture	Gravel Content	Soil Organic Content	Dry bulk density	Porosity	Thermal Conductivity	Soil Water Retention Curve	Hydraulic Conductivity	Sampling Depths		
									Maqu	Naqu	Ngari
Plastic bag	√	√	√						5cm 10cm 20cm 40cm 80cm	5cm 10cm 20cm 40cm 50cm	5cm 10cm 20cm 40cm
Standard sample rings				√	√	√					
Small sample rings							√				
Profile Auger								√	10cm 20cm 40cm 80cm	10cm 20cm 40cm 50cm	10cm 20cm 40cm

Within the Maqu network, soil samples were collected at eight stations located in areas to the east, west and southeast of the ELEBARA-III radiometer location and in the southwest corner of the Maqu network (Figure 2.1a). K_s was measured at three locations near the ELBARA station and one location (CST05-near) in the southwest corner. Within the Naqu network, soil samples were obtained at eight sites along the southwest branch of the CTP-SMTMN network (Figure 2.1b), and K_S was measured at seven sites at BJ, Naqu_west, NQ01-04 and MS3608. Within the Ngari network, soils were sampled at 14 stations (Figure 2.1c). Eight sites at Ali02, SQ03, SQ07, SQ10, SQ17, SQ18, SQ20 and SQ21 were chosen for K_S measurement. In total, 155 soil samples were collected and loaded in plastic bags, 101 samples were collected in standard rings, and another 96 samples were collected in small-sample rings. Due to the remoteness and harsh environment on the TP, the locations chosen for soil sampling and fieldwork necessitates certain practical considerations, such as 1) the location should be accessible by railway, local roads or national roads; 2) the surrounding area should be flat enough to be representative of the local area.

2.2.2 Laboratory Experiments

Three categories of soil samples were processed. Among the 155 samples (59 from Ngari, 45 from Naqu and 51 from Maqu) in plastic bags, the collected soils were first separated into gravel and fine minerals (size < 2 mm) by using a 2-mm diameter mesh sieve and separately weighed to obtain gravimetric gravel fractions (GGFs). Sand (0.05 mm < size < 2 mm), silt (0.002 mm < size < 0.05 mm) and clay (size < 0.002 mm) percentages and the mean particle diameter of fine minerals (FD) were determined with a Malvern Mastersizer 2000 particle size analyzer (http://www.malvern.com), and the SOC was determined with a total organic content analytical instrument, Multi N/C 3100 (http://www.analytik-jena.de/). In regard to gravel, a set of sieves with diameters of 2, 2.5, 4, 5, 7, 10, 16, 20, 25, 31.5, 40 and 50 mm was used to obtain the particle size distribution and the mean particle diameter of gravel particles(GD).

The 101 undisturbed soil samples (35 from Ngari, 21 from Naqu and 45 from Maqu) contained in standard sample rings were saturated and then dried in an oven (105 °C) for 24 hours. The difference between the wet and dry weights with a known volume was used to calculate the porosity and BD. A KD2Pro thermal property analyzer connected to an SH-1 sensor (Decagon Devices) was employed to measure the heat capacity C_s and thermal conductivity λ during soil drying, providing drying C_s-SM and λ-SM curves.

The 96 samples contained in small-sample rings were reserved for SWRC experiments via the pressure cell method, but completing the entire task was considered highly time- and labor-consuming. Therefore, instead of utilizing all soil samples, only 30 out of the 96 samples were used among the E-east, E-west, E-southwest, CST05-near, NST30 and NST33 sites in the Maqu network. As the structure of the samples collected at the Naqu and Ngari networks was highly unconsolidated resulting in the material not remaining enclosed within the rings, only 25 undisturbed samples contained in standard rings were analyzed, which were obtained at the Naqu_north, SQ17, SQ18 and SQ21 sites.

The quality of the measured soil property dataset was evaluated based on four quality indicators (i.e., observation date, level of trust, data quality rating and accuracy) recommended by the World Soil Information (WoSIS) Institute (Ribeiro et al., 2015). These four indicators provide measures that allow investigators to recognize factors that may compromise the quality of certain data and hence their suitability for use (Ribeiro et al., 2015). The results indicate that the dataset is of trust level C, which is the highest level of this subjective measure inferred from soil expert knowledge. The entered data (level A) were standardized (level B), i.e., the data numbers were correspondingly aligned with the measured soil properties involved in GlobalSoilMap specifications (GlobalSoilMap, 2009) and with the measurement methods and units (please refer to the above paragraphs in section 2.2.2). The level B dataset was further harmonized (C) to enable sorting in a reference table (Ribeiro et al., 2015). For instance, tables of profile data (please refer to the raw data in the data repository (Zhao et al., 2018b)) describe soil profiles and their attributes (e.g., land cover, position) and constituent layers with their respective soil properties. These collated raw data included error-checking steps for possible inconsistencies. Furthermore, the values of the measured soil properties and SHPs & STPs were compared to those available in the literature to cross-check if they occurred within a reasonable range.

The compiled basic soil property and SHP & STP dataset, denoted as the Tibet-Obs dataset, will be further used to evaluate the existing soil datasets of the FAO-UNESCO Soil Map of the World (2007) (hereafter referred to as FAO-UNESCO), the Harmonized World Soil Database (hereafter referred to as HWSD) (FAO/IIASA/ISRIC/ISSCAS/JR, 2012), a Chinese dataset of soil properties (Shangguan et al., 2012; Shangguan et al., 2013) and soil hydraulic parameters using PTFs (Dai et al., 2013) released by the Beijing Normal University (hereafter referred to as BNU), SoilGrids1km (Hengl et al., 2014a) and the updated version of SoilGrids250m (Hengl, 2017) released by the International Soil Reference and Information Center (ISRIC) - WoSIS Institute, and hydraulic parameters based on SoilGrids1km and Schaap et al. (2001) PTFs (hereafter referred to as HPSS) (Montzka et al., 2017). A description of these existing datasets is listed in Table A1.1 of Appendix A. All datasets were linearly interpolated to match the measured dataset at specific depths to ensure (inter) comparability.

2.2.3 Parameterization Schemes

Many basic soil property-dependent schemes have been proposed for porosity estimation. The Cosby et al. (1984) univariate PTF that uses the sand percentage (hereafter referred to as the Cosby-S scheme, please refer to equation (A1.1) in

Appendix A) has been widely used, and it should be noted that there exists a multivariate PTF (Cosby et al. (1984)) that considers both clay and sand (Van Looy et al., 2017). The porosity can be inversely related to the soil dry bulk density (Hillel, 2003) and calculated from *in situ* BD data (hereafter referred to as the BD scheme, please refer to equation (A1.2)). In most cases, these schemes perform well. However, with the SOC in soils, the soil porosity tends to increase. Another factor affecting the porosity is the gravel content. With increasing gravel content, the porosity tends to decrease. Chen et al. (2012) parameterized the impact of SOC and gravel content via a porosity estimation scheme (hereafter referred to as the SocVg scheme, as expressed in equations (A1.3-A1.6)). Zhang et al. (2011) proposed a mixing-coefficient model to calculate the porosity of a binary mixture consisting of coarse (gravel) and fine components within a certain gravel content range (hereafter refer to as the BM scheme, as expressed in equations (A1.7-A1.10)). In this study, as shown in Figure 2.2, the Cosby-S, BD, SocVg and BM schemes were evaluated for their applicability in the three climate zones.

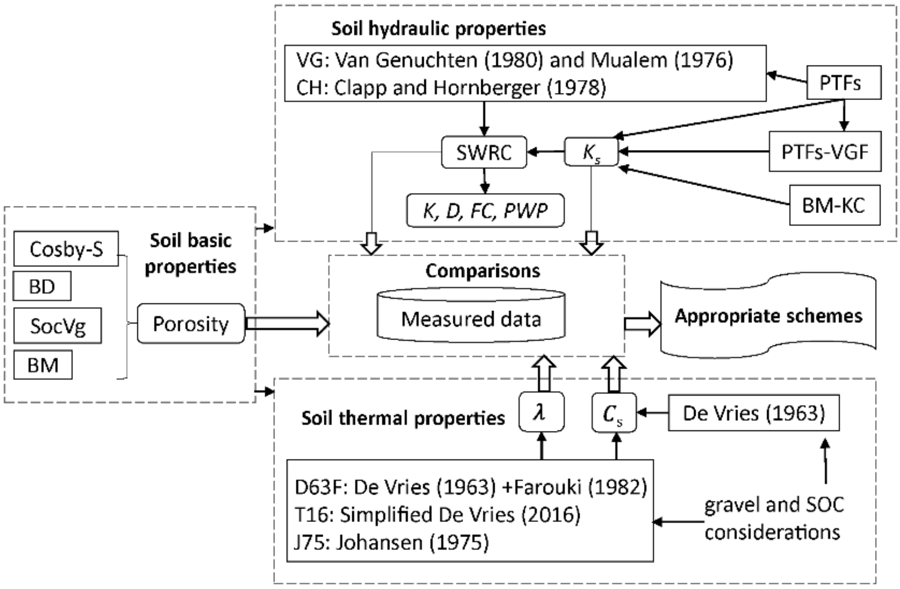


Figure 2.2 Flowchart of the implementation of the different porosity and SHP & STP schemes by using in situ basic soil property data. Dashed boxes indicate the various categories of the parameterization schemes and comparisons to measurements. Black arrows indicate the main data flow for these comparisons. Single arrows indicate the steps that occur internally for each part or they connect various parts. Rectangles represent the schemes. Rounded rectangles denote the porosity and SHP & STP parameters. K and D denotes the hydraulic conductivity and diffusivity, respectively.

Regarding SHP estimation, we selected the Clapp and Hornberger (1978) (hereafter denoted as CH) and the van Genuchten (1980) - Mualem (1976) (hereafter denoted as VG) schemes. Based on the measured SWRCs, we used the scaling method (please refer to equation (A1.15)) (Montzka et al., 2017) to determine the hydraulic parameters of the saturated soil moisture (θ_s), soil water potential at air-entry (φ_s), and b, which is an empirical parameter related to the pore-size distribution of the soil matrix in the CH function (please refer to equations (A1.11-A1.12)), and the parameters of θ_s, residual soil moisture θ_r, α, which is a parameter corresponding approximately to the inverse of the air-entry value, and n, which is a shape parameter of the VG model (as expressed in equations (A1.13-A1.14)). The field capacity (FC) and the permanent wilting point (PWP) (regarded as the SM under matric pressures of approximately -33 kPa and -1500 kPa, respectively) were also derived, as they are the main parameters of the soil water budget. Furthermore, the selected PTFs (please refer to Table A1.2 in Appendix A) were used to estimate the hydraulic parameters of SWRC-CH and SWRC-VG. Given that a good θ_s estimate will improve SWRC prediction, the optimal porosity scheme will be preselected to predict SWRC-CH and SWRC-VG. The estimated SWRCs based on PTFs were further compared to measurement-determined SWRCs to indicate the uncertainty in the application of the different PTFs.

The saturated hydraulic conductivity, K_s, combined with SWRCs-CH or SWRCs-VG is adopted to calculate the unsaturated hydraulic conductivity (K) and diffusivity (D). The PTFs used for SWRC-CH and SWRC-VG estimation also have corresponding equations (please refer to the footnotes in Table A1.2, Appendix A) to predict K_s, and most PTFs were developed based on fine minerals. To estimate K_s of a mixture containing gravel, Peck and Watson (1979) applied a heat-flow analogy correlating the K_s value of the mixture with the K_s values of both fine minerals and volumetric gravel fraction (VGF) (hereafter referred to as the PTFs-VGF scheme, as expressed in equation (A1.16)). The PTFS-VGF scheme can be applied to soils with a low gravel content (Zhang et al., 2011). It is noted that the PTFS-VGF scheme requires input ($K_{sat,f}$, as described in A.3) from K_s estimation based on PTFs. Furthermore, Koltermann (1995) used the Kozeny–Carman equation to estimate the hydraulic conductivity of binary mixtures, and suitable grain diameter estimation was deemed important (Kamann et al., 2007). To improve the performance of the Kozeny–Carman equation, Zhang et al. (2011) introduced the BM scheme for porosity estimation and a power-averaging method to calculate the representative grain diameter (hereafter referred to as the BM-KC scheme, as expressed in equations (A1.17-A1.19)). In this study, the standard PTFs (Table A1.2), PTFS-VGF and BM-KC schemes were employed, as shown in Figure 2.2.

Several (semi-) empirical models have been developed to estimate soil thermal conductivity λ. De Vries (1963) developed a Maxwell equation analogous to a physics-based model to describe λ (please refer to equation (A1.22)). This model can predict λ accurately, although this is complicated by the fact that at least five soil mineral components and their separate shape features must be considered (Tarnawski & Wagner, 1992). Furthermore, the effect of vapor movement caused by the temperature gradient is parameterized in the De Vries (1963) model. It should be noted that the consideration of soil vapor flow is critical to accurately investigate the simultaneous transfer of moisture and heat, particularly in semi-arid and arid environments (Zeng & Su, 2013; Zeng et al., 2011b; Zeng et al., 2011c). Farouki (1981) proposed an alternative method and regarded liquid water as the continuous medium and soil minerals as uniform particles in the De Vries (1963) model. In this model, λ of soil minerals was estimated with a geometric mean equation according to the quartz content in soil minerals and the λ values of quartz and other soil minerals (please refer to equation (A1.23)). The λ value of vapor and the shape factor of air pores were calculated in terms of the water content and porosity, respectively (please refer to equations (A1.24-A1.25)) (hereafter referred to as the D63F scheme). Tian et al. (2016) developed a simple and generalized De Vries-based model, which assumed that the λ values and shape features of soil minerals are determined by the soil texture (sand, clay and silt), and that the effect of vapor movement is negligible (hereafter referred to as the T16 scheme, equations (A1.26-A1.29)). The empirical model proposed by Johansen (1975) used the Kersten (1949) number and λ under dry and saturated conditions to estimate λ (hereafter referred to as the J75 scheme, equations (A1.30-A1.35)). In this study, as shown in Figure 2.2, the D63F, T16 and J75 schemes were adopted. For each λ scheme, a comparison was made using parameters (i.e., the λ value of soil minerals) with (equation (A1.35)) and without (equation (A1.23)) gravel/SOC consideration. The De Vries (1963) model was applied to calculate C_s (equations (A1.20-A1.21)). Details on the porosity and SHP & STP schemes are listed in Appendix A (A1.1–A1.5).

2.3 Tibet-Obs Dataset

2.3.1 Data availability

The soil physical dataset is available at the 4TU.Center for Research Data at https://data.4tu.nl/repository/uuid:c712717c-6ac0-47ff-9d58-97f88082ddc0 (Zhao et al., 2018b). The data are stored as Excel files. A readme file describes the structure of Excel files, measurement devices and contact information. The download links of the existing soil property datasets considered in this chapter are included in the .txt file. The sampling locations are stored in a .kmz file. The raw data for each sampling site are also provided.

2.3.2 Basic analyses of the Tibet-Obs dataset

Soil texture

Figure 2.3 shows the mean sand, clay and silt percentages, GGF, SOC, FD and GD at different depths across the three climate zones on the TP. Within the Ngari network in the arid zone (Figure 2.3a), the mean sand content was approximately 80%, with higher values in the 5- and 10-cm surface layers than that in deeper layers. The silt and gravel contents ranged from 10-20%, and the percentages increased with the depth. The clay content and SOC were 3% and 0.8%, respectively, and remained constant along the profile. The FD and GD ranged from 0.19-0.24 mm and 4-8 mm, respectively, and showed a tendency to increase from the top to a depth of 20 cm but decreased in the deeper layers. It can be concluded that the soil texture in the arid zone consists of a high proportion of coarse sand accompanied by gravel and that the gravel content increases until 20 cm and then slightly decreases in the deeper layers.

Within the Naqu network in the semi-arid zone (Figure 2.3b), the mean sand fraction ranged from 70-80% with a slight decrease with the depth. The silt and clay contents ranged from 15-25% and 4-8%, respectively, and increased with the depth. GGF exceeded 50% at soil depths of 40 and 50 cm, while it was much lower in shallow layers. Mean FD and GD ranged 0.18-0.22 mm and 4-8 mm, respectively. GD at deep layers was larger than that at shallow layers. SOC approached 10% in the surface layers but quickly declined in the deep layers. In summary, the soil texture in the semi-arid zone is dominated by a high percentage of sand mixed with a low proportion of gravel, but with a high SOC in shallow layers and mainly mixed with large gravel particles in deep layers.

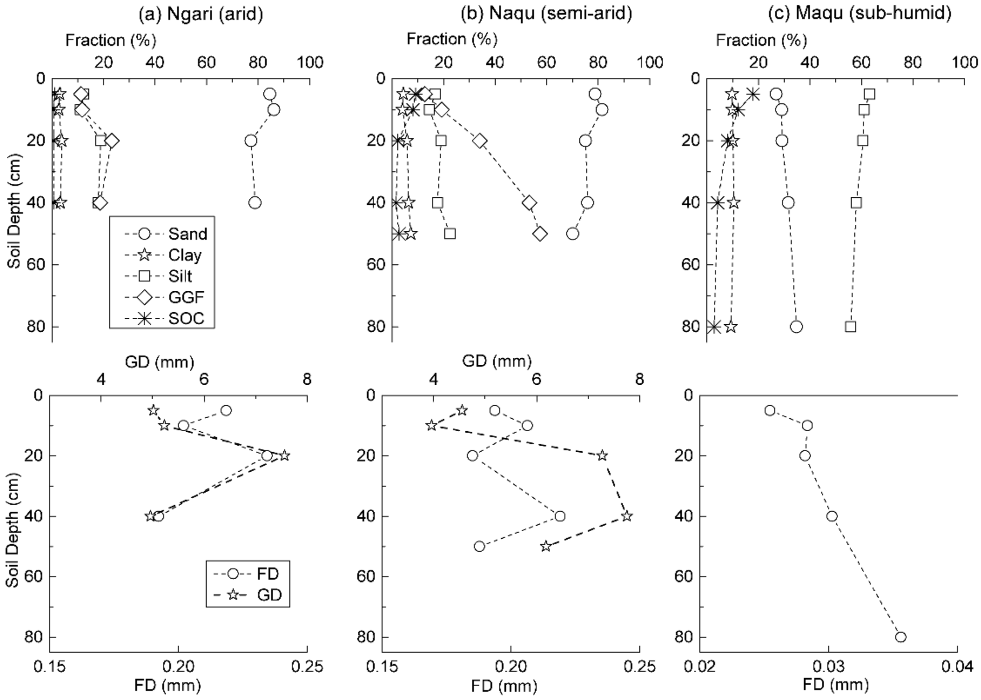


Figure 2.3 Profiles of the mean basic soil properties in the three climate zones. Top panel: Variations in the sand, clay, silt, GGF, and SOC at the various depths. Bottom panel: Variations in GD and FD at the different depths. GGF is the gravimetric gravel fraction. SOC is the soil organic matter content. FD is the mean particle diameter of fine minerals. GD is the mean particle diameter of gravel particles.

Within the Maqu network in the sub-humid zone (Figure 2.3c), the mean silt and clay contents were approximately 60% and 10%, respectively, with a smoothly decreasing trend along the profile. The mean sand fraction ranged from 28-40% and increased with the depth. No gravel was found. The mean FD value ranged from 0.024-0.036 mm, and the fine soil mineral particles in the deep layers (40 and 80 cm) were larger than those in the shallow layers (as shown in the lower panel of Figure 2.3c). Similar to the SOC profile distribution in the Naqu network, the SOC reached almost 20% in the surface soil layers and declined to 2.8% at 80 cm. The soil texture in the sub-humid zone is characterized as being dominated by a high percentage of silt content with a relatively high SOC in the shallow layers and mainly fine sand in the deep layers.

Dry bulk density and Porosity

Within the Ngari network in the arid zone (Figure 2.4a), BD varied slightly (between 1.55 and 1.65 g/cm^3) with the depth, showing a peak at 10 cm. The porosity of the surface layer was slightly higher than that of the deeper layers, with a mean profile porosity of 0.33. The porosity at 20 cm was the lowest along the profile, which might be caused by this layer containing the highest proportion of gravel and the largest GD and FD values (please refer to Figure 2.3a,). Within the Naqu network in the semi-arid zone (Figure 2.4b), BD increased continuously with the depth, with a minimum of 1 g/cm^3 in the top layer and a maximum of 2.1 g/cm^3 in the bottom layer. The porosity peaked at approximately 0.6 in the top layer, and monotonously decreased to 0.25 in the bottom layer. Combined with soil texture analysis (please refer to Figure 2.3b), variations in BD and porosity along the profile were inferred to be related to the high SOC in the surface layer and the high gravel content in the bottom layer. Within the Maqu network in the sub-humid zone (Figure 2.4c), BD ranged from 0.8 to 1.5 g/cm^3 and increased with the depth, while the porosity decreased with the depth and ranged from 0.72 to 0.45. The profile pattern of BD and porosity might be the result of SOC variation in the surface layer and soil texture fraction variation in the deeper layers as shown in Figure 2.3c. In summary, the profiles of BD and porosity differed with the soil texture between the three climate zones, and both SOC and gravel content affected the porosity. The overall porosity of the shallow layers (5, 10 and 20 cm) increased from the arid to the semi-arid and sub-humid zones, while that of the deeper layers (>= 40 cm) exhibited an increase from the semi-arid to the arid and sub-humid zones.

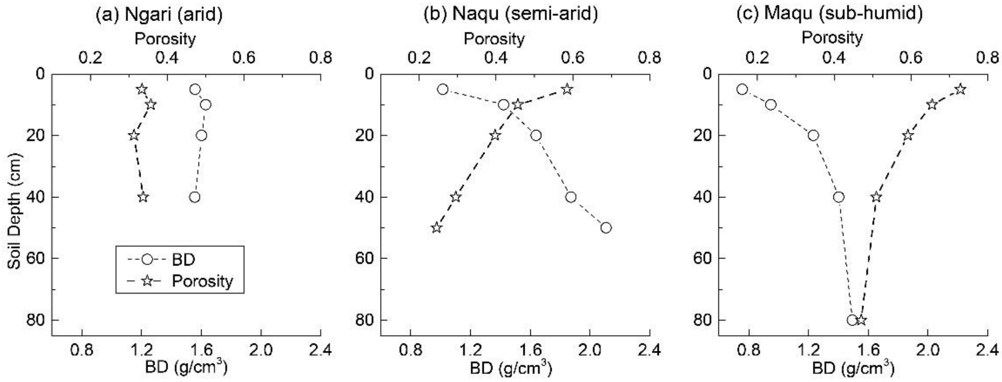


Figure 2.4 Profiles of the mean dry bulk density (BD) and porosity in the three climate zones.

Soil water retention curve (SWRC) and saturated hydraulic conductivity (K_s)

Figure 2.5 shows that the pressure-cell measured SWRCs (markers in the figure) differed across the three climate zones. Within the Ngari network in the arid zone, soil water retention decreased with increasing suction (Figure 2.5a). The same situation occurred in the deeper layers in the Naqu network in the semi-arid zone (Figure 2.5b). Within the Maqu network in the sub-humid zone, soil water retention was high and gradually decreased with increasing suction (Figure 2.5c). Figure 2.5 also shows that the CH and VG models captured the retention characteristics of soil water (lines in the figure) well across the three climate zones. Determined parameters of $[\theta_s, b, \varphi_s]$ of the CH model and $[\theta_r, \theta_s, \alpha, n]$ of the VG model based on the measured SWRCs and scaling method are listed in Table 2.2.

Table 2.2 Pressure-cell determined parameters of the CH and VG models in the three climate zones. The scaling method used for the determination is equation (A1.15).

Region	Depth (cm)	CH				
		b	φ_s	θ_s	FC	PWP
		-	cm	$m^3 m^{-3}$	$m^3 m^{-3}$	$m^3 m^{-3}$

Ngari (arid)	5	0.19	4.35	0.3	0.21	0.1
	10	0.16	5.83	0.32	0.24	0.13
	20	0.16	2.02	0.27	0.17	0.09
	40	0.18	2.45	0.28	0.17	0.09
Naqu (semi-arid)	5	0.07	0.02	0.51	0.3	0.23
	10	0.1	11.21	0.43	0.39	0.27
	20	0.16	4.59	0.39	0.29	0.15
	40	0.13	1.64	0.39	0.27	0.16
	50	0.19	0.58	0.39	0.18	0.09
Maqu (sub-humid)	5	0.28	39.04	0.79	0.75	0.29
	10	0.25	39.17	0.72	0.7	0.29
	20	0.24	37.89	0.66	0.65	0.27
	40	0.2	33.13	0.54	0.53	0.25
	80	0.27	36.61	0.56	0.56	0.21

Region	Depth (cm)	VG					
		θ_r	θ_s	α	n	FC	PWP
		$m^3\ m^{-3}$	$m^3\ m^{-3}$	cm^{-1}	-	$m^3\ m^{-3}$	$m^3\ m^{-3}$
Ngari (arid)	5	0.03	0.25	0.02	1.39	0.22	0.09
	10	0.03	0.29	0.02	1.31	0.26	0.12
	20	0.03	0.2	0.02	1.34	0.18	0.08
	40	0.03	0.21	0.03	1.37	0.18	0.08
Naqu (semi-arid)	5	0.03	0.33	0.05	1.1	0.3	0.23
	10	0.04	0.44	0.04	1.15	0.4	0.25
	20	0.04	0.35	0.04	1.29	0.3	0.14
	40	0.04	0.3	0.02	1.27	0.28	0.14
	50	0.04	0.23	0.03	1.43	0.2	0.08
Maqu (sub-humid)	5	0.05	0.77	0.02	1.33	0.61	0.3
	10	0.05	0.6	0.01	1.27	0.58	0.3
	20	0.05	0.54	0.01	1.25	0.53	0.28
	40	0.05	0.47	0.01	1.23	0.45	0.25
	80	0.05	0.49	0.01	1.31	0.49	0.21

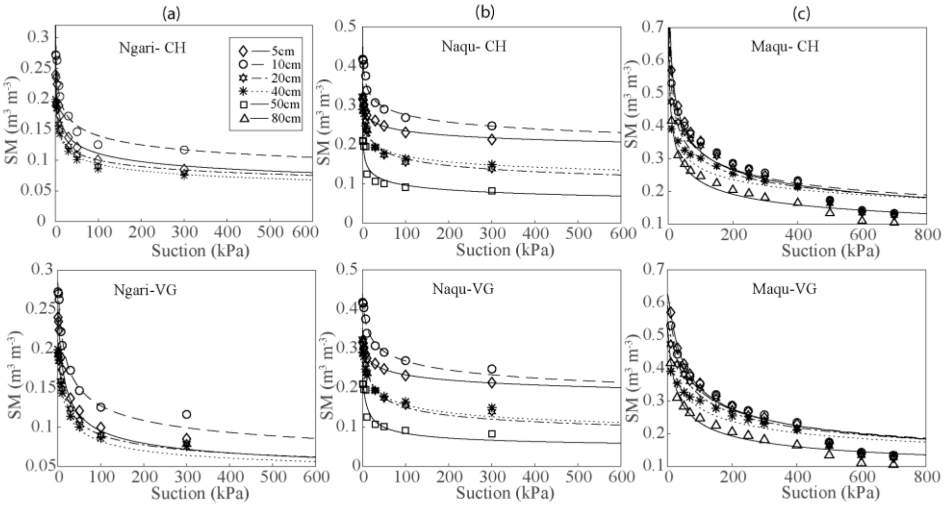


Figure 2.5 Average observational SWRCs and determined SWRCs-CH and SWRCs-VG from the scaling method at the different depths in the three climate zones: Ngari in the arid zone, Naqu in the semi-arid zone and Maqu in the sub-humid zone. Dots indicate the average observed soil moisture content under a specific suction. Lines represent the determined SWRCs-CH and SWRCs-VG.

Within the Ngari network in the arid zone (Figure 2.6a), the magnitude of the mean K_S value was on the order of 10^{-5} (m/s). K_S at 20 cm was lower than that at the other depths, which might be attributed to the lowest values of the porosity in this layer (please refer to Figure 2.4a). Within the Naqu network in the semi-arid zone (Figure 2.6b), the mean K_S value exhibited a variation of one order of magnitude with the depth, namely, 10^{-6} (m/s) at depths of 10, 20 and 50 cm and 10^{-5} (m/s) at a depth of 40 cm. In the Maqu network under the sub-humid zone (Figure 2.6c), K_S also differed by one order of magnitude: 10^{-6} (m/s) at depths of 5, 10, 20 and 80 cm and 10^{-7} (m/s) at a depth of 40 cm. It should be noted that the K_S profiles of both the semi-arid and sub-humid zones revealed a lower K_S value in the shallow layers than that in the deeper layers. This mainly occurs due to the negative correlation between the saturated hydraulic conductivity and soil organic carbon in soils where the hydrophobic functional group might dominate the organic carbon composition and reduce the soil wettability (Ellerbrock et al., 2005; Nemes et al., 2005; Wang et al., 2009). K_S varies with the soil texture in the three climatic zones, and both SOC and the gravel content yield an effect. At a certain depth, where the basic soil properties undergo a transition (please refer to Figure 2.3), K_S reaches a minimum. The mean and standard deviation of the soil properties of the profiles in the three climate zones are listed in Appendix A (Tables A1.3-A1.5).

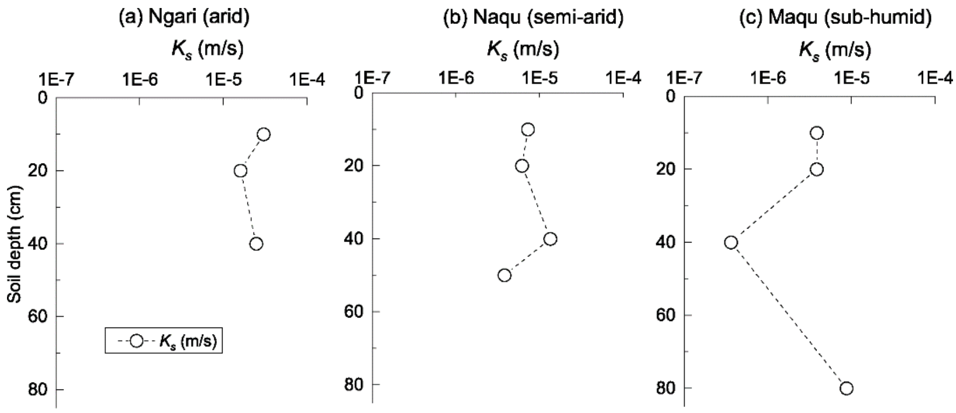


Figure 2.6 Profiles of the mean saturated hydraulic conductivity (K_S) in the three different climate zones.

Gravel impact on porosity and K_S

Figures 2.7a&b show that the porosity did not change with GGF increasing from 0-0.3 in the shallow layers, while at GGF > 0.4, the porosity tended to decline with increasing GGF, especially in the deeper layers. For example, the porosities of the layers with GGF values of 0.6 and 0.72 at 20 cm and 40 cm depths were lower than those with GGF < 0.3 at 5 cm and 10 cm depths (Figure 2.7a). With more gravel particles embedded in the matrix, the flow paths in the soil would become blocked and the porosity reduced (Zhang et al., 2011). However, the porosity did not always decrease with increasing GGF. The porosity of the layer with a GGF value of 0.84 in the semi-arid zone was higher than that of the layers with a GGF varying between 0.4 and 0.6 at the 50 cm depth (Figure 2.7b). The porosity of the layer with a GGF value of 0.7 at the 20 cm depth in the arid zone was also higher than that with a GGF value of 0.6 at the 40 cm depth (Figure 2.7a). The porosity tended to increase with increasing GGF, because when the GGF is relatively high (> 1 minus the porosity of gravel), connected pores can form among gravel particles and thus increase the porosity (Zhang et al., 2011).

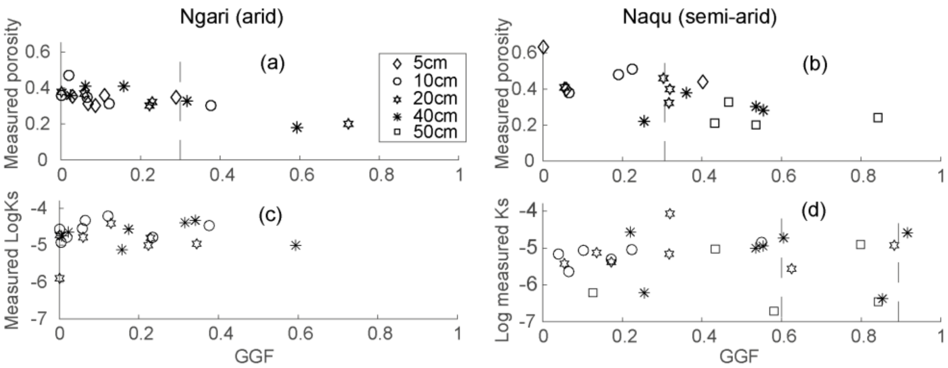


Figure 2.7 Scatter points of the measured porosity (top panel) and K_s (bottom panel) with the GGF in the different depths in the arid and semi-arid zones.

Figures 2.7c&d show a slight decrease in K_s at 10 cm for GGF < 0.62 and a slight increase in K_s at 20 cm and 40 cm for GGF > 0.8, which is consistent with the changes in porosity. These observations clearly show that gravel exerts a distinct impact on the porosity and K_s in the arid and semi-arid zones. It should be noted that although the *in situ* K_s measurements were conducted at locations adjacent to soil sampling sites, heterogeneity may have affected the values of the soil properties and parameters throughout our sampling procedures, similar to any soil field experimentation. Nevertheless, the current findings based on field experiments are in line with the reported findings based on laboratory experiments (Koltermann, 1995; Sakaki & Smits, 2015; Zhang et al., 2011).

Heat capacity C_s and thermal conductivity λ

Figures 2.8a&b show that the heat capacity C_s increased and decreased with increasing SM within the Ngari and Naqu networks in the arid and semi-arid zones, respectively. Some samples collected from these two networks are fine-grained soils mixed with gravel particles. In regard to these samples, it was not easy to vertically insert the needle probes of the KD2Pro device (please refer to section 2.2). Instead, the needle probes were buried in the surface of the sample as the alternative for the measurement. Additionally, the KD2 needle probes might experience a slight deviation upon contact with gravel. All these factors might cause the fluctuation in C_s with increasing SM, while an overall rising trend was still observed in Figures 2.8a&b. Figure 2.8c shows that C_s almost steadily increased with SM within the Maqu network in the sub-humid zone. The samples collected from the sub-humid zone are all fine-grained soils, and the needle probes of the KD2 device were easily inserted, thus forming a steady environment for the measurements. Figures 2.8a&b&c show that no distinct layering occurred at the different depths for C_s with SM in the three climate zones. C_s ranged from 1 MJ m^{-3} K^{-1} in the oven-dry state to 2.5 MJ m^{-3} K^{-1} as the soil reached complete saturation in the arid zone, 0.5 to 3 MJ m^{-3} K^{-1} in the semi-arid zone, and 0.5 to 2.4 MJ m^{-3} K^{-1} in the sub-humid zone.

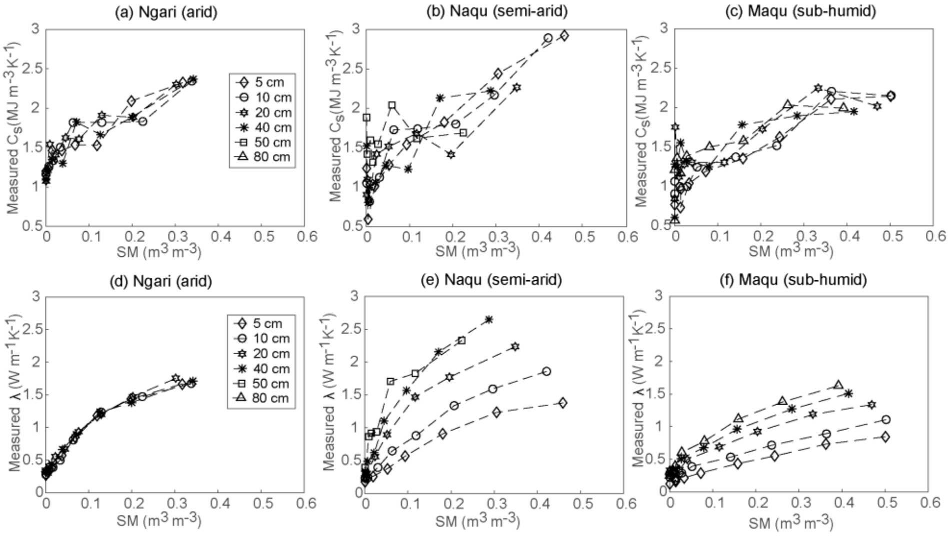


Figure 2.8 Mean soil heat capacity (C_s) and thermal conductivity (λ) with the water content (SM) at the different depths in the three climate zones.

Figures 2.8d&e&f show the λ-SM relationship with the depth. In the arid zone (Figure 2.8d), the λ-SM curves were very similar at each depth due to the nearly homogenous sandy soils across the whole profile (please refer to Figure 2.3a). The mean λ value ranged from 1.8 W m^{-1} K^{-1} at full saturation to 0.2 W m^{-1} K^{-1} as the soils reached the oven-dry state. In the semi-arid zone (Figure 2.8e), the λ-SM curves were stratified, and the soils with gravel in the deeper layers (please refer to Figure 2.3b) clearly attained a higher λ value (>2 W m^{-1} K^{-1}) than that of the other layers and other climate zones. In the sub-humid zone (Figure 2.8f), the λ-SM curves also presented variation with the depth, albeit within a much narrower range than that in the semi-arid zone. This variation is mainly caused by the sand distribution along the profile, which increased slightly with the depth (please refer to Figure 2.3c). The mean λ value in the sub-humid zone ranged from 1.6-0.2 W m^{-1} K^{-1} as soils dried out. Furthermore, the surface layers in the semi-arid and sub-humid zones exhibited lower λ values (Figures 2.8e&f) because of the SOC influence.

2.4 Applications of the Tibet-Obs Dataset

2.4.1 Assessing the parameterization schemes for LSMs

Porosity estimation

With the use of the basic soil property data of Tibet-Obs and four schemes, the porosity was estimated. Comparisons to measured porosities (please refer to Table A1.6) indicated that the BD scheme performs the best in the estimation of porosity in the different depths in the three climate zones, as this bulk estimation scheme considers both gravel and fine minerals. The Cosby-S scheme overestimated the porosity in the arid zone and provided constant porosity values in the semi-arid and sub-humid zones. The SocVg scheme also overestimated the porosity because the assumed porosity of gravel soils with a theoretical minimum value (0.363) was higher than the observed maximum porosity (0.31) (Wu & Wang, 2006). The BM scheme suitably estimated the porosity of soils with more gravel, especially in the deeper layers of the arid and semi-arid zones.

SWRC and K_s estimation

With the basic soil property data of Tibet-Obs and the selected PTFs (please refer to Table A1.2), the parameters of SWRC-CH and SWRC-VG were estimated (please refer to Tables A1.7-A1.8). Figure 2.9 shows comparisons of the estimated SWRCs based on PTFs combined with the BD porosity scheme to the measurement-determined SWRCs at 5 cm (please refer to section 2.3.2). The Saxton et al. (1986) PTFs overestimated the SWRCs-CH in the arid zone (Ngari), while the PTFs given by Campbell and Shiozawa (1992) and Saxton and Rawls (2006) underestimated them (Figure 2.9a), whereas the Cosby et al. (1984) PTFs (1 and 2) yielded good SWRC-CH predictions with lower absolute biases (please refer to Table A1.9) over the measurements. In the semi-arid zone (Naqu) (Figure 2.9b), all PTFs underestimated the SWRCs-CH at 5 cm, while the Cosby et al.

(1984) PTFs (1 and 2) and the Saxton et al. (1986) and Saxton and Rawls (2006) PTFs captured them well with lower biases (please refer to Table A1.9) over the measurements. In the sub-humid zone (Maqu) (Figure 2.9c), the Cosby et al. (1984) PTFs (1) and Saxton et al. (1986) PTFs predicted the SWRCs-CH well. It should be noted that in combination with the BD scheme, the Cosby PTFs (1) performed much better in SWRC-CH estimation than did the estimates obtained with the Cosby PTFs (1) combined with the Cosby-S porosity scheme (please refer to section 2.2.2). In contrast, without the BD scheme, the Saxton and Rawls (2006) PTFs were found to perform better (as indicated in Table A1.10 in Appendix A) in the semi-arid and semi-humid zones.

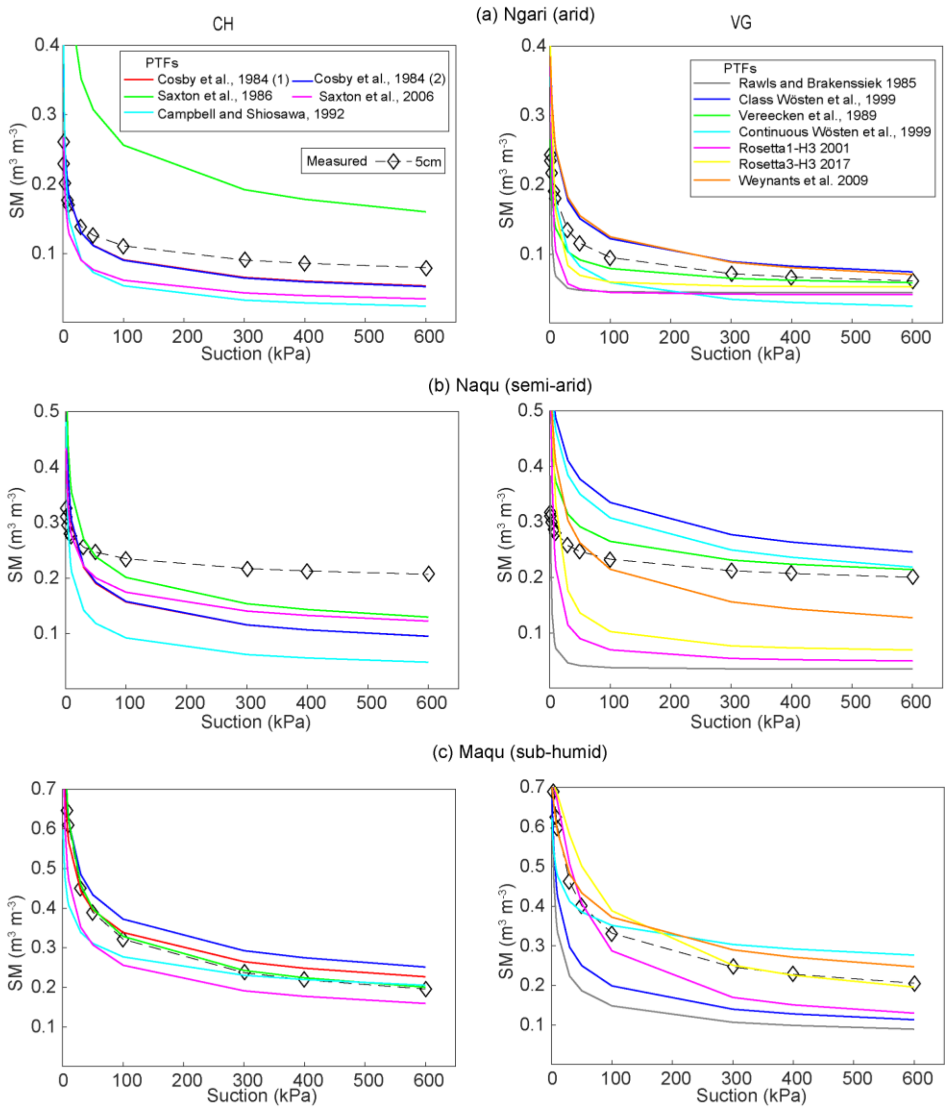


Figure 2.9 Comparison of the estimated SWRCs via PTFs combined with the BD scheme and the measurement-determined SWRCs at 5 cm in the three climate zones. It should be noted that the SWRC estimated with the Vereecken et al. (1989) PTFs is beyond the range in the sub-humid zone and not considered (right figure in Figure 9c).

In regard to SWRCs-VG estimation, the Rosetta1-H3 and Rosetta3-H3 PTFs were developed based on a mixed database (Schaap et al., 2001). Figure 2.9 (right panel) shows that they underestimated the SWRCs-VG across the three climate zones, as did the Rawls and Brakensiek (1985) PTFs. The Weynants et al. (2009) PTFs underestimated the SWRCs-VG in the semi-arid zone, while the Class Wösten et al. (1999) PTFs overestimated them (Figure 2.9b). The Vereecken et al. (1989) PTFs, which were developed based on a database containing hydraulic properties measured with the same measurement techniques across all samples (Vereecken et al., 2010), performed well when m was set to 1. However, these PTFs did not perform well for m=1-1/n in the VG model and estimated SWRCs-VG that were out of the range in the sub-humid zone. The continuous Wösten et al. (1999) PTFs were developed based on the Hydraulic Properties of European Soils (HYPRES) database and as such were more akin to the database of Vereecken et al. (1989). The Weynants et al. (2009) PTFs were developed based on the Vereecken et al. (1989) database and included BD as a variable. These two PTFs predicted the SWRCs-VG well in the three climate zones. Comparisons of the estimated SWRCs via PTFs to the measurements at 10, 20 and 40 cm were shown in Figures A1.1-A1.3 in Appendix A. Accordingly, the Cosby et al. (1984) PTFs (1) and the continuous Wösten et al. (1999) PTFs combined with the BD porosity scheme are suggested to be more applicable for the prediction of the SWRC-CH and SWRC-VG, respectively, across the three climate zones.

Adopting the basic soil property data of Tibet-Obs as input, K_s was estimated by using the PTF scheme (please refer to the footnotes in Table A1.2 in Appendix A), the empirical PTFS-VGF scheme (please refer to equation (A1.16)) and the semi-physical BM-KC scheme (as expressed in equations (A1.17-A1.19)). Comparing the estimated values against *in situ* measured K_s values, Figures 2.10a&d show that the PTF scheme yielded a lower bias in $\mathrm{Log}_{10}K_s$ prediction than that yielded by PTFS-VGF and BM-KC schemes in the arid zone (Ngari). In particular, the PTFs given by Cosby et al. (1984) (1 and 2) predicted good K_s values that could be used in the CH model to estimate the hydraulic conductivity

(K), similar to the Rosetta1-H3 PTFs, Rosetta3-H3 PTFs and Rawls and Brakensiek (1985) PTFs did with the VG model. The K_S values derived based on BM-KC scheme had a smaller RMSE with measurements at 40 cm depth, indicating the gravel impact on K_S.

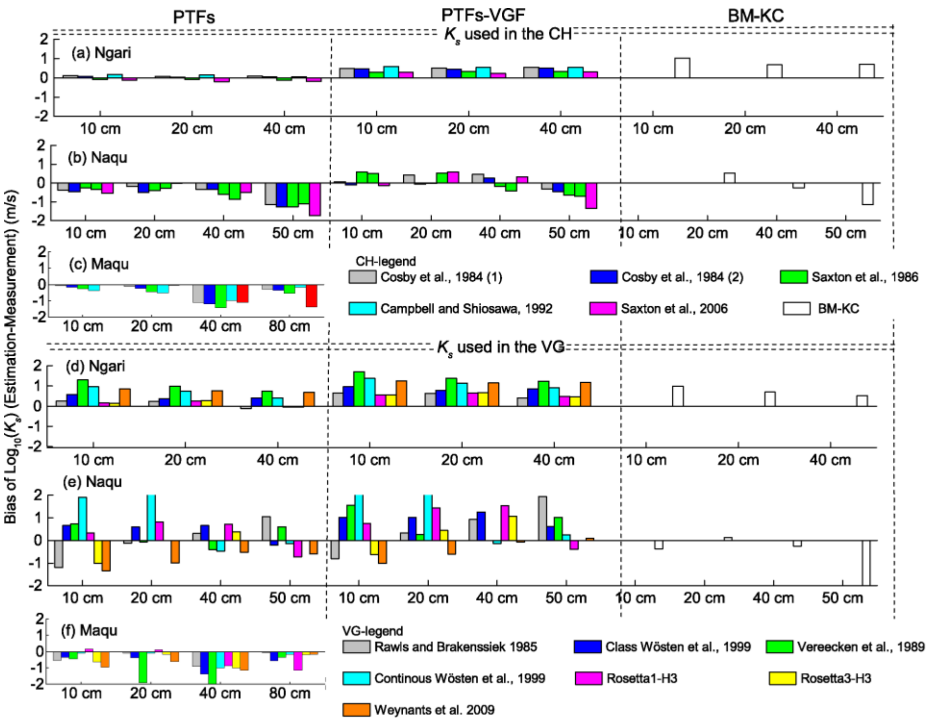


Figure 2.10 Comparisons of K_S, derived from the PTFs, PTFs-VGF and BM-KC schemes with the CH and VG models, to field measurements at the different depths in the three climate zones.

Figures 2.10b&e show that the BM-KC scheme predicted better K_S at depths of 10, 20 and 40 cm in the semi-arid zone (Naqu) than did most PTFs and PTFs-VGF. Regarding the K_S estimates used in the CH model, the Cosby et al. (1984) (1) PTFs performed the best at shallow depths, while the PTFs-VGF among these PTFs performed better in the deeper layers at 40 and 50 cm. In terms of usage in the VG model, the K_s values derived from the PTFs and PTFs-VGF schemes were almost the same, indicating that the estimated K_s used in the VG model is less affected by gravel. The Rosetta1-H3 PTFs better predicted K_s than did the other PTFs. Figures 2.10c&f show that most of the PTFs underestimated K_s, while the selected PTFs (i.e., Cosby (1) and Rosetta1-H3) in the arid zone also predicted K_s values close to the measurements in the sub-humid zone (Maqu). In summary, the Cosby et al. (1984) (1) and Rosetta1-H3 PTFs are appropriate for K_s estimation, which are used in the CH and VG models, respectively, across the three climate zones. The PTFs-VGF of the Saxton and Rawls (2006) scheme should be applied in the deeper layers in the semi-arid zone, where gravel is abundant in the soil.

C_s and λ estimation

With the use of the basic soil property data of Tibet-Obs, C_s was estimated through the De Vries (1963) model. Compared to the C_s measurements, this scheme performed well in the three climate zones. Furthermore, considering the SOC impact, it improved the C_s estimates (as indicated in Table A1.11) for soils at top layers in the semi-arid and sub-humid zones.

Based on the Tibet-Obs basic soil property data, the D63F, T16 and J75 schemes combined with the BD porosity scheme were used to estimate λ. In the arid (Ngari) and semi-arid (Naqu) regions, the estimation of λ considered two scenarios: with (Case 1) and without (Case 2) gravel impact. In the sub-humid region (Maqu), λ estimations with (Case 1) and without (Case 2) SOC impact were considered. Table 2.3 revealed that the λ values derived from the D63F model had a lower bias in all cases compared to the measurement over the other schemes in the three climate zones. The T16 scheme overestimated λ, which may be attributed to its ideal assumption that the λ value of soil minerals is totally determined by sand, clay and silt particles. The J75 scheme generally underestimated λ.

Table 2.3 Biases of the λ estimates based on the D63F, T16 and J75 schemes combined with the BD scheme along the different depths in the three climate zones and the measurements. Case 1 is the bias (listed in the upper part of the table) derived from the schemes not considering gravel impact parameterization in the arid and semi-arid zone or SOC impact parameterization in the sub-humid zone. Case 2 is the bias (listed in the lower part of the table) with these parameterizations considered. The unit of the listed value is W m^{-1} K^{-1}.

Schemes		D63F	T16	J75	D63F	T16	J75
Ngari (arid)	5 cm	0.06	0.3	-0.26	0	0.26	-0.29
	10 cm	0.18	0.32	-0.24	0.08	0.25	-0.23
	20 cm	0.02	0.3	-0.42	-0.13	0.18	-0.43
	40 cm	-0.01	0.07	-0.37	-0.14	-0.02	-0.39
Naqu (semi-arid)	5 cm	-0.09	0.02	-0.38	-0.2	-0.13	-0.34
	10 cm	-0.02	0.25	-0.21	-0.11	0.18	-0.28
	20 cm	-0.03	0.31	-0.35	-0.16	0.2	-0.43
	40 cm	0.01	0.54	-0.35	-0.26	0.27	-0.53
	50 cm	0.2	0.99	-0.23	-0.2	0.44	-0.49
Maqu (sub-humid)	5 cm	-0.01	0.05	-0.13	-0.14	-0.16	-0.22
	10 cm	-0.04	0.1	-0.2	-0.19	-0.16	-0.31
	20 cm	-0.06	0.18	-0.29	-0.22	-0.14	-0.41
	40 cm	-0.05	0.23	-0.32	-0.13	0.01	-0.41
	80 cm	-0.08	0.2	-0.37	-0.12	0	-0.45

Table 2.3 also indicates that the D63F scheme improved the λ estimates for surface layers in the arid zone and at a depth of 50 cm when incorporating gravel impact parameterization (lower biases in Case 2). The improvement also occurred with the T16 scheme, while the biases tended to be higher under the J75 scheme. In the sub-humid zone, the biases also increased for all schemes when SOC impact parameterization was considered. Although the parameterization of the

SOC impact was demonstrated to improve the λ estimate for the top layer (SOC > 12%) in the eastern TP (Chen et al., 2012; Zheng et al., 2015b), it should be noted that in these studies of porosity estimation, the Cosby-S scheme was used instead of the BD scheme as adopted in this chapter. The comparisons in Table 2.3 indicate that the D63F scheme combined with the BD porosity scheme can predict λ well in the three climate zones. It should be noted that combined with the Cosby-S scheme, the D63F scheme also performs well (as shown in Figure A1.4 in Appendix A).

2.4.2 Evaluation of the existing soil datasets

The current existing global and regional soil datasets, including FAO-UNESCO (FAO/UNESCO, 2007), HWSD (FAO/IIASA/ISRIC/ISSCAS/JR, 2012), BNU (Shangguan et al., 2012; Shangguan et al., 2013), SoilGrids1km (Hengl et al., 2014a), SoilGrids250m and HPSS (Montzka et al., 2017), were extracted on the TP and compared to the *in situ* and laboratory measurements of Tibet-Obs.

Basic soil properties

Figure 2.11 shows that all datasets underestimated both the sand fraction and BD in the arid and semi-arid regions, but overestimated them in the sub-humid region. Regarding the silt fraction, the pattern was reversed. Almost all datasets overestimated the silt fraction in the arid and semi-arid regions (only FAO-UNESCO underestimated the silt fraction very slightly in the semi-arid region) and underestimated the silt fraction in the sub-humid region. All datasets overestimated the clay fraction in the three climate zones.

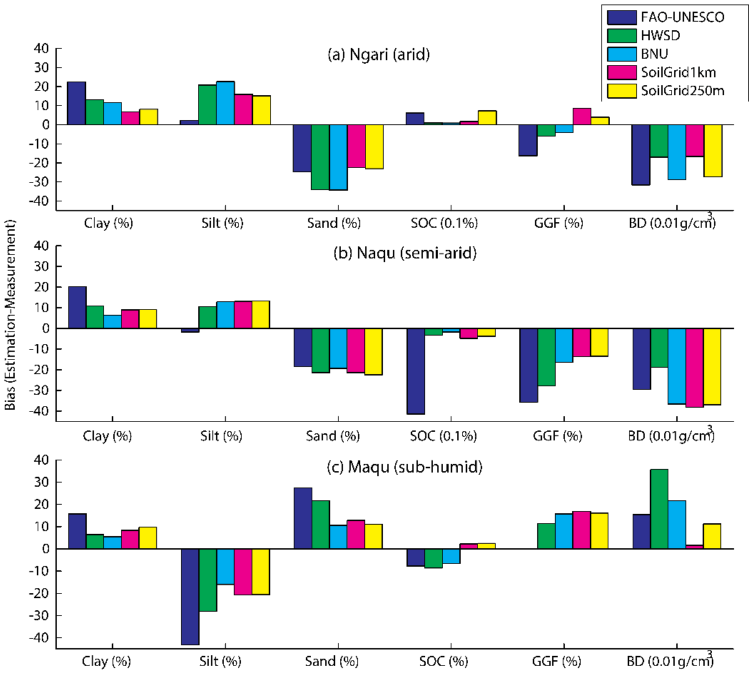


Figure 2.11 Average bias in the basic soil properties between the existing products and the laboratory measurements in the three climate zones. To enable the comparison of BD with the same order of magnitude as that of the other properties, the original BD value multiplied by 100 (unit × 100 g/cm^3). Likewise, a multiplication (% × 10) is applied to the SOC data in the semi-arid zone. FAO-UNESCO is the FAO-UNESCO Soil Map of the World (2007). HWSD is the Harmonized World Soil Database (FAO/IIASA/ISRIC/ISSCAS/JR, 2012). BNU is a Chinese data set of soil properties (Shangguan et al., 2012; Shangguan et al., 2013) and soil hydraulic parameters using PTFs (Dai et al., 2013) released by the Beijing Normal University. The SoilGrids1km (Hengl et al., 2014a) and the updated version of SoilGrids250m (Hengl, 2017) datasets are released by the International Soil Reference and Information Center (ISRIC) - WoSIS Institute. HPSS is the hydraulic parameter set of the van Genuchten (1980)-Mualem (1976) model based on SoilGrids1km and Schaap et al. (2001) PTFs (Montzka et al., 2017).

The estimates of SOC based on all the datasets were within a 1% range of the measurements in the arid and semi-arid zones and within 10% in the sub-humid zone, except for the FAO-UNESCO data, which heavily underestimated SOC in this region. Most of the GGF estimates in the arid zone were within 10%, with the FAO-UNESCO data underestimating GGF by 20%. In the semi-arid and sub-humid regions, all datasets consistently underestimated and overestimated, respectively, GGF.

The BD scheme was used to derive the porosity from the existing datasets. Figure 2.12a shows that the estimations of the porosity were higher than the *in situ* measurement in the arid zone, with SoilGrids1km and HWSD providing the closest approximations. In the semi-arid zone (Figure 2.12b), all datasets underestimated the porosity of the top layer but overestimated it at the other depths. It should be noted that SoilGrids1km and SoilGrids250m yielded an almost constant porosity in each profile, which is not representative of the conditions in the field. The porosity estimations based on FAO-UNESCO, HWSD and BNU did show profile variation, although much less than the *in situ* measurements did. In the sub-humid region (Figure 2.12c), all datasets underestimated the porosity in the surface layers at 5, 10 and 20 cm, and either underestimated or overestimated the porosity of the deeper layers.

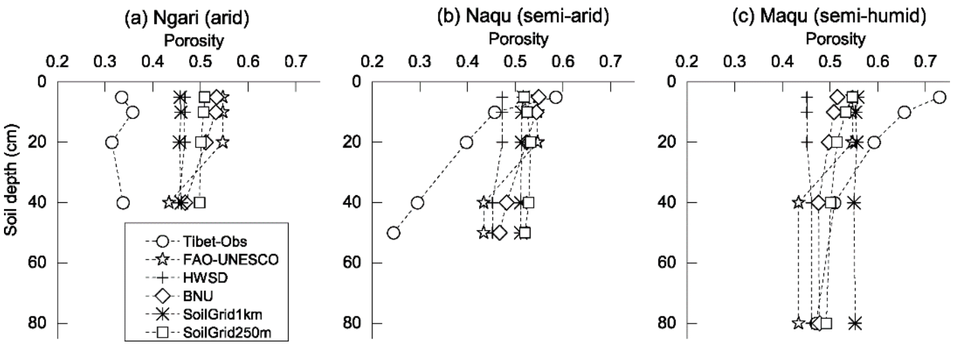


Figure 2.12 Comparisons between the porosity estimated with the various existing datasets based on the BD scheme and the in situ measurements.

SWRC and K_s

As the previous analyses of the PTFs (please refer to section 4.1) suggested, the Cosby et al. (1984) and continuous Wösten et al. (1999) PTFs were used with basic soil properties (i.e., only texture, BD and SOC) obtained from the independent datasets (e.g., SoilGrids) to estimate SWRCs-CH and SWRCs-VG, respectively. Given the relatively homogenous soil profile derived from the existing products (Figure 2.13), the averaged SWRCs derived from the existing datasets over the different depths were compared to the laboratory measurements.

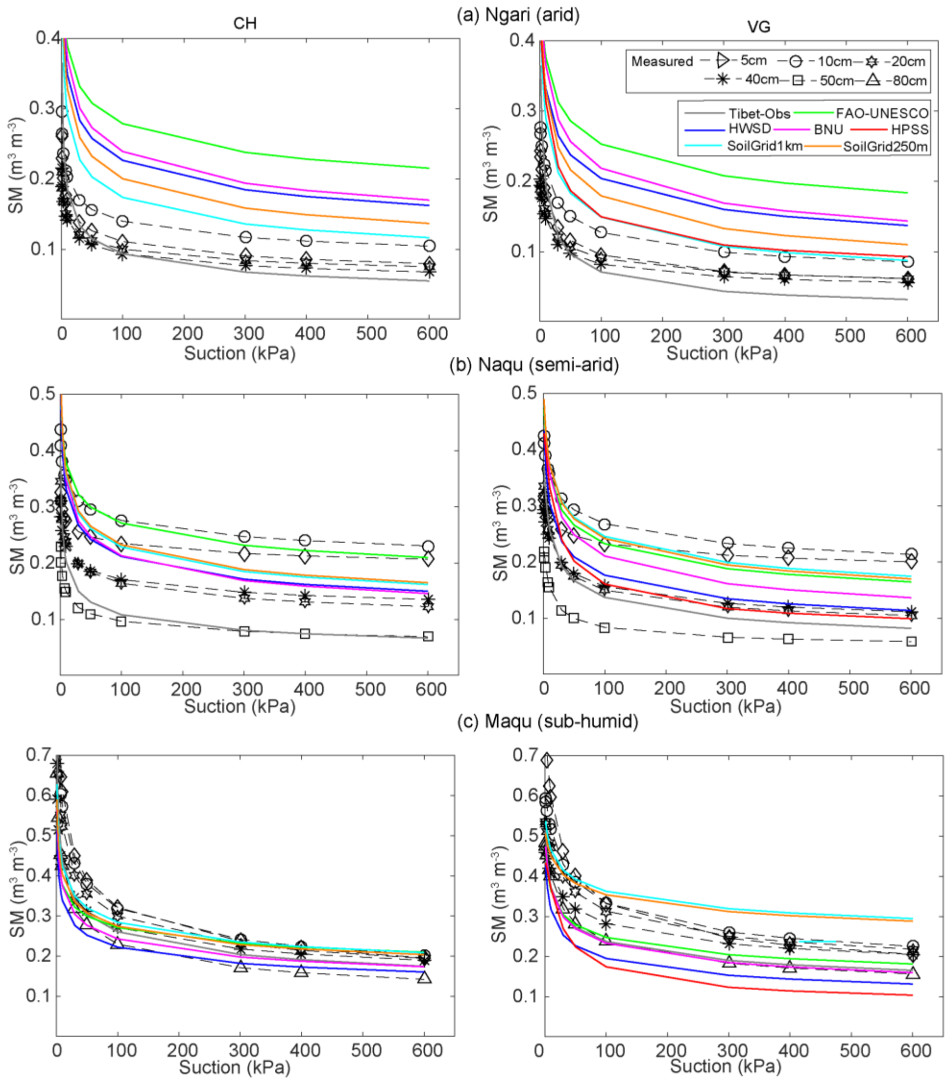


Figure 2.13 Comparisons of the SWRCs derived from the applicable PTFs based on the various datasets, to the laboratory measurements. The left panels show the SWRCs obtained with the CH model based on the six datasets. The right panel shows the SWRCs obtained with the VG model based on the seven datasets, of which HPSS only provides hydraulic parameters for the VG model.

Figure 2.13a shows that all datasets overestimated the SWRCs in the arid zone, in the order of FAO-UNESCO > BNU > HWSD > SoilGrids250m > HPSS (for the VG model) > SoilGrids1km > Tibet-Obs. In the semi-arid zone (Figure 2.13b), all datasets underestimated the SWRCs for the surface layers at 5 and 10 cm but overestimated the SWRCs for the deeper layers. FAO-UNESCO captured the SWRCs-CH for the surface layers well, and BNU presented the closest estimations for the deeper layers. Regarding the SWRCs-VG, SoilGrids250m and HWSD matched the measurements for the surface and deeper layers well. In the sub-humid zone (Figure 2.13c), all datasets showed similar SWRCs-CH, slightly underestimating them under a low suction (< 100 kPa) but then eventually becoming consistent with the measurements. The results for the SWRCs-VG were quite diverse. HWSD and HPSS showed consistent underestimations. FAO-UNESCO and BNU closely matched the measurements for the deeper layers. The SoilGrids1km and SoilGrids250m results were within the range of the measurements across the whole profile, although their mean values were larger in the high-suction range (> 300 kPa). Furthermore, it should be noted that the averaged profile SWRCs derived from Tibet-Obs tended to reflect the SWRCs for the deeper layers in the three climate zones. Additionally, the SoilGrids1km-, HWSD- and SoilGrids1km-derived FC (0.37, 0.41 and 0.51 m^3 m^{-3}) and PWP (0.16, 0.20 and 0.27 m^3 m^{-3}) were found to be close to the mean measured values in the three respective climate zones (please refer to Table A1.12).

With basic soil properties (i.e., soil texture, BD and SOC) obtained from the independent datasets (e.g., SoilGrids, etc.) with the Cosby et al. (1984) (1) and Rosetta1-H3 PTFs, Table 2.4 provides the mean predicted K_S (10^{-6} m/s) values based on the existing datasets across the three climate zones. The values were of a smaller order than that of some of the field measurements in the arid and semi-arid zones but of a larger order than some of the field measurements in the sub-humid zone. The Tibet-Obs dataset as input for the applicable PTFs predicted K_S well. The existing datasets for SWRCs estimation, i.e., SoilGrids1km, HWSD, and SoilGrids1km, also performed well in estimating K_s in the three climate zones.

Table 2.4 Comparisons of the mean derived K_s values obtained with the applicable PTFs and the CH and VG models based on the various soil datasets, to the measurements. The unit of the listed value is m/s.

Region	Ngari (arid)		Naqu (semi-arid)		Maqu	(sub-humid)
	CH	VG	CH	VG	CH	VG
Measured	2.53E-05	2.53E-05	2.50E-05	2.50E-05	4.21E-06	4.21E-06
Tibet-Obs	1.81E-05	2.41E-05	1.49E-05	2.38E-05	3.00E-06	1.65E-05
FAO-UNESCO	7.16E-06	6.09E-06	7.05E-06	5.25E-06	7.05E-06	5.25E-06
HWSD	5.74E-06	4.14E-06	7.98E-06	5.39E-06	6.72E-06	4.85E-06
BNU	4.92E-06	4.26E-06	7.91E-06	9.59E-06	3.91E-06	3.07E-06
SoilGrids1km	7.51E-06	4.96E-06	6.34E-06	5.84E-06	4.24E-06	5.62E-06
SoilGrids250m	7.29E-06	6.95E-06	6.22E-06	5.99E-06	3.97E-06	3.15E-06
HPSS		7.42E-06		6.65E-06		4.33E-06

SHPs & STPs in LSMs

Most LSMs use Richards' equation in soil water flow modeling (please refer to A.6 in Appendix A) with the hydraulic conductivity. In certain LSMs (e.g., Noah and H-TESSEL), the soil diffusivity (D) is used. When soil dries down, with the largest soil pores draining, K and D are reduced many orders of magnitude from complete saturation to dryness (Bittelli et al., 2015). A lower K (higher D) value results in slower water transport and thereby a higher SM estimated with LSMs than determined from soil moisture measurements, and vice versa.

LSMs use the thermal diffusion equation in soil heat transport modeling (please refer to A.6 in Appendix A). The soil heat capacity (C_s) and thermal conductivity

(λ) are the most important thermal parameters in the equation. Lower λ values with higher C_s values lead to reduced soil heat fluxes and thereby the higher soil temperature estimated with LSMs than measured soil temperature, and vice versa. The curves of K, D, C_s and λ based on basic soil properties obtained from the independent datasets (SoilGrids etc.) with the recommended parameterization schemes were compared to the measurements (as shown in Figures A1.5-A1.7 in Appendix A) to quantify the LSM uncertainty inherited from the soil dataset. A special case is formed by the FAO-UNESCO dataset, which slightly overestimated VG-K for the surface layers and heavily underestimated it for the deeper layers, while it heavily overestimated VG-D for the surface layers and slightly underestimated it for the deeper layers. These factors led to overestimation of the derived SM values, similar to the ECMWF SM analyses in this region (Su et al., 2013). The uncertainty stemming from the soil dataset also propagates to soil temperature estimation. The FAO-UNESCO dataset underestimated C_s-SM for the surface layers, but overestimated λ-SM, while at the other depths, this dataset estimated C_s-SM well but underestimated λ-SM. These factors led to the underestimation of the simulated soil temperature, which is also consistent with the findings of previous ECMWF soil temperature analyses (Su et al., 2013).

2.5 Conclusions

In this study, an *in situ* measurement dataset of soil physical properties was set up across the arid (Ngari), semi-arid (Naqu) and sub-humid (Maqu) climate zones on the Tibetan Plateau. This dataset can fill geographical gaps in global profile data on the Third Pole region. Analyzing this *in situ* dataset shows that the soil texture within the Ngari network in the arid zone consists of a high proportion of coarse sand with gravel and that the gravel content increases until 20 cm and then decreases slightly in the deeper layers. BD and the porosity slightly vary with the depth. The soil texture within the Naqu network in the semi-arid zone is dominated by a high percentage of sand mixed with a small proportion of gravel, but with a high SOC in the shallow layers and mainly large gravel particles in the deeper layers. BD reached a minimum in the top layer and a maximum in the bottom layer, and the porosity presents the opposite trend. The soil texture within the Maqu network in the sub-humid zone is dominated by a high percentage of silt with a relatively high SOC in the shallow layers and mainly fine sand in the deeper layers. BD increases with the depth, and the porosity decreases. Depending on the basic soil properties in the three climate zones, the soil hydraulic properties (SHPs, i.e., soil water retention curve, hydraulic conductivity) and thermal properties (STPs, i.e., heat capacity and thermal conductivity) differ in each climate zone and vary within each profile (e.g., presenting layering in the semi-arid and sub-humid zones), and gravel was found to affect the porosity and SHPs & STPs in the arid zone and in the deeper layers of the semi-arid zone.

Various schemes to estimate the porosity and SHPs & STPs on the TP were examined. The Cosby et al. (1984) PTFs proved more applicable for SHP estimation with the Clapp and Hornberger (1978) (CH) model, and the continuous Wösten et al. (1999) PTFs for SHP estimation with the van Genuchten (1980) - Mualem (1976) (VG) model. The original formulation of the De Vries (1963) model could be deployed successfully to estimate the heat capacity along the profile. Furthermore, the De Vries (1963) model combined with the Farouki (1981) scheme (D63F) under the implementation of the BD porosity scheme proved superior in thermal conductivity estimation.

Referenced by the measurements, the uncertainties in the existing basic soil property datasets and their derived SHPs & STPs were quantified on the TP. This information is of significance for the assessment of the LSM uncertainty inherited from soil datasets to screen the proper soil datasets for LSMs on the TP.

Furthermore, the existing soil property datasets can also be used as ancillary data for SM retrieval. For example, the composited datasets of the FAO and HWSD were used in Soil Moisture and Ocean Salinity (SMOS) and Soil Moisture Active Passive (SMAP) SM product generation. Therefore, the information is also valuable to better understand the uncertainties in satellite SM products inherited from soil maps. Based on dataset comparison, this chapter indicates that SoilGrids1km can reduce this uncertainty and is therefore recommended for use in the arid and sub-humid zones, while the combination of FAO-UNESCO in shallow layers and HWSD in deeper layers is recommended in the semi-arid zone on the TP.

In summary, this chapter provides a comprehensive *in situ* measured dataset of soil physical properties on the TP and presents applicable schemes to use for porosity and SHP & STP estimation in LSM on the TP. The dataset contributes significantly to the generation of spatial-temporally consistent soil moisture and temperature estimates with LSMs. Furthermore, the evaluation of the existing soil property datasets is crucial to quantify the uncertainty arising from the soil data considered in LSMs and soil moisture retrieval from microwave remote sensing.

Chapter 3. Observation Operator – A Discrete Scattering-Emission Model Incorporating An Air-to-Soil Transition Model

This chapter is based on:

Zhao, H., Zeng, Y., Wen, J., Wang, X., Wang, Z., Meng, X., & Su, Z. (2021). An Air-to-Soil Transition Model for Discrete Scattering-Emission Modeling a*t L-Band, Journal of Remote Sensing, vol. 2021, Article ID 3962350, 20 pages*.

Abstract: Topsoil structures and inhomogeneous distribution of moisture in a given soil volume induce dielectric discontinuities from air to the bulk soil, which in turn may induce multiple and volume scattering and affects microwave surface emission. *In situ* ELBARA-III L-band radiometer observations of brightness temperature T_B^p (p =H or V polarization) at the Maqu site on the Eastern Tibetan Plateau are exploited to better understand the effect of surface roughness on coherent and incoherent emission processes. Assisted by *in situ* soil moisture (SM) and temperature profile measurements, this study develops an air-to-soil transition (ATS) model that incorporates the dielectric roughness (i.e., resulting from fine-scale topsoil structures and the soil volume) characterized by SM and geometric roughness effects, and demonstrates the necessity of the ATS model for L-band T_B^p modeling. The Wilheit (1978) coherent model and Lv et al. (2014) incoherent model are compared in the determination of the dielectric constant of bulk soil in the ATS zone and in the calculation of soil effective temperature T_{eff}. The Tor Vergata discrete scattering model (TVG) integrated with the advanced integral equation model (AIEM) is used as the baseline model configuration to simulate L-band T_B^p. Then, the ATS model is integrated with the foregoing model to assess its performance. The results show that the ATS-based models reduce the underestimation of T_B^p (≈ 20-50 K) generated in the baseline simulations. As a dynamic parameter in nature, the proposed dielectric roughness parameterization in the ATS model significantly improves the ability to interpret T_B^p dynamics, which is important for SM retrieval enhancement at the global scale.

Keywords: Air-to-soil transition model, L-band radiometry, dielectric roughness dynamics, soil effective temperature, discrete scattering model.

3.1 Introduction

Soil moisture (SM) is highly important for weather and climate predictions by controlling the partition of heat and water fluxes across the land-atmosphere interface (Babaeian et al., 2019; Seneviratne et al., 2010; Taylor et al., 2012). Passive L-band microwave remote sensing has become the most promising technique for near-surface SM measurement by properly quantifying the contributions of vegetation and ground surface (Entekhabi et al., 2010; Kerr et al., 2010; Wigneron et al., 2017). Independent L-band brightness temperature observations and radiative transfer models (e.g., the community microwave emission model (de Rosnay et al., 2009)), if integrated with land surface models within a data assimilation framework, can be used to estimate soil physical properties (Bandara et al., 2015; Dimitrov et al., 2014; Han et al., 2014a; Montzka et al., 2011; Yang et al., 2016), which are crucially important to better understand SM dynamics (Prentice et al., 2015; Su et al., 2013).

The efforts related to microwave remote sensing of the land surface may be traced back to the work of Peake (1959), which demonstrated the complementary relationship between emission and scattering and verified it against data obtained from Straiton et al. (1958). This may be called the scattering-emission radiative transfer approach. More recent works include those of Fung (1994) and Chen et al. (2003a) on an advanced integral equation model (AIEM) for a rough bare soil surface, and those of Ferrazzoli and Guerriero (1996) and Bracaglia et al. (1995) on a discrete scattering model (Tor Vergata model) for a vegetated surface. These approaches consider uniform soil moisture and soil temperature (SMST) profiles and use the surface value of the dielectric constant with roughness parameters in the calculation of the surface reflectivity by integrating bistatic scattering coefficients over the half-space above the surface. The other line of work includes the studies of Njoku and Kong (1977) and Wilheit (1978), which relied on stratified coherent radiative transfer approaches to calculate microwave emission of a medium with a nonuniform temperature profile (i.e., coherent models) to account for the nonuniform SMST profile of natural soil. In particular, the SMST profile was used to determine the smooth surface reflectivity and soil effective temperature T_{eff}. Due to its simpler formulation, the Wilheit (1978) model is widely adopted, followed by simplified semi-empirical models (Choudhury et al., 1982; Parrens et al., 2014; Schmugge & Choudhury, 1981; Wang & Choudhury, 1981; Wigneron et al., 2017) in applications involving airborne and satellite microwave remote sensing. Generally, these models use Fresnel equations to obtain the surface reflectivity with roughness corrections, which is continued into the zeroth-order radiative transfer model used for Soil Moisture and Ocean Salinity (SMOS) (Kerr et al., 2010) and Soil Moisture Active Passive (SMAP) (Entekhabi et al., 2010) SM retrieval (Das et al., 2010; Kerr et al., 2012). In contrast, these models do not retain the coherent character as similar to the Wilheit model, mainly due to the simplified T_{eff} parameterization scheme (Choudhury et al., 1982; Lv et al., 2014). To investigate the impact of T_{eff} on the microwave radiometry, Lv et al. (2014) used an analytical formulation to physically explain various T_{eff} schemes, all of which have their roots in the scheme of Choudhury et al. (1982), and proposed the Lv's T_{eff} scheme (e.g., incoherent model). In this study, we investigate the effects of the Wilheit (1978) coherent model and Lv et al. (2014) incoherent model on T_{eff} and associated brightness temperature (T_B^p, with p = H or V polarization) simulations.

Another research focus of microwave L-band radiometry is the surface roughness effect. The geometric roughness resulting from the variation in surface height influences surface scattering and is modeled with the physically-based AIEM. However, the AIEM assumes isotropic roughness properties for a homogenous dielectric half-space and does not consider the dielectric effects due to any heterogeneities in the soil characteristics (e.g., composition, moisture content, bulk density). On the other hand, the fact that lateral structures (e.g., the unfilled surface composed of organic matter and clods) are much smaller than the observation wavelength (e.g., λ_0 = 21 cm, in the L-band) influences the manner of wave propagation and induces an impedance mismatch of the rough surface between air and soil. The aforementioned heterogeneities produce the dielectric roughness (namely, the notable dielectric discontinuities at the soil surface and within the soil volume) and may induce the volume and multiple scattering processes, which affects microwave surface emission. The model parameters used in zeroth-order radiative transfer models may implicitly account for both the geometric and dielectric roughness effects. However, they are site-specific empirical parameters, obtained by using the best-fit approaches based on limited field observations and model simulation results. An air-to-soil transition (ATS) model (Mätzler, 2006)—an intermediate modeling approach between physical and semi-empirical approaches—is suggested to describe the roughness effects of topsoil structures on L-band radiation through an impedance matching between the dielectric constants of air and bulk soil. In the original ATS model (Schneeberger et al., 2004; Schwank et al., 2004), the structured topsoil is adopted as a transition layer with a geometric thickness h, considering that the volume fraction of soil materials increases with the depth in the ATS zone. Moreover, h is related to s (the height standard deviation with a Gaussian distribution, centered on lateral separation) as $s(h) = 0.2489h$ (Mätzler, 2006; Schneeberger et al., 2004). As the geometric surface roughness (i.e., s) does not notably change, h is a fixed peak-to-trough transition layer thickness induced by topographical effects and independent of soil moisture. However, regarding h to be constant is questionable due to the fact that the dielectric properties of the topsoil and the soil volume below may be modulated by inhomogeneity related to moisture. This study develops an enhanced air-to-soil transition (ATS) model with a new h parameterization scheme to investigate the soil moisture-dependent dielectric roughness of the topsoil structures and the soil volume below on L-band radiation. The Maqu site (33.91°N, 102.16°E) in the eastern Tibetan Plateau meadow region, which provides comprehensive field observations (Su et al., 2011; Su et al., 2020a; Zeng et al., 2016; Zhuang et al., 2020), is chosen as the study area to validate the model.

With the ATS dielectric layer obtained, an equivalent homogenous dielectric entity acting as the ground scattering-emission medium can be assumed with a given dielectric constant and surface geometric roughness. The AIEM (Chen et al., 2003a), which uses a more complete expression of the single scattering terms to maintain the acceptable energy conservation when calculating emissivity according to bistatic scattering coefficients (Zeng et al., 2017), is employed to simulate soil surface scattering. As the research object is a natural grassland, the Tor Vergata model simulating vegetation scattering is coupled with the AIEM (TVG+AIEM) for overall vegetation-soil scattering-emission modeling. The coupled model including the ATS model (TVG+AIEM+ATS) is further employed to investigate the impacts on h estimations and T_B^p simulations by adapting the Wilheit stratified model or Lv model or by applying SM at the single layer depth of 2.5 cm (*in situ* measurements) given its topmost role in surface emission (Escorihuela et al., 2010; Wilheit, 1978). Finally, the applicability and uncertainty in the enhanced ATS model on L-band radiometry modeling are discussed.

This chapter is organized as follows: *in situ* SMST profiles and ELBARA-III T_B^p observations at the Maqu site are described in section 3.2. In section 3.3, a brief description of the TVG model is introduced. The improved ATS model is presented with the Wilheit (1978) and Lv et al. (2014) models. The configurations of the different simulation experiments are also explained. The results regarding the performance of the enhanced ATS model and seasonal T_B^p simulations are provided in section 3.4, in addition to the influences of the above coherent and incoherent models and SM at 2.5 cm on T_B^p simulations. Discussions in applicability and uncertainty in the ATS models on T_B^p simulations are described in section 3.5, as well as discussions in the impacts of the geometric roughness on the performance of the ATS model. The potential advantage of the ATS model in terms of satellite-based SM retrieval improvement is also discussed in section 3.5. Conclusions are drawn in section 3.6.

3.2 *In Situ* Measurements at the Maqu site

ELBARA-III (Schwank et al., 2010) is a Dicke-type radiometer and equipped with a dual-polarized conical horn antenna with a -3 dB full-beamwidth of 12°. The ELBARA-III radiometer at the Maqu site is mounted on a 4.8-m high scaffold tower, which makes the center of rotation in elevation at 6.5 m above ground (upper panel in Figure 3.1). The direction of the antenna beam is toward the south. Daily measurements include elevation scanning sequences toward the ground and sky measurements. The angular range of ground scans at every 15 min is between 40° and 70° (relative to nadir) in steps of 5°. Sky measurements are performed at the local time of 23:55 every day with an observation angle of 155°. The half axes a and b (the lower panel in Figure 3.1) are estimated corresponding to the -3dB beamwidth, the installation height, and the incidence angle θ_i . The horizontal distances d_{min} and d_{max} (measured from the radiometer to the closet and the farthest border of the elliptic footprints) and the footprint areas A are also estimated following Schwank et al. (2005) and described in Table 3.1. The diagram of footprint areas is shown in Figure 3.1 (lower panel).

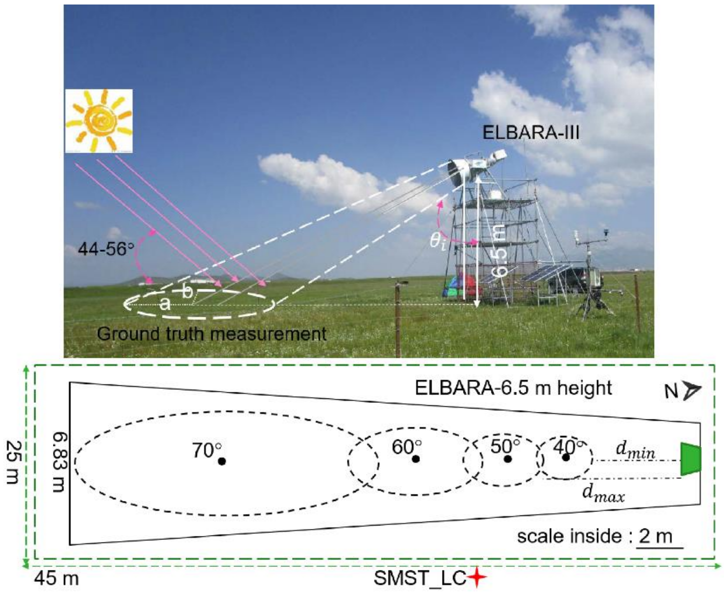


Figure 3.1 Sketch of the ELBARA-III T_B^p footprint and SMST_LC site at the Maqu site. The upper panel shows a diagram of the ELBARA-III observation geometry, and the lower panel shows the location of the installed in situ soil moisture and soil temperature sensors (SMST_LC) and the surface projection of the observation footprints at different incidence angles. Note that the scale is only applied to the elliptic footprints described inside the blackline box. The half-axes a and b of the elliptic footprint with the incidence angle θ_i are given in Table 3.1.

Table 3.1 Footprint dimensions at the different incidence angles θ_i

θ_i (°)	a (m)	b (m)	d_{min} (m)	d_{max} (m)	A (m^2)
40	1.17	0.90	4.38	6.73	3.31
45	1.38	0.98	5.26	8.03	4.24
50	1.68	1.08	6.28	9.64	5.70
55	2.12	1.22	7.48	11.73	8.13
60	2.83	1.41	8.95	14.6	12.55
65	4.03	1.70	10.82	18.88	21.57
70	6.37	2.18	13.33	26.07	43.64

During the 2016 summer campaign, soil moisture and soil temperature (SMST) 5TM ECH2O probes (Decagon Devices, Inc., USA) were installed in the SMST_LC pit (Lv et al., 2018; Su et al., 2020a) near the ELBARA-III location (shown in the lower panel in Figure 3.1). The 5TM probe—a capacitance sensor operated at 75 MHz measures the real part of the dielectric constant of the surrounding soil, and the Topp' model (Topp et al., 1980) is used to convert real dielectric constant values to SM values. As the Topp' model does not consider the impact of the soil texture on the soil dielectric constant, the site-specific calibration conducted by Dente et al. (2012) is used to improve the accuracy of the probe output. A dense SMST profile (at 2.5, 5, 10, 20, 35, 60 and 80 cm below the soil surface) was collected in the period between 08/08/2016 and 20/03/2017. Due to power outages (for example, snow cover on the solar panels), the ELBARA-III radiometer was occasionally out of commission in December, and T_B^p data were only available during the period from 08/08/2016 to 30/11/2016 in 2016. The analyses of this study focus on *in situ* data and T_B^p simulations during the late-monsoon (August to September) and post-monsoon (October to November) periods in 2016. To keep consistent with the SMAP incidence angle, the T_B^p data analysis is confined to the angle of 40°, although the location of the SMST_LC sensors is close to 50° (please refer to the lower panel in Figure 3.1). The T_B^p data contaminated by solar beams reflected into the ELBARA-III antenna horn due to the solar elevation in the range of 44-56º (please refer to the upper panel in Figure 3.1 for an illustration of such geometry) (Su et al., 2020a) is regarded as outliers and removed with interquartile range filtering.

Additionally, leaf area index (LAI) is an important input parameter in vegetation modeling with the Tor Vergata model, which determines the number of grass leaves as discrete dielectric scatters. LAI time series is obtained from MCD15A2H – Moderate Resolution Imaging Spectroradiometer (MODIS/Terra+Aqua) LAI (500-m resolution) (https://lpdaac.usgs.gov/products/mcd15a2hv006/). Affected by atmospheric conditions and issues from spectrum—the sensor itself, the MODIS LAI contains noise and is discontinuous in time series. Moreover, due to the cloud impact caused by water vapor absorption in the red band area, some of the original MODIS LAI data might be lower than real values. Therefore, the harmonic analysis of the time series (HANTS) algorithm (Verhoef, 1996) is used to filter MODIS LAI data during the period of 2016. The results in Figure 3.2 show a reliable interpreted LAI. The processed MODIS LAI data are further compared to *in situ* ones measured in the 2018 filed campaign (Su et al., 2019), and a good match between them is confirmed. Furthermore, meteorological observation data in the Maqu site (Su et al., 2019) are used to support analysis. The data mainly involve precipitation intensity, air temperature, and albedo with ground surface temperature, which are derived from *in situ* four components radiation measurement (i.e., up- and down-welling shortwave and longwave radiation).

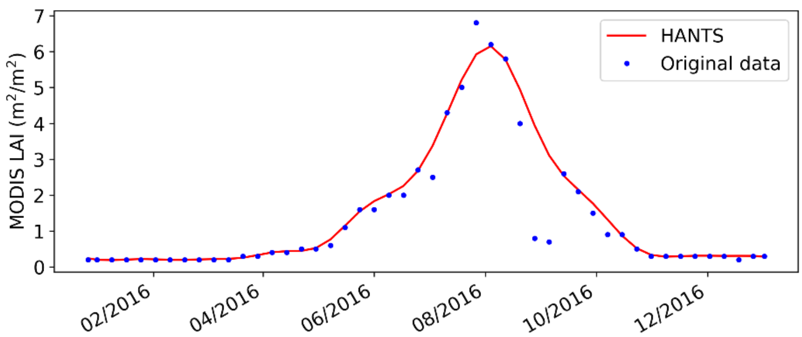


Figure 3.2 MODIS/Terra+Aqua leaf area index (500-m resolution) original data with filtered data using the HANTS algorithm.

3.3 Methods

3.3.1 Tor Vergata discrete scattering model

The TVG model (Bracaglia et al., 1995; Ferrazzoli & Guerriero, 1996) assumes the soil as a homogeneous infinite half-space with a rough interface, and the overlying vegetation is represented as an ensemble of discrete dielectric scatters. The scattering modeled by the TVG model involves three components: vegetation volume scattering, soil surface scattering and the scattering component resulting from vegetation-surface interactions. The TVG model has been investigated in the Maqu area (Bai et al., 2019; Dente et al., 2014; Wang et al., 2018; Zheng et al., 2017). As reported in Dente et al. (2014), the grass leaves at the Maqu site are described by dielectric thin discs with a random orientation distribution. Bistatic scattering and extinction (absorption plus scattering) cross sections of these dielectric discs are computed by applying the Rayleigh-Gans approximation in the L-band (Eom & Fung, 1984) (① in Figure 3.3), in which the Mätzler (1994) model is applied to calculate the vegetation dielectric constant. Subsequently, the contributions of all discs (scatters) are integrated using the matrix doubling algorithm (② in Figure 3.3). Therefore, the scattering and transmission matrices are computed for the whole vegetation (③ in Figure 3.3). Values of vegetation parameters such as the disc radius, disc thickness, number of discs (calculated as LAI/the disc area) and plant moisture content used in this study are calibrated ones from Dente et al. (2014) and Wang et al. (2018), and they are found insensitive to the emissivity in the L-band (Bai et al., 2019).

Soil surface scattering is computed with the AIEM, in which the soil dielectric constant and various surface roughness parameters (i.e., the standard deviation of surface height s, correlation length of surface height L, and assumed exponential autocorrelation function for natural surfaces) are needed as inputs. In the previous version of the TVG model adopted for the Maqu site, a grassland litter component was included (Dente et al., 2014). The litter part, however, is not implemented in this study. One reason is that the grassland on the Tibetan Plateau is grazed by sheep and yaks (Su et al., 2011), and the litter in most areas is decomposed (as shown in two snapshots in Appendix B). In contrast, for natural land surfaces, there does exist a more gradual transition zone of the dielectric constant from air to the bulk soil than that of an abrupt surface, such as in calm seas (Mätzler, 2006). Therefore, our concern in this chapter is to improve transition zone modeling. Ignoring the litter part in the modeling system is for simplicity but also reduces the numerous required input parameters that may degrade the performance of the model.

In this study, the enhanced air-to-soil transition (ATS) model (section 3.3.2) is used to obtain the effective dielectric constant (④ in Figure 3.3). Two components, incoherent bistatic scattering coefficients (computed with the AIEM, ⑤ in Figure 3.3) and coherent specular reflection coefficients (computed with Fresnel equations corrected by the roughness factor, ⑤ in Figure 3.3), are computed to determine composite air-to-soil medium scattering. The same matrix doubling algorithm is then used to combine the calculated vegetation contribution with that of the air-soil medium (⑥ in Figure 3.3). Subsequently, the emissivity $e_p(\theta_i)$ under p polarization (i.e., H or V) at an incidence angle θ_i is obtained by applying the energy conservation law and integrating the bistatic scattering coefficients over the half-space above the surface (⑦-⑧ in Figure 3.3). Due to the low vegetation emission in the L-band, the physical temperature of vegetation is assumed to be the same as that of soil in this study. The soil effective temperature T_{eff} can be estimated with either the Wilheit (1978) coherent model or Lv et al. (2014) incoherent model (section 3.3.3) (⑨ in Figure 3.3). Finally, T_B^p (=$e_p(\theta_i) \cdot T_{eff}$) is computed using the emissivity $e_p(\theta_i)$ multiplied by T_{eff} (⑩ in Figure 3.3). Figure 3.3 shows a flowchart of the forward T_B^p simulations with the coupled TVG model and AIEM including the ATS model. The parameter values used in the TVG+AIEM+ATS simulations are listed in Table A2.1 in Appendix B for reference.

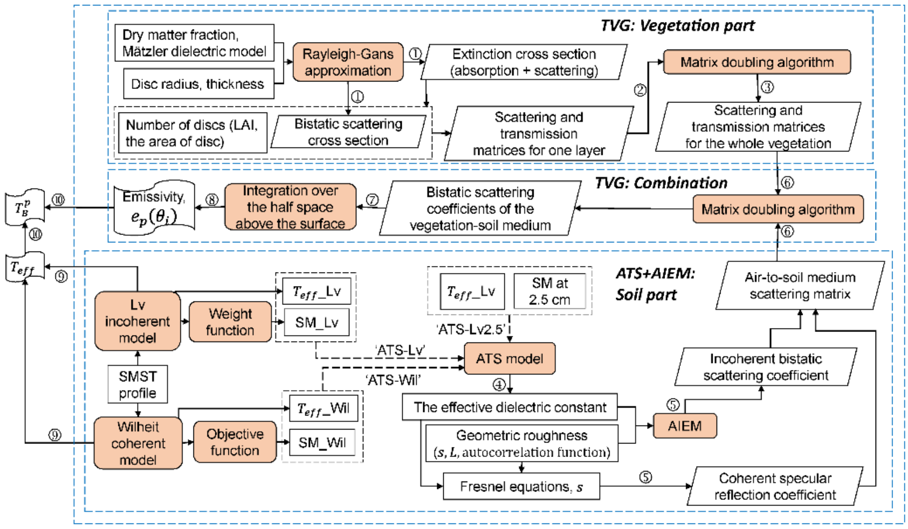


Figure 3.3 Flowchart of the procedure for the forward T_B^p simulations with the coupled TVG model and AIEM including the ATS model. ①-⑩ represent the workflow. The square rectangle indicates inputs and parameters, and the rounded rectangle in orange refers to models and algorithms. The outermost dashed blue box encloses elements of the TVG+AIEM+ATS model. Three dashed boxes in blue enclose elements of scattering modeling of the vegetation and soil parts and their combination. The black dashed box inside the upper blue dashed box contains inputs used to calculate scattering and transmission matrices. The black dashed box inside the lower blue dash box is with inputs used for calculating the dielectric constant, and the dashed arrow points to the inputs used for the baseline and ATS-based simulations (section 3.3.4). Detailed descriptions of the ATS model are provided in section 3.3.2, and the Wilheit coherent model and Lv incoherent model are described in section 3.3.3.

3.3.2 Air-to-soil transition (ATS) model

The vegetated soil medium is composed of a substantial amount of loose dirt, plant debris and crumbs scattered on the surface and a much denser soil entity below (Figure 3.4a). Driven by changing weather systems such as after the dry and sunny conditions following rainfall events, wetted plant debris and large clods on the surface dry out more quickly than those of the bulk soil beneath the surface. Affected by roots and air pockets present in the soil volume, an inhomogeneous layer is produced between the soil structures near the surface and the bulk soil at the bottom. All these effects may lead to high spatial variability in SM at the soil surface and within the soil volume (Mätzler, 2006). Consequently, this induces different (e.g., wet-dry layers) interfaces occurring in the gradual transition zone of the dielectric constant from air to the bulk soil (ATS). The ATS zone may extend across the peak-to-trough geometric thickness for natural smooth surfaces, especially when the soil surface is dry and electromagnetic waves from deeper layers transmit toward the surface.

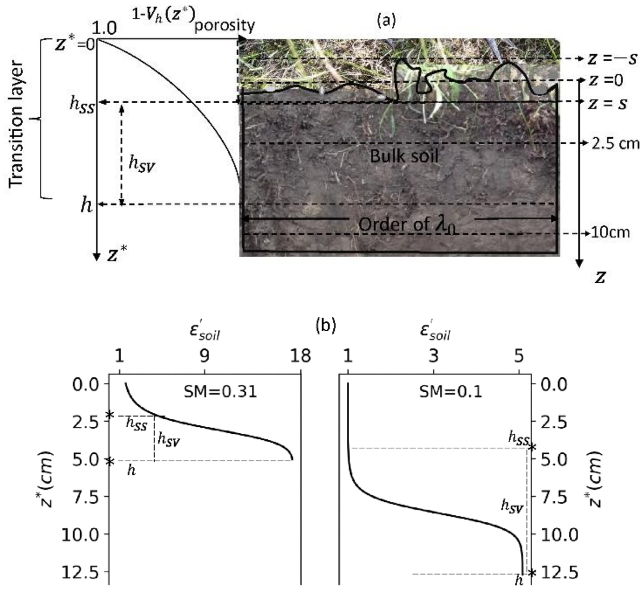


Figure 3.4 Sketch of the air-to-soil transition (ATS) model (a) and the dielectric depth profile in the ATS zone (b). $V_h(z^*)$ is the cumulative probability of the density of $S_h(z^*)$, and 1- $V_h(z^*)$ corresponds to the total air volume fraction of the ATS zone. h is the total dielectric roughness thickness (unit: cm) of a soil area with the order of λ_0 by λ_0 ($\lambda_0 = 21$ cm in the L-band). h_{SS} (unit: cm) refers to the dielectric roughness thickness of the topsoil structures. h_{SV} (unit: cm) is the dielectric roughness thickness of the soil volume. $z = 0$ is the average surface geometric height with standard deviation s (unit: cm). The ECH2O 5TM probes (Decagon Devices, Inc., USA) are located at different depths (e.g., 2.5 and 10 cm). (b) describes h under wet and dry soil conditions with a s value of 0.9 cm, in which ε'_{soil} is the real part of the dielectric constant of bulk soil, $\varepsilon_{soil} = 17.3 + i\ 2.2$ for SM = 0.31 m^3/m^3, and $\varepsilon_{soil} = 5.1 + i\ 0.4$ for SM = 0.1 m^3/m^3.

Assuming L is smaller than the wavelength λ_0 in free space, the concept of the dielectric roughness (h, L) is proposed for the ATS zone, in which h is the dielectric roughness thickness characterizing the depth of interfaces, not only resulting from topsoil structures (h_{SS}) affected by both irregularities (i.e., the geometric roughness) of the soil surface and inhomogeneous distribution of moisture but also attributed to the inhomogeneity within the soil volume (h_{SV})

that is related to the soil porosity and moisture. The dielectric roughness thickness h for the ATS zone is the sum of h_{SS} and h_{SV} (as shown in Figure 3.4a).

$$h = h_{SS} + h_{SV} \tag{3.1}$$

Due to the difficulty in obtaining detailed volumetric information on inhomogeneous mediums (e.g., loose dirt, plant debris, bulk soil mixed with roots) along the ATS zone, the Fermi-Dirac distribution function (Kittel, 1976) is used in this study to construct the dielectric depth profile, following an exponential dependence of the roughness thickness on SM (Schneeberger et al., 2004; Wigneron et al., 2001). Subsequently, an equivalent homogenous dielectric ATS zone with a given dielectric constant is produced as a consequence of impedance matching over the ATS zone, which is used to calculate the scattering of the ATS medium by the AIEM (please refer to section 3.3.1).

In this study, h is chosen as the SM-dependent roughness parameter. Modulated by SM, h varies, and the probability density function $S_h(z^*)$ for the dielectric roughness height is assumed to have an exponential distribution with a rate parameter α, considering exponential attenuation in regard to the water content and (physical) height of surface emission (Mätzler, 2006). $S_h(z^*)$ is expressed in equation (3.2):

$$S_h(z^*) = \alpha \exp(-\alpha z^*), \quad z^* \geq 0 \tag{3.2}$$

As h is also affected by the geometric roughness, the depth dependence of the volume fraction of soil materials underlying the variations in the dielectric profile is consequently related to the dielectric height distribution and can be described as the cumulative distribution of $S_h(z^*)$. Specifically, the integral of $S_h(z^*)$ over depth z^*— $V_h(z^*)$ represents the volume fraction of soil materials and 1-$V_h(z^*)$ represents the air volume fraction (equation (3.3)):

$$1 - V_h(z^*) = \exp(-\alpha z^*), \quad z^* \geq 0 \tag{3.3}$$

At the soil surface, h_{SS} is impacted by topographic effects and is related to $2s$ as the surface geometric height (as shown in Figure 3.4a). h_{SS} in this study is also assumed to depend on both the incidence angle and polarization as reported in previous investigations, e.g., Wang and Choudhury (1981), Escorihuela et al. (2007) and Bircher et al. (2013). If we assume that the air volume fraction is 1 at an arbitrary position above the surface structures where $z^* = 0$ (z^* is away from the average surface geometric height; $\mathrm{z} = 0$ for nadir observation), since more soil particles occupy voids in topsoil structures at the lateral scale with increasing depth ($z^* > 0$), the air volume fraction in the topsoil structures decreases (Figure 3.4a). The decreased air volume is filled by an increased volume of soil materials, and the resultant effect of moisture on h_{SS} can be described by a logarithmic SM (equation (3.4)) function. While h_{SS} decreases when the soil surface is wet, the surface may become 'saturated' when it is sufficiently wetted, namely, the soil moisture reaches the field capacity (FC), at which $z^* = 0$ moves to $z = -s$ and ensures that h_{SS} approaches $2s$. h_{SS} is then given as follows:

$$h_{SS} = \begin{cases} 2 \cdot s \cdot (-\ln(SM)) \cdot \cos^{Np}(\theta_i) & \text{SM} < \text{Field capacity} \\ 2 \cdot s & \text{SM} \geq \text{Field capacity} \end{cases} \tag{3.4}$$

where θ_i is the incidence angle, Np is a polarization modulation parameter, and Np is set to 0 for H polarization and -1 for V polarization in this study. SM is the volumetric soil moisture.

With z^* deepening in the ATS zone, when the topsoil structures within the whole lateral scale range tend to become fully filled with soil materials (Figure 3.4a), the soil texture (i.e., the porosity) and SM profile become the dominant factors whose influences can be represented by h_{SV}, which is the depth where the air volume fraction in the ATS zone equals the maximum volume of the pore space in the soil (porosity) (equation (3.5)).

$$h_{SV} = \frac{-\ln(porosity)}{\alpha} \tag{5}$$

Affected only by moisture in the soil volume (this is where SM is measured), the parameter of the distribution α can be estimated by the power attenuation coefficient, similar to Choudhury et al. (1982) and Lv et al. (2018), which is determined by λ_0 and the complex dielectric constant of bulk soil in the ATS zone as follows:

$$\alpha = \frac{2\pi}{\lambda_0} * \frac{\varepsilon_{soil}''}{\sqrt{\varepsilon_{soil}'}} \tag{3.6}$$

where $\varepsilon_{soil} = \varepsilon_{soil}' + i\varepsilon_{soil}''$, ε_{soil}' is the real part, and ε_{soil}'' is the imaginary part of the soil dielectric constant of the ATS zone. In this study, the soil porosity is set to 0.62 according to laboratory measurements (Zhao et al., 2018a), and FC is valued at 0.35 m^3/m^3 for silt loam soil. A sketch of the improved ATS model is shown in Figure 3.4a.

The Fermi-Dirac distribution function (Kittel, 1976) expressed in equation (3.7) is used to describe the dielectric depth profile $\varepsilon(z^*)$ in the ATS zone. The steepness parameter k_{AS} in equation (3.8) is related to s and SM effects.

$$\varepsilon(z^*) = \varepsilon_{air} + \frac{1}{1 + \exp\left(-\frac{z^* - h_{SV}}{k_A}\right)}(\varepsilon_{soil} - \varepsilon_{air}) \tag{3.7}$$

$$k_{AS} = \exp(-\alpha z^*) \cdot s \tag{3.8}$$

Figure 3.4b shows two estimated h profiles and the correspondingly derived $\varepsilon(z^*)$ values under wet and dry soil conditions. The same temperature can be assumed for all layers in the ATS zone due to the small influences of temperature change on the dielectric constant of (organic) soils (Mironov et al., 2015). As such, a coherent radiative transfer model can be used to compute the overall coherent reflectivity for the ATS zone considering $\varepsilon(z^*)$, which is based on a matrix formulation of the boundary conditions of dielectric discontinuities derived from Maxwell's equations (Bass et al., 1995). The coherent model is applied over a total depth of h with the thickness of each layer set to 1 mm, which is less than one-tenth of the wavelength λ_0. This coherent model predicts reflectivity trend as a function of the layer thickness but is characterized by enhanced oscillations due to the coherent interactions among the multiple reflected waves, and this process can be smoothed by the natural variations in the layer thickness around its average value and an averaging procedure (Della Vecchia, 2006). Considering the impacts of both the surface geometric roughness and SM at the bottom of the ATS zone, the average dielectric surface ($z^*_{avg} = h/2$ along z^*, as shown in Figure 3.4a) is assumed to decrease with the depth measured by s multiplied by the natural logarithm of SM (equations (3.9-3.10)).

$$z^*_{avg} = h/2 \tag{3.9}$$

$$z^*_{lb} = h/2 - \ln(SM) \cdot s \tag{3.10}$$

Consequently, the reflectivities obtained for this layer thickness ($z^*_{avg} \leq z^* \leq z^*_{lb}$) are averaged, and the effective dielectric constant of an equivalent homogenous dielectric ATS zone (used for the calculation of scattering by the AIEM) is computed by minimizing an objective function between the obtained reflectivities and those computed for the ATS zone using Fresnel equations (Della Vecchia, 2006).

The effective roughness parameters obtained via model calibration are recommended for use in physically-based surface backscattering models (Mattia et al., 2006; Su et al., 1997). In this study, s is regarded as 0.9 cm and L as 9.0

cm in the Maqu case (Dente et al., 2014). These two calibrated values consider the high correlation between s and the roughness slope (s/L) (Benninga et al., 2018; Su et al., 1997). The SM of the lower boundary of the ATS zone used to calculate h_{SS} (equation (3.4)), ε_{soil} and associated h_{SV} (refer to equations (3.5-3.6)) in the L-band is difficult to obtain, but can be regarded as the representative SM. The measured SM at 2.5 cm is considered as the representative SM, because the reflectivity of a stratified dielectric medium is primarily determined by changes in the real part of the refractive index over a depth interval of approximately 1/10 to 1/7 wavelengths (~2.5 cm in the L-band) (Wilheit, 1978). Notably, the representative SM is also obtained by considering the impact of the SMST profile on soil microwave emissions through either the Wilheit (1978) model or Lv et al. (2014) stratified model (please refer to section 3.3.3). The Mironov et al. (2015) dielectric model is used to calculate ε_{soil} throughout the whole study period, and the considered soil texture (i.e., clay fraction and bulk density) information is based on laboratory measurements (Zhao et al., 2018a).

3.3.3 Wilheit coherent model and Lv incoherent model

In this study, we regard the Wilheit (1978) model as a coherent model because the electromagnetic wave considered in this model is formulated with the amplitude and phase, and electromagnetic energy flow through the plane is given by the Poynting vector with the retained coherent character. When rapid surface drying occurs, there are dry and wet layers at various depths. Reflections from the air-soil interface and the dry-wet soil interface may interfere, resulting in the wave stemming from the deeper layers adding to the surface energy density either constructively (in phase) or destructively (out of phase), or in between (i.e., the addition of two waves). In contrast, the Lv et al. (2014) model is regarded as an incoherent model because the model derivation is based on the radiation intensity (i.e., with only amplitude considerations but without phase considerations)

(Choudhury et al., 1982). T_{eff} and associated T_B^p simulations with these two models are expected to provide physical insights into the interactions of microwaves with soil medium. From an application perspective, this can help determine whether coherent and incoherent effects should be considered for emission modeling of natural surfaces, whose status changes with meteorological and hydrological conditions.

3.3.3.1 Wilheit coherent model

In the Wilheit (1978) model, soils are treated as a layered plane dielectric medium. The basic assumption is that there is a reflection for the incident radiation on the air-soil interface and the thermal equilibrium in each following layer (i.e., beneath the interface) of this stratified medium. That is to say, only the absorption and transmission of electromagnetic waves are considered in each layer. The fraction of absorption (f_i^p) can be calculated by solving Maxwell's equations with the aid of boundary conditions at the interfaces for a coherent electromagnetic wave propagating through the layered soil (Wilheit, 1978). If T_i is the temperature of the ith layer, under thermodynamic equilibrium, the layer radiates energy equal to the product of the fractional absorption f_i^p and the temperature T_i. In terms of the conservation of energy, the reflectivity of smooth air-soil interface R_s^p is described by equation (3.11):

$$R_s^p = 1 - \sum_{i=1}^{N} f_i^p \tag{3.11}$$

where N represents the total number of discrete soil layers. As such, the representative SM (SM_Wil) used for determining h is the one resulting in a minimum root mean square error difference between the obtained reflectivities and those computed for a set of SM through the Fresnel equations (Della Vecchia, 2006). Wilheit (1978) also defined the thermal sampling depth δ_T as the average depth, at which the upwelling thermal radiation from the soil originates. δ_T is a function of integrals over the imaginary part of the refraction index but calculated using an approximation (equation (3.12)).

$$\delta_T = \frac{\sum x_i f_i}{\sum f_i} \tag{3.12}$$

where x_i is the depth of the ith layer and f_i^p (p = H, V polarization) is the weighting function (e.g., the fraction of absorption) for that layer as previously defined. The average soil temperature over the δ_T is regarded as the soil effective temperature T_{eff} and calculated by equation (3.13).

$$T_{eff} = \frac{\sum f_i T_i}{\sum f_i} \tag{3.13}$$

3.3.3 Lv incoherent model

The soil effective temperature is defined as the net intensity at the soil surface, which is a superposition of intensities emitted at various depths (Choudhury et al., 1982). The formula is as follows:

$$T_{eff} = \int_0^\infty \alpha(z) T_s(z) \exp\left[- \int_0^z \alpha(z') \quad '\right] dz \tag{3.14}$$

where $T_s(z)$ is the soil temperature at depth z. $\alpha(z)$ is the attenuation coefficient related to the complex soil dielectric constant $\varepsilon = \varepsilon' + i \,.\, \varepsilon''$ at depth z. For low-loss dielectric and nonmagnetic soil medium, $\alpha(z)$ can be expressed as (Choudhury et al., 1982; Ulaby et al., 2014),

$$\alpha(z) = \frac{2\pi}{\lambda_0} . \frac{\varepsilon''(z)}{\sqrt{\varepsilon'(z)}} \tag{3.15}$$

Assuming uniform SM and texture in each layer, a discrete formulation of (14) is derived by Lv et al. (2014),

$$T_{eff} = T_{s,1}\left(1 - e^{-\alpha_1 \cdot \Delta z_1}\right) + \sum_{i=2}^{N-1} T_{s,i}\left(1 - e^{-\alpha_i \cdot \Delta z_i}\right) \prod_{j=1}^{i-1} e^{-\alpha_j \cdot \Delta z_j} + T_{s,N} \prod_{j=1}^{N-1} e^{-\alpha_j \cdot \Delta z_j} \tag{3.16}$$

Where 1, i, and j represent the soil layers. N shares the same meaning as in equation (3.11). T_s is soil temperature. α is the attenuation coefficient given in equation (3.15) and Δz is soil thickness. $1 - e^{-\alpha_i \cdot \Delta z_i}$ is defined as the weight function for the ith layer (Lv et al., 2016b). By assuming uniform dielectric properties of soils throughout the emitting layer and a linear soil temperature gradient along the soil optical depth, the soil temperature at one time of the soil optical depth is proved equivalent to T_{eff}. The depth corresponding to one soil optical thickness is defined as the penetration depth of soil effective temperature δ_{PD} (Lv et al., 2018) as follows:

$$\begin{cases} \int_0^{\delta_{PD}} \Delta z_i \, \alpha(z) = 1 \\ \delta_{PD} = \dfrac{\lambda_0 \sqrt{\varepsilon'}}{2\pi\varepsilon''} \end{cases} \tag{3.17}$$

where Δz, α and ε share the same meanings as in equations (3.15-3.16). The *in situ* SM at 2.5 cm is used for calculating δ_{PD} in this case. Measured SM at depths above δ_{PD} are integrated using the weight function (in equation (3.16)) for obtaining the representative SM (SM_Lv, equation 3.18) that considers the impacts of profile SMST. The δ_{PD} (Lv's model) and δ_T (Wilheit's model) are referred to as sampling depths of soil effective temperature in the following analysis.

$$\text{SM\_Lv} = SM_1\left(1 - e^{-\alpha 1\,.\Delta z1}\right) + \sum_{i=2}^{\delta_{PD}-1} SM_i\left(1 - e^{-\alpha i\,.\Delta zi}\right)\prod_{j=1}^{i-1} e^{-\alpha j\,.\Delta zj} + SM_{\delta_{PD}}\prod_{k=1}^{N-1} e^{-\alpha k\,.\Delta zk} \tag{3.18}$$

where SM_i refers to SM at i layer. The other symbols in equation (3.18) share the same meanings as in equation (3.16).

3.3.4 Configuration of the simulation experiments

To assess the importance of the ATS model in seasonal T_B^p simulations, five experiments involving the Wilheit coherent model and the Lv incoherent model for the soil part are carried out. The first two experiments are called “baseline” simulations, only using the AIEM+TVG model without integration of the ATS model. The baseline experiments are configured with both the Wilheit and Lv models, which can reflect the impacts of the effective soil temperature and effective soil moisture, respectively, on T_B^p simulations. Considering that emission is sensitive to SM in the top layers (Zheng et al., 2017), the *in situ* measured SM at 2.5 cm is used to calculate the dielectric constant of bulk soil in the second experiment, which is equivalent to a concept of representative SM based on the Wilheit model. Furthermore, a total soil depth of 60 cm is considered in these two stratified layer models since the *in situ* measured SM at 60 cm remains almost constant during the study periods.

The third and fourth experiments integrate the ATS model combined with the combination of the Wilheit and Lv models separately (namely, “ATS-Wil” and “ATS-Lv”, respectively) to reflect the effects of the dielectric roughness on T_B^p simulations. Specifically, SM_Wil and SM_Lv are used respectively, to determine the dielectric constant of the bulk soil and T_{eff} (i.e., T_{eff}_Wil, T_{eff}_Lv, respectively) (as shown in Figure 3.3). Furthermore, the fifth experiment is considered with the combination of the Lv model and the 2.5cm SM to calculate the dielectric constant of the bulk soil and to evaluate the effectiveness of the Lv weighted SM approach (please refer to Figure 3.3). Table 3.2 summarizes the configurations of the simulation experiments.

Table 3.2 Configuration of the simulation experiments.

Experiments	AIEM	ATS	T_{eff}_Lv	T_{eff}_Wil	Emissivity
1) Base-Wil	+			+	SM_Wil, computed based on *in situ* SMST at 2.5, 5, 10, 20, 35, 60 cm
2) Base-Lv	+		+		*In situ* SM at 2.5 cm
3) ATS-Wil	+	+		+	SM_Wil, computed based on *in situ* SMST at 2.5, 5, 10, 20, 35, 60 cm
4) ATS-Lv	+	+	+		SM_Lv, computed based on *in situ* SMST at 2.5, 5, 10, 20, 35, 60 cm
5) ATS-Lv2.5	+	+	+		*In situ* SM at 2.5 cm

With the Wilheit (1978) model, properties such as the layer thickness *d* and the interpolation method to estimate the SMST in each layer based on a limited number of observations, affect model simulations. Similar to Raju et al. (1995) linear interpolation is used in this study. Considering the high sensitivity of coherent models to the optical thickness, a preliminary test was carried out to investigate the sensitivity of the Wilheit model to the soil layer thickness *d*. The results confirm the use of *d* value of 1 mm in the Wilheit model simulations, which is consistent with that considered by Raju et al. (1995).

3.4 Results

3.4.1 Late-monsoon period

3.4.1.1 Dielectric roughness thickness (h) and the sampling depths of the soil effective temperature (δ_{PD} and δ_T)

As a constant difference in the dielectric roughness thickness (h) between H and V polarizations is assumed (as expressed in equations (3.1, 3.4, 3.5)), only the estimated h value for H polarization is analyzed. Figure 3.5 shows comparisons of h estimated by the third (h_Wil), fourth (h_Lv) and fifth (h_Lv_SM2.5cm) experiments (bottom panel), together with the representative SM derived from Wilheit's (SM_Wil) and Lv's (SM_Lv) models and the *in situ* SM at 2.5 cm (SM_2.5cm) (upper panel). SM_Wil is found changing coincidently with SM_2.5cm, which might be attributed to the sampling depth of SM determined by the Wilheit model being approximately one-tenth the wavelength (approximately 2.5 cm in the L-band). SM_Wil is also slightly higher than SM_2.5cm when soils experience dry-wet-dry transition during the mid-monsoon period. Comparatively, a slight variation in SM_Lv is observed during this period, and this is the consequence of the Lv incoherent model with an assumed uniform SM distribution along the profile. Correspondingly, h_Lv does not increase as much as h_Wil and h_Lv_SM2.5cm do when upon soil drying. However, the Wilheit model can simulate h with obvious variations when soils experience dry and wet conditions due to its capability of considering the effect of the SM profile in calculating the dielectric constant of bulk soil. Compared to h_Lv_SM2.5cm, h_Wil is slightly lower in the soil drying process (Figure 3.5). The wet soil in the deeper layers considered in the Wilheit model leads to a smooth increase in h when the soil surface dries. h is estimated to exceed 10 cm under dry conditions (e.g., SM ≈ 0.1 m^3/m^3) (Figure 3.5). When SM increases and exceeds 0.3 m^3/m^3, the estimated h decreases under all schemes (Figure 3.5).

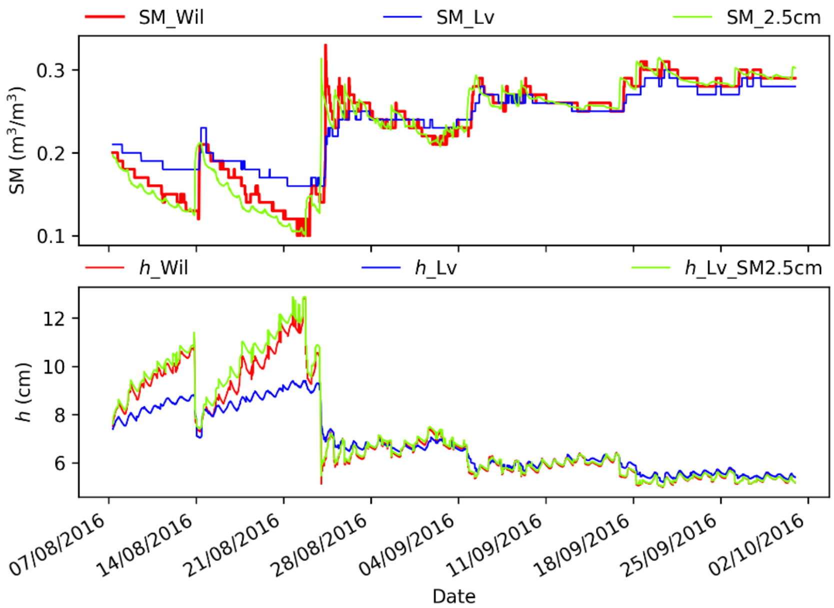


Figure 3.5 Comparisons of the estimated dielectric roughness thickness (h) in the third (ATS-Wil), fourth (ATS-Lv) and fifth (ATS-Lv2.5) experiments, supported by comparisons of SM during the late-monsoon period. SM_Wil is the representative SM derived from the Wilheit coherent model using Fresnel equations (e.g., by minimizing the objective function of the reflectivity difference). SM_Lv is the weighted SM derived from the Lv incoherent model, and SM_2.5cm refers to the in situ measured SM at 2.5 cm.

Figure 3.6 shows that the sensing depths of the effective temperature derived from the Lv model (δ_{PD}) and Wilheit model (δ_T) exhibit the same variations and approach each other during the whole late-monsoon period. Due to the consideration of the coherent effect in the Wilheit model, constructive interference might occur in reflections from the air-soil interface, and the dry-wet soil interfaces (drying front) and δ_T are found higher (~2.3 cm) than δ_{PD} (Figure 3.6). T_{eff}_Wil is close to T_{eff}_Lv, but both reveal a phase lag reflecting the propagation of periodic temperature waves from the deeper soil, over the *in situ* soil temperature at 2.5 cm (ST_2.5cm). T_{eff}, as a result of the superposition of the foregoing waves at the various depths within the soil, does not show as much variation as does ST_2.5 cm because the latter experiences rapid diurnal variations due to direct solar radiation, and this kind of variation is attenuated with increasing soil depth. As such, ST_2.5 cm is higher than T_{eff}_Wil and T_{eff}_Lv at midday but lower at midnight. T_{eff}_Wil is slightly higher (~0.2 K) than T_{eff}_Lv, especially at midday and midnight (Figure 3.6), whereas their differences decrease when soils are wet following rainfall events (as shown by the sharp jumps in SM_2.5cm in Figure 3.5). Figure 3.6 shows a negligible difference (~0.2 K) between T_{eff}_Wil and T_{eff}_Lv, while the varied thermal sampling depths of the soil temperature are important to determine the optimal mounting depth for observations Lv et al. (2016b).

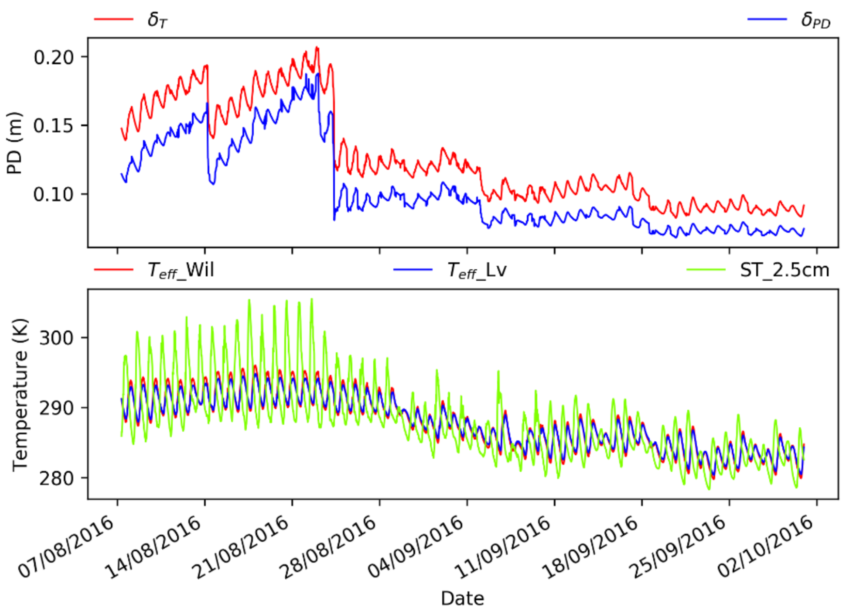


Figure 3.6 Comparisons of the sampling depths of the soil effective temperature estimated by the Wilheit coherent model and Lv incoherent model (δ_T and δ_{PD}) and T_{eff} (T_{eff}_Wil

and T_{eff}_Lv) during the late-monsoon period. ST_2.5cm is the in situ soil temperature measured at 2.5 cm.

3.4.1.2 T_B^p simulation

Figure 3.7 shows that the two baseline simulations underestimate T_B^H throughout the whole late-monsoon period, signifying that considering only the impacts of the effective soil temperature and effective soil moisture does not mitigate the discrepancy between simulations and observations. However, the underestimation of T_B^H (≈ 30-50 K) is obviously compensated by integrating the ATS model. The T_B^H simulated by the ATS-based models is close to the ELBARA-III observations in magnitude before late August, in which the ATS-Wil and ATS-Lv2.5 simulations match the observations, while the ATS-Lv model yields underestimations. The T_B^H simulated with the ATS-based models continues to be consistent with the observations in September when the soil surface is wet. This indicates the necessity of the ATS model for surface emission modeling under H polarization during the late-monsoon period.

Regarding V polarization, Figure 3.7 shows that the two baseline models simulate T_B^V well in August but underestimate T_B^V (≈ 20 K) in September when the soil surface is wet. By integrating the dielectric roughness in the ATS model, the underestimation degree is reduced, similar to that under H polarization. Compared to the ELBARA-III observations, the ATS-based models yield slight underestimations before late August but yield values closer to the observations than those of the baseline simulations. The ATS-Wil model performs better than the ATS-Lv model in T_B^V modeling during this period. All ATS-based models attain the same performances and capture T_B^V well during the end-monsoon period (September), despite certain discrepancies occurring after high-rainfall events (e.g., on 25/08/2016). The ATS model improves surface emission modeling under V polarization during the late-monsoon period.

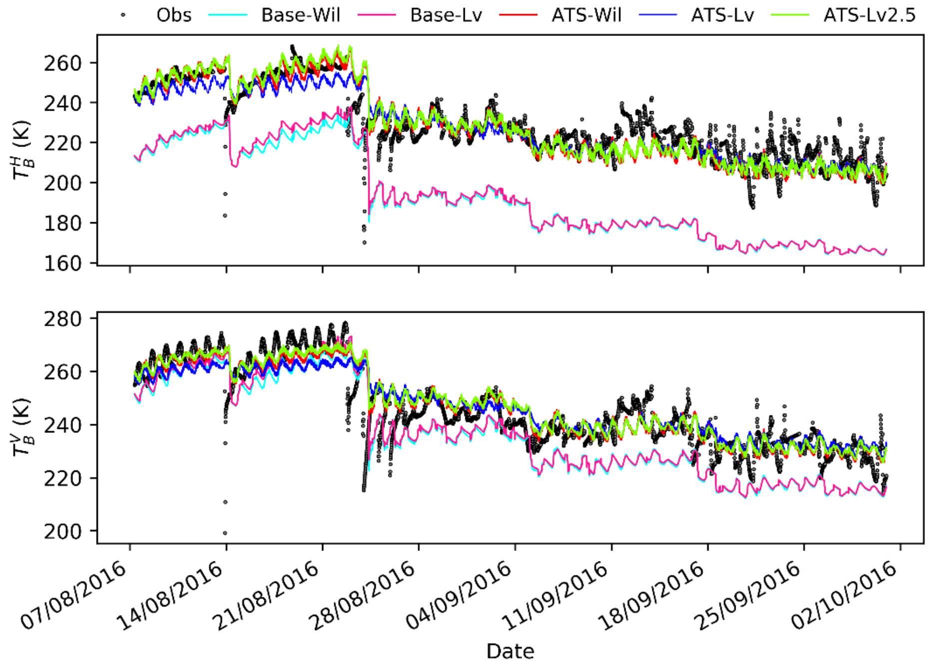


Figure 3.7 Comparisons of T_B^p simulated by the different models to the ELBARA-III T_B^p observations during the late-monsoon period. p denotes the H or V polarization mode. Base-Lv and Base-Wil denote the experiments using AIEM+TVG in combination with the Wilheit and Lv models, respectively. ATS-Lv and ATS-Wil denote the experiments using ATS+AIEM+TVG in combination with Wilheit and Lv models, respectively. ATS-Lv2.5 denotes the experiments using ATS+AIEM+TVG in combination with the Lv model and the soil moisture at 2.5 cm to calculate the dielectric constant of bulk soil.

3.4.2 Post-monsoon period

3.4.2.1 Dielectric roughness thickness (h) and the sampling depths of the soil effective temperature (δ_{PD} and δ_T)

Figure 3.8 shows that the SM_Wil and SM_Lv and associated h_Lv and h_Wil, respectively, are consistent with the *in situ* SM at 2.5 cm and h_Lv_SM2.5cm during the post-monsoon period before the soil freezing-dominated period, where the surface soil temperature is below 0 ° C (e.g., from 26/11/2016 to 30/11/2016, as shown in Figure 3.9). SM_Wil and SM_Lv and associated h_Lv, h_Wil and h_Lv_SM2.5cm (Figure 3.8) are found to experience fewer diurnal variations than those in SM_2.5cm during the surface freeze-thaw transition period, in which the surface soil temperature fluctuates around the freezing level 0 °C (e.g., from 12/11/2016 to 25/11/2016, Figure 3.9). At the beginning of the freezing-dominated period, the *in situ* SM at 2.5cm, SM_Wil and SM_Lv drop rapidly, resulting in all estimated h increasing. Notably, the estimated h value from 4-6 cm corresponds to the surface SM changes from 0.24-0.32 m^3/m^3 (Figure 3.8).

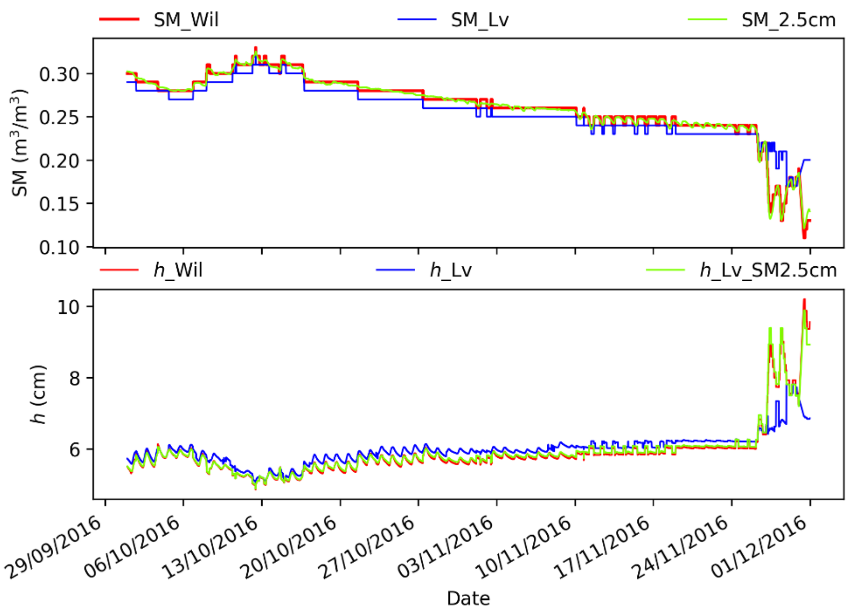


Figure 3.8 Same as Figure 3.5 but for the post-monsoon period.

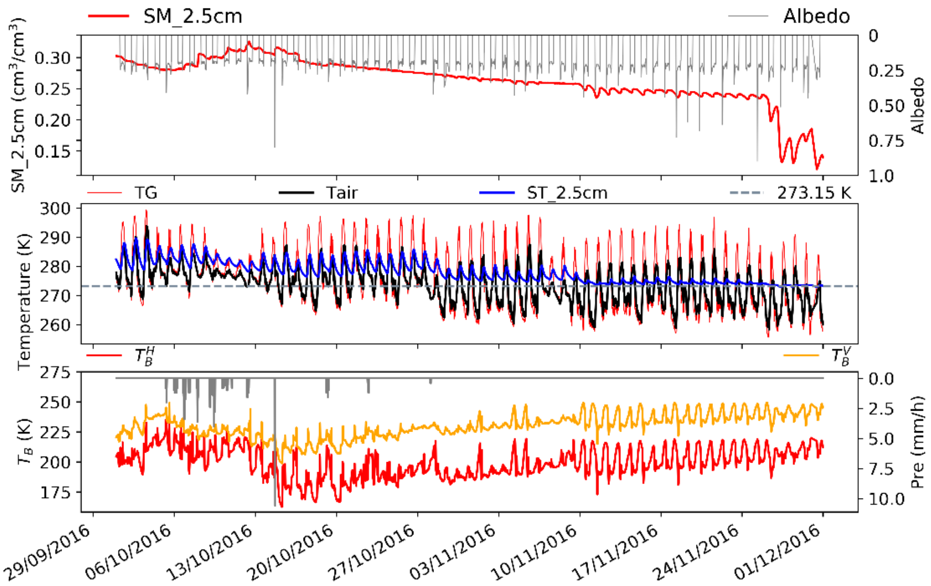


Figure 3.9 In situ measurements of the atmospheric variables, soil moisture and soil temperature at 2.5 cm (SM_2.5cm and ST_2.5cm, respectively) and the ELBARA-III T_B^p observations at the Maqu site during the post-monsoon period. The albedo (dimensionless) is calculated using the measured down- and up-welling solar radiation. TG refers to ground surface temperature, derived from the measured down- and up-welling longwave radiation. Tair refers to air temperature and Pre is the precipitation intensity.

Figure 3.10 shows that δ_T is higher (~1.6 cm) than δ_{PD} most of the time during the post-monsoon period with δ_T ranging from 7.2-14.0 cm and δ_{PD} ranging from 6.1-12 cm. The values of T_{eff}_Wil and T_{eff}_Lv are almost the same, and the difference between T_{eff}_Wil and T_{eff}_Lv with ST_2.5 cm as shown in Figure 3.10 is smaller than that during the late-monsoon period (as shown in Figure 3.6), which is related to seasonal variations in solar radiation. Similar to Figure 3.6, Figure 3.10 also shows a negligible difference between T_{eff}_Wil and T_{eff} _Lv, while the important varied thermal sampling depths of the soil temperature are used to determine the optimal mounting depth for observations as previously mentioned.

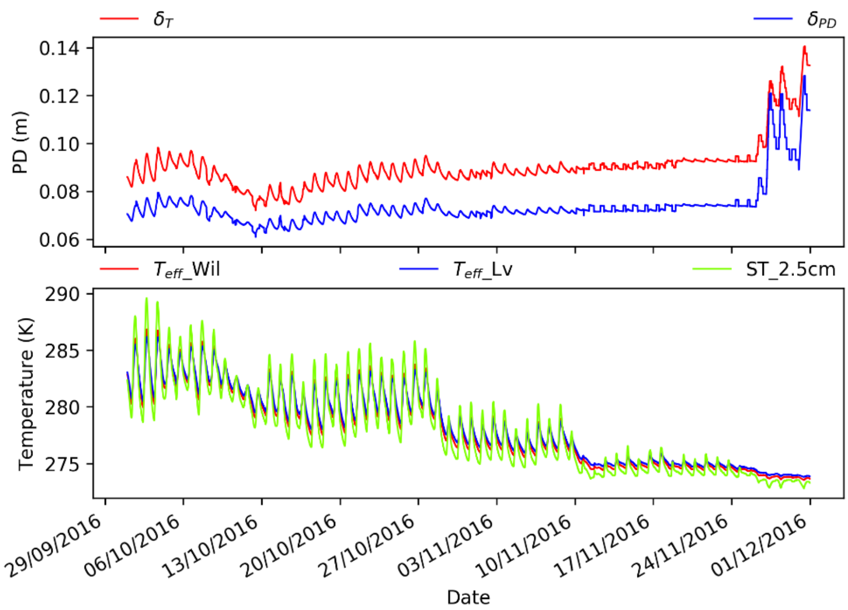


Figure 3.10 Same as Figure 3.6 but for the post-monsoon period.

3.4.2.2 T_B^p simulation

Figure 3.11 shows that the two baseline simulations underestimate T_B^p (≈ 20-50 K) during the post-monsoon period, while the T_B^p simulated by the ATS-based models is much closer to the observations. The T_B^p simulated by the ATS-based models deviates in October (Figure 3.11) when the weather system changes and the soils start to experience freeze-thaw processes. For instance, a low air temperature (Tair) and ground surface temperature (TG) occurred on approximately around 13/10/2016 (Figure 3.9), below the freezing point during the nighttime, followed by heavy precipitation with intensity over 10.0 mm/hour on 14/10/2016. As both TG and Tair are below the freezing point during the nighttime on the following days (as shown in Figure 3.9), the surface soil water might freeze at night and melt during the daytime, and the accumulated surface water might pond atop a frozen soil layer. The T_B^p simulated by the ATS-based models exhibits deviations but are much closer to the observations than those simulated by the baseline models in this time window (Figure 3.11). When soils experience steady surface freeze-thaw processes (e.g., without rainfall and snowfall during the period from 12/11/2016 to 25/11/2016, as shown in Figure 3.9), the T_B^H simulated by the ATS-based models is close to the observations but contain weak diurnal variations (① in Figure 3.11). The T_B^V simulated by the ATS-based models is underestimated (① in Figure 3.11). However, both T_B^H and T_B^V simulated by the ATS-based models coincide with the observations during the freezing-dominated period (e.g., from 26/11/2016 to 30/11/2016, ② in Figure 3.11) during the study period.

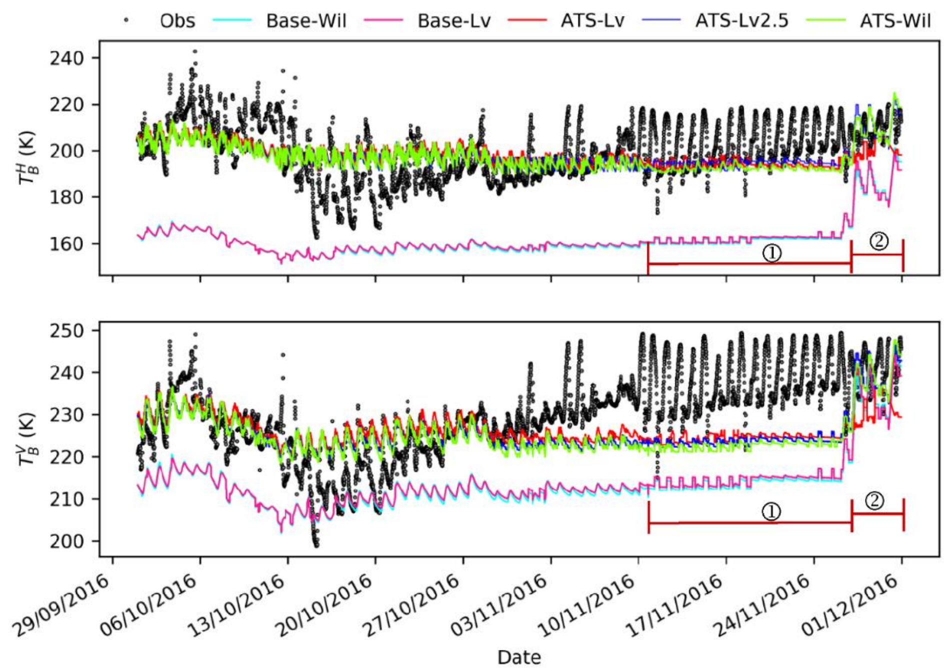


Figure 3.11 Same as Figure 3.7 but for the post-monsoon period.

3.5 Discussion

3.5.1 Applicability and uncertainty in the ATS model

The enhanced ATS model in this study stems from the original ATS model (Schneeberger et al., 2004; Schwank et al., 2008; Schwank et al., 2004), considering the effect of roughness components within the observed λ_0 range, and finding the impedance match in the ATS zone (Mätzler, 2006). The enhanced ATS model with the above new parameterizations of the dielectric roughness effects maintains the original physical considerations and helps improve T_B^p simulations.

The dielectric roughness thickness h is a key parameter in the ATS model, which is parameterized as comprising two components. One component is dielectric roughness within the soil volume (h_{SV}), and the other is the dielectric roughness induced by both SM in the surface and the geometric roughness effects (h_{SS}). Figure 3.12 shows that during the study period except for the soil freeze-thaw transition period, with SM decreasing (refer to SM decreasing from 0.2 to 0.1 m^3/m^3 in Figure 3.5), the contribution induced by SM in the soil volume to the dielectric roughness thickness (h_{SV}/h) increases (Figure 3.12a), while that of the surface (h_{SS}/h) decreases, and the opposite occurs when SM increases (refer to SM increasing from 0.15 to 0.3 m^3/m^3 in Figure 3.5). This is reasonable because as the soil dries, emissions originate from the deeper layers, the spatial heterogeneity in the dielectric constant of the soil volume notably increases and the dielectric roughness effects are enhanced, and the opposite is observed when soils become wet. This phenomenon was also reported by Escorihuela et al. (2007), Mo et al. (1987), Schneeberger et al. (2004), and Wigneron et al. (2001), in which the site-specific empirical soil roughness parameter H_R (Choudhury et al., 1979; Wang & Choudhury, 1981) was obtained for zeroth-order radiative transfer models.

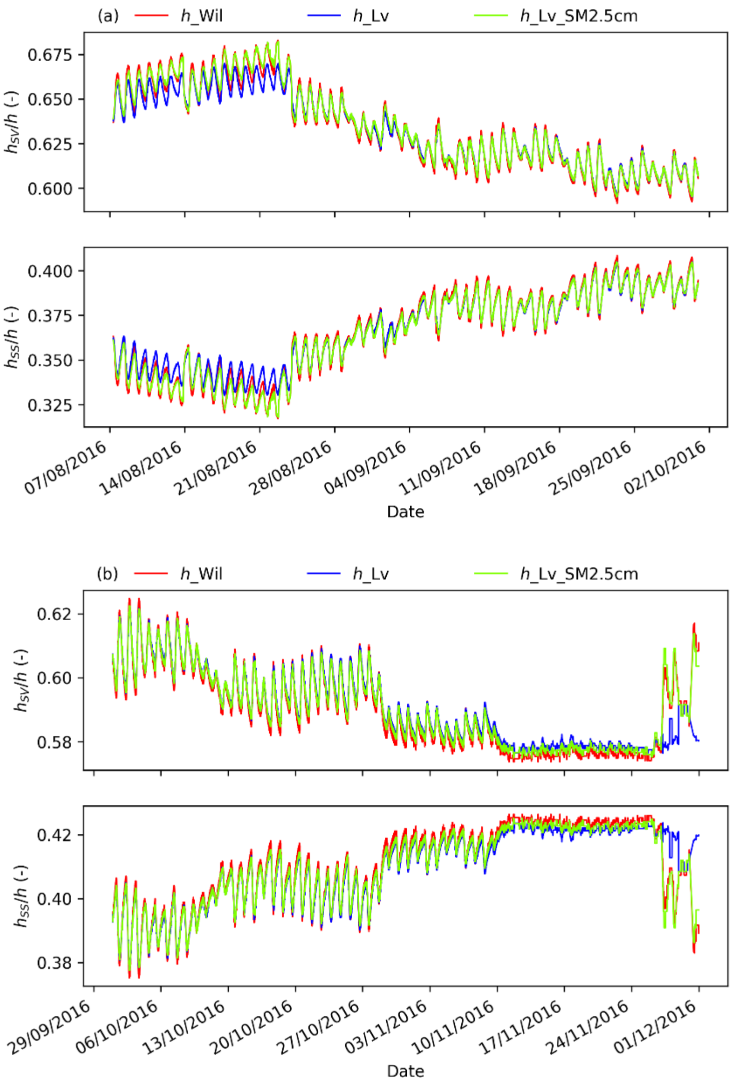


Figure 3.12 Comparisons of the ratio of h_{SV} and h_{SS} to h based on the third (ATS-Wil), fourth (ATS-Lv), and fifth (ATS-Lv2.5) experiments. (a) is for the late-monsoon period and (b) the post-monsoon period.

With decreasing SM, the scattering medium (e.g., loose dirt, plant debris, and clogs) at the soil surface increasingly dries and becomes more transparent for electromagnetic wave transmission. Hence, the contribution to the dielectric roughness of the soil surface (h_{SS}/h) decreases as shown in Figure 3.12a. Conversely, with increasing SM, the scattering medium including senescent vegetation (please refer to the decreased LAI in Figure 3.2) atop soils becomes wet and may trigger litter effects, leading to an increase in roughness (h_{SS}/h) (Figure 3.12). This is in line with findings reported by Grant et al. (2008) and Saleh et al. (2006), in which higher values of the calibrated H_R parameter are used to account for surface effects related to litter in grasslands. Moreover, given the parameterization of h_{SS} related to both the geometric roughness and SM, the scaled h_{SS} (namely h_{SS}/h) parameter can be comparable to H_R, which implicitly accounts for both the geometric and dielectric roughness effects. h_{SS}/h ranges from 0.31 to 0.43 (Figure 3.12), and these values are close to H_R ($= 1.3972 \cdot (s/L)^{0.5879}$) from Wigneron et al. (2001) and H_R ($= (0.9437 \cdot s/(0.8865 \cdot s + 2.29143))^6$) from Wigneron et al. (2011), with the same s and L values used in this study. In another study on grasslands in the Goulburn River catchment, Australia (Saleh et al., 2009), H_R was approximately 0.4.

Correspondingly, for the associated T_B^p modeling during the late-monsoon period, the two baseline simulations have high Pearson correlation coefficients (R ≥ 0.87) but yield consistently underestimated results. The T_B^p simulated by the ATS-based model only considering h_{SS} is higher than those of the two baseline simulations but still lower than the observed T_B^p values (please refer to section B.1 in Appendix B). However, the ATS-based models considering the impacts of both h_{SS} and h_{SV} greatly compensate for the underestimations of the foregoing simulations, with more simulation results closely aligned to the 1:1 line as shown in Figure 3.13a-b. The simulated T_B^p has similarly high R values (≥ 0.85) but much smaller root mean square errors (RMSEs ≤ 9.2 K for T_B^H and 8.0 K for T_B^V) than those of the baseline simulations (RMSEs over 37 K for T_B^H and 12 K for T_B^V). The ATS-Wil and ATS-Lv2.5 models perform better than the ATS-Lv model during this period, as more clustered points aligned along the 1:1 line are observed in Figure 3.13a-b that better match the observations shown in Figure 3.7. This underlines the importance of obtaining a realistic SM value that can reflect the moisture status of the ATS zone, and suggests that coherent effects can be considered during the late-monsoon season. With changing weather systems during the post-monsoon period, the ATS-based models maintain their performances with smaller RMSEs ($\leq$ 12.5 K for T_B^H and 10.9 K for T_B^V) than those of the baseline simulations (RMSEs over 39 K for T_B^H and 18 K for T_B^V), in which the ATS-Lv2.5 and ATS-Lv models with smaller RMSEs perform better than the ATS-Wil model (Figure 3.13c-d). This may indicate that the coherent effects occurring during the late-monsoon period may be disrupted due to the freeze-thaw processes during this period.

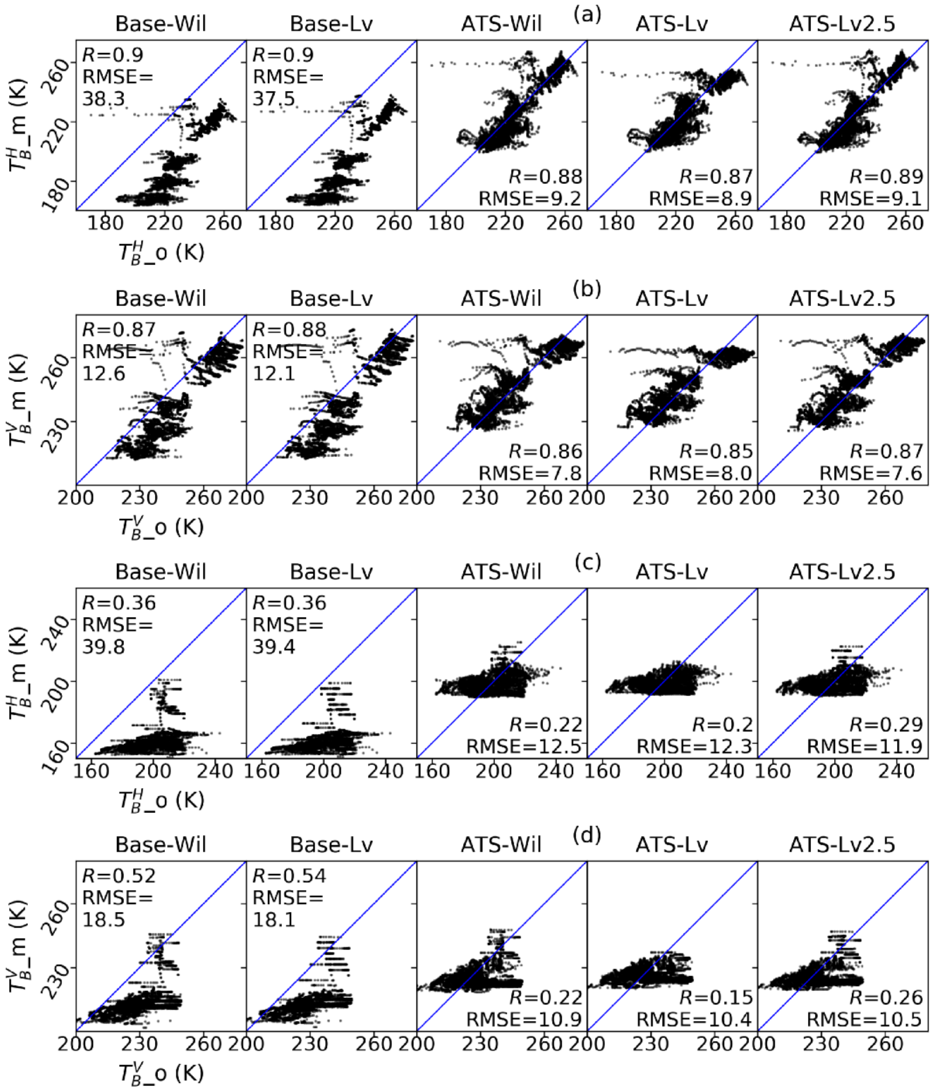


Figure 3.13 Simulated T_B^p (_m) from the Base-Lv, Base-Wil, ATS-LV, ATS-Wil and ATS-Lv2.5 experiments against the ELBARA-III observed T_B^p (_o) at the Maqu site. (a) and (b) are for the late-monsoon period (from 08/08/2016 to 30/09/2016), and (c) and (d) are for the post-monsoon period (from 01/10/2016 to 30/11/2016). R is the Pearson correlation coefficient, and RMSE is the root mean square error and is calculated as $\sqrt{\frac{1}{n}(T_B^p\_m - T_B^p\_o)^2}$ (where n is the number of observations). Base-Lv and Base-Wil

denote the experiments using AIEM+TVG in combination with the Wilheit and Lv models, respectively. ATS-Lv and ATS-Wil denote the experiments using ATS+AIEM+TVG in combination with the Wilheit and Lv models, respectively. ATS-Lv2.5 denotes the experiment using ATS+AIEM+TVG in combination with the Lv model and the soil moisture at 2.5 cm to calculate the dielectric constant of bulk soil.

However, the ATS-based models underestimate T_B^p for soils undergoing surface freeze-thaw processes, and the simulated T_B^p exhibits weak diurnal variations (please refer to Figure 3.11). The estimated dielectric roughness derived from the ATS-based models during this period (as shown in Figure 3.12b) is also found exhibiting slight variations for the ATS-based models, and the stable δ_T and δ_{PD} values and associated T_{eff}_Wil and T_{eff}_Lv (please refer to Figure 3.8 and Figure 3.10, respectively) may partially account for this variation. As the air temperature and ground surface temperature play the topmost roles in affecting the soil surface freeze-thaw process, appropriate temperature information is necessary to refine h estimation for soils during the freeze-thaw transition period. TG and Tair as shown in Figure 3.9 exhibit strong diurnal variations, and they should be investigated prior to soil moisture measurement at 2.5 cm, which does not reveal obvious diurnal variations during this period, but this is beyond the scope of this study. On the other hand, freeze-thaw processes exaggerate the inhomogeneity in soil media (e.g., composed of ice in pores mixed with pre-existing cracks, or melted liquid water mixed with ice, organic matter and soil solids). The formed ice affects the dielectric constant of the bulk soil during the nighttime and early morning, and the melted surface (soil) water affects that during the daytime. Without the soil ice content and surface (soil) water information considered in the ATS model parameterizations, the ATS model does not capture similar mixtures nor accurately models h and associated T_B^p. If we further consider surface (soil) liquid water and ice content information in h

estimation, the performance of the ATS model for the L-band radiometry of frozen-thawed soil is expected to improve. Notably, the soil ice content cannot be measured directly *in situ* but can be retrieved indirectly by assimilating proximal sensing signals (Mwangi et al., 2020). The inclusion of the soil ice content, surface liquid water fraction and ground surface temperature in the ATS model will be explored in further studies. A similar improvement can be expected to be implemented in T_B^p modeling considering rainfall events, such as with the ATS model considering surface water effects when the accumulated intensities of rainfall become higher than the soil infiltration capacity of the surface and the formed surface water blocks soil emission from the deeper layers.

Nevertheless, the correspondences between T_B^p estimated by the ATS-based models and the observations indicate that the ATS model is necessary for L-band T_B^p modeling. The h variable parameterized in this study acting as a dynamic parameter can describe the dielectric roughness of the soil surface and within the soil volume well, which is significant for the interpretation of observed T_B^p dynamics. The h_{SS} variable in this study is also related to the incidence angle and polarization, and N_p, similar to the empirical roughness parameterization (Wang & Choudhury, 1981), is found with a N_H value of 0 and N_V value of -1 in the best T_B^p simulations at the Maqu site. This result supports the finding that different N values should be used for horizontal and vertical polarizations (Escorihuela et al., 2007; Lawrence et al., 2013; Wigneron et al., 2011). The applicability of N_p in other climate regimes should be further confirmed, but this is beyond the scope of this study.

3.5.2 Impacts of the geometric roughness on the performance of the ATS model

Geometric surface roughness parameters (s and L) exert great impacts on surface scattering. s is considered in the dielectric roughness h parameterization and affects the depth that determines the variations in the effective dielectric constant of the air-to-soil medium (please refer to section 3.3.2). Considering the difficulty in determining the "true" values of s and L for natural grasslands and their importance in calculating the backscattering coefficients with the AIEM (Su et al., 1997), the effective geometric roughness parameters (s and L) obtained from satellite measurements in the Maqu area (Dente et al., 2014) are used in this study, as described in section 3.3.2. To investigate the impacts of the geometric roughness on the performance of the baseline and ATS-based model T_B^p simulations, sensitivity analyses by applying a varying s value in the range of [0.75, 0.9, 1.2, 1.5, 2.5 cm] with constant a constant L value of 9 cm (considering its lower impacts than those of s) are carried out. A s value of 0.9 cm is regarded as a smooth natural surface, and a s value of 2.5 cm is regarded as a rough surface in the L-band in this case. The results presented in section 3.4 and the aforementioned discussions confirm that both the ATS-Wil and ATS-Lv2.5 models perform the best in terms of T_B^p simulations during the late-monsoon period and that both the ATS-Lv and ATS-Lv2.5 models do so during the post-monsoon period except for the freeze-thaw transition period. Figures reflecting the impacts of the geometric roughness are thus only shown based on the ATS-Wil and ATS-Lv models together with the corresponding baseline models, and the discussions are focused on the whole study period except for the freeze-thaw transition period. Error metrics including R and RMSE are listed in Table 3.3 to quantitatively describe the performances of these models with varying s values.

The simulated T_B^p (especially T_B^H) by the baseline models (Figure 3.14a,c) is very sensitive to s variations, and the simulation results with large s values (e.g., 2.5 cm) are closer to the observations than those with small s values during the late-monsoon and post-monsoon periods. Please also refer to the reduced RMSEs with larger s values in Table 3.3. However, most simulations do not capture diurnal variations of T_B^p over the observations. With the ATS model integrated, the variations of T_B^p (especially T_B^H) can be captured, and the impacts of s variations on T_B^H during the late-monsoon period are reduced (i.e., refer to the narrow range of the variations of T_B^H with the different s settings in Figure 3.14d-1), although this is less apparent for T_B^V (Figure 3.14b-2,d-2). T_B^p simulations with small s values (e.g., 0.75 cm) present overestimations (Figure 3.14b-1), which is the opposite to those based on the Base-Wil model. T_B^p simulations with a s value of 2.5 cm are found consistent with those using a s value of 0.9 cm (similar higher R and small RMSE values in Table 3.3) and match the observations except during the soil freeze-thaw transition period (Figure 3.14a-d). However, the calculated microwave polarization difference index (MPDI=$(T_B^V - T_B^H)/(T_B^V + T_B^H)$) with a s value of 2.5 cm does not match the observed diurnal variations, while it does so with a s value of 0.9 cm (Figure 3.14e,f). This indicates that the positive effects imposed by high geometric surface roughness (e.g., s = 2.5 cm) on surface emission may become dominant and balance the negative effects of SM in the ATS model. In contrast, the ATS model with a s value of 0.9 cm can continuously capture the dynamic variations in dielectric roughness, not only at the soil surface related to the distribution of water and geometric roughness but also within the soil volume. Based on these analyses, the surface geometric roughness parameters (s = 0.9 cm and L = 9 cm) used in the ATS model are proved sufficient in this study. The surface geometric roughness may experience slight changes due to soil freeze-thaw processes, such as soil water freezing causing volume expansion and surface meltwater might smooth the surface, but this is beyond the scope of this study.

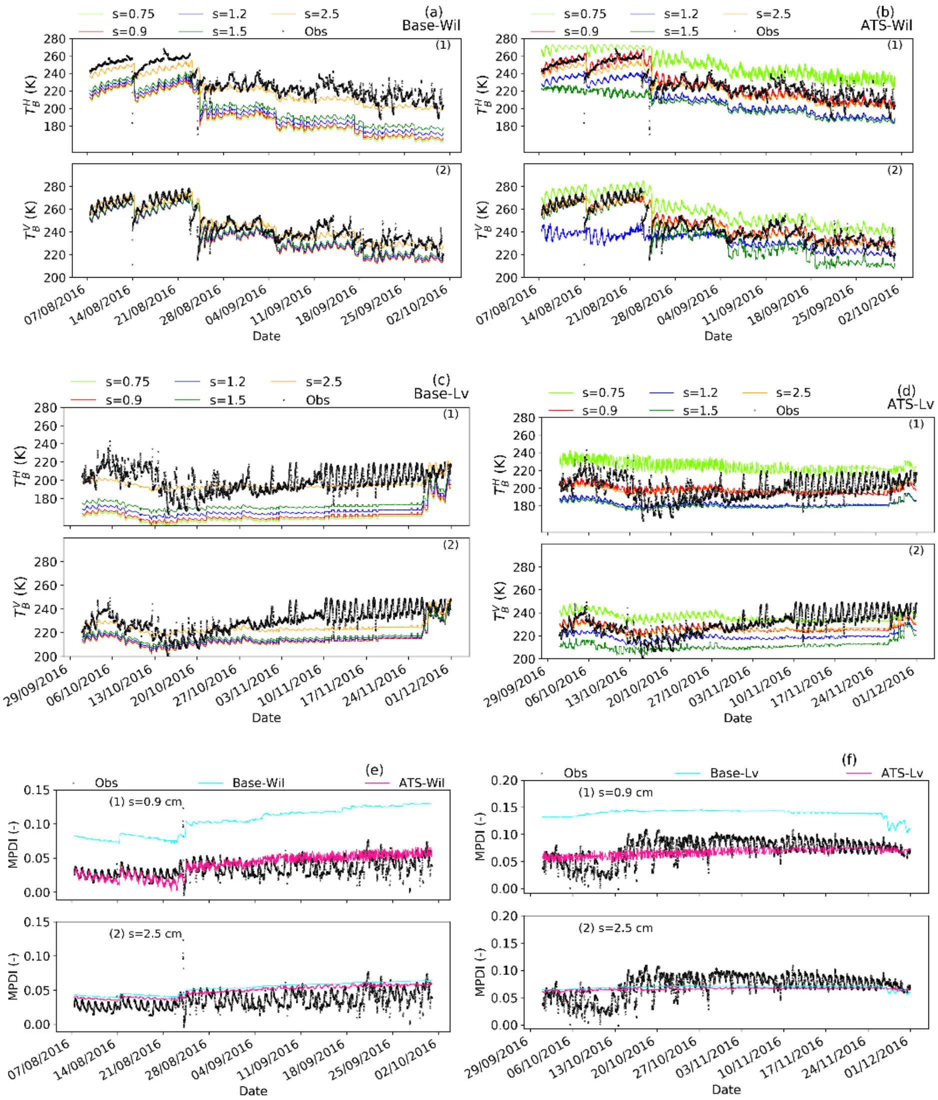


Figure 3.14 Comparisons of T_B^p simulated by the different models with the different s [0.75, 0.9, 1.2, 1.5, 2.5 cm] settings with a constant L (9 cm). (a) and (b) are by the Base-Wil and ATS-Wil models respectively, for the late-monsoon period. (c) and (d) are by the Base-Lv and ATS-Lv models respectively, for the post-monsoon period. (e) and (f) are the

calculated MPDI ($= (T_B^V - T_B^H)/(T_B^V + T_B^H)$*) with a s value of 0.9 cm and 2.5 cm for the late-monsoon and post-monsoon periods, respectively. Base-Lv and Base-Wil denote the experiments using AIEM+TVG in combination with the Wilheit and Lv models, respectively. ATS-Lv and ATS-Wil denote the experiments using ATS+AIEM+TVG in combination with the Wilheit and Lv models, respectively.*

Table 3.3 Comparisons of T_B^p *simulated by different models with different s [0.75, 0.9, 1.2, 1.5, 2.5 cm] settings with constant L (9 cm) to ELBARA-III observations.*

Polarization	Period	Late-monsoon				Post-monsoon			
	Experiments	Base-Wil		ATS-Wil		Base-Lv		ATS-Lv	
	s (cm)	R	RMSE (K)	R	RMSE (K)	R	RMSE (K)	R	RMSE (K)
H	0.75	0.9	40.1	0.85	23.0	0.37	41.4	0.13	28.8
	0.9	0.9	38.4	0.88	9.2	0.37	39.4	0.21	12.2
	1.2	0.9	34.0	0.89	20.8	0.37	34.5	0.37	20.6
	1.5	0.9	29.0	0.86	27.3	0.37	28.9	0.41	22.0
	2.5	0.9	11.3	0.89	9.8	0.38	12.1	0.36	11.6
V	0.75	0.88	12.8	0.84	14.8	0.55	18.3	0	12.3
	0.9	0.88	12.5	0.86	7.7	0.55	18.1	0.16	10.3
	1.2	0.88	11.7	0.73	17.6	0.54	17.1	0.48	13.9
	1.5	0.88	10.6	0.87	14.1	0.53	15.7	0.56	20.5
	2.5	0.87	7.5	0.87	7.4	0.48	10.2	0.45	9.6

3.5.3 Impacts of a fixed h_{SS}/h analogous to H_R in the SMAP and SMOS-CMEM systems on T_B^p simulations

SMAP and SMOS brightness temperature forward modeling approaches use a fixed soil roughness parameter H_R (Choudhury et al., 1979; Wang & Choudhury, 1981) for SM retrieval at the global scale. Given the similarity between h_{SS}/h derived from the ATS model and H_R (please refer to section 3.5.1), this section attempts to investigate the impacts of fixed roughness parameters analogous to those used in SMAP and SMOS retrievals on T_B^p modeling. In the SMAP SM retrieval algorithms, parameter H_R is assumed to be linearly related to s as $H_R = 0.1cm^{-1} \cdot s$ with a s value of 1.56 cm for grasslands (O'Neill et al., 2015). The default SMAP H_R value of 0.156 is reported too low for T_B^p modeling on the Tibetan Plateau in comparison to the recommended Wigneron et al. (2011) soil roughness model ($H_R = (0.9437 \cdot s/(0.8865 \cdot s + 2.29143))^6$) with the same s value (=1.56 cm) adopted (Zheng et al., 2018b). Regarding SMOS T_B^p long-term monitoring at the ECMWF (European Center for Medium-Range Weather Forecasts), the simple Wigneron et al. (2001) soil roughness model ($H_R = 1.3972 \cdot (s/L)^{0.5879}$) with a s value of 2.2 cm and L value of 6 cm is used in the CMEM (Community Microwave Emission Modeling Platform) (de Rosnay et al., 2020). The Choudhury et al. (1982) soil roughness model ($H_R = (2ks)^2$) used in SMOS SM retrieval (Kerr et al., 2012) adopts $H_R = 1.73$ if a s value of 2.2 cm is used, similar to CMEM, and this parameterization is found inferior to the simple Wigneron et al. (2001) roughness model (de Rosnay et al., 2020). Notably, this H_R ($= 1.73$) is different from the scaled h_{SS} (i.e., h_{SS}/h), which attains a maximum value of 1 in this study. An alternative SMOS soil moisture product (SMOS-IC) (Fernandez-Moran et al., 2017) uses globally mapped H_R values decoupled from the optimized combined vegetation and roughness parameter TR (= $\tau_{nad} + \frac{2}{H_R}$, where τ_{nad} is the vegetation optical depth at the nadir (Parrens et al., 2016)) by assuming a linear relationship between TR and LAI obtained from MODIS. As such, the uncertainties in the obtained H_R are more related to vegetation properties than to surface roughness which is the primary interest of this chapter. In contrast, the obtained H_R is directly applied in SMOS-IC retrieval, and there is no quantified relationship between H_R and geometric roughness parameters (s and L). Given the great impact of s on ATS+AIEM+TVG modeling (please refer to section 4.2), SMOS-IC H_R is not a good choice for comparisons in this study. To facilitate comparison, h_{SS}/h is set to a constant to match $H_R = 0.77$ (used in SMOS-CMEM) and $H_R = 0.58$ (suggested in SMAP on the Tibetan Plateau) during the study period. The same s and L values adopted in SMOS-CMEM and SMAP are used in the ATS-based model simulations. The Np parameter is set the same to that used in this study (i.e., 0 for H polarization and -1 for V polarization).

The T_B^H simulated by the ATS-based models with the SMAP setting (i.e., h_{SS}/h = 0.58) is found lower than the observations during the late-monsoon (Figure 3.15a, RMSE of 28.7 K in Table 3.4) and post-monsoon periods (Figure 3.15b, RMSE of 26.9 K in Table 3.4). In contrast, moderate underestimations are attained in the T_B^V simulations (Figure 3.15) with RSME values of 9.7 K and 13.3 K (Table 3.4) for the late-monsoon and post-monsoon periods, respectively. This may explain why only T_B^V is used in the SMAP soil moisture retrieval algorithms (O'Neill et al., 2020), which might be attributed to its lower sensitivity to changes in the surface roughness (refer to the narrower dynamic range of the T_B^V simulations than that of T_B^H, as shown in Figure 3.14, and the T_B^p(50°) simulation results presented in section B.2 in Appendix B). A similar finding was also reported by Zeng et al. (2017). The T_B^p simulated by the ATS-based models with the SMOS-CMEM setting (i.e., h_{SS}/h = 0.77) is found close to the observations during the late-monsoon and post-monsoon periods (RMSEs ≤ 13.3 K in Table 3.4) but do not match the strong diurnal variations (Figure 3.15). Furthermore, the fixed roughness parameter does not capture the temporal variations in the roughness characteristics related to changing surface conditions driven by the weather system, especially during the soil freeze-thaw transition period. In contrast, the proposed ATS model in this study (please refer to section 3.4) has the potential to reflect the dynamics of the dielectric roughness related to surface conditions, which is important to improve SM retrieval at the global scale.

Table 3.4 Comparisons of T_B^p simulated by the ATS-based models using fixed roughness parameters to the ELBARA-III observations. The ATS-Wil model is for the late-monsoon period and the ATS-Lv model for the post-monsoon period.

Polarization	Period	Late-monsoon		Post-monsoon	
	Experiments	R	RMSE (K)	R	RMSE (K)
H	SMOS-CMEM	0.89	10.7	0.40	11.5
	SMAP	0.89	28.7	0.29	26.9
V	SMOS-CMEM	0.87	7.6	0.47	9.2
	SMAP	0.87	9.7	0.51	13.3

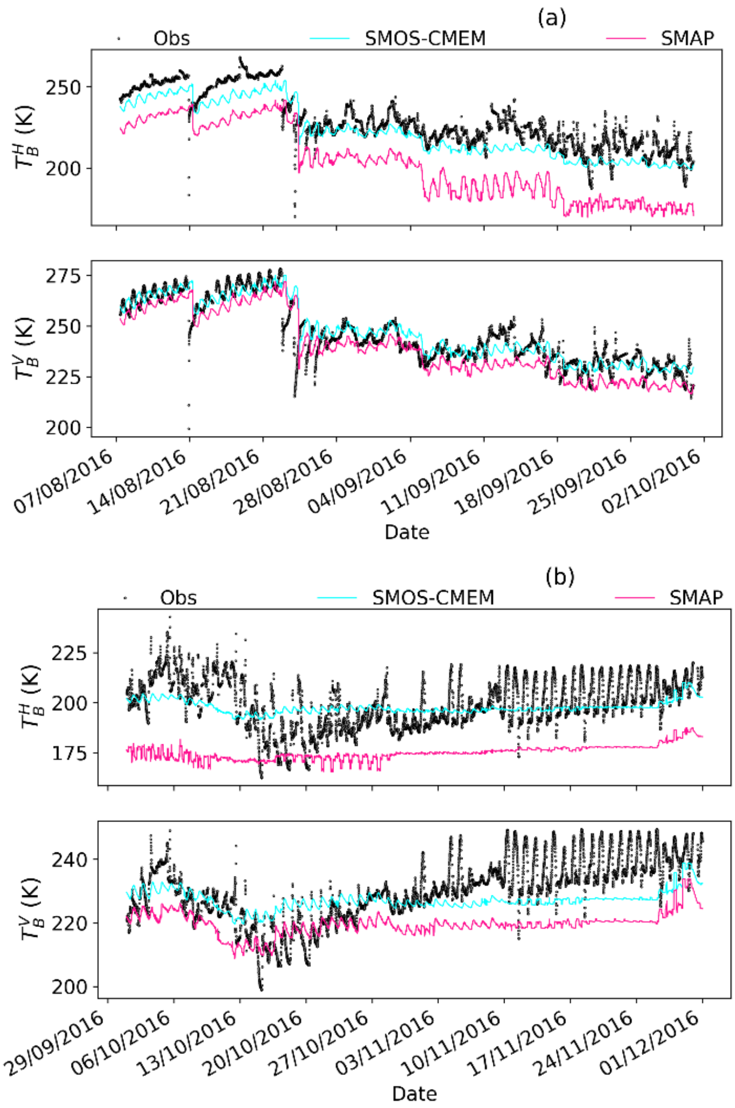


Figure 3.15 Simulated T_B^p values by the ATS-based models using fixed roughness parameters compared to the ELBARA-III observations. (a) and (b) show the values simulated by the ATS-Wil model for the late-monsoon period and by the ATS-Lv model for the post-monsoon period, respectively. SMOS-CMEM is run with $h_{SS}/h = 0.77$, and SMAP is run with $h_{SS}/h = 0.58$.

In summary, the enhanced ATS model coupled with the AIEM+TVG model is validated for natural grasslands. The proposed ATS model, as a physically-based model is expected to be applicable once all parameters become available for the area of interest (e.g., bare soil, cropland and forest). Parameters such as the wavelength and polarization mode are obtained from the sensor configuration. The effective geometric roughness parameters (i.e., s, L and Np) can be obtained by calibration using backscatter observations and soil moisture measurements as introduced by Su et al. (1997). When *in situ* measurements are not available, the most consistent input of soil moisture and temperature profiles in a consistent manner can be estimated by using land surface models (LSMs), such as the community land model (CLM) (Oleson et al., 2013), Noah LSM (Chen & Dudhia, 2001) and Hydrology-Tiled European Center for Medium-Range Weather Forecasts (ECMWF) Scheme for Surface Exchanges over Land (H-TESSEL) (Balsamo et al., 2009), while the simulation results should be validated to ensure the accuracy. Soil texture information can be obtained from global and regional soil maps, for instance, SoilGrids1km (Hengl et al., 2014b) suggested for the Tibetan Plateau (Zhao et al., 2018a). It is to note that the AIEM used in this study only involves single-scattering terms. Once a multiple scattering term is incorporated, the ATS model should be coupled with the updated version to ensure a more realistic scattering calculation.

Finally, we highlight the potential uses of the ATS model in microwave multi-frequency (i.e., commonly used 1-10 GHz) applications because it considers the wavenumber factor when parameterizing the dielectric roughness h_{SV} induced by the inhomogeneity in the soil volume (please refer to equations (3.5)and (3.6) in this chapter) and causes h scaling with the wavelength. Moreover, the developed ATS model can be applied to the active microwave case. Radar backscattering depends not only on soil moisture dynamics but also on the surface roughness, and better quantification of the latter can contribute to substantial improvements in soil moisture retrieval results (Lievens et al., 2011; Moran et al., 2004; Su et al., 1997). When the surface roughness issue is better approached with our proposed method, a better understanding of the vegetation scattering-emission can be focused, which will further contribute to soil moisture retrieval of vegetated regions.

3.6 Conclusions

In this study, the Tor Vergata discrete scattering model (TVG) coupled with the advanced integral equation model (AIEM) is used as a basis to investigate the effect of the surface roughness on coherent and incoherent emission processes. The developed air-to-soil transition (ATS) model with the proposed dielectric roughness parameterization is integrated with the TVG+AIEM model to investigate seasonal T_B^p signals as observed by the ELBARA-III radiometer on an eastern Tibetan alpine meadow. The Wilheit (1978) coherent and Lv et al. (2014) incoherent models are compared in terms of quantification of the dielectric constant of bulk soil in the ATS zone together with the *in situ* SM at 2.5 cm and calculation of the soil effective temperature T_{eff} (T_{eff}_Wil and T_{eff}_Lv). The penetration depths representing the sensing depth of T_{eff} derived from the coherent (δ_T) and incoherent (δ_{PD}) models are also compared.

The reduced discrepancy (≈ 20-50 K) between the modeled and observed T_B^p values demonstrates that the ATS model is necessary in seasonal L-band T_B^p simulations. The dielectric roughness thickness h parameterized in the ATS model can suitably describe the surface roughness resulting from fine-scale topsoil structures characterized by SM and geometric roughness effects and resulting from the soil volume attributed to a heterogeneous SM distribution. The proposed dielectric roughness can replace the fixed roughness parameter H_R and capture the dynamics of the surface roughness related to hydro-meteorological conditions. The consideration of the dynamic dielectric roughness is important for the improvement of SM retrieval at the global scale, in which a fixed H_R is used in current state-of-the-art processing. Furthermore, the soil porosity and logarithmic SM considered in the h parameterization enhance the physical coupling between microwave emission models and land surface models, which might improve retrievals of soil physical properties and contribute to the development of a soil monitoring system utilizing space-based Earth observation data with *in situ* data and modeling, especially in remote areas such as the Third Pole region.

The ATS model combined with the Wilheit coherent model (ATS-Wil) can be applied in T_B^p simulations for soils that occur under quasi-equilibrium conditions, such as thawed soils with a vegetation cover during the late-monsoon season and the beginning of the post-monsoon period. The ATS model combined with the Lv incoherent model (ATS-Lv) is applicable in T_B^p simulations of soils undergoing complex physical processes driven by rapidly changing weather systems, such as freeze-thaw processes after heavy rainfall events during the post-monsoon season. The ATS model using the *in situ* SM measured at 2.5 cm (ATS-Lv2.5) can be applied during the whole study period except the soil freeze-thaw transition period. The discrepancy between the modeled and observed T_B^p values during the soil freeze-thaw transition period suggests a potential enhancement of the ATS model by considering the effects of the surface temperature, surface water fraction and liquid water-ice mixtures in the calculation of h.

Chapter 4. Retrieving Soil Physical Properties via Assimilating SMAP Brightness Temperature Observations in the Community Land Model

This chapter is based on:

Zhao, H., Zeng, Y., Han, X., & Su, Z. (2021). Retrieving Soil Physical Properties via Assimilating SMAP Brightness Temperature Observations in the Community Land Model. *Journal of Advances in Modeling Earth System* (under review).

Abstract: Basic soil physical properties (i.e., soil texture and organic matter) and associated soil hydraulic properties (i.e., soil water retention curve and hydraulic conductivity) play an essential role in land surface models (LSMs) for soil moisture estimation. With the physical link between soil properties, LSMs and radiative transfer models (RTMs), soil physical properties can be retrieved using a LSM coupled with a microwave L-band emission observation model in a data assimilation framework. For this purpose, this chapter couples an enhanced physically-based discrete scattering-emission model with the community land model (CLM) v4.5 to retrieve soil physical properties with the local ensemble transform Kalman filter (LETKF) algorithm, assimilating Soil Moisture Active and Passive (SMAP) Level-1C (L1C) brightness temperature under H or V polarization (T_B^H and T_B^V), assisted with *in situ* measurements at the Maqu site on the eastern Tibetan Plateau. The results show improved estimates of the soil properties of the topmost layer by assimilating SMAP T_B^p ($p =$ H or V), as well as of the profile using retrieved top-layer soil properties and a prior depth ratio. The use of T_B^H and T_B^V exhibits varied sensitivities to the retrieval of different soil compositions (i.e., sand, clay, organic) and soil moisture estimates. However, analyses indicate that the obtained (retrieved) soil properties with high accuracy are not sensitive factors that could lead to the improvement of soil moisture estimates. Instead, the uncertainties in the CLM model structures, such as the fixed PTFs (pedotransfer functions), the hydraulic function describing the soil water retention curve and the water stress function determining root water uptake, should be considered.

Keywords: Soil physical property retrieval; Data assimilation; Enhanced discrete scattering-emission model; Community land model; Uncertainties in model structures

4.1 Introduction

Soil moisture, as one of the lower boundary conditions of the atmosphere, is a crucial land surface state that controls the partitioning of incoming energy into latent and sensible heat fluxes, and of rainfall into soil drainage and surface runoff, particularly in semi-arid and arid regions, where strong coupling between soil moisture and precipitation occurs (Koster et al., 2004; Seneviratne et al., 2010). Spatiotemporally consistent soil moisture information can be obtained using land surface models (LSMs) by assimilating *in situ* and remote sensing observations (De Lannoy & Reichle, 2016; Reichle et al., 2017; Reichle et al., 2007; Rodell et al., 2004; Yang et al., 2020). However, soil moisture simulations with LSMs are found more dependent on the specification of soil hydraulic (i.e., soil water retention characteristic and hydraulic conductivity) and thermal properties (i.e., soil heat capacity and thermal conductivity) (SHPs and STPs, respectively) than on the specification of atmospheric forcing or surface conditions (Gutmann & Small, 2005; Pitman, 2003; Santanello et al., 2001), because they govern the partitioning of soil moisture between infiltration and evaporation fluxes, and soil heat transport processes (Garcia Gonzalez et al., 2012; Zeng & Decker, 2009; Zeng et al., 2011a; Zeng et al., 2011c; Zeng et al., 2009a; Zhao et al., 2018a). Parameters of SHP & STP functions required by LSMs can be derived via pedotransfer functions (PTFs) (Van Looy et al., 2017; Vereecken et al., 2010), which use basic soil properties (i.e., soil texture and organic matter content) as input data. This consideration guarantees soil physical consistency in the land-atmosphere process (Cooper et al., 2020; Santanello et al., 2007; Soet & Stricker, 2003).

Many global and local efforts have been made to compile and develop soil databases, but uncertainties in these soil datasets might also cause a bias in SHP & STP predictions, and hence introduce uncertainties in representing the land surface states with LSMs (Su et al., 2013; Zhao et al., 2018a). To obtain soil properties and associated SHPs & STPs on a large scale (e.g., LSMs at the field and larger spatial scales), a retrieval approach, where the information on the spatial and temporal distributions of state variables (e.g., near-surface soil moisture) stemming from remote sensing observations, can be used to estimate soil properties by inversely solving the equations of water flow and energy balance (Mohanty, 2013).

With the availability of soil moisture derived from proximal-, air- or satellite-based observations (Su et al., 2020b), soil moisture has become the main state variable used for the retrieval of soil physical properties through improved LSMs (Bandara et al., 2014; Bandara et al., 2015; Montzka et al., 2011; Peters-Lidard et al., 2008; Qin et al., 2009; Santanello et al., 2007), and passive L-band microwave remote sensing has become the most promising technique to measure near-surface soil moisture (Lv et al., 2014; Lv et al., 2016a; Zheng et al., 2019). This trend is further accelerated by the launch of two innovative space missions equipped with L-band radiometers, the Soil Moisture and Ocean Salinity (SMOS) (Kerr et al., 2010) and the Soil Moisture Active and Passive (SMAP) (Entekhabi et al., 2010) missions, which have provided global soil moisture products at the daily scale. The zeroth-order radiative transfer model (RTM), the so-called $\tau - \omega$ model that contains a zero scattering phase function, is commonly applied as the forward model in L-band brightness temperature modeling (T_B^p, with p = H or V polarization) during soil moisture retrieval (de Rosnay et al., 2020; Kerr et al., 2012; O'Neill et al., 2018), in which the surface reflectivity is calculated through Fresnel equations (based on the plane wave assumption) combined with a semi-empirical surface roughness correction model (Choudhury et al., 1979; Wang & Choudhury, 1981).

SMOS soil moisture products have been used to retrieve soil hydraulic properties (Bandara et al., 2015; Lee et al., 2014; Shellito et al., 2016). However, current soil moisture products contain considerable biases, especially in high-latitude regions (Su et al., 2013; Su et al., 2011; Zeng et al., 2015; Zeng et al., 2016; Zhuang et al., 2020), due to uncertainties in specified parameters (e.g., surface roughness and vegetation optical parameters) in regard to microwave remote sensing (Su et al., 2020a). Any uncertainties in remotely sensed data at the calibration and retrieval stages can propagate to retrieved soil physical properties (Ines & Mohanty, 2009). Moreover, the soil dielectric mixing model (DMM) implemented in the RTM depends on soil moisture and temperature profiles and basic soil properties (i.e., soil texture, organic matter content and bulk density), which contain conditional uncertainties. These uncertainties degrade the consistency of soil physics and further affect the consistency between soil (moisture and temperature) states and boundary layer fluxes. Therefore, it may be challenging to retrieve soil physical properties with currently available soil moisture products (Bandara et al., 2014; Corbari et al., 2015; Yang et al., 2016).

Nevertheless, a LSM implements SHP & STP functions to predict soil water and heat fluxes and soil moisture and temperature profiles, with the latter outputs driving RTMs. As such, the coupled LSM and RTM approach naturally enables forward dynamic observation simulation of T_B^p. By assimilating T_B^p observations within a data assimilation framework with the coupled system involved, soil properties and associated SHPs & STPs can be retrieved (Burke et al., 1997; Camillo et al., 1986; Han et al., 2014a; Liu & Gupta, 2007; Reichle et al., 2013). Dedicated studies on soil property retrieval via T_B^p assimilation have been reported; for example, synthetic L-band T_B^H data have been assimilated into the coupled community land model (CLM) (Oleson et al., 2013) and community microwave emission model (CMEM) (de Rosnay et al., 2020) (Han et al., 2014a), ground-based ELBARA-II L-band T_B^H observations of bare tilled soil have been assimilated into a coupled coherent RTM with the HYDRUS-1D model rather than a LSM (Dimitrov et al., 2015; Dimitrov et al., 2014), and basic soil properties have been calibrated through a similar coupled system (Yang et al., 2016) with satellite-based AMSE-R (Advanced Microwave Scanning Radiometer for Earth Observing System) T_B^V observations at 6.9, 10.7 and 18.7 GHz assimilated. All observation operators used in current studies are zeroth-order RTMs with semi-empirical parameterizations.

The complementary relationship between emission and scattering (Peake, 1959) has allowed the development of scattering (formulated with the involved scattering phase function) – emission physically-based models through Maxwell's equations, such as the integral equation model (IEM) (Fung, 1994) and its advanced version (AIEM) (Chen et al., 2003b) for rough bare soil surfaces, and discrete scattering model (notably Tor Vergata model) (Ferrazzoli & Guerriero, 1996) for vegetated surfaces. The Tor Vergata model simulating vegetation scattering coupled with the AIEM can be used to model scattering-emission of the overall vegetation-soil continuum (Bai et al., 2019; Dente et al., 2014; Wang et al., 2018; Zheng et al., 2017). As the surface roughness greatly impacts L-band T_B^H simulations (Schneeberger et al., 2004; Schwank et al., 2004), a physical process-based surface roughness model, an air-to-soil transition (ATS) model (Zhao et al., 2021), was developed, which incorporates the dielectric roughness, resulting not only from topsoil structures affected by both irregularities (i.e., geometric roughness) of the soil surface and inhomogeneous moisture distribution, but also attributed to the inhomogeneity within the soil volume related to the soil porosity and moisture. The above development has enabled the application of an enhanced physically-based observation operator, namely, the integrated TVG-AIEM-ATS-DMM model, for soil property retrieval.

This chapter adopts the integrated TVG-AIEM-ATS-DMM (abbreviated as TVG) model as the observation operator. The TVG model coupled with CLM 4.5 (Oleson et al., 2010) in the open-source multivariate land data assimilation framework (DasPy) (Han et al., 2014a; Han et al., 2015) is deployed as the T_B^p simulator, and basic soil properties are retrieved by means of the local ensemble transform Kalman filter (LETKF) (Hunt et al., 2007) algorithm by assimilating SMAP Level-1C (L1C) T_B^p data. The Maqu site (33.91°N, 102.16°E) on the eastern Tibetan Plateau, providing comprehensive field observations (Su et al., 2011; Su et al., 2020a; Zeng et al., 2016; Zhuang et al., 2020), is selected as the study area to investigate: 1) whether the SMAP T_B^p observation assimilation improves estimates of soil properties and their vertical descriptions? Consequently, do the retrieved soil properties improve estimates of soil moisture and temperature profiles, and land surface heat fluxes with the CLM? 2) Is the retrieval of soil property polarization-dependent? Since T_B^H is claimed to be sensitive to soil moisture changes (Njoku et al., 2002), and comparably, T_B^V observations are less affected by surface roughness changes, this results in a higher accuracy when applied in satellite soil moisture retrieval results (O'Neill et al., 2015). As such, the assimilation of T_B^p signals on different polarizations may yield different soil property retrievals. This study assimilates SMAP T_B^H and T_B^V data separately and analyzes the effect of the polarization configuration on soil property retrieval. 3) Are the model physics and structure adequate for soil property retrieval and associated land state estimates over *in situ* observations?

The chapter is organized as follows: section 4.2 presents the Maqu site observations, a brief description of methods (the CLM, TVG and LETKF algorithms), and the experimental design for soil property retrieval. Results and discussions are provided in sections 4.3 and 4.4, respectively. Conclusions are drawn in section 4.5.

4.2 Materials and Methods

4.2.1 Maqu site observations

The Maqu area has a cold climate with dry winters and warm summers (Dwb) according to the updated Köppen-Geiger climate classification (Peel et al., 2007). The winter season ranges from late November to late March, in which soils undergo freeze-thaw cycles (Su et al., 2020a). The land cover mainly comprises alpine meadows with grass heights ranging from 5 to 15 cm throughout the growing season due to intensive grazing by livestock (e.g., yaks and sheep). The prevailing soil types are sandy loam, silt loam, and organic soil with average proportions of 30.3% sand, 9.9% clay, and a maximum of 39.0% organic matter (Zhao et al., 2018a).

In this chapter, collected half-hourly wind speed, near-surface air temperature, near-surface relative humidity and air pressure, liquid precipitation, and incident solar and longwave radiation data at the Maqu site (Su et al., 2020a) are used for atmospheric forcing. Basic soil properties (i.e., soil texture, organic matter content) measured in the laboratory (Zhao et al., 2018a), and soil moisture and temperature profiles combined with turbulent heat fluxes measured *in situ* are adopted to evaluate the performance of T_B^p assimilation in regard to soil property retrieval. MCD15A2H - MODIS/Terra+Aqua leaf area index (LAI, 500m resolution) (https://lpdaac.usgs.gov/products/ mcd15a2hv006/) products are extracted to obtain LAI time series (please refer to section 3.2), which is used to define the plant functional type in CLM model and determine the number of grass leaves regarded as scatters parameterized in the TVG model.

The SMAP Level-1C T_B^p data acquired by the L-band radiometer at 2 to 3-day intervals (Entekhabi et al., 2014) are also available at the Maqu site (Su et al., 2020a). SMAP T_B^p at near 6:00 AM local time (the time of the SMAP descending pass) is considered for assimilation, because the temperature within one model grid cell can be regarded as homogeneous at this time, and the vegetation temperature is the same as soil temperature. The different spatial scales of the SMAP T_B^p and field observations were not considered in this study. The SMAP T_B^p observations used in the study are the arithmetic average of fore- and aft-looking data. Furthermore, T_B^p at the field scale measured by the ELBARA-III radiometer installed at the Maqu site is adopted for validation. Detailed descriptions of the instrumentation and data at the Maqu site are given in Su et al. (2020a).

Data on the soil non-frozen period are more favorably used for soil property retrieval, because the soil freeze-thaw processes occurring during the winter period (from November afterward) might alter the topsoil particle composition and alter soil physical properties (Xie et al., 2015; Zhang et al., 2016). Moreover, snowfall and soil freeze-thaw processes complicate microwave emission and render T_B^p model unreliable. Nevertheless, the basic soil texture does not change dramatically due to these complications, although the soil hydraulic properties can be altered due to the presence of soil ice (Yu et al. 2018, Yu et al. 2020, Mwangi et al. 2020). Soil property retrieval by assimilating SMAP T_B^p is therefore concentrated during the non-frozen period.

4.2.2 Methods

4.2.2.1 Community land model

Community land model (CLM) v4.5 (Oleson et al., 2013) uses a modified Richards equation to predict the one-dimensional multi-layer vertical soil water flow and heat transport. The Monin-Obukhov similarity theory is used to derive land surface fluxes. In the CLM, soils are divided into 15 layers, in which the depths of the soil layers exhibit an exponential relationship along the profile, and the first 10 layers are used for soil moisture estimations. Soil layer node depths that define where the volumetric soil water and temperature are estimated by the CLM are listed in Table 4.1 for the ten layers of the soil column. The upper boundary conditions for soil water flow and heat transfer are the infiltration flux and ground heat flux into the topsoil layer, respectively, and the lower boundary condition depends on the depth of the water table and the zero heat flux at the bottom of the soil column. The CLM uses the Campbell (1974) power function to describe soil water retention and soil hydraulic characteristics (SHPs). Four hydraulic parameters—saturated soil moisture content θ_s (m^3/m^3), saturated matric potential φ_s (mm), the pore size distribution index B (dimensionless) and saturated hydraulic conductivity K_s (mm/s) characterize the function, and they are estimated through Cosby et al. (1984) PTFs by using the percentages of sand and clay as inputs and organic properties of the soil (Lawrence & Slater, 2008). The De Vries (1963) thermal parameterization scheme instead of the default Johansen (1975) scheme is used to estimate soil thermal properties (i.e., soil heat capacity and thermal conductivity), given its physical considerations and higher performance based on *in situ* investigations (Yu et al., 2018; Zhao et al., 2018a). Other default model physics/parametrizations/constant values are applied in this study. Detailed equations for soil water flow and heat transfer modeling by the CLM are listed in Appendix C.

Table 4.1 Soil layer node depth, thickness and depth at layer interface for the ten layers in the CLM.

Layer	Layer node depth z_i (cm)	Thickness (cm)	Depth $z_{h,i}$ (cm)
1	0.71	1.75	1.75
2	2.79	2.76	4.51
3	6.23	4.55	9.06
4	11.89	7.5	16.55
5	21.22	12.36	28.91
6	36.61	20.38	49.29

7	61.98	33.6	82.89
8	103.8	55.39	138.28
9	172.76	91.33	229.61
10	286.46	150.58	380.19

4.2.2.2 Tor Vergata model

A detailed flowchart of the forward T_B^p simulation with the integrated TVG-AIEM-ATS-DMM model can be found in section 3.3.1. The DMM recently developed by Park et al. (2017) is used in this study, which considers the effect of the organic matter (by linking it with the dry bulk density) on soil dielectric constant. The input of the Park DMM includes the soil volumetric water content, soil temperature, sand and clay fractions, and organic matter content.

4.2.2.3 Local Ensemble Transform Kalman Filter (LETKF)

The DasPy data assimilation framework (Han et al., 2014a; Han et al., 2015) was developed to integrate observations from multiple sources with the CLM to improve predictions of the water and energy cycles of the soil-vegetation-atmosphere continuum. The LETKF (Hunt et al., 2007) is incorporated as the main data assimilation algorithm in DasPy, which uses the Gaussian approximation and follows the time evolution of the mean and covariance by propagating an ensemble of states. This section follows the derivation given in Hunt et al. (2007). Given an ensemble $\boldsymbol{x}_{t-1}^{a(i)}$ of m-dimensional model state vectors at time $t-1$, a nonlinear model ($\boldsymbol{M}_{t-1,t}$) is applied to drive the evolution of each ensemble member to form a background ensemble $\boldsymbol{x}_t^{b(i)}$ at time t (equation (4.1)). $\boldsymbol{y}_t^o$ is a vector of observations at time t. These observations are related to the state vector by equation (4.2).

$$\boldsymbol{x}_t^{b(i)} = \boldsymbol{M}_{t-1,t}\left(\boldsymbol{x}_{t-1}^{a(i)}\right) + \boldsymbol{\epsilon}_t \tag{4.1}$$

$$\boldsymbol{y}_t^o = \boldsymbol{H}_t\left(\boldsymbol{x}_t^{b(i)}\right) + \boldsymbol{\varepsilon}_t \tag{4.2}$$

where $i = 1,2 \ldots . k$ with k as the ensemble member size. The subscripts $a(i)$ and $b(i)$ denote the analysis (posterior) and background (prior), respectively, of member i. $\boldsymbol{H}_t$, as the observation operator, maps the state vector onto the observations. $\boldsymbol{\epsilon}_t$ and $\boldsymbol{\varepsilon}_t$ denote the model and observation errors respectively. $\boldsymbol{\varepsilon}_t$ are assumed as unbiased Gaussian and uncorrelated errors in time with known covariance matrices $\boldsymbol{R}$. In the following, the operations and results at the posterior t are concentrated and the subscript t is dropped. Regarding the prior state estimate and its covariance, the sample mean and covariance of the prior ensemble are used (equations (4.3-4.4)).

$$\overline{\boldsymbol{x}}^b = k^{-1}\sum_{i=1}^{k}\boldsymbol{x}^{b(i)} \tag{4.3}$$

$$\begin{aligned}\boldsymbol{P}^b &= (k-1)^{-1}\sum_{i=1}^{k}(\boldsymbol{x}^{b(i)} - \overline{\boldsymbol{x}}^b)(\boldsymbol{x}^{b(i)} - \overline{\boldsymbol{x}}^b)^T \\ &= (k-1)^{-1}\boldsymbol{X}^b(\boldsymbol{X}^b)^T\end{aligned} \tag{4.4}$$

Formally, the LETKF aims to find the initial state and/or parameters that minimize the distance to the prior estimate, weighted by $\boldsymbol{P}^b$, while also minimizing the distance of the model trajectory to the observations, weighted by $\boldsymbol{R}^{-1}$ at time t. The Kalman filter cost function to be minimized to determine the posterior mean $\overline{\boldsymbol{x}}^a$ is formulated in equation (4.5).

$$\begin{aligned}\boldsymbol{J}(\boldsymbol{x}) = &\left(\boldsymbol{x} - \overline{\boldsymbol{x}}^b\right)^T\left(\boldsymbol{P}^b\right)^{-1}\left(\boldsymbol{x} - \overline{\boldsymbol{x}}^b\right) \\ &+ [\boldsymbol{y}^o - \boldsymbol{H}(\boldsymbol{x})]^T\boldsymbol{R}^{-1}[\boldsymbol{y}^o - \boldsymbol{H}(\boldsymbol{x})]\end{aligned} \tag{4.5}$$

The $m \times m$ background covariance matrix $\boldsymbol{P}^b$ is noted to be invertible, as it exhibits a rank of at most $k-1$. However, as a symmetric matrix, it is one-to-one on its column space S (i.e., the column space of $\boldsymbol{X}^b$) spanned by the prior ensemble perturbations. $\boldsymbol{P}^b$ and $\boldsymbol{J}(\boldsymbol{x})$ are therefore well defined on S, and minimization is carried out in this space. To obtain the posterior ensemble on S, a linear transformation is assumed for $\boldsymbol{X}^b$ from a k-dimensional space $\tilde{S}$ onto S, and local analysis is performed in $\tilde{S}$. Let $\boldsymbol{w}$ denote a vector in $\tilde{S}$, $\boldsymbol{X^b w}$ then belongs to the space S spanned by the prior ensemble perturbations, and $\boldsymbol{x} = \overline{\boldsymbol{x}}^b + \boldsymbol{X}^b\boldsymbol{w}$ is the corresponding model state. Let l be the number of scale observations used in the analysis. Assuming $\boldsymbol{w}$ is a Gaussian random vector with mean 0 and covariance $(k-1)^{-1}\boldsymbol{I}$, $\boldsymbol{x} = \overline{\boldsymbol{x}}^b + \boldsymbol{X}^b\boldsymbol{w}$ is Gaussian with mean $\overline{\boldsymbol{x}}^b$ and covariance $\boldsymbol{P}^b = (k-1)^{-1}\boldsymbol{X}^b(\boldsymbol{X}^b)^T$. Then, equation (4.5) becomes equation (4.6):

$$\tilde{\boldsymbol{J}}(\boldsymbol{w}) = (k-1)\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{w} + [\boldsymbol{y}^o - \boldsymbol{H}(\overline{\boldsymbol{x}}^b + \boldsymbol{X}^b\boldsymbol{w})]^T\boldsymbol{R}^{-1}[\boldsymbol{y}^o - \boldsymbol{H}(\overline{\boldsymbol{x}}^b + \boldsymbol{X}^b\boldsymbol{w})] \tag{4.6}$$

An ensemble $\boldsymbol{y}^{b(i)} = \boldsymbol{H}(\boldsymbol{x}^{b(i)})$ of prior observation vectors is defined with mean $\overline{\boldsymbol{y}}^b$, and the $l \times k$ matrix $\boldsymbol{Y}^b = \boldsymbol{y}^{b(i)} - \overline{\boldsymbol{y}}^b$ as the perturbation matrix. The linear approximation is assumed by equation (4.7):

$$\boldsymbol{H}\left(\overline{\boldsymbol{x}}^b + \boldsymbol{X}^b\boldsymbol{w}\right) \approx \overline{\boldsymbol{y}}^b + \boldsymbol{Y}^b\boldsymbol{w} \tag{4.7}$$

Then the cost function (equation (4.6)) yields a quadratic form and is formulated by equation (4.8):

$$\tilde{\boldsymbol{J}}^*(\boldsymbol{w}) = (k-1)\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{w} + [\boldsymbol{y}^o - \overline{\boldsymbol{y}}^b - \boldsymbol{Y}^b\boldsymbol{w})]^T\boldsymbol{R}^{-1}[\boldsymbol{y}^o - \overline{\boldsymbol{y}}^b - \boldsymbol{Y}^b\boldsymbol{w}] \tag{4.8}$$

This cost function (equation (4.8)) occurs in the form of the Kalman filter cost function, using the prior mean $\overline{\boldsymbol{w}}^b = 0$, prior covariance $\widetilde{\boldsymbol{P}}^b = (k-1)^{-1}\boldsymbol{I}$ and

$\boldsymbol{Y}^b$ acting as the observation operator. Then, analogous to the updates in the Kalman filter:

$$\overline{\boldsymbol{w}}^a = \widetilde{\boldsymbol{P}}^{\boldsymbol{a}}(\boldsymbol{Y}^b)^T\boldsymbol{R}^{-1}(\boldsymbol{y}^o - \overline{\boldsymbol{y}}^b) \tag{4.9}$$

$$\widetilde{\boldsymbol{P}}^{\boldsymbol{a}} = [(k-1)\boldsymbol{I} + (\boldsymbol{Y}^b)^T\boldsymbol{R}^{-1}\boldsymbol{Y}^{\boldsymbol{b}}]^{-1} \tag{4.10}$$

In model space, the posterior mean and covariance are formulated by equations (4.11-4.12):

$$\overline{\boldsymbol{x}}^a = \overline{\boldsymbol{x}}^b + \boldsymbol{X}^b\overline{\boldsymbol{w}}^a \tag{4.11}$$

$$\boldsymbol{P}^a = \boldsymbol{X}^b\widetilde{\boldsymbol{P}}^a(\boldsymbol{X}^b)^T \tag{4.12}$$

Then, the posterior ensemble described by $\overline{\boldsymbol{X}}^a = \boldsymbol{X}^b\boldsymbol{W}^a$ also must be updated, in which the symmetric square root is used to determine $\boldsymbol{W}^a$ based on $\widetilde{\boldsymbol{P}}^{\boldsymbol{a}}$.

$$\boldsymbol{W}^a = [(k-1)\widetilde{\boldsymbol{P}}^{\boldsymbol{a}}]^{1/2} \tag{4.13}$$

Finally, $\overline{\boldsymbol{w}}^a$ is added to each column of $\boldsymbol{W}^a$ to form the vector $\boldsymbol{w}^{a(i)}$. $\boldsymbol{w}^{a(i)}$, as the weight vector, is used to obtain the posterior ensemble in the model space:

$$\boldsymbol{x}^{a(i)} = \overline{\boldsymbol{x}}^b + \boldsymbol{X}^b\boldsymbol{w}^{a(i)} \tag{4.14}$$

To ensure stable ensemble Kalman filter operation, localization and inflation are required to address sampling errors (Carrassi et al., 2018). To reduce residual sampling errors and make the ensemble spread (Carrassi et al., 2018), the multiplicative inflation algorithm (Whitaker & Hamill, 2012) is implemented in DasPy and applied to the soil properties and soil moisture ensemble, in which the inflation of the posterior ensemble is proportional to the reduction amount of the ensemble spread attributed to observations. Because there is only one observation site in this study, spatial localization implemented in the LETKF for spurious spatial error correlation reduction is not used. While time localization is implemented as assimilation is conducted during a specific period.

4.2.2.4 Experimental design

In this study, the model physics and structure/parameterization involving the formulation of the PTFs described in section 4.2.2.1 are assumed to be ideal (i.e., the $\boldsymbol{\epsilon}_t$ term in equation (4.1) is disregarded). Uncertainties that affect the model prediction performance are assumed due to errors in the basic soil properties and atmospheric forcing data. The SoilGrids1km dataset (Hengl et al., 2014b) is used to provide basic soil properties given its higher accuracy over other global and regional datasets on the Tibetan Plateau (Zhao et al., 2018a). To carry out investigations on whether SMAP T_B^p data assimilation improves estimates of soil properties, four experiments including a reference, an open loop and two types of data assimilation strategies are designed for evaluations and comparisons, in which the time step for CLM integration is half an hour.

The reference (unperturbed, single-ensemble) run (denoted as Ref) is driven by *in situ* atmospheric forcing over the period from 01/05/2016 to 30/10/2016 to generate soil moisture and temperature profiles and heat fluxes, which are used to evaluate the accuracy of the data assimilation experiments. A one-month spin-up run is conducted in advance to drive the land states close to equilibrium with the simulated climate. The six layers of soil data of SoilGrids1km are linearly interpolated to generate soil properties for the ten CLM layers (please refer to Table 4.1).

To accomplish the open loop and data assimilation experiments, atmospheric forcing data and basic soil properties are perturbed to generate 30 ensemble members. An ensemble size of 30 is chosen due to its effectiveness in balancing the ensemble performance and the cost of computational resources (Han et al., 2014a; Ma et al., 2012). The thickness of the soil layer whose dielectric properties contribute to soil emission in the L-band can exceed 10 cm under dry conditions (Mätzler, 2006; Zhao et al., 2021). However, as the near-surface (~2.5 cm) soil contributes the most to soil emission in L-band (Wilheit, 1978; Zheng et al., 2019), soil properties of the first layer in the CLM are perturbed and retrieved in this study. Soil properties at the other depths are obtained by using a prior depth ratio considering the typical development of soil and its profiles, namely pedogenesis (Buol et al., 2011). The soil profile at shallow depths (until 40 cm) investigated on the Tibetan Plateau (Zhao et al., 2018a) typically consists of two horizons (i.e., layers). The top layer (~2.5 - 10 cm), the A mineral horizon, especially on the humid cold climate regime, is enriched with organic matter, which results from the decomposition of the plant (roots) and animal residues. The organic matter deposited on the surface usually acts as cement and mixes with fine mineral materials to form the aggregated topsoil structure (Hillel, 2003). Some clay particles formed in this layer due to mineral weathering tend to migrate downward. The B mineral horizon (~20 - 40 cm) beneath the surface contains less organic matter and concentrated sand and silt particles (Buol et al., 2011). Due to the fact that the organic matter content decreases and sand fraction increases along the depth, the prior depth ratio is determined in terms of the exponential form adopted by the CLM to obtain fine soil layers near the soil surface. The resultant depth ratios in terms of organic matter and sand fractions at the Maqu site are [1.0, 0.98, 0.95, 0.45, 0.28, 0.18, 0.12, 0.07, 0, 0] and [1.0, 1.02, 1.06, 1.12, 1.14, 1.16, 1.18, 1.19, 1.21, 1.23] respectively for the ten CLM layers (Table 4.1). Clay fraction experiences very small changes (within 2%) with the depth (Zhao et al., 2018a). Therefore, the prior depth ratio for the clay fraction is set 1.

Given the good accuracy of SoilGrids1km's at the Maqu site, sand fraction and clay fraction of the first layer are both perturbed by adding a small uniformly distributed noise in the range of [-2%, +2%], and the perturbation range for the organic matter density is [-1.0 (kg/m^3), 1.0 (kg/m^3)]. In terms of the determined variances of these uniform distributions, a Gaussian random noise field is generated through the geoR package (https://rdrr.io/cran/geoR/man/grf.html) of statistical data analysis software R. The perturbed values of the first layer are the sum of the original ones extracted from the SoilGrids1km dataset and the defined Gaussian noise. The range of [14%~60%] is set for sand fraction, [3%~20%] for clay fraction and [1~40 kg/m^3] for organic matter density in this study. Accordingly, the perturbed soil properties (sum of the sand fraction and clay fraction and range of the organic matter density), as well as the retrieved ones at each assimilation step, are rechecked and adjusted. *In situ* atmospheric forcing inputs such as precipitation, air temperature and radiation are perturbed because they are the dominant forcing data for soil moisture and temperature and T_B^p estimates, and the perturbation parameters according to Reichle et al. (2008) are listed in Table 4.2. More detailed descriptions of atmospheric driving perturbations can be found in Han et al. (2012). Driven by the generated 30 ensembles of *in situ* atmospheric forcing and soil properties, and the initial soil moisture and temperature condition data (unperturbed) after the aforementioned spin-up period, the open loop run (denoted as OL as a prior) is conducted without data assimilation.

Table 4.2 Perturbations in the atmospheric forcing data used in this study.

Variables	Noise	Distribution	Mean	Standard deviation
Air temperature	additive	normal	0	1.0 K
Precipitation	multiplicative	lognormal	1.0	0.5
Shortwave radiation	multiplicative	normal	1.0	0.3
Longwave radiation	additive	normal	0	36.0 W/m^2

The third and fourth experiments are to retrieve soil properties using the data assimilation technique over the period from 01/05/2016 to 31/08/2016 with the following two months (from 01/09/2016 to 31/10/2016) as the verification period.

To identify whether the retrieved soil properties alone are sufficient to help improve estimates of land state variables (i.e., soil moisture and temperature) and land heat fluxes by the CLM, the third experiment updates soil properties only, i.e., without updating soil moisture (denoted as Only_Para). In particular, the 30 ensembles and the initial unperturbed soil moisture and temperature condition data (the same as for the open loop experiment) are used to run 30 realizations of the CLM to produce soil moisture and temperature at the ten soil layers. When T_B^p observation is available at time t, the simulated soil moisture and temperature at t together with the mean soil properties acquired from the soil property ensemble are fed into the TVG model to produce an ensemble of prior estimates of T_B^p at t. The mean soil properties rather than the perturbed ones for each member are used because the former is found appropriate for parameter retrieval (Han et al., 2014a). Moreover, the soil moisture and temperature of the second layer (2.79 cm in Table 4.1) are adopted as those of the first layer by the TVG model because the soil moisture content at depth of 1/10 of the wavelength (~2.5 cm for the L-band) affects the sensing depth of soil moisture (Wilheit 1978). Soil moisture and temperature of the following four layers (i.e., 36.61cm) are also fed into the TVG model for effective soil temperature calculations with the Wilheit model (please refer to section 3.3.3) since the *in situ* measured soil moisture at 40 cm remains almost constant during the study periods (Zhao et al. 2020). In the following, the soil property vector ($\boldsymbol{z}^b = [sand\ \ clay\ \ organic]^T$, similar to $\boldsymbol{x}^b$ in equation (4.1)) is updated through the LETKF algorithm by assimilating SMAP T_B^p observations, and posterior estimates of the soil properties $\boldsymbol{z}^a$ (similar to $\boldsymbol{x}^a$ in equation (4.14)) are obtained. Subsequently, the updated soil properties $\boldsymbol{z}^a$ are utilized by the CLM to run 30 realizations to produce an ensemble of prior land state variables (e.g., soil moisture and temperature) at the next time step $t + 1$. The LETKF is then recursively implemented to sequentially assimilate observations once they are available until the end of the assimilation period. These described steps are also shown in schematic form in Figure 4.1.

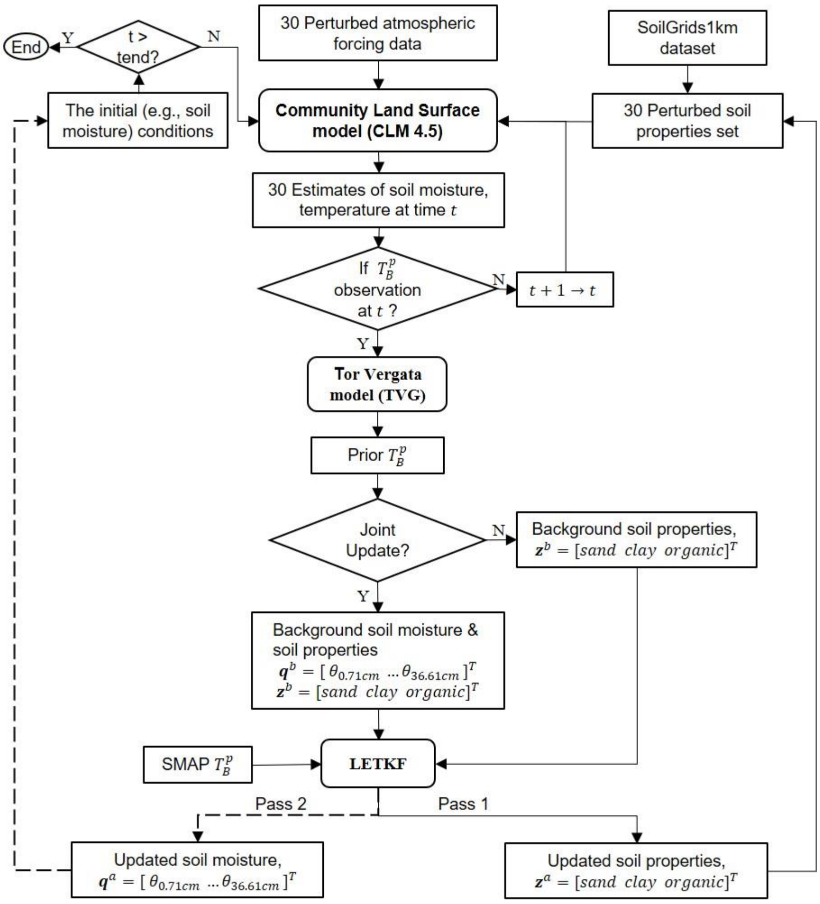


Figure 4.1 Flowchart for the retrieval of basic soil physical properties under the DasPy data assimilation framework. Rounded rectangles indicate the following three parts: the system model CLM, observation operator TVG, and LETKF data assimilation algorithm. Black arrows refer to forward flow in soil property retrieval and dashed arrows denote soil moisture update. t refers to time and $t+1$ is the next time step. 30 represents the ensemble size. (note: in the 'only_Para' experiment, only 'Pass1' is used; in the "Joint_Updt' experiment, 'Pass1' and 'Pass2' are used to update both soil properties and soil moisture.)

The fourth experiment (denote as Joint_Updt) updates both soil properties and soil moisture for the assimilation period (from 01/05/2016 to 31/08/2016) in comparison to the third experiment. To this end, the state augmentation method (Gelb, 1974) is used by DasPy, in which the aforementioned soil property vector $\boldsymbol{z}^b$ and soil moisture vector $\boldsymbol{q}^b$are adjoined as an extended prior 'state vector' ($\boldsymbol{x}^b = \binom{\boldsymbol{z}^b}{\boldsymbol{q}^b}$) and updated simultaneously. To save computational memory, a dual-pass approach (Han et al., 2014a; Yang et al., 2007) is implemented (as shown in Figure 4.1). In the parameter estimation pass (Pass 1), the same $\boldsymbol{z}^b$ described above is updated first. Then, in the following state estimate pass (Pass 2), the soil moisture at the first six layers $\boldsymbol{q}^b = [\,\theta_{0.71cm}\ \theta_{2.79cm}\ \theta_{6.23cm}\ \theta_{11.89cm}\ \theta_{21.22cm}\ \theta_{36.61cm}\,]^T$ is updated. Subsequently, both the updated soil properties and soil moisture are fed into the CLM for simulations at the next time step, and the rest of the sequence remains the same as that in the third experiment.

Finally, the retrieved soil properties obtained at the final assimilation step are compared to *in situ* measurements using the root mean square error (RMSE). Clay particles are generally plate-like and some exhibit internal surface areas, while sand and silt particles tend to have a smooth surface. As such, clay particles contribute more to the overall specific surface of soil than sand and silt particles do, and the larger the specific surface is, the higher the soil water retention (Hillel, 2003). On the other hand, clay particles typically carry a net negative electrostatic charge due to ions replacements occurring during the incomplete charge neutralization of terminal ions on lattice edges (Hillel, 2003). When hydrated, polar water molecules are attached to clay surfaces and form an electrostatic double layer. The adsorbed water does not move freely and is retained by soils under high suctions. It is also called bound water in the soil dielectric modeling, which is assumed to exhibit a different dielectric constant to that of free soil water (Mironov et al., 2004; Park et al., 2019; Wang & Schmugge, 1980). Therefore, the clay fraction is the key textural fraction affecting soil physical properties (e.g., soil moisture, dielectric constant). Moreover, sensitivity analyses also indicate that the variation of the clay fraction results in estimated soil moisture values changing greater than those due to the variations of the sand and silt fractions. The clay fraction retrieved at high accuracy, to a great degree, is assumed to reflect the success of soil property retrieval. Furthermore, the soil moisture and temperature at the different depths and the latent heat and sensible heat fluxes simulated by the three experiments are compared to the reference results, and Pearson correlation coefficient (R) and RMSE values are calculated. Better simulation results are obtained with higher R and smaller RMSE values.

4.3 Results

4.3.1 Retrieved soil properties

Figure 4.2 shows the prior (gray) and posterior (blue) distributions of the retrieved sand fraction (%), clay fraction (%) and organic matter density (kg/m^3) of the first layer obtained at the final assimilation step in the Only_Para and Joint_Updt experiments by assimilating SMAP T_B^H, with the truth based on laboratory measurements of a 0-5 cm soil layer sampled in the field. The statistics listed in Table 4.3 reveal the mean prior and posterior soil properties and their standard deviations. When only the soil properties are updated, the posterior distribution of the retrieved sand fraction remains almost the same to that of the prior (the overlapped gray and light blue lines in Figure 4.2a), and both of their mean values are larger than the truth (the solid black vertical line in Figure 4.2) within 8% (Table 4.3). The posterior distribution of the retrieved organic matter content exhibits a slight shift toward the truth (Figure 4.2a). The posterior distribution of the retrieved clay fraction is obviously shifted and narrowed (a standard deviation of 1.21 vs. 0.97, as indicated in Table 4.3) to approach the truth (Figure 4.2a).

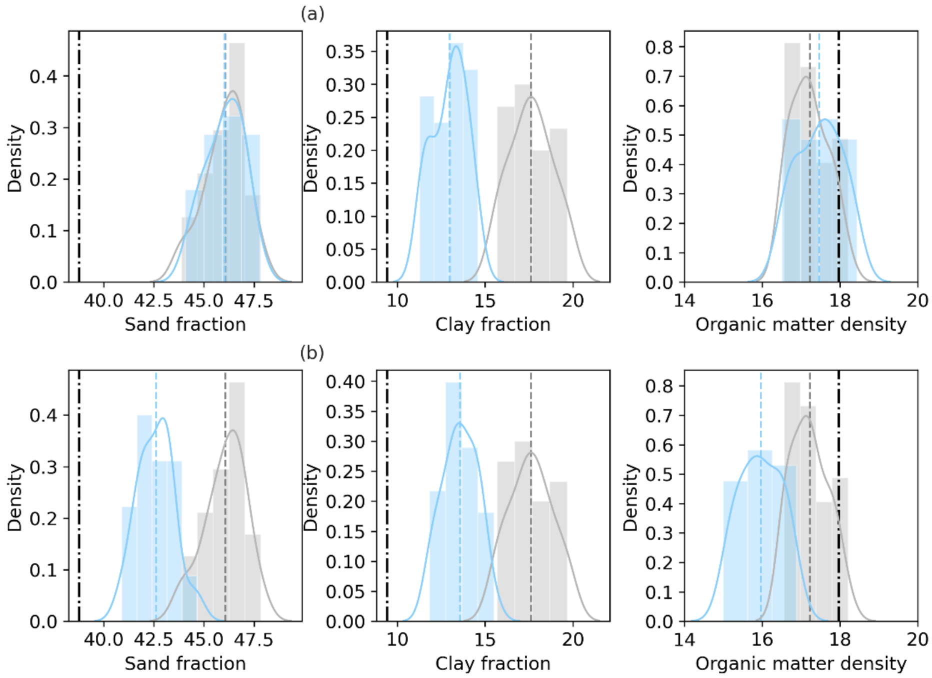


Figure 4.2 Prior and posterior distributions of the sand fraction (%), clay fraction (%) and organic matter density (kg/m^3) of the first layer, with the truth based on laboratory measurements of the 0-5 cm soil layer sampled in the field. Gray indicates the prior and light blue indicates the posterior, and black dash-dotted line indicates the laboratory measurements. (a) Only_Para and (b) Joint_Updt experiments by assimilating SMAP T_B^H with an ensemble of size 30.

Table 4.3 Two data assimilation results and their calculated standard deviations (_std) for the retrieved soil properties of the first layer when SMAP T_B^p is assimilated with an ensemble of size 30 in the experiments.

Soil property	True	$\boldsymbol{z}^b$	$\boldsymbol{z}^b$_std	T_B^H				T_B^V			
				Only_Para		Joint_Updt		Only_Para		Joint_Updt	
				$\boldsymbol{z}^a$	$\boldsymbol{z}^a$_std	$\boldsymbol{z}^a$	$\boldsymbol{z}^a$_std	$\boldsymbol{z}^a$	$\boldsymbol{z}^a$_std	$\boldsymbol{z}^a$	$\boldsymbol{z}^a$_std
Sand fraction (%)	38.79	46.06	1.01	46.04	0.95	42.6	0.88	43.34	1.11	43.33	1.1

Clay fraction (%)	9.42	17.62	1.21	12.96	0.97	13.56	0.97	15.27	1.14	15.05	1.1
Organic matter density (kg/m^3)	17.97	17.22	0.47	17.47	0.57	15.96	0.54	17.19	0.53	16.88	0.57

When both the soil properties and soil moisture are updated (refer to Figure 4.2b), the posterior distributions of both the retrieved sand fraction and clay fraction shift toward the truth. The improvement of sand fraction retrieval is observed in the Joint_Updt experiment but not in Only_Para (as shown in Figure 4.2a). While the distribution of the retrieved organic matter density deviates from the prior and the truth (shown in Figure 4.2b, a standard deviation of 0.47 vs. 0.54, as listed in Table 4.3), which goes against the Bayesian theorem stating that the posterior normally falls between the prior and the truth. Table 4.4 shows that the Joint_Updt experiment greatly reduces the RMSE for the sand fraction (46.6%) over the Only_Para does (0.4%). The Only_Para experiment reduces the RMSE for the organic matter density (14.6%), and the Joint_Updt experiment increases the RMSE (2.08 vs. 0.89), indicating a high negative efficiency in terms of organic matter density retrieval. Nevertheless, both data assimilation experiments by assimilating SMAP T_B^H result in the largest reduction in RMSE for the clay fraction (> 48%).

Table 4.4 RMSE values of the retrieved soil properties of the first layer in the two data assimilation experiments by assimilating SMAP T_B^p with an ensemble size of 30.

Soil property	$\boldsymbol{z}^b$ _ RMSE	T_B^H				T_B^V			
		Only_Para		Joint_Updt		Only_Para		Joint_Updt	
		$\boldsymbol{z}^a$ _ RMSE	Reduction	$\boldsymbol{z}^a$ _ RMSE	Reduction	$\boldsymbol{z}^a$ _ RMSE	Reduction	$\boldsymbol{z}^a$ _ RMSE	Reduction
Sand fraction (%)	7.34	7.31	0.40%	3.92	46.60%	4.68	36.20%	4.67	36.40%
Clay fraction (%)	8.29	3.67	55.70%	4.25	48.70%	5.96	28.10%	5.74	30.70%

Organic matter density (kg/m^3)	0.89	0.76	14.60%	2.08	-133%	0.94	-5.60%	1.23	-38%

Figure 4.3 shows that the posterior distributions of both the retrieved sand fraction and clay fraction shift toward the truth when the two experiments assimilate SMAP T_B^V. Although neither experiment achieves a positive efficiency regarding organic matter density retrieval (negative values in Table 4.4), they reduce the RMSE by ~36% for the sand fraction and RMSE by ~28% for the clay fraction. In the Only_Para experiment, the use of T_B^H is found to be more sensitive to retrieval of the clay fraction and organic matter than the application of T_B^V assimilation, which shows sensitivity to sand fraction retrieval. In contrast, the Joint_Updt experiment can retrieve both the sand and clay fractions when assimilating either T_B^H or T_B^V. Nevertheless, all reductions in RMSE values for the clay fraction indicate the improvement of soil property estimates by assimilating SMAP T_B^p, as the clay fraction is the key textural fraction over the other fractions as reasoned in section 4.2.2.4. Furthermore, utilizing the prior depth ratio (section 4.2.2.4), the posterior distributions of the retrieved clay fraction and organic matter density of the third layer (i.e., 11.89 cm) (as shown in Figure A3.1 and Tables A3.1-A3.2 in Appendix C as an example) also shift toward the truth. This indicates that by updating the soil properties of the first layer, the descriptions of soil properties at the other depths can be enhanced through the prior depth ratio.

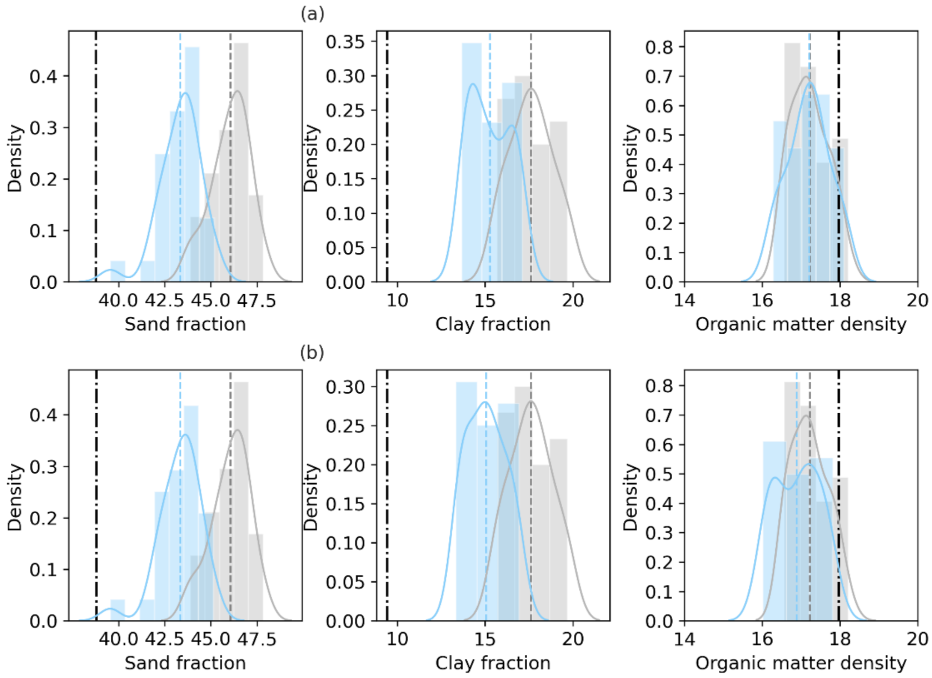


Figure 4.3 Same as Figure 4.2 but based on the experiments by assimilating SMAP T_B^V.

4.3.2 Comparison of the land state variables

Since the clay fraction retrieved by the experiments assimilating T_B^H indicates more improvements than those assimilating T_B^V (please refer to section 4.3.1), evaluation analysis of the soil states (soil moisture and temperature) and land surface fluxes (LE and H) in the OL, Only_Para and Joint_Updt experiments was only implemented considering the assimilation of T_B^H. Figure 4.4b and Figure 4.5b show that the assimilation of SMAP T_B^H slightly improves the soil moisture estimate, and the results of Joint_Updt (with higher R and smaller RMSE values) are slightly better than those obtained with Only_Para. Compared to the soil moisture of the first layer, the estimated soil moisture of the deep layers near the surface (e.g., 11.89 and 21.22 cm) has lower R and larger RMSE values (Figures 4.4b and 4.5b). In CLM, soil moisture and temperature are connected through heat fluxes. The Joint_Updt experiment with the slight improvement in soil moisture estimate yields in better LE estimate, correspondingly, better H and soil heat flux estimates in terms of the energy balance (higher R and lower RMSE values in Figures 4.4a and 4.5a, respectively). Accordingly, the soil temperature estimated by the Joint_Updt experiment also yields a smaller RMSE (Figure 4.5c).

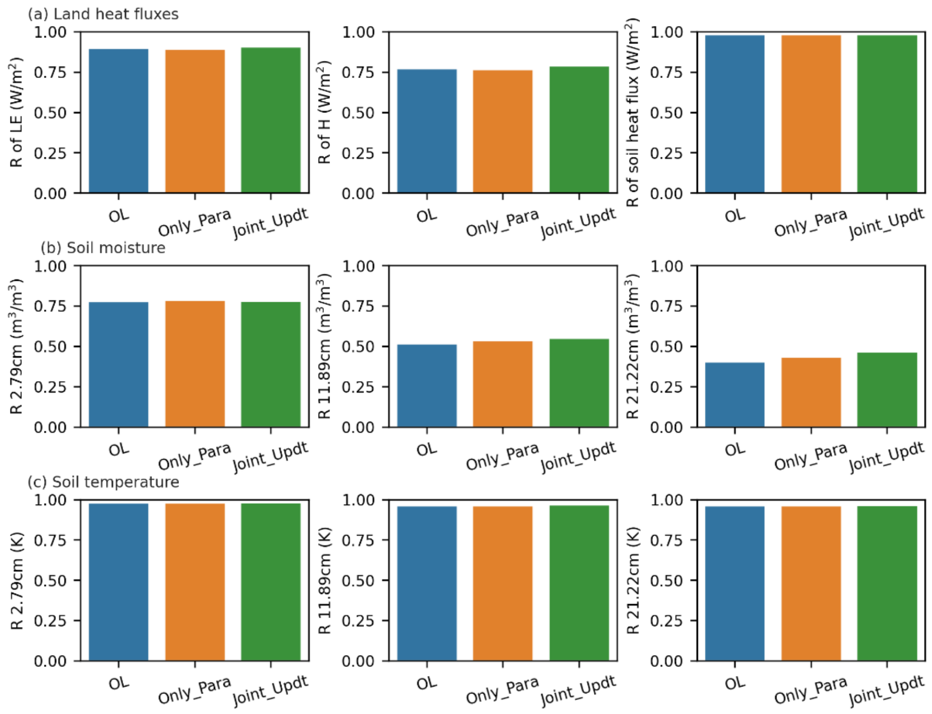


Figure 4.4 R values of the land state and flux variables in the open loop (OL) experiment, the data assimilation experiment with only soil properties updated (Only_Para), and the data assimilation experiment with both the soil properties and soil moisture estimates (Joint_Updt) by assimilating SMAP T_B^H over the assimilation period (from 01/05/2016 to 31/08/2016). (a) is for the land latent heat flux (LE), sensible heat flux (H) and soil heat fluxes, and (b) and (c) are for the soil moisture and soil temperature, respectively, at the different depths.

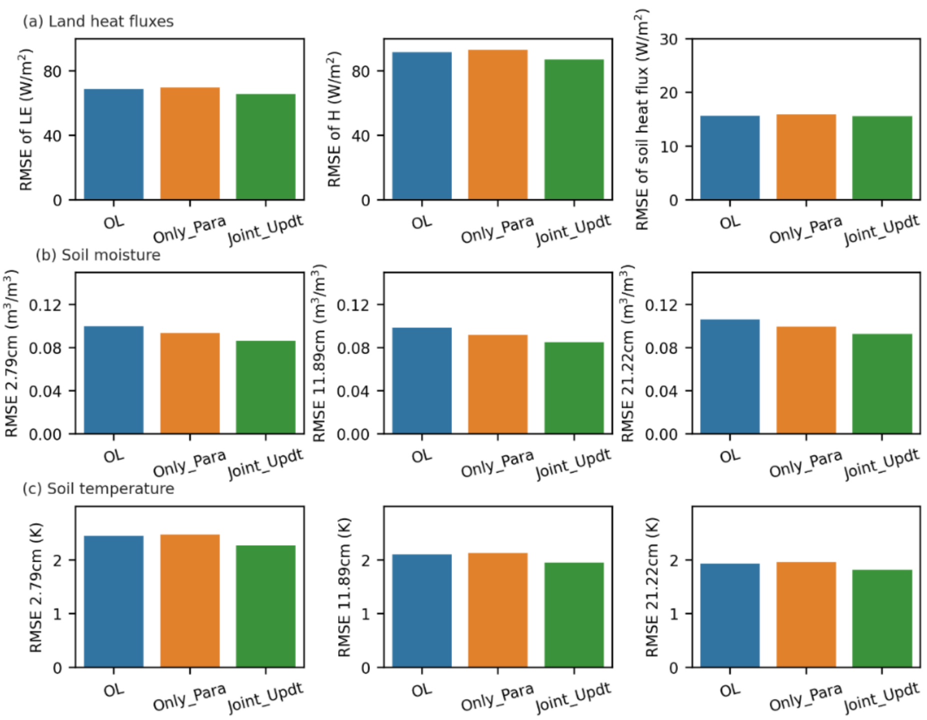


Figure 4.5 Same as Figure 4.4 but for the RMSE values.

During the verification period, the soil moisture biases are again slightly reduced in the Only_Para and Joint_Updt experiments (Figures 4.6b, 4.7b). The Only_Para experiment yields a slightly smaller RMSE than that of the Joint_Updt experiment. No improvements are found in the land heat flux (the same R and RMSE values in Figures 4.6a and 4.7a) and soil temperature (Figures 4.6c and 4.7c) estimates over the verification period.

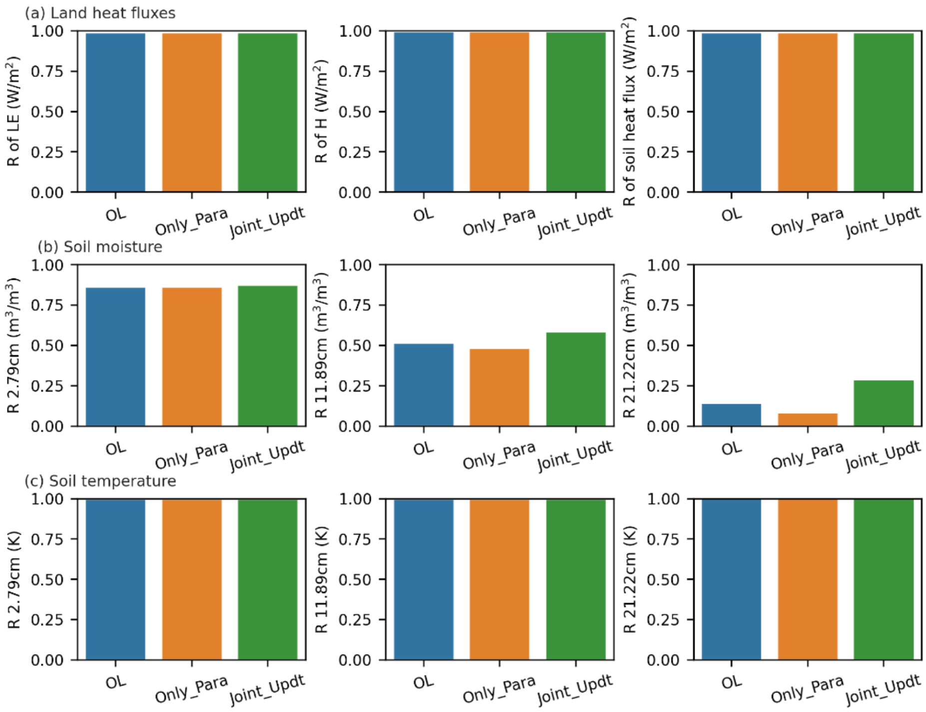


Figure 4.6 Same as Figure 4.4 but over the verification period (from 01/09/2016 to 31/10/2016).

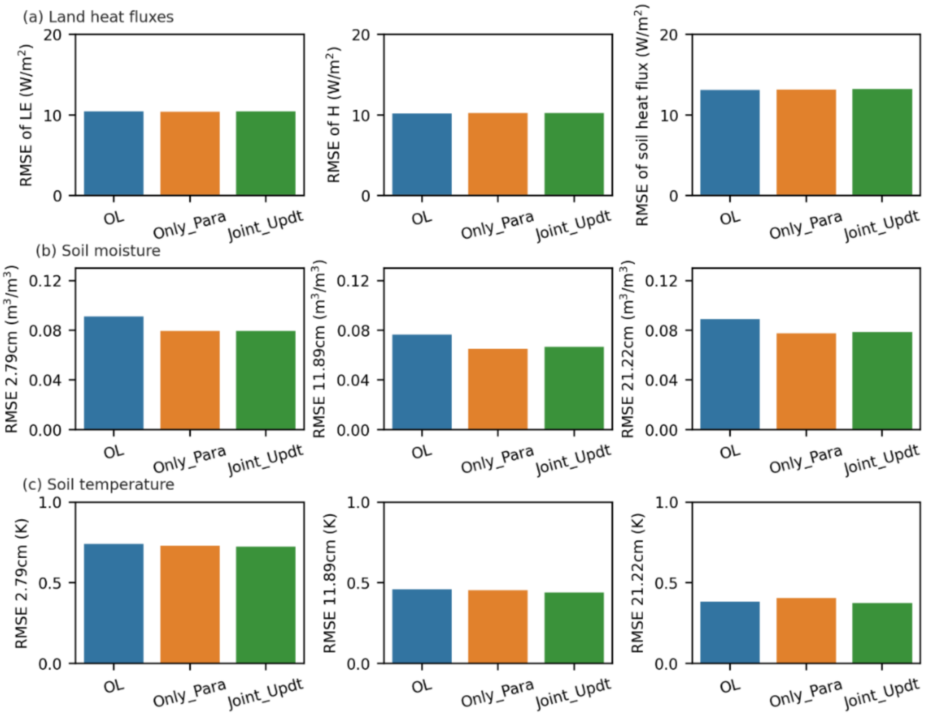


Figure 4.7 Same as Figure 4.5 but over the verification period (from 01/09/2016 to 31/10/2016).

4.4 Discussion

The results in section 4.3 indicate the possibility of correcting the SoilGrids1km soil property product by assimilating SMAP T_B^p at the Maqu site. However, only slight improvements in the simulations of the land surface state and flux variables are achieved, with large RMSE values obtained for the land heat fluxes (> 40 W/m^2) and soil moisture (> 0.04 m^3/m^3, the required RMSE value of satellite soil moisture accuracy (Reichle et al., 2017)) over the assimilation period. This indicates that reducing the errors in the basic soil properties and atmospheric forcing data may not be sufficient to account for the unsatisfactory model performance. To reveal whether the model physics and structure are adequate, comparisons are made between the four experimental results and (available) *in situ* observations over the assimilation period for discussion purposes.

4.4.1 Estimates of the near-surface soil moisture

Figure 4.8 shows that the soil moisture at 2.79 cm simulated in the reference run is close to the *in situ* observations (R value of 0.94 in Figure 4.9b and RMSE < 0.04 m^3/m^3 in Figure 4.10b) despite overestimations when the soil is wet. The soil moisture at 2.79 cm simulated in the OL run shown in Figure 4.8 is much higher than the observations and yields a large RMSE (approximately 0.08 m^3/m^3 in Figure 4.10b). After data assimilation, the soil moisture at 2.79 cm estimated in the Only_Para experiment is still higher than *in situ* measurements but exhibits slight improvements over the OL run (Figure 4.8, RMSE of 0.07 m^3/m^3 vs. 0.08 m^3/m^3 in Figure 4.10b). This implies that the soil property with a fine accuracy is not a sensitive factor impacting the accuracy of soil moisture estimates by CLM. The soil moisture at 2.79 cm estimated in the Joint_Updt experiment is closer to the observations than that estimated in the OL and Only_Para experiments (RMSE of 0.05 m^3/m^3 vs. 0.08 m^3/m^3 vs. 0.07 m^3/m^3 in Figure 4.10b), especially when the soil undergoes the drying process (e.g., soil moisture < 0.28 m^3/m^3), but soil moisture is overestimated in wet soil (Figure 4.8). This signifies that the model structure relating to surface soil moisture estimates may contain uncertainties since its performance is not consistent between dry and wet conditions.

Figure 4.8 Soil moisture time series from 07/08/2016 to 08/31/2016 of the second (2.79 cm) and fourth (11.89) layers in the reference (Ref) experiment, open loop (OL) experiment, experiment with only soil properties updated (Only_Para) and experiment with both the soil properties and soil moisture updated (Joint_Updt) by assimilating SMAP T_B^H.

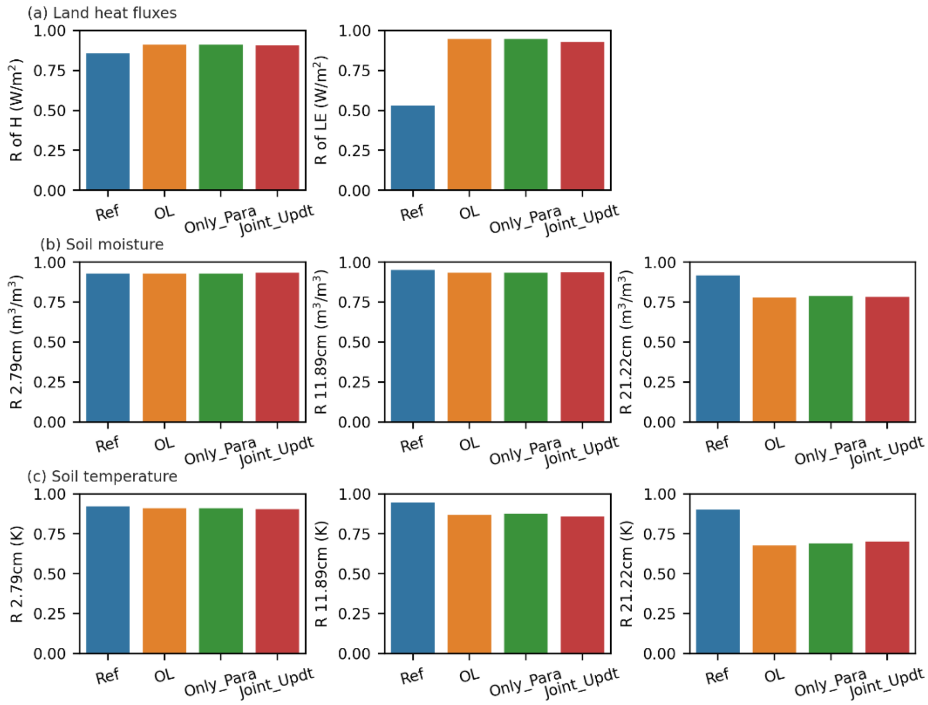


Figure 4.9 R values of the land state and flux variables for the reference (Ref) experiment, open loop (OL) experiment, experiment with only soil properties updated (Only_Para), and experiment with both the soil properties and soil moisture updated (Joint_Updt) by assimilating SMAP T_B^H over the assimilation period (from 08/08/2016 to 31/08/2016), in comparison to the in situ observations. (a) is for the land latent heat flux (LE) and sensible heat flux (H), and (b) and (c) are for the soil moisture and soil temperature, respectively, at the different depths.

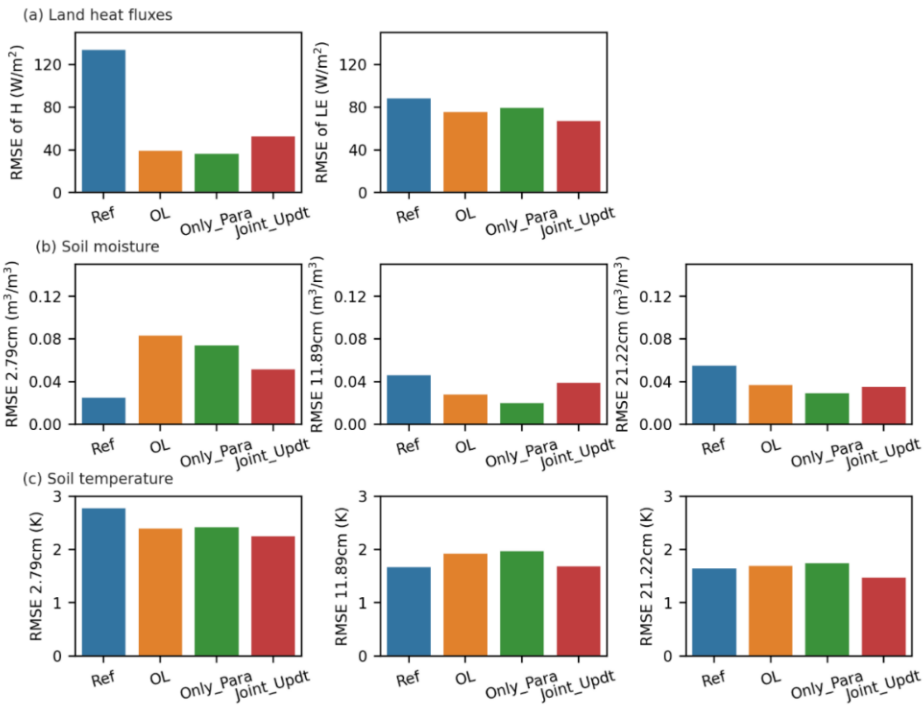


Figure 4.10 Same as Figure 4.9 but for the RMSE values.

In the CLM, soil water flow in the unsaturated zone is modeled by Darcy's law (refer to equation (A2.1) in Appendix C). The soil liquid water in the node layer z_i depends on the hydraulic conductivity of the interface of two adjacent layers $z_{h,i}$ and the soil matric potential gradient. The soil node layer depth z_i (refer to Table 4.1) is known, and the soil matric potential at z_i can be calculated from the soil moisture content at z_i through the Campbell (1974) soil water retention function parameterized by three hydraulic parameters θ_s, φ_s and B (refer to equation (A2.6) in Appendix C). As such, the change in volumetric soil liquid water over the z_{i-1} and z_i layers is mainly determined by the hydraulic conductivity in the $z_{h,i}$. The hydraulic conductivity at $z_{h,i}$ is parameterized as a function of k_s in the $z_{h,i}$, θ_s and the soil moisture content in the z_{i-1} and z_i layers, and the hydraulic parameter B at z_i (refer to equation (A2.3) in Appendix

C). As the (prior and posterior) soil properties are close to the measurements (shown in Figure 4.2 and Figure 4.3), our initial guess of the uncertainties focuses on the soil hydraulic parameters calculated through the Cosby PTFs with fixed structures.

Figure 4.11 and Figure 4.12 show that the prior and posterior values, respectively, of the soil hydraulic parameters deviate from the *in situ* measurements. Both the estimated θ_s and K_s values at 2.79 cm are lower than the measurements, and both φ_s and B are overestimated. The mutual corroboration of these four parameters may lead to the soil moisture at 2.79 cm being slightly overestimated in the reference run. When only the soil properties are updated in the Only_Para experiment (Figure 4.11a), B decreases due to the reduced retrieved clay fraction, and the other three parameters retain the same values as the posteriors of the sand fraction and organic matter density with no appreciable differences from the prior (shown in Figure 4.2a). This implies that the change in clay fraction mainly affects B. In contrast, the Joint_Updt experiment assimilating T_B^H (shown in Figure 4.2b) and both assimilation experiments assimilating T_B^V (shown in Figure 4.3) estimate reduced posterior sand fractions, resulting in K_s and φ_s decreasing (Figures 4.11b, 4.12). However, due to the fixed PTF structures, the slight changes in the values of the soil hydraulic properties do not correct the biased soil moisture estimates due to the uncertainties in the precipitation forcing data, which in reality is always the case, especially in large-scale applications (Koster et al., 2018).

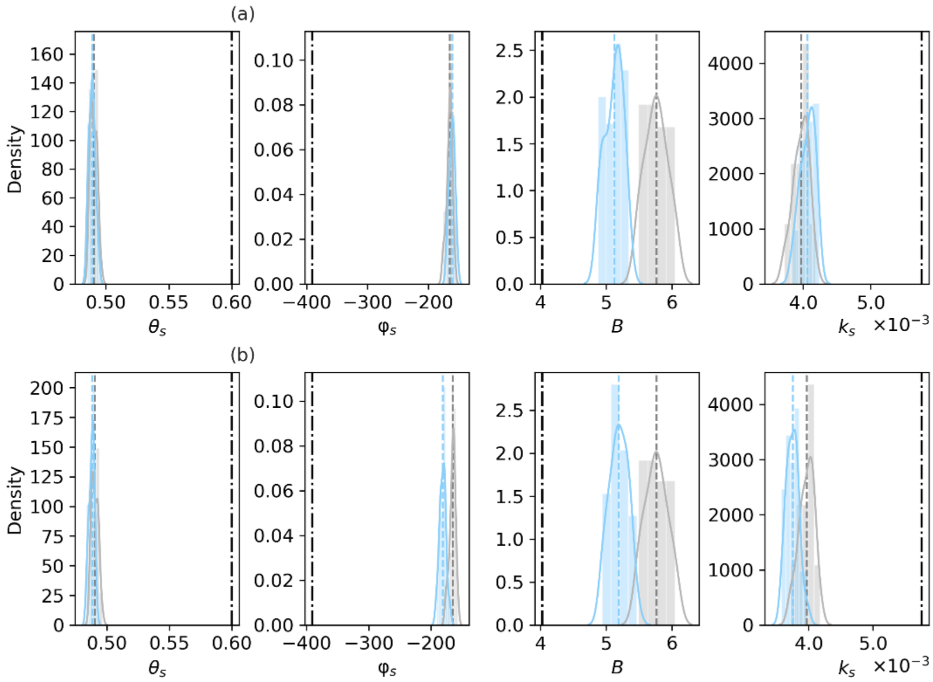


Figure 4.11 Prior and posterior distributions of the soil hydraulic parameters at 2.79 cm, with the truth based on laboratory measurements of the 0-5 cm soil layer sampled in the field. Gray shows the prior, and light blue indicates the posterior, and the black dash-dotted line indicates the laboratory measurements. (a) Only_Para and (b) Joint_Updt experiments by assimilating SMAP T_B^H with an ensemble size of 30. The hydraulic parameters are the saturated soil moisture content θ_s (m^3/m^3), saturated matric potential φ_s (mm), pore size distribution index B (dimensionless) and saturated hydraulic conductivity K_s (mm/s).

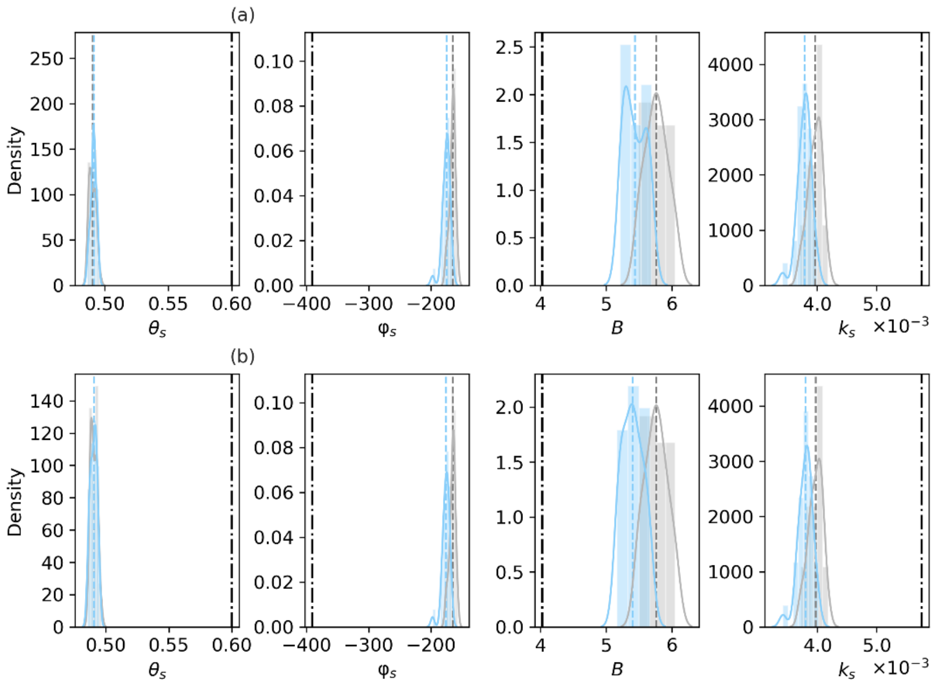


Figure 4.12 Same as Figure 4.11 but based on the experiments assimilating SMAP T_B^V with an ensemble size of 30.

The Cosby PTFs are derived by fitting to results retrieved from laboratory experiments on 1448 small soil samples (cm dimensions) collected in 23 states in the United States (for further details of the soil samples and sampling methods, please refer to Rawls (1976) and Holtan (1968)). Hence, the estimated soil hydraulic parameters are also based on a small-scale (cm), which may differ from the effective parameters at the field scale (~m). Cooper et al. (2020) optimized the constants in the underlying PTFs to obtain soil hydraulic parameters representing the field scale by assimilating daily-averaged COSMOS-UK soil moisture data, and showed that the performances of LSMs in soil moisture simulations were improved with the optimized PTFs. In this study, *in situ* atmospheric forcing data are measured at the field scale while the assimilated SMAP T_B^p data exhibit a large spatial resolution of 36 km. The scale problem is beyond the scope of this study. However, obtaining physically consistent parameter sets is an alternative way to optimize PTF structures and obtain soil hydraulic parameters with a matched scale to that of the inputs (i.e., atmospheric forcing and T_B^p). The use of optimized PTFs in LSMs may help improve temporally continuous soil moisture estimates at large scales (e.g., hydrometeorology at ~10 km and hydroclimatology at ~ 40 km scale), as reflected by the improved performance obtained in Cooper et al. (2020).

The soil moisture at 2.79 cm estimated in all experiments is overestimated (shown in Figure 4.8) when the soil is undergoing the wetting process due to rainfall events. Our second presumption is that the Campbell (1974) function describing soil water retention may also contain uncertainties. This function leads to an air-entry pressure in the soil water retention curve above which the soil is assumed to be saturated. As such, there exists a discontinuity in the slope of the curve at the air-entry value, and the transition zone near saturation is ignored for natural fine-textured soils (Clapp & Hornberger, 1978), resulting in moisture overestimations when soils occur near saturation. The van Genuchten (1980) soil water retention function parameterized with five independent hydraulic parameters was demonstrated to suitably simulate the soil water retention curve for soils near saturation (van Genuchten & Nielsen, 1985). Compared to the Campbell (1974) function, the van Genuchten (1980) function is more widely used in vadose zone research due to its flexibility in describing a large range of soil water retention curve shapes, in which the hydraulic parameters can be estimated by PTFs (Van Looy et al., 2017; Vereecken et al., 2010). LSMs such as H-TESSEL (Hydrology-Tiled ECMWF Scheme for Surface Exchanges over Land) used by the ECMWF (European Center for Medium-Range Weather Forecasts) (Balsamo et al., 2009) employ the van Genuchten (1980) function to replace the previously used Campbell (1974) function. In the CLM used in this study, the update of the soil hydraulic function can be implemented in future studies, and comparisons will be made to evaluate the uncertainty in soil property retrieval by using different soil hydraulic functions.

4.4.2 Estimates of the soil moisture of the deeper layers near the surface

The soil moisture at greater depths (i.e., 11.89 and 21.22 cm) simulated in the reference run presents underestimations (shown in Figure 4.8, and a RMSE value of 0.05 m^3/m^3 in Figure 4.10b) over the observations. While the soil moisture at depths of 11.89 and 21.22 cm simulated in the OL run is close to the observations despite slight overestimations (shown in Figure 4.8), it yields a smaller RMSE (< 0.04 m^3/m^3 in Figure 4.10b). The soil moisture at depths of 11.89 and 21.22 cm estimated in the Only_Para experiment is close to that estimated in the OL run (shown in Figure 4.8). This is expected because the updated soil properties of the first layer experience only a small change over the prior soil properties (shown in Figure 4.2), and this results in slight changes in the updated soil properties of the deeper layers through the prior depth ratio. Combined with a fixed PTF structure, the estimated soil moisture does not differ from that estimated in the OL run. However, in the Joint_Updt experiment, the soil moisture at depths of 11.89 and 21.22 cm is jointly updated by using the calculated surface increment (due to the soil moisture sensing depth of the SMAP radiometer in the near-surface zone) through the LETKF algorithm. The updated soil moisture values are close to those obtained by the OL and Only_Para experiments except when the surface soil becomes dry (e.g., 18/08/2016 to 24/08/2016 in Figure 4.8), but yield consistencies with the observations when the soil occures under wet conditions (e.g., 25/08/2016 to 31/08/2016 in Figure 4.8).

To a certain degree, this can reflect the improvement in surface information propagating downward to deeper layers through assimilation. However, the improvement may be impeded by deficiencies in the modeled subsurface physical process, since the soil moisture in the deeper layers estimated by the reference experiment tends to 'stick' at a water content of 0.1 during the dry period, and soil moisture assimilation is not consistent between dry and wet conditions. For instance, the (saturated) soil moisture at the interface of two adjacent layers $z_{h,i}$, calculated as half of the sum of the soil moisture in the z_{i-1} and z_i layers, does not rest on a theoretical basis justifying this average approach (Yang et al., 2009b). Moreover, root water uptake in the CLM is estimated by using the water stress function, which is calculated by the product of the root fraction and the plant wilting factor parameterized by using the soil matric potential (Oleson et al., 2013). However, the water potential in the root collar that drives the water flux from a given soil layer z_i is not considered, nor is the absolute root biomass, while the relative root fraction is used instead (Kennedy et al., 2019). This may result in biases in the soil water availability and thus in estimates of the soil moisture in the root zone (deep soil layers). The plant hydraulic stress (PHS) (Kennedy et al., 2019), i.e., a new plant water stress parameterization based on the hydraulic theory considered in the recently released CLM v5, has been developed. The PHS implementing a physical model of the vegetation water potential is demonstrated to improve estimates of the root water uptake and thereby the vertical distribution of soil water (Kennedy et al., 2019). With the consideration of the PHS and a physically-based root growth model, the soil-plant-atmosphere continuum (SPAC) system is expected to be modeled consistently. As such, the coupling strength of land state information (i.e., soil moisture and temperature) will coherently depend on the root zone (deeper soil layers). Additionally, a highly discretized profile near the surface is used in the CLM, which may lead to a weak coupling strength from the surface to deeper layers in the CLM, as claimed by Kumar et al. (2009), thus constraining the efficiency of the abovementioned improvements through data assimilation. Changing the layering structure (especially near the surface) of the CLM should be tested in future research to verify if the coupling strength can be enhanced, and thereby the soil moisture and even basic soil properties of the deeper layers updated through assimilation.

4.4.3 Estimates of the land heat fluxes and soil temperature

Affected by underestimations of the soil moisture in the deeper layers near the surface (shown in Figure 4.8), soil evaporation is confined, and the resulting LE simulated by the reference run is lower than the observations (Figure 4.13) and yields a large RMSE (over 80 W/m^2 in Figure 4.10a). The reference run shows small RMSE values (< 3 K in Figure 4.10c) for the estimates of the soil temperature at the different depths (Figure 4.14). Therefore, the ground flux that enters the soil column does not deviate as much as the simulated LE. In terms of energy balance equations and simulated net radiation with limited accuracy (RMSE of 82 W/m^2 in this case), H simulated by the reference run is larger than the observations (Figure 4.13) and exhibits a large RMSE similar to that of the simulated LE (shown in Figure 4.10a). The LE and H simulated by the OL run are closer to the observations (Figure 4.13, with smaller RMSE values in Figure 4.10a) than is the reference run, and this occurs due to the unexpected match between the soil moisture at the different depths simulated by the OL run and the observations (shown in Figure 4.8). The Only_Para experiment simulates better H than do the other experiments (smaller RMSE values in Figure 4.10a). In contrast, the Joint_Updt experiment simulates better LE (smaller RMSE values in Figure 4.10a). The performance of these two data assimilation experiments in LE estimation differs when the soil occurs under dry conditions (e.g., the *in situ* soil moisture is close to 0.1 m^3/m^3 from 18/08/2016 to 23/08/2016 in Figure 4.13). This may reflect the deficiency in dry soil water simulations by the CLM, in which vapor transport is not considered while deemed predominant under dry conditions (Zeng et al., 2009a; Zeng et al., 2009b).

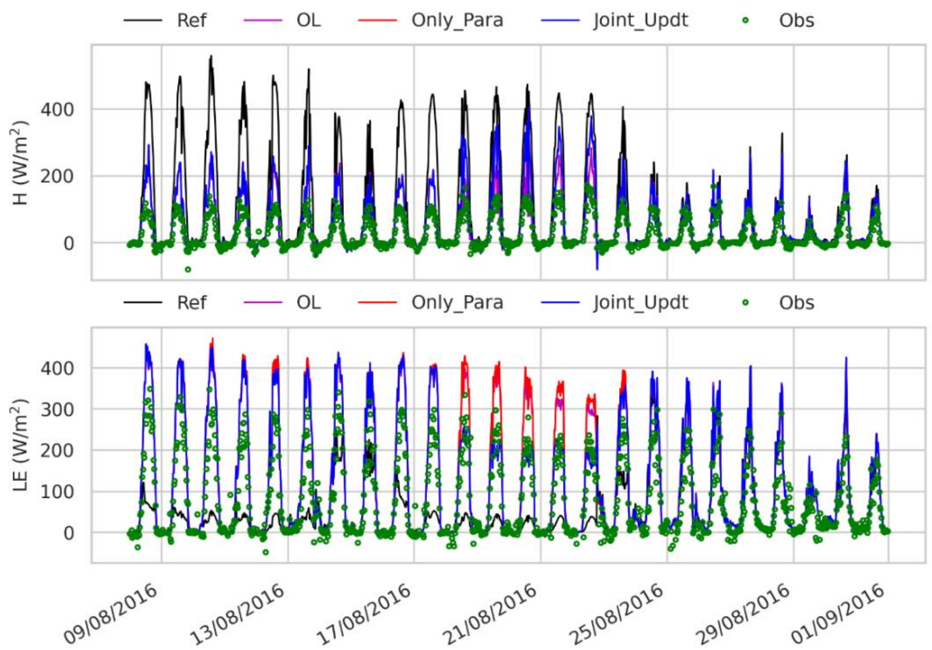


Figure 4.13 Land sensible heat flux (H) and latent heat flux (LE) time series from 07/08/2016 to 08/31/2016 in the reference (Ref) experiment, open loop (OL) experiment, the scenario with only soil properties updated (Only_Para) and the scenario with both the soil properties and soil moisture estimates (Joint_Updt) updated by assimilating SMAP T_B^H.

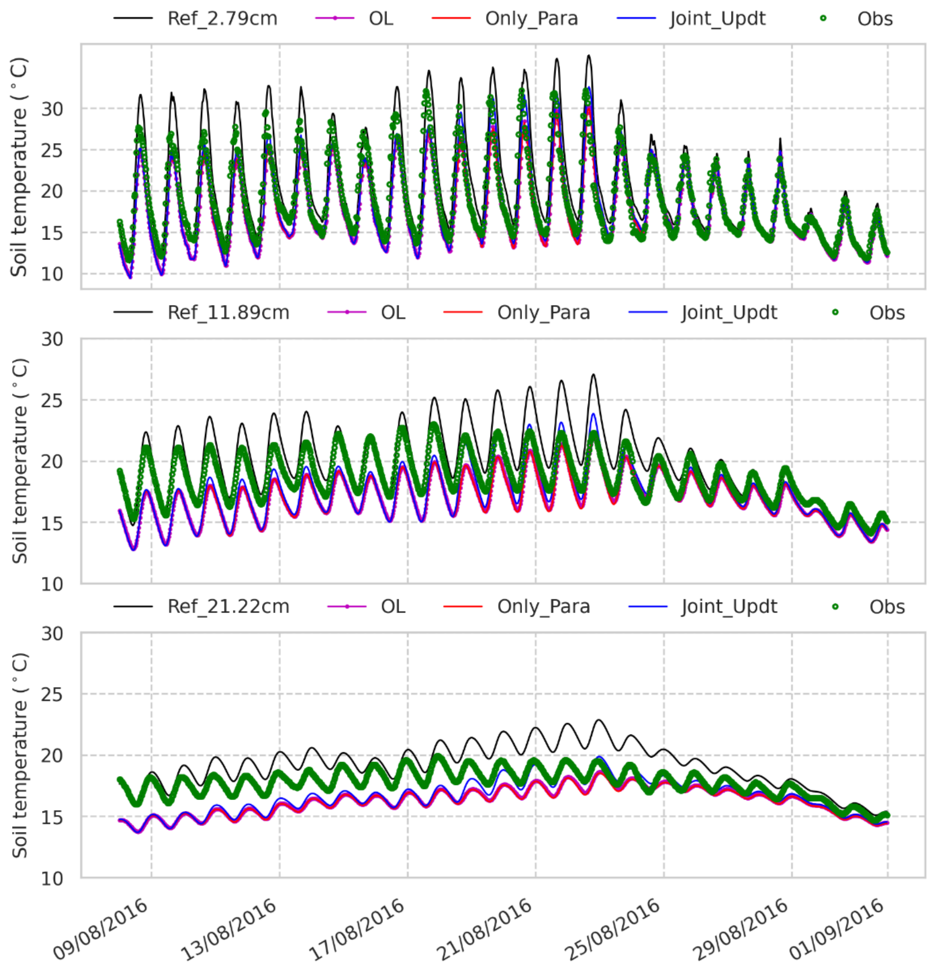


Figure 4.14 Same as Figure 4.13 but for the soil temperature time series.

In contrast, the large discrepancies in H and LE (RMSE over 70 W/m^2 in Figure 4.10a) in these four experiments might relate to the inadequate parameterizations used in the heat flux simulations by the CLM. Previous studies (Yang et al., 2009a) have shown that an excess resistance must be introduced to estimate H from the ground and air temperature difference. Neglecting this resistance was found to lead to the overestimation of H during the daytime (Yang et al., 2008). The current parameterization of the heat transfer resistance in CLM does not consider this excess resistance, and H simulated by the reference experiment as shown in Figure 4.13, is clearly overestimated during the day. The excess resistance parameterized by Yang et al. (2009a) can be added in future studies to improve the H estimate.

The soil surface resistance to evaporation with respect to soil moisture is found to be very sensitive and results in the simulated LE changing drastically, especially for dry soil surfaces (Yang et al., 2009a). The CLM uses a ‘soil beta’ parameterization to represent the effect of the soil resistance on soil evaporation (Oleson et al., 2013). This empirical function depends on the CLM top layer soil moisture and the field capacity of the top layer, and the latter is parameterized by θ_s, K_s and B of the topsoil layer. As θ_s and K_s are underestimated and B is overestimated (shown in Figures. 8-9), the field capacity is underestimated. This leads to an increase in soil resistance and consequently a decrease in soil evaporation. As shown in Figure 4.13, the LE simulated by the reference run during the dry period (e.g., from 09/08/2016 to 15/08/2016) is lower than the observations during the daytime and quickly reaches a peak at the diurnal scale. When the soil surface becomes wet, net radiation dominates soil evaporation rather than soil moisture, and the soil resistance is negligible (Yang et al., 2009a). Figures 4.13 and 4.14 show the consistency between the simulated LE, H and soil temperature with the observations when the soil is wet (e.g., 25/08/2016 to 31/08/2016). Based on the above analyses, a better estimate of LE can be anticipated when the estimates of the soil hydraulic properties and soil moisture are improved. Additionally, the incorporation of the PHS parameterization (refer to section 4.3) will also help improve LE estimates.

4.4.4 Estimates of T_B^p

We compare T_B^p estimated by the assimilation experiments to the SMAP and ELBARA-III observations. Figures 4.15a and 4.15b show the underestimations in T_B^p by the experiments compared to the SMAP observations. Moreover, to correct and obtain superior estimates of T_B^p through assimilation, the Only_Para and Joint_Updt experiments tend to yield lower values of soil moisture (shown in Figure 4.8). Correspondingly, when the uncertainties of the soil properties are assumed to account for overestimations in soil moisture, the clay fraction is preferentially reduced through assimilation as shown in Figures 4.2 and 4.3.

However, the values of the posterior (updated) T_B^H through assimilation (Figure 4.15a) are observed consistently lower (~10-30 K) than the SMAP observed values, especially during the soil drying period (e.g., approximately 19/08/2016 in Figure 4.8). The T_B^p estimate for the soil part is mainly affected by the soil temperature, which determines the soil effective temperature (T_{eff}), and the soil moisture, which determines the surface dielectric roughness (please refer to section 3.3.2) and soil emissivity. The aforementioned uncertainties in the soil hydraulic property estimates in this study (refer to section 4.4.1) result in biased soil moisture estimates, and the resulting overestimated soil moisture (shown in Figure 4.8) leads to underestimations of the surface dielectric roughness and thereby overestimated effective dielectric constant and associated underestimated emissivity values. In contrast, the estimates of T_{eff} obtained through assimilation show coincidences with the *in situ* soil temperature at 2.5 cm on 08/2016 (Figure 4.15c). As *in situ* soil temperature observations are not available during the period from 05/2016 to 07/2017, the ground surface temperature (TG in Figure 4.15c) derived from *in situ* longwave radiation measurements is used as a surrogate in comparison to assess the uncertainty in the T_{eff} estimates. Figure 4.15c shows that the difference between T_{eff} and TG is confined within 5 K except for the sharply decreased TG values before rain events (e.g., the sudden drop in SMAP T_B^p in Figures 4.15a,b). Combined with smaller RMSE values (< 3 K) for the soil temperature estimates through assimilation (shown in Figures 4.5c, 4.7c, 4.9, 4.14), it is observed that the T_{eff} estimates contain acceptable uncertainties. The uncertainty in the soil moisture estimates contributes the most to the large gap between the posterior T_B^H values and SMAP observations in this study. Furthermore, vegetation during the assimilation period experiences phenological changes, and the uncertainties in vegetation modeling in the TVG model also affect the T_B^p estimates, which is beyond the scope of this study.

In contrast to the T_B^H estimates, a better match is shown between the posterior T_B^V values and SMAP observations (Figure 4.15b). Because the lateral topsoil structures are significantly smaller than the observation wavelength (λ_0 = 21 cm, in the L-band), the heterogeneities in the topsoil structures and within the soil volume (e.g., composition and soil moisture) may impose greater impacts on the brightness temperature at H polarization (T_B^H). T_B^V variations are less affected by this kind of surface roughness change. Therefore, this may be the reason that only T_B^V is used in existing SMAP soil moisture retrieval algorithms (O'Neill et al., 2020). The results presented in section 4.3.1 indicate that T_B^H is more applicable to the assimilation of clay fraction and organic matter retrievals than T_B^V. This may be related to the plate-like structure of clay and may account for the general claim (based on regression analysis using observations) that T_B^H is sensible to soil moisture changes (Njoku et al., 2002).

Last but not the least, the SMAP T_B^p at the Maqu site is consistent with the ELBARA-III observations (Figures 4.15a,b), indicating good-quality SMAP T_B^p data. Due to their different spatial resolutions, the SMAP T_B^H data with a spatial resolution of 36 km are lower (~15 K) than the ELBARA-III observations at the field scale (~m), similar to SMAP T_B^V (lower ~5 K). The difference in vegetation dynamics between the SMAP scale (grazed) and the field scale (fenced off) may be another factor contributing to the differences in T_B^H and T_B^V.

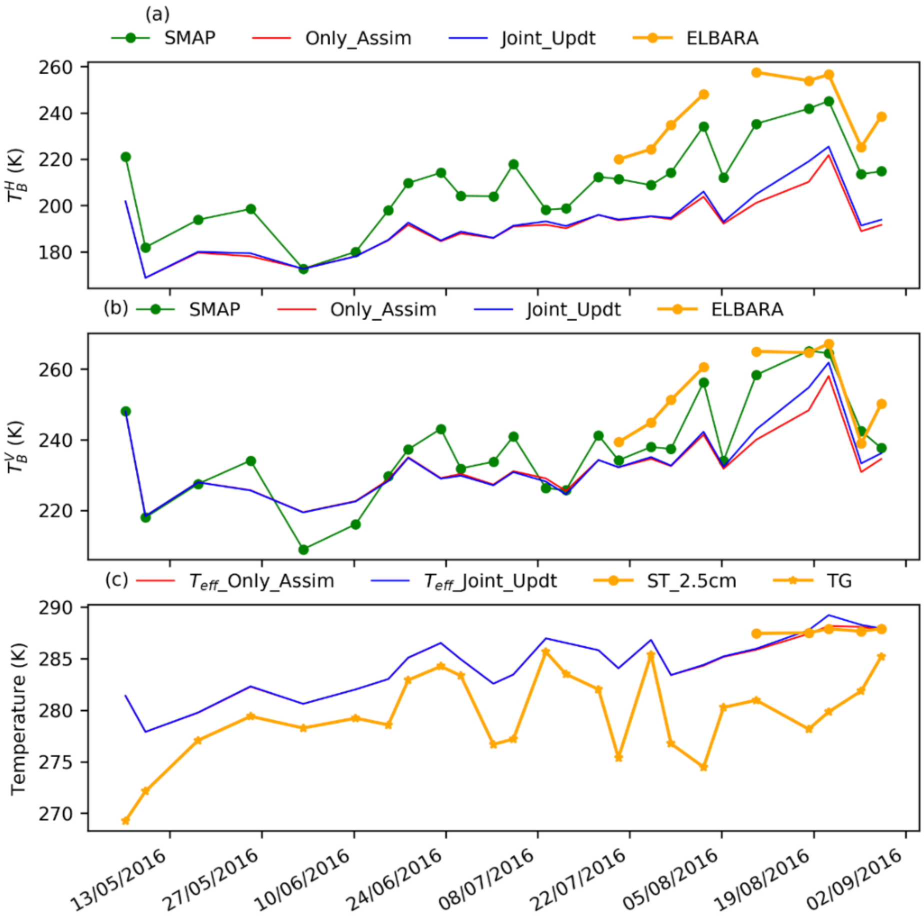


Figure 4.15 Comparisons of SMAP T_B^p, T_B^p estimated by the OL and assimilation experiments and field T_B^p observed by the ELBARA-III radiometer during the assimilation period (from 01/05/2016 to 31/08/2016), as well as T_{eff} comparisons. The in situ soil temperature at 2.5 cm (ST_2.5cm) is available after 07/08/2016. TG denotes the ground surface temperature, which is derived based on the in situ measured downward and upward longwave radiation using the Stefan-Boltzmann equation.

4.5 Conclusions

In this study, a physically-based discrete scattering-emission model (the Tor Vergata model, or TVG model) is for the first time coupled with CLM 4.5 in the DasPy data assimilation framework with the LETKF algorithm implemented. To investigate whether SMAP T_B^p data assimilation improves estimates of soil properties and associated land states (i.e., soil moisture and temperature) and land surface heat fluxes, four experiments including a reference, an open loop and two types of data assimilation strategies are designed. One assimilation experiment updates only the soil properties (Only_Para), and the other updates both the soil properties and soil moisture (Joint_Updt). *In situ* observations at the Maqu site on the eastern Tibetan Plateau are utilized to help with the investigation. To assess the effect of the different polarization configurations on the retrieval results, SMAP T_B^H and T_B^V are assimilated separately.

The results show the improvement of the soil property estimates by assimilating SMAP T_B^p, as both assimilation experiments reduce the RMSE for the retrieved clay fractions over the *in situ* measurements. The descriptions of the soil properties along the profile are also improved through the retrieved soil properties of the first layer and prior depth ratio. In the Only_Para experiment, the use of T_B^H is more sensitive to clay fraction and organic matter retrieval, and T_B^V to sand fraction retrieval. Comparatively, the Joint_Updt experiment can retrieve both the sand and clay fractions when assimilating either T_B^H or T_B^V. The Joint_Updt experiment also provides better estimates of the soil moisture, soil temperature and land heat fluxes during the assimilation period than those provided by the Only_Para experiment. However, they perform almost the same during the verification period.

Comparing the assimilation results to the *in situ* observations indicates that the retrieved soil properties with a finer accuracy are not sensitive factors affecting the accuracy of the soil moisture estimates by the CLM. Uncertainties in the model structures relating to soil moisture estimates should be considered. The above discussions reveal that optimizing PTF structures may be an alternative way to improve soil hydraulic property estimates and thereby soil moisture estimates. On the other hand, the van Genuchten (1980) function, which yields good estimates of soil moisture near saturation may be used to replace the Campbell (1974) function in the CLM. To enhance the surface soil moisture information propagating downward to the deeper layers through assimilation, the developed parameterization of the plant hydraulic stress within CLM 5 can be incorporated, and the highly discretized layering structure of the CLM may need to be adjusted. For better heat flux simulations, the parameterizations of land heat fluxes in the CLM also need to be improved, as the estimates of the reference run exhibit large discrepancies over the *in situ* observations. Improvement in model structures is expected to improve soil moisture estimates and thereby the L-band brightness temperature.

Chapter 5. Synthesis

5.1 Summary

As the lower boundary condition of the atmosphere, soil moisture plays an important role in land and atmosphere interactions and thereby in weather/climate predictions. Basic soil properties (i.e., soil texture and organic matter content) and associated soil hydraulic properties (SHPs) (i.e., soil water retention curve and hydraulic conductivity) and soil thermal properties (STPs) (i.e., soil heat capacity and thermal conductivity) are essential in estimating soil moisture-temperature profiles with land surface models (LSMs). Due to a lack of detailed soil property maps, the soil parameterization schemes considered in LSMs may be unrepresentative and hence introduce uncertainties in land surface states and heat fluxes estimates. With the use of the physical link between soil physical properties, soil moisture and temperature and the soil dielectric constant, soil physical properties can be retrieved with a LSM coupled with a microwave emission observation model in a data assimilation (DA) framework.

Passive L-band microwave remote sensing has become the most promising technique for the retrieval of near-surface soil moisture based on the measured brightness temperature (T_B^p, p =H, V). This trend has been further accelerated by the launch of two innovative space missions equipped with L-band radiometers, the Soil Moisture and Ocean Salinity (SMOS) and the Soil Moisture Active and Passive (SMAP) missions, which have provided global T_B^p and soil moisture products at nearly daily scales. The retrieval of soil physical properties by assimilating T_B^p observations into a coupled LSM and L-band emission model can thus be regarded, especially in remote areas, as a soil monitoring system utilizing space-based Earth observation (EO) data with *in situ* data and modeling.

This research aims to improve our understanding of soil physical property retrieval by using the aforementioned DA system. The LSM as a system model and the emission model as an observation operator are the two main components involved. In LSMs, soil moisture-temperature estimates are primarily determined according to basic soil properties and associated SHPs & STPs. In microwave L-band emission modeling, the surface roughness imposes significant impacts, especially on H polarization emission, and the resulting uncertainties in emission will propagate into the soil physical property retrieval process. Therefore, before carrying out the retrieval of soil properties, this thesis first focuses on soil physical properties for LSM modeling (Chapter 2) and the effect of surface roughness on L-band T_B^p estimates (Chapter 3). The Maqu site (33.91°N, 102.16°E) on the eastern Tibetan Plateau (TP), which provides comprehensive field observations, is chosen as the study area to help conduct the investigation.

To validate the retrieved results and analyze basic soil properties and SHPs & STPs for accurate land surface modeling on the TP, we conducted *in situ* and laboratory measurements of soil physical property profiles across different climate zones of the TP. With these collected profiles on the arid (Ngari), semi-arid (Naqu) and sub-humid (Maqu) zones, we compiled the Tibet-Obs soil property dataset. Based on this dataset, in Chapter 2, we 1) analyzed the variations in basic soil properties and SHP & STP across these three climate zones; 2) examined various schemes for estimation of the porosity and SHPs & STPs on the TP; 3) quantified the uncertainties in existing basic soil property datasets and their derived SHPs & STPs on the TP. We found that 1) the basic soil properties and SHPs & STPs differed in each climate zone and varied along the soil profile; 2) the Cosby et al. (1984) pedotransfer functions (PTFs) proved more applicable for SHP estimation with the Clapp and Hornberger (1978) (CH) model, and the continuous Wösten et al. (1999) PTFs with the van Genuchten (1980) - Mualem (1976) (VG) model. The De Vries (1963) semi-empirical model proved superior in the estimation of soil thermal properties. 3) We recommend the SoilGrids1km dataset in the arid and sub-humid zones, and a combination of FAO-UNESCO in the shallow layers and HWSD in the deeper layers of the semi-arid zone is recommended on the TP.

We have published this dataset at the 4TU.Center for Research Data (https://data.4tu.nl/articles/Soil_Hydraulic_and_Thermal_Properties_for_Land_Surface_Modeling_over_the_Tibetan_Plateau_version_1_/12721418/2). We hope that the compiled dataset and our findings contribute to modeling and research of the Third Pole environment in the hydro-climatology community, and fill geographic gaps amidst the already published global soil databases by the soil community.

To account for the effect of topsoil structures and inhomogeneous moisture distribution in the soil volume on L-band radiation, we developed an air-to-soil transition (ATS) model with a newly proposed dielectric roughness parameterization scheme in Chapter 3. The Tor Vergata discrete scattering model (TVG) integrated with the advanced integral equation model (AIEM) was adopted as the baseline model configuration to simulate the L-band brightness temperature (T_B^p, p =H, V). Then, the ATS model was coupled with the foregoing model to assess its performance. The comparison results indicate that the ATS model was necessary, as it compensated for the underestimation of T_B^p (≈ 20-50 K) in the baseline simulations. The proposed dielectric roughness is suggested as a replacement of the fixed roughness parameter H_R currently considered in state-of-the-art SMAP and SMOS soil moisture retrieval methods, as it captures the dynamics of surface roughness related to hydro-meteorological conditions. However, the discrepancy between the modeled and observed T_B^p values during the soil freeze-thaw transition period suggests that the ATS model still requires improvement by also incorporating the effects of surface temperature, surface water fraction and liquid water-ice mixtures in the calculation of the dielectric roughness thickness.

In Chapter 4, we coupled the enhanced physically-based discrete scattering-emission model, namely, the ATS+AIEM+TVG model (Chapter 3) with the community land model (CLM) v4.5 to retrieve soil physical properties by using the local ensemble transform Kalman filter (LETKF) algorithm assimilating SMAP Level-1C (L1C) T_B^H and T_B^V data separately. To identify whether the retrieved soil properties alone were sufficient to improve estimates of soil moisture and thereby land heat fluxes with the CLM, we conducted two comparative DA experiments. One experiment updated only the soil properties (Only_Para), and the other experiment updated both the soil properties and soil moisture (Joint_Updt). The results revealed an improvement of the estimates of the soil properties of the topmost layer by assimilating SMAP T_B^p ($p =$ H, V), as well as of the profile using the retrieved top-layer soil properties and prior depth ratio. In the Only_Para experiment, the use of T_B^H was very sensitive to the retrieval of the clay fraction and organic matter, whereas T_B^V was sensitive to sand fraction retrieval. In contrast, the Joint_Updt experiment could retrieve both the sand and clay fractions when assimilating either T_B^H or T_B^V. The Joint_Updt experiment provided better estimates of soil moisture, soil temperature and land heat fluxes during the assimilation period than those provided by the Only_Para experiment. However, we found that our obtained (retrieved) soil properties, with improved accuracy, were not the sensitive factors for improving soil moisture estimates. Therefore, in future studies, we should change our focus to the uncertainties in CLM model structures, such as the fixed PTF structures, the hydraulic function describing the soil water retention curve, the water stress function determining root water uptake and the soil layering structure, as well as the parameterizations describing land heat fluxes.

As a final note, we would like to highlight that the DA system for soil property retrieval developed in this study has the potential to obtain regional and even global soil parameter sets consistent not only in physics but also at different scales. Moreover, we should further investigate the uncertainties in LSM and L-band radiometry modeling, as this may improve not only the retrieval of soil properties but also the estimation of land surface states/fluxes.

5.2 Outlook

Future works should be carried out from the perspective of improving the performance of the observation operator, process model and DA algorithm. The retrieval of soil physical properties can be applied across the whole Tibetan Plateau, and extended to the global scale.

5.2.1 Enhancement of the integrated ATS+AIEM+TVG model for seasonal T_B^p simulation

Based on the results in section 3.4.2 and the discussions in section 3.5.1, we postulate that the integrated ATS+AIEM+TVG scattering-emission model can be improved considering soils experiencing freeze-thaw cycles, if the ground surface temperature, surface water fraction and soil ice content in the soil mixture is appropriately incorporated. To obtain an intuitive sense of how these three elements may affect surface T_B^p modeling during the freeze-thaw period, we recommend assessing the results reported in Su et al. (2020a). To simulate similar diurnal variations, we may attempt to obtain surface water fraction information. It is obviously difficult to acquire this kind of information, but there are three potential solutions. The first solution is to derive an index to reflect the information of the surface water fraction in terms of the microwave polarization difference index based on statistical values of T_B^p observations (communication with Prof. Bob Su and Dr. Yijian Zeng). For instance, considering that the surface water is highly related to H polarization due to geometric considerations, variations in the observed T_B^H can be explored. If assuming that the maximum value of T_B^H is characteristic for frozen soils, which also applies to the minimum

T_B^H value for thawed soils, their difference can indicate the amount of frozen/thawed water, and the ratio of T_B^H variation to this difference may reflect surface water fraction information. The second solution is to apply the (surface runoff) outputs of the CLM or STEMMUS-FT model (the Simultaneous Transfer of Energy, Mass and Momentum in Unsaturated Soil (STEMMUS) with freeze-thaw (FT) components) (Yu et al., 2020; Yu et al., 2018) with the enhanced consideration of soil freeze-thaw processes (please refer to section 5.2.2). The third solution is to adopt *in situ* observations, for example, mounting a high-end camera near the surface to monitor the change in surface status.

The current approach for the acquisition of soil ice content (SIC = TSWC − USWC) involves the determination of the unfrozen soil water content (USWC) and total soil water content (TSWC). To obtain SIC, two approaches can be considered. One is to use simulation results from the CLM with the enhanced parameterizations (please refer to section 5.2.2). The other is to obtain enhanced TSWC estimates by assimilating *in situ* cosmic-ray neutron probe (CRNP) observations, which will consequently lead to improved estimates of SIC. Mwangi et al. (2020) investigated the retrieval of SIC at the Maqu site utilizing *in situ* soil moisture (i.e., liquid phase, USWC) and CRNP observations (i.e., total water including liquid and ice, TSWC) with Observing System Simulation Experiments (OSSE). The OSSE in their study involved the STEMMUS-FT model (Yu et al., 2018; Zeng et al., 2011c) as the physically-based process model, and the cosmic-ray soil moisture interaction code (COSMIC) model as the observation operator (i.e., forward neutron simulator). Assimilating CRNP observations into the coupled CLM with COSMIC was implemented based on the DasPy framework by Han et al. (2014b). Therefore, this is a promising approach to obtain SIC. With the enhanced integrated ATS+AIEM+TVG model, we anticipate that this system can help monitor real-time changes in the surface status due to weather system changes.

5.2.2 CLM 5.0 to enhance the consideration of soil physical processes

Version 5.0 of the community land model (CLM 5.0) (Lawrence et al., 2019) is the latest in a series of global land models developed by the Community Earth System Model (CESM) Land Model Working Group (LMWG) and maintained at the National Center for Atmospheric Research (NCAR). Compared to CLM 4.5, CLM 5.0 introduces a dry surface layer-based soil evaporation resistance parameterization scheme (Swenson & Lawrence, 2014). CLM 5.0 applies a varying soil thickness in space (Brunke et al., 2016; Swenson & Lawrence, 2015), which allows a realistic retrieved soil physical property profile across different climate zones. CLM 5.0 also introduces an adaptive time-stepping solution to Richard's equation to improve the accuracy and stability of the numerical soil water solution. Moreover, the plant hydraulic stress (PHS) (Kennedy et al., 2019) approach to the modeling of water transport through vegetation according to a hydraulic framework is implemented in CLM 5.0, in which water supply equations are used to determine the vegetation water potential forced by the transpiration demand and a set of layer-by-layer soil water potentials. As described in section 4.4.2, using the PHS improves the estimation of root water uptake and thereby the vertical distribution of soil water (Kennedy et al., 2019). The performance of CLM 5.0 at the Maqu site should be assessed, and CLM 4.5, implemented in DasPy can be replaced by CLM 5.0.

In CLM 5.0, in regard to the freezing process in soil layers, the concept of supercooled soil water proposed by Niu and Yang (2006) is adopted. In contrast, STEMMUS (Zeng et al., 2011c) is a coupled water and energy model considering the gaseous phase (water vapor and dry air) flow mechanism. As soil freezing processes can be analogous to the drying process (Farouki, 1981; Koopmans & Miller, 1966; Rautiainen et al., 2014), water vapor during the freezing period can be transported from beneath the freezing front to the land surface (Yu et al., 2020; Yu et al., 2018). The STEMMUS model using the van Genuchten (1980) -

Mualem (1976) (VG) model can be incorporated into soil part modeling in CLM 5.0. As such, the uncertainties in the different parameterizations of soil hydraulic properties as well as soil water dynamic processes (i.e., freeze-thaw process) and their impacts on soil property retrieval can be investigated.

5.2.3 Use of the four-dimensional ensemble variational (4DEnVar) DA algorithm

Sequential ensemble methods such as the LETKF algorithm are easy to implement, which performs a linear update and propagates the statistics of errors. However, the ensemble Kalman filter (EnKF)-based approach is claimed to lead to the retrieval of time-varying parameter sets physically inconsistent with the behavior of the land surface (Pinnington et al., 2020). The variational method designed for the estimation of the model trajectory that best fits all the observations within a prescribed observing window can help with this issue, which relies on time-independent statistics of errors and searches for a nonlinear estimation of the maximum posterior of the underlying probability distribution function using nonlinear optimization techniques (Carrassi et al., 2018). The derivation of adjoint models is generally required to obtain the variational cost function and can be minimized using nongradient-based optimization routines but comes at the cost of many more model runs to achieve convergence and a loss of accuracy (Pinnington et al., 2018).

For theoretical reasons (nonlinear analysis and error accumulation) and technical reasons (no adjoint model), new hybrid methods combining both ensemble and four-dimensional variational techniques (Bannister, 2017; Desroziers et al., 2014; Liu et al., 2008) have been developed, presenting a way to retrieve time-invariant parameters within a specific time window with improved background-error covariance matrices and efficiency. Pinnington et al. (2020) implemented the hybrid technique of four-dimensional ensemble variational (4DEnVar) DA with the Joint UK Land Environment Simulator (JULES) land surface model to estimate parameters (e.g., LAI, gross primary productivity) controlling crop behaviors. With the use of the same 4DEnVar algorithm, Cooper et al. (2020) optimized constants in the underlying Cosby PTFs adopted in JULES and thereby improved the estimation of soil moisture by JULES. Similarly, the LETKF algorithm implemented in DasPy can be upgraded into 4DEnVar to improve the background-error covariance matrices and DA efficiency (Bannister, 2017) and thereby the robust retrieval of soil physical properties.

5.2.4 Retrieval of soil physical properties on the Tibetan Plateau

Once the retrieval system is improved, soil physical properties can be estimated across the whole Tibetan Plateau. The first high-spatiotemporal resolution gridded near-surface meteorological dataset, namely, the China Meteorological Forcing Dataset (CMFD) (He et al., 2020), can be adopted as the atmospheric forcing to drive CLM 5.0. The dataset has a temporal resolution of three hours and a spatial resolution of 0.1° (He et al., 2020). The SoilGrids1km data can be used as the prior in the retrieval system. The SMAP L1C T_B^p dataset with a spatial resolution of 36 km and the SMAP L1C enhanced T_B^p dataset with a spatial resolution of 9 km can be assimilated and compared. The retrieved soil physical properties can be validated against the available *in situ* measured Tibet-Obs soil dataset (Zhao et al., 2018a) at the point scale. FAO-UNESCO (FAO/UNESCO, 2007), HWSD (FAO/IIASA/ISRIC/ISSCAS/JR, 2012), BNU (Shangguan et al., 2012; Shangguan et al., 2013) and HPSS (Montzka et al., 2017) datasets can be used for cross-validation in terms of spatial distribution patterns. The estimated soil moisture-temperature and land heat fluxes can be evaluated via a comparison to the Tibet-Obs soil moisture dataset (Su et al., 2011; Su et al., 2020a; Zeng et al., 2016; Zhang et al., 2020; Zhuang et al., 2020). The period from 01/2016 to the present can be specifically investigated, in which the *in situ* Tibet-Obs dataset and ELBARA-III T_B^p observations are available. The retrieved soil physical properties, which are consistent with atmospheric forcing data and SMAP T_B^p observations, are expected to be applied in the Earth system models to acquire better estimates of initial conditions and thereby accurate weather/climate predictions.

Appendix

Appendix A. Soil Water and Heat Flow Modeling

Table A1.1 Information of the exiting soil datasets adopted in this study.

Name	Soil property	Spatial resolution	Profile information	Data source	Reference	Date
FAO-UNESCO	Soil texture fractions, SOC	5km	2 layers 0-30cm, 30-100cm	1:5 million Soil Map of the World	FAO-UNESCO Digital Soil Map of the World, 2007	2007
HWSD	Soil texture fractions, SOC	1km	2 layers 0-30cm, 30-100cm	1:1 million soil map of China; 1:5 million Soil Map of the World; 7292 profiles in China	FAO/IIASA/ISRIC/ISSCAS/JRC, 2012. Harmonized World Soil Database (version 1.2). FAO, Rome, Italy and IIASA, Laxenburg, Austria.	2012
BNU	Soil texture fractions, SOC, BD, GGF, Porosity	1km	8 layers 0-4.5cm, 4.5-9.1cm 9.1-16.6cm, 16.6-28.9cm 28.9-49.3cm, 49.3-82.9cm 82.9-138.3cm, 138.3-229.6cm	1:1 million soil map of China; 8979 profiles in China	Shangguan et al. (2012, 2013)	2012
SoilGrids1km	Soil texture fractions, SOC, BD, GGF	1km	7 layer 0, 5, 15, 30, 60, 100 and 200cm.	Chinese soil profile database (Shangguan et al. 2013) Covariables: MODIS images,	Hengl et al. (2014)	2014

				SRTM DEM etc.		
SoilGrids250m	Soil texture fractions, SOC, BD, GGF	250m	7 layer 0, 5, 15, 30, 60, 100 and 200cm.	Chinese soil profile database (Shangguan et al. 2013) Covariables: MODIS images, SRTM DEM etc.	Hengl et al. (2017)	2017
HPSS	Parameters in the van Genuchten and Mualem model	25km	7 layer 0, 5, 15, 30, 60, 100 and 200cm.	SoilGrids 1km	Montzka et al. (2017)	2017

Where SRTM denotes the Shuttle Radar Topography Mission and DEM denotes the digital elevation model.

A.1 Porosity estimation schemes

Cosby-S scheme (univariate)

The Cosby et al. (1984) PTF is used to estimate porosity based on the sand percentage of the soil texture,

$$\phi = 0.489 - 0.001268 \times (\%\mathrm{sand}) \qquad \text{(A1.1)}$$

where ϕ is the soil porosity, *%sand* is the sand proportion.

BD scheme

The BD scheme for porosity calculation (Hillel, 2003) is as follows,

$$\phi = 1 - \frac{\rho_b}{\rho_s} \qquad \text{(A1.2)}$$

where ρ_b is the dry bulk density (g/cm^3). ρ_s is the mineral particle density valued at 2.65g/cm^3. In the soil mixture, the BD scheme assumes that the coarse and fine components share the same particle density.

SocVg scheme

Regarding soil as a mixture of organic and fine minerals, Chen et al. (2012) conceptualized the porosity as expressed in equation (A1.3). Through the determination of the volumetric SOC, the gravel impact was considered (equations (A1.4-A1.5)) and was assumed to be equal to the impact of sand particles. The effective sand proportion was calculated with equation (A1.6).

$$\phi_m = (1 - V_{soc})\phi_F + V_{soc}\phi_{soc,sat} \tag{A1.3}$$

$$V_{soc} = \frac{\rho_s(1-\phi_F)m_{soc}}{\rho_{soc}(1-m_{soc}) + \rho_s(1-\phi_F)m_{soc} + (1-\phi_F)\frac{\rho_{soc}GGF}{(1-GGF)}} \tag{A1.4}$$

$$\mathrm{VGF} = \frac{\rho_{soc}(1-\phi_F)\mathrm{GGF}}{(1-GGF)(\rho_{soc}(1-m_{soc}) + \rho_s(1-\phi_F)m_{soc} + (1-\phi_F)\frac{\rho_{soc}GGF}{(1-GGF)}} \tag{A1.5}$$

$$\%sand_e = \%sand * (1 - VGF) + VGF \tag{A1.6}$$

where ϕ_m is the porosity of the soil mixture. V_{soc} and VGF are the volumetric fractions of SOC and gravel, respectively. ϕ_F is the porosity of the fine component and was calculated with equation (A1.1), where %sand is determined with equation (A6). GGF and m_{soc} are the gravimetric fractions of gravel particles and SOC, respectively. ρ_{soc}=0.13 g/cm^3 is BD of peat. $\phi_{soc,sat}$=0.9 is the porosity of peat.

Binary mixture (BM) scheme

Zhang et al. (2011) proposed a mixing-coefficient model to estimate the porosity of the binary mixture,

$$\phi_m = \begin{cases} (VGF - \beta_m * VGF + \beta_m)\phi_g + VFF * \phi_F - \beta_m * VFF \\ \quad if\ VFF < \phi_g \\ (1 - \beta_m) * VGF * \phi_g + VFF * \phi_F \\ \quad if\ VFF \geq \phi_g \end{cases} \tag{A1.7}$$

where *VFF* is the volumetric fraction of the fine mineral. *VGF* is determined with equation (A1.8). ϕ_F shares the same definition as in the SocVg scheme. ϕ_g is the porosity of gravel particles, which is mainly affected by the median grain size (Frings et al., 2011). In this study, ϕ_g was calculated by using empirical equation (A1.9) given by Wu and Wang (2006). β_m is the mixing coefficient related with the grain size (equation (A1.10)).

$$VGF = \frac{GGF(1 - \phi_F)}{GGF(1 - \phi_F) + (1 - GGF)\left(1 - \phi_g\right)} \tag{A1.8}$$

$$\phi_g = 0.13 + \frac{0.21}{(GD + 0.002)^{0.21}} \tag{A1.9}$$

$$\beta_m = \begin{cases} 0.0363\frac{GD}{FD} + 0.2326 & for\frac{GD}{FD} \leq 21 \\ 1 & for\frac{GD}{FD} > 21 \end{cases} \tag{A1.10}$$

where GD and FD are the mean grain sizes of gravel particles and the fine minerals, respectively, and the unit is mm.

A.2 Functions modeling soil water retention curve (SWRC)

The function of Clapp and Hornberger (1978) (i.e. CH) modeling soil water retention is written as:

$$\varphi = \varphi_s({}^{\theta}/_{\theta_s})^{-1/b} \quad \varphi \leq \varphi_i \tag{A1.11}$$

where φ_s is the saturated capillary potential (cm). b is the pore size distribution index (dimensionless). θ is SM ($m^3\ m^{-3}$) and θ_s is the saturated SM. φ_i denotes an inflection point near saturation. The corresponding soil conductivity and diffusivity estimations are formulated as:

$$\begin{cases} K(\theta) = K_s\left({}^{\theta}/_{\theta_s}\right)^{3+2/b} \\ D(\theta) = D_s\left({}^{\theta}/_{\theta_s}\right)^{2+1/b} \\ D_s = {}^{1}/_{b} * K_s({}^{\varphi_s}/_{\theta_s}) \end{cases} \tag{A1.12}$$

where K and D are the soil hydraulic and thermal conductivity, respectively. K_s and D_s are the saturated hydraulic conductivity (m/s) and diffusivity (m^2/s), respectively.

The Van Genuchten (1980)-Mualem (1976) (i.e. VG) model describes the water retention curve as shown in equation (A1.13),

$$\theta(\boldsymbol{h}) = \theta_r + \frac{\theta_s - \theta_r}{(1 + (a\boldsymbol{h})^n)^{1-1/n}} = f(\boldsymbol{h}, \theta_r, \theta_s, \alpha, n) \tag{A1.13}$$

where $\theta(h)$ is SM (m^3 m^{-3}) at pressure head h (cm). θ_r is the residual SM (m^3 m^{-3}). θ_s shares the same meaning as the above. α is the inverse of the air entry value (cm^{-1}). n is the shape parameter (dimensionless). The corresponding soil conductivity and diffusivity estimations are formulated as:

$$\Theta = \frac{\theta - \theta_r}{\theta_s - \theta_r} \tag{A1.14}$$

$$K = K_s \Theta^{1/2}\left[1 - (1 - \Theta^{1/(1-1/n)})^{1-1/n}\right]^2$$

$$D(\Theta) = \frac{(1-m)K_s}{\alpha m(\theta_s - \theta_r)}\Theta^{1/2-1/m}\left[\left(1 - \Theta^{1/m}\right)^{-m} + \left(1 - \Theta^{1/m}\right)^{m} - 2\right]$$

$$m = 1 - 1/n$$

where Θ is the effective saturation.

Based on the measured soil water potential and SM, we adopted the scaling method proposed by Montzka et al. (2017) to estimate the hydraulic parameters of the CH and VG models. The expected-scale (representative) parameters ($\widehat{\theta_s}, \hat{b}, \widehat{\varphi_s}$) and ($\widehat{\theta_r}, \widehat{\theta_s}, \hat{\alpha}, \hat{n}$) were obtained with the damped least-squares method of Levenberg-Marquardt (Marquardt, 1963), which generates the minimum of the sum of squares of the deviations between the water retention curves of $f(h, \theta_s, b, \varphi_s)$ and $f(h, \theta_r, \theta_s, \alpha, n)$ with all respective observations $i = 1 \ldots N$ (equation (A1.15)). The initial values adopted in the parameter fitting algorithm were the means of (θ_s, α, n) and $(\theta_r, \theta_s, \alpha, n)$ based on each observation.

$$\left(\widehat{\theta_s}, \hat{b}, \widehat{\varphi_s}\right) = \operatorname{argmin} \sum_{i=1}^{N} [\theta_i - f(\boldsymbol{h}, \theta_{s,i}, b_i, {\varphi_s}_i)]^2 \tag{A1.15}$$

$$\left(\widehat{\theta_r}, \widehat{\theta_s}, \hat{\alpha}, \hat{n}\right) = \operatorname{argmin} \sum_{i=1}^{N} [\theta_i - f(\boldsymbol{h}, \theta_{r,i}, \theta_{s,i}, \alpha_i, n_i)]^2$$

A.3 PTFs for SWRC estimation

Various PTFs have been developed to determine soil hydraulic properties. In terms of criteria described in Dai et al. (2013), five PTFs (No. 1-5 in Table A1.2) were selected to estimate parameters (θ_s, φ_s, b) of the CH model, and seven PTFs (No. 6-12 in Table A1.2) were selected to estimate parameters (θ_r, θ_s, α, n) of the VG model.

Table A1.2 List of PTFs to estimate soil water retention curve.

No.	PTF	Retention/ K_s model	Sand	Silt	Clay	Organic Carbon	Dry bulk density	Depth
			%	%	%	%	g cm^{-3}	-
1	Cosby et al., 1984 (1)	CH, K_s[1]	√		√			

[1] $K_s = 60.96 * 10^{-0.884+0.0153*sand}$ (unit: cm/day).

2	Cosby et al., 1984 (2)	CH, K_s[2]	√	√	√			
3	Saxton et al., 1986	CH, K_s[3]	√		√	√		
4	Campbell and Shiosawa, 1992	CH, K_s[4]	√	√	√		√	
5	Saxton et al., 2006	CH, K_s[5]	√		√	√		
6	Rawls and Brakenssiek 1985	VG, K_s[6]	√		√		√	
7	Class Wösten et al., 1999	VG, K_s[7]	√	√	√			√
8	Vereecken et al., 1989,1990	VG, K_s[8]	√		√	√	√	
9	Continuous Wösten et al., 1999	VG, K_s[9]		√	√	√	√	√
10	Rosetta1-H3	VG, K_s[10]	√	√	√		√	
11	Rosetta3-H3	VG, K_s[11]	√	√	√		√	
12	Weynants et al. 2009	VG, K_s[12]	√		√	√	√	

where 10 and 11 were developed by Schaap et al. (2001) and Zhang and Schaap (2017), respectively.

[2] $K_s = 60.96 * 10^{-0.6+0.0126*sand-0.0064*clay}$ (multi-variate) (unit: cm/day).

[3] $K_s = 24 * \{\exp[12.012 - 0.0755 * sand + [-3.8950 + 0.03671 * sand - 0.1103 * clay + 8.7546 * 10^{-4} * clay^2](1/\theta_s)]\}$ (unit: cm/day).

[4] $K_s = 339.0 * \left(\frac{1.3}{BD}\right)^{1.3b} \exp(-0.06888 * clay - 0.03638 * silt - 0.025)$ (unit: cm/day).

[5] $x = 0.00251 * \mathrm{sand} + 0.00195 * \mathrm{clay} + 0.00011 * \mathrm{SOC} + 0.00006 * \mathrm{sand} * \mathrm{SOC}\ 0.00027 * \mathrm{clay} * \mathrm{SOC} + 0.0000452 * \mathrm{sand} * \mathrm{clay} + 0.299$; $K_s = 4632 * (\theta_s - x)^{3-b}$ (unit: cm/day).

[6] $\theta_s = \phi = 1 - BD/2.65$; $K_s = 24.0 * \exp(19.52348 * \phi - 8.96847 - 0.028212 * clay + 0.00018107 * sand^2 - 0.0094125 * clay^2 - 8.395215 * \phi^2 + 0.077718 * sand * \phi - 0.00298 * sand^2 * \phi^2 - 0.019492 * clay^2 * \phi^2 + 0.0000173 * sand^2 * clay + 0.02733 * clay^2 * \phi + 0.001434 * sand^2 * \phi - 0.0000035 * sand * clay^2$ (unit: cm/day).

[7] The K_s for the FAO textural classes (Pachepsky & Rawls, 2004). (unit: cm/day).

[8] $\mathrm{Log}(K_s) = 20.62 - 0.96 * \log(clay) - 0.66 * \log(sand) - 0.46 * \log(clay) - 8.43 * BD$ (unit: cm/day).

[9] $K_s = \exp(7.755 + 0.0352 * \mathrm{silt} + 0.93 * \mathrm{itop} - 0.967 * BD^2 - 0.000484 * clay^2 - 0.000322 * silt^2 + 0.001/\mathrm{silt} - 0.0748/\mathrm{SOC} - 0.643 * \ln(\mathrm{silt}) - 0.01398 * \mathrm{BD} * \mathrm{clay} - 0.1673 * \mathrm{BD} * \mathrm{SOC} + 0.02986 * \mathrm{itop} * \mathrm{clay} - 0.03305 * \mathrm{itop} * \mathrm{silt}))$, where topsoil is an ordinal variable having the value of 1 (depth 0–30 cm) or 0 (depth 30 cm). (unit: cm/day).

[10] H3 hierarchical pedotransfer function in Schaap et al. (2001) (unit: cm/day).

[11] Updated H3 hierarchical pedotransfer function in Zhang and Schaap (2017) (unit: cm/day).

[12] $K_s = \exp(1.9582 + 0.0308 sand - 0.6142 BD - 0.01566 SOC * 1.72)$ (unit: cm/day).

A.4 Saturated hydraulic conductivity estimation schemes

PTFs-VGF scheme

The PTFs-VGF scheme estimates K_s of the soil mixture (Peck & Watson, 1979) as follows:

$$K_{sm} = K_{sat,f}\frac{2(1-VGF)}{2+VGF} \quad \text{(A1.16)}$$

where K_{sm} is K_s of the soil mixture. $K_{sat,f}$ is K_s of the fine mineral and was calculated using PTFs listed in Table A1.2. VGF shares the same definition as in equation (A1.8).

BM-Kozeny-Carman equation (BM-KC scheme)

The Kozeny-Carman equation (A1.17), originally developed to quantitatively describe hydraulic conductivity vs. the mean grain size in the capillary flow, was used to estimate K_s of the binary mixture. The porosity was obtained by using the BM scheme described in A.1 section. The representative grain diameter was estimated using the power-averaging method (equation (A1.18)) proposed by Zhang et al. (2011). This method introduces an empirical coefficient (equation (A1.19)), which is parameterized considering the critical fraction of gravel particles.

$$K_{sm} = \left(\frac{\rho g}{\mu}\right)\left[\frac{d_m^2\phi_m^3}{180(1-\phi_m)^2}\right] \quad \text{(A1.17)}$$

$$d_m = (VGF * GD^p + VFF * FD^p)^{1/p} \quad \text{(A1.18)}$$

$$p = \frac{1}{1+\exp[(\alpha(VGFc-VGF))]} - 1 \quad \text{(A1.19)}$$

where ϕ_m shares the same definition as in equation (A1.7). d_m is the representative grain diameter of the soil mixture. ρ is the fluid density. g is the gravitational acceleration, and μ is the dynamic viscosity. VGF, VFF, GD and FD share the same definitions as in the BM scheme described in A.1 section. p is the coefficient that varies sigmoidally from −1 to 0 as VGF increases from 0 to 1. The $VGFc$ is the critical fraction of gravel particles and is calculated by $VGFc = 1 - \phi_g$ (ϕ_g from equation (A1.9)). α is the shape factor with a value of 20 as adopted by Zhang et al. (2011).

A.5 Heat capacity and thermal conductivity modeling

Heat capacity estimation

The soil heat capacity C_s depends on the heat capacities of all constituents of soil, and is calculated with equation (A1.20) given by De Vries (1963),

$$C_s = \theta C_w + (1 - \theta_s) C_{soil} + (\theta_s - \theta) C_{air} \tag{A1.20}$$

where θ and θ_s share the same meanings as in equation (A1.11). *C* represents the heat capacity (MJ m^{-3} K^{-1}), and the subscripts ‘w’, ‘soil’ and ‘air’ refer to water, soil solid and air, respectively. C_w, C_{soil} and C_{air} are valued at 4.2, 2.0 and 0.001 MJ m^{-3} K^{-1}, respectively. If considering SOC impact, C_s is calculated with equation (A1.21) shown as follows,

$$C_s = \theta C_w + (1 - \theta_s) * ((1 - Vsoc) * C_{soil} + V_{soc} * C_{soc}) + (\theta_s - \theta) C_{air} \tag{A1.21}$$

where $Vsoc$ shares the same definition as in equation (A1.4). C_{soc} is the heat capacity of the organic matter and valued at 2.5 MJ m^{-3} K^{-1}.

The De Vries (1963) model revised by the Farouki (1981) model (D63F)

The De Vries (1963) model was developed from the Maxwell equation, which was used to estimate the electrical conductivity of a mixture of the granular materials dispersed in a continuous fluid (Eucken, 1932). Farouki (1981) set liquid water as the continuous medium and regarded soil minerals as the uniform particles. Considering soil as the binary mixture of fine minerals and coarse gravel particles, λ is estimated as follows:

$$\lambda = \frac{\theta\lambda_w + w_a x_a(\lambda_a + \lambda_v) + w_m x_m \lambda_m + w_g x_g \lambda_g + w_{soc} x_{soc} \lambda_{soc}}{\theta + w_a x_a + w_m x_m + w_g x_g + w_{soc} x_{soc}} \quad \text{(A1.22)}$$

where w is the weighting factor, x is the volume fraction, λ is the thermal conductivity, and the subscripts 'w', 'a', 'v', 'm', 'g' and 'soc 'refer to water, air, vapor, fine minerals, gravel particles and SOC compositions of soil, respectively. λ_w =0.57 W m^{-1} K^{-1}, λ_a =2.0 W m^{-1} K^{-1}, λ_g =2.54 W m^{-1} K^{-1} and λ_{soc} =0.25 W m^{-1} K^{-1}. λ_m is calculated with equation (A1.23),

$$\lambda_m = \lambda_q^q \lambda_o{}^{(1-q)} \quad \text{(A1.23)}$$

where λ_q is the thermal conductivity of quartz (λ_q=7.7 W m^{-1} K^{-1}), and λ_o is the thermal conductivity of other minerals (λ_o=2.0 W m^{-1} K^{-1}). $Vsoc$ shares the same meaning as in equation (A1.4). In this study, the volumetric quartz q is assumed to equal half of the sand fraction ($q = 1/2V_{sand}$) in terms of investigations by Chen et al. (2012).

w in equation (A1.24) is estimated empirically by:

$$w_i = \frac{1}{3}\left[\frac{2}{1 + \left(\frac{\lambda_i}{\lambda_w} - 1\right) g_a} + \frac{1}{1 + \left(\frac{\lambda_i}{\lambda_w} - 1\right)(1 - 2g_a)}\right] \quad \text{(A1.24)}$$

where g_a is the shape factor of the ellipsoidal particles. A uniform shape factor g_a value of 0.125 is used for the fine minerals (Farouki, 1981), a g_a value of 0.33 in gravel particles and a g_a value of 0.5 in SOC (De Vries, 1963).

Regarding the estimations of λ_v and g_a of air, Farouki (1981) provided the following equations (A1.25-A1.26).

For 0.09 m^3 m^{-3} $\leq \theta \leq \phi$,

$$\lambda_v = \lambda_v^s \quad \text{and} \quad g_{a(air)} = 0.333 - (0.333 - 0.035)x_a/\phi \quad \text{(A1.25)}$$

And for $0 \leq \theta \leq$ 0.09 m^3 m^{-3},

$$\lambda_v = \frac{\theta}{0.09}\lambda_v^s \quad \text{and} \quad g_{a(air)} = 0.013 + 0.944\theta \tag{A1.26}$$

where λ_v^s is the value of λ_v of the saturated vapor. ϕ shares the same meaning as in equation (A1.2).

The Simplified De Vries-based model (T16)

The T16 scheme (Tian et al., 2016) assumed a negligible effect of the vapor movement (i.e. $\lambda_v = 0$) in the De Vries-based model (equation (A22)). The soil texture was assumed to determine the physical properties of the soil minerals. λ of the fine minerals (λ_m) and the shape parameters of the minerals and air were computed using equations (A1.27-A1.29),

$$\lambda_m = \lambda_{sand}^{V_{sand}} {\lambda_{clay}}^{V_{clay}} {\lambda_{silt}}^{V_{silt}} \tag{A1.27}$$

$$g_{a(m)} = g_{a(sand)}V_{sand} + g_{a(silt)}V_{silt} + g_{a(clay)}V_{clay} \tag{A1.28}$$

where λ_{sand} =7.7 W m^{-1} K^{-1}, $g_{a(sand)}$ =0.782, λ_{silt} =2.74 W m^{-1} K^{-1}, $g_{a(silt)}$=0.0534, λ_{clay}=1.93 W m^{-1} K^{-1}, and $g_{a(clay)}$=0.00775.

$g_{a(air)}$ is assumed to linearly vary with the air fraction and is estimated using equation (A1.29),

$$g_{a(air)} = 0.333 * (1 - x_a/\phi\ ) \tag{A1.29}$$

where ϕ shares the same meaning as in equation (A1.2) and x_a shares the same meaning as in equation (A1.22).

In regard to dry soil, the calculation of λ_{dry} follows equation (A1.30) proposed by De Vries (1963),

$$\lambda_{dry} = 1.25 * \frac{w_a x_a \lambda_a + w_m x_m \lambda_m + w_g x_g \lambda_g + w_{soc} x_{soc} \lambda_{soc}}{w_a x_a + w_m x_m + w_g x_g + w_{soc} x_{soc}} \tag{A1.30}$$

where the parameters share the same definitions as in equation (A1.22).

Johansen model (J75)

The Johansen (1975) model estimates λ in terms of the combination of dry and saturated state values and a weight factor known as the Kersten number as shown in equation (A1.31),

$$\lambda = K_e\left(\lambda_{sat} - \lambda_{dry}\right) + \lambda_{dry} \quad \text{(A1.31)}$$

where λ_{dry} and λ_{sat} are the dry and saturated thermal conductivity, respectively. K_e is the Kersten number, a normalized thermal conductivity, which relates to the logarithm of SM (Kersten, 1949) as shown in equation (A1.32),

$$K_e = \frac{\lambda - \lambda_{dry}}{\lambda_{sat} - \lambda_{dry}} \quad \text{(A1.32)}$$

$$\begin{cases} K_e = log_{10}(S_r) + 1.0 & S_r > 0.05 \\ K_e = 0.7 * log(S_r) + 1.0 & S_r > 0.1 \\ K_e = 0 & S_r \le 0.05 \end{cases}$$

where S_r is the saturation degree and is defined as equation (A1.33),

$$S_r = {}^{\theta}/_{\theta_s} \quad \text{(A32)} \quad \text{(A1.33)}$$

where θ is SM (m^3 m^{-3}). θ_s is the saturated SM (m^3 m^{-3}) calculated using equation (A1.2).

The saturated thermal conductivity is calculated using equation (A1.34),

$$\lambda_{sat} = \lambda_m^{1-\theta_s} {\lambda_w}^{\theta_s} \quad \text{(A1.34)}$$

where λ_m shares the same definition as in equation (A1.23). If considering the SOC impact, λ_m is calculated using equation (A1.35),

$$\lambda_m = \lambda_q^{q(1-Vsoc)} {\lambda_o}^{(1-q)(1-Vsoc)} {\lambda_{soc}}^{Vsoc} \quad \text{(A1.35)}$$

The estimation of the thermal conductivity of dry soil is given by equation (A1.36):

$$\lambda_{dry} = \frac{0.135\rho_b + 64.7}{2700 - 0.947\rho_b} \tag{A1.36}$$

where ρ_b is the dry bulk density (kg/m^3).

A.6 Soil water flow and heat transport modeling

The vertical movement of water in the unsaturated zone of the soil matrix is modeled by the Richards equation (Richards, 1931):

$$\frac{\partial\theta}{\partial t} = \frac{\partial}{\partial z}\left(D(\theta)\frac{\partial\theta}{\partial z} - K(\theta)\right) + S_\theta \tag{A1.37}$$

where $D(\theta)$ (m^2/s) and $K(\theta)$ (m/s) are the hydraulic diffusivity and hydraulic conductivity, respectively, and S_θ is a volumetric sink term associated with the root uptake (m^3 m^{-3} s^{-1}), which depends on the surface energy balance and the root profile.

The soil heat transfer is modeled by the Fourier law of diffusion:

$$C_s\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left(\lambda\frac{\partial T}{\partial z}\right) \tag{A1.38}$$

where C_S is the soil thermal heat capacity (J m^{-3} K^{-1}), λ is the thermal conductivity (W m^{-1} K^{-1}) and T is soil temperature (°C).

Table A1.3 Means and standard deviations of the measured soil properties at the different depths in the Ngari area

Parameter	5 cm	10 cm	20 cm	40 cm
Sand (%)	84.54±8.28	86.19±8.63	77.21±17.55	78.81±17.93
Clay (%)	3.05± 1.99	2.73± 2.03	3.83± 2.67	3.27± 2.71
Silt (%)	12.42± 6.54	11.08± 6.93	18.96± 14.98	17.93± 15.41
GGF (%)	11.26± 10.76	11.77± 15.71	23.17± 25.05	18.75± 20.17
SOC (%)	1.02± 0.67	0.7± 0.49	0.73± 0.49	0.79± 0.67
Porosity (%)	33.49± 2.78	35.8± 6.76	31.4± 7.2	33.8± 9.47
GD (mm)	5.02± 2.92	5.23± 2.47	7.56± 5.07	4.96± 1.7
FD (mm)	0.22± 0.1	0.2± 0.08	0.23± 0.19	0.19± 0.09

BD (g/cm^3)	1.56± 0.13	1.63± 0.26	1.6± 0.26	1.56± 0.18
LogK_s (m/s)		-4.57± 0.24	-4.94± 0.47	-4.68± 0.30

Table A1.4 Means and standard deviations of the measured soil properties at the different depths in the Naqu area

Parameter	5 cm	10 cm	20 cm	40 cm	50 cm
Sand (%)	78.79± 6.86	81.48± 13.49	75.13± 14.67	75.93± 10.64	70.15± 20.28
Clay (%)	4.41± 1.63	4.02± 3.04	5.84± 3.87	6.43± 4.17	7.29± 6.39
Silt (%)	16.8± 5.79	14.5± 10.46	19.03± 10.85	17.64± 7.09	22.56± 14.14
GGF (%)	12.69± 13.11	19.3± 15.91	34± 25.97	53.29± 24.05	57.43± 22.43
SOC (%)	9.18± 3.55	8.17± 3.95	2.25± 1.11	1.61± 0.93	2.68± 3.24
Porosity (%)	58.5± 21.49	45.67± 6.81	39.75± 5.8	29.5± 6.61	24.5± 5.92
GD (mm)	4.55± 1.78	3.96± 1.2	7.28± 4.57	7.75± 4.99	6.18± 2.6
FD (mm)	0.19± 0.04	0.21± 0.07	0.19± 0.08	0.22± 0.05	0.19± 0.12
BD (g/cm^3)	1.01± 0.48	1.42± 0.08	1.64± 0.17	1.87± 0.21	2.11± 0.18
LogK_s (m/s)		-5.20± 0.25	-5.09± 0.50	-5.20± 0.77	-6.12± 0.99

Table A1.5 Means and standard deviations of the measured soil properties at the different depths in the Maqu area

Parameter	5 cm	10 cm	20 cm	40 cm	80 cm
Sand (%)	26.95± 10.55	29.03± 13.08	29.21± 12.61	31.6± 12.47	34.83± 17.06
Clay (%)	9.86± 0.89	9.95± 0.91	10.15± 0.61	10.43± 0.89	9.35± 2.68
Silt (%)	63.19± 10.08	61.02± 12.52	60.65± 12.48	57.97± 12.18	55.82± 14.95
SOC (%)	17.88± 9.05	12.16± 6.23	8.05± 5.05	4.13± 3.14	2.87± 2.89
Porosity (%)	72.92± 7.55	65.57± 7.57	59.21± 6.22	50.96± 7.5	47.06± 6.5
FD (mm)	0.03± 0.01	0.03± 0.01	0.03± 0.01	0.03± 0.01	0.04± 0.02
BD (g/cm^3)	0.76± 0.22	0.95± 0.25	1.23± 0.19	1.4± 0.12	1.49± 0.18
LogK_s (m/s)		-5.5± 0.32	-5.55± 0.44	-6.52± 0.3	-5.65± 0.97

Table A1.6 Values of the Bias and RMSE between the estimated porosities with the measurements on the three climate zones. The unit of the listed values is m^3 m^{-3}.

Schemes		Cosby-S		BD		SocVg		BM	
Region	Depth	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Ngari (arid)	5 cm	0.05	0.05	0.06	0.06	0.1	0.11	0.03	0.03
	10 cm	0.02	0.07	0.04	0.04	0.07	0.1	0.06	0.06
	20 cm	0.07	0.09	0.05	0.05	0.09	0.09	0.06	0.06
	40 cm	0.05	0.09	0.07	0.07	0.11	0.15	0.07	0.07
Naqu (semi-arid)	5 cm	0.18	0.25	0.04	0.05	0.14	0.17	0.06	0.07
	10 cm	0.08	0.1	0.03	0.03	0.15	0.16	0.13	0.15
	20 cm	0.05	0.05	0.03	0.04	0.08	0.09	0.07	0.08
	40 cm	0.09	0.11	0.02	0.02	0.16	0.17	0.05	0.06
	50 cm	0.14	0.15	0.04	0.04	0.2	0.21	0.05	0.06
Maqu (sub-humid)	5 cm	0.27	0.28	0.02	0.03	0.04	0.05		
	10 cm	0.2	0.21	0.03	0.05	0.06	0.08		
	20 cm	0.14	0.15	0.07	0.08	0.07	0.09		
	40 cm	0.07	0.09	0.06	0.06	0.08	0.09		
	80 cm	0.06	0.07	0.05	0.06	0.08	0.1		

Table A1.7 Values of the absolute bias of the estimated SWRCs based on PTFs without the combination of the BD scheme, with the measurements at 5 cm in the three climate zones.

PTFs	Ngari (arid)	Naqu (semi-arid)	Maqu (sub-humid)
	Absolute Bias (m^3 m^{-3})	Absolute Bias (m^3 m^{-3})	Absolute Bias (m^3 m^{-3})
Cosby et al., 1984	0.03	0.09	0.23
Cosby et al., 1984	0.03	0.09	0.20
Saxton et al., 1986	0.09	0.08	0.21
Campbell and Shiosawa, 1992	0.06	0.11	0.17
Saxton et al., 2006	0.05	0.07	0.07
Rawls and Brakenssiek 1985	0.06	0.16	0.30

Wösten et al., (Class PTF)	0.05	0.05	0.36
Vereecken et al., 1989	0.02	0.11	0.36
Wösten et al., 1999	0.05	0.07	0.25
Rosetta1-H3	0.05	0.14	0.26
Rosetta3-H3	0.04	0.12	0.40
Weynants et al. 2009	0.05	0.07	0.23

Table A1.8 Estimated values of the parameters of the CH model based on PTFs on the three climate zones of the TP. The θ_s aligned with each PTFs was estimated in terms of the estimation scheme parameterized in PTFs. The values of θ_s listed in the last line was calculated from the in situ BD measurements.

		Ngari (arid)				Naqu (semi-arid)					Maqu (sub-humid)				
PTFs	Parameter	5 cm	10 cm	20 cm	40 cm	5 cm	10 cm	20 cm	40 cm	50 cm	5 cm	10 cm	20 cm	40 cm	80 cm
	b(-)	0.30	0.30	0.29	0.30	0.28	0.29	0.27	0.26	0.26	0.22	0.22	0.22	0.22	0.23
Cosby et al., 1984	φ_s (cm)	-6.22	-5.93	-8.67	-8.37	-7.31	-7.23	-8.84	-8.18	-10.97	-35.30	-33.92	-33.55	-31.16	-29.42
	θ_s ($m^3 m^{-3}$)	0.38	0.38	0.39	0.39	0.39	0.39	0.39	0.39	0.40	0.46	0.45	0.45	0.45	0.45
	λ(-)	0.30	0.31	0.29	0.30	0.28	0.29	0.27	0.26	0.26	0.22	0.22	0.22	0.22	0.23
Cosby et al., 1984	φ_s (cm)	-6.77	-6.43	-10.18	-9.86	-8.09	-8.05	-9.95	-8.94	-12.49	-51.00	-48.74	-47.99	-43.64	-41.15
	θ_s ($m^3 m^{-3}$)	0.38	0.38	0.39	0.39	0.39	0.39	0.40	0.39	0.40	0.46	0.46	0.46	0.46	0.45
Saxton et al., 1986	b(-)	0.26	0.27	0.26	0.27	0.24	0.25	0.24	0.23	0.24	0.28	0.27	0.27	0.27	0.27

	φ_s (cm)	-164.28	-10.67	-14.66	-18.59	-10.27	-9.94	-10.47	-8.87	-13.25	-59.90	-55.70	-55.04	-50.33	-50.61
	θ_s (m^3 m^{-3})	0.31	0.31	0.34	0.32	0.35	0.34	0.37	0.37	0.37	0.44	0.44	0.44	0.44	0.43
	λ(-)	0.44	0.47	0.40	0.43	0.36	0.41	0.34	0.33	0.34	0.17	0.17	0.17	0.17	0.19
Campbell and Shiosawa, 1992	φ_s (cm)	-10.13	-9.49	-12.20	-12.85	-5.01	-7.26	-9.94	-15.21	-19.03	-3.42	-7.63	-17.68	-28.02	-37.95
	θ_s (m^3 m^{-3})	0.41	0.38	0.39	0.41	0.62	0.46	0.38	0.29	0.20	0.71	0.64	0.54	0.47	0.44
	b(-)	0.32	0.45	0.42	0.40	0.20	0.20	0.25	0.26	0.30	0.26	0.27	0.27	0.29	0.31
Saxton et al., 2006	φ_s (cm)	-2.60	-7.02	-11.42	-12.62	-1.68	-1.77	-3.91	-2.87	-8.49	-20.46	-21.83	-29.39	-37.87	-44.39
	θ_s (m^3 m^{-3})	0.43	0.43	0.42	0.43	0.65	0.63	0.46	0.44	0.47	1.00	0.85	0.69	0.53	0.49
θ_s estimated with the BD scheme (m^3 m^{-3})		0.41	0.38	0.39	0.41	0.62	0.46	0.38	0.29	0.20	0.71	0.64	0.54	0.47	0.44

Table A1.9 Estimated values of the parameters of the VG model based on PTFs on the three climate zones of the TP. The θ_s aligned with each PTFs was estimated in terms of the estimation scheme parameterized in PTFs. The values of θ_s listed in the last line was calculated from the in situ BD measurements.

		Ngari (arid)				Naqu (semi-arid)					Maqu (sub-humid)				
PTFs	Parameters	5 cm	10 cm	20 cm	40 cm	5 cm	10 cm	20 cm	40 cm	50 cm	5 cm	10 cm	20 cm	40 cm	80 cm

Rawls and Brakenssiek 1985	θ_r (m^3 m^{-3})	0.04	0.04	0.05	0.04	0.03	0.04	0.05	0.05	0.06	0.04	0.04	0.05	0.05	0.05
	θ_s (m^3 m^{-3})	0.41	0.38	0.39	0.41	0.62	0.46	0.38	0.29	0.2	0.71	0.64	0.54	0.47	0.44
	a (cm^{-1})	0.11	0.1	0.08	0.09	0.13	0.14	0.1	0.06	0.04	0.06	0.06	0.04	0.03	0.03
	n (-)	2.16	2.21	2.01	2.03	2.07	2.31	2.18	2.07	2.1	1.45	1.48	1.48	1.5	1.55
Wösten et al., (Class PTF)	θ_r (m^3 m^{-3})	0.03	0.03	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.03	0.03	0.03	0.03	0.03
	θ_s (m^3 m^{-3})	0.4	0.4	0.4	0.37	0.44	0.44	0.44	0.39	0.39	0.4	0.4	0.4	0.37	0.37
	a (cm^{-1})	0.04	0.04	0.04	0.04	0.03	0.03	0.03	0.02	0.02	0.04	0.04	0.04	0.04	0.04
	n (-)	1.38	1.38	1.38	1.52	1.18	1.18	1.18	1.17	1.17	1.38	1.38	1.38	1.52	1.52
Vereecken et al., 1989	θ_r (m^3 m^{-3})	0.04	0.03	0.04	0.04	0.12	0.08	0.05	0.04	0.04	0.21	0.16	0.13	0.1	0.09
	θ_s (m^3 m^{-3})	0.37	0.35	0.37	0.37	0.56	0.41	0.35	0.28	0.22	0.61	0.55	0.47	0.42	0.4
	a (cm^{-1})	0.32	0.34	0.18	0.25	0.16	0.35	0.25	0.19	0.28	0.01	0.01	0.01	0.01	0.02
	n (-)	1.38	1.45	1.22	1.27	1.25	1.49	1.41	1.31	1.37	0.82	0.83	0.82	0.82	0.85
Wösten et al., 1999	θ_r (m^3 m^{-3})	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	θ_s (m^3 m^{-3})	0.55	0.36	0.37	0.38	0.57	0.41	0.35	0.28	0.21	0.64	0.57	0.48	0.42	0.4
	a (cm^{-1})	0.06	0.05	0.05	0.04	0.04	0.05	0.06	0.03	0.02	0.2	0.09	0.03	0.02	0.02
	n (-)	1.48	1.46	1.4	1.44	1.19	1.35	1.39	1.3	1.2	1.13	1.16	1.19	1.25	1.28
Rosetta1-H3	θ_r (m^3 m^{-3})	0.04	0.03	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.07	0.07	0.05	0.05	0.04

θ_s (m^3 m^{-3})	0.36	0.35	0.34	0.35	0.46	0.41	0.34	0.31	0.29	0.55	0.51	0.43	0.38	0.37
a (cm^{-1})	0.04	0.04	0.03	0.03	0.04	0.04	0.04	0.05	0.04	0	0.01	0.01	0.01	0.01
n (-)	2.26	2.3	1.95	2.07	1.84	2.41	2.13	1.84	2.45	1.73	1.7	1.67	1.61	1.54
θ_s estimated with the BD scheme (m^3 m^{-3})	0.41	0.38	0.39	0.41	0.62	0.46	0.38	0.29	0.20	0.71	0.64	0.54	0.47	0.44

Table A1.10 Values of the bias of the estimated SWRCs based on PTFs with the combination of the BD scheme and the measurements at 5 cm in the three climate zones.

PTFs	Ngari (arid)	Naqu (semi-arid)	Maqu (sub-humid)
	Absolute Bias (m^3 m^{-3})	Absolute Bias (m^3 m^{-3})	Absolute Bias (m^3 m^{-3})
Cosby et al., 1984	0.03	0.09	0.15
Cosby et al., 1984	0.04	0.09	0.12
Saxton et al., 1986	0.15	0.08	0.12
Campbell and Shiosawa, 1992	0.06	0.11	0.17
Saxton et al., 2006	0.05	0.10	0.18
Rawls and Brakenssiek 1985	0.06	0.16	0.30
Wösten et al., (Class PTF)	0.05	0.04	0.29
Vereecken et al., 1989	0.01	0.10	0.38
Wösten et al., 1999	0.05	0.07	0.22
Rosetta1-H3	0.06	0.14	0.20
Rosetta3-H3	0.06	0.12	0.16
Weynants et al. 2009	0.06	0.07	0.17

Table A1.11 Values of the biases of the estimated C_s based on the De Vries (1963) model and the measurements on the three climate zones. The upper part of the table lists bias values based on estimations without considering SOC impact, and the lower part of the *table lists bias values based on estimations with considering SOC impact in the Maqu and Naqu regions. The unit of listed values is MJ m^{-3} K^{-1}.*

Region	5 cm	10 cm	20 cm	40 cm	50 cm	80 cm
Ngari (arid)	-0.04	0.01	-0.05	0.00		
Naqu (semi-arid)	-0.29	0.00	0.22	0.31	0.14	
Maqu (sub-humid)	-0.10	-0.02	0.00	0.10		0.13
Naqu (semi-arid) +SOC	-0.22	0.11	0.26	0.33	0.15	
Maqu (sub-humid) +SOC	0.00	0.08	0.10	0.16		0.19

Table A1.12 Comparisons of the mean derived FC and PWP based on SWRC-CH and SWRC-VG models using the various soil datasets, with the laboratory measurements. FC is the field capacity, and PWP is the permanent wilting point.

Region	Parameters	Measured	Tibet-Obs	FAO-UNESCO	HWSD	BNU	SoilGrids1km	SoilGrids250m
Ngari (arid)	FC (m^3 m^{-3})	0.20	0.26	0.46	0.43	0.46	0.37	0.41
	PWP (m^3 m^{-3})	0.10	0.08	0.26	0.21	0.22	0.16	0.18
Naqu (semi-arid)	FC (m^3 m^{-3})	0.28	0.27	0.44	0.41	0.44	0.44	0.45
	PWP (m^3 m^{-3})	0.18	0.10	0.26	0.20	0.20	0.21	0.22
Maqu (sub-humid)	FC (m^3 m^{-3})	0.68	0.56	0.44	0.41	0.46	0.51	0.48
	PWP (m^3 m^{-3})	0.26	0.24	0.26	0.21	0.23	0.27	0.26

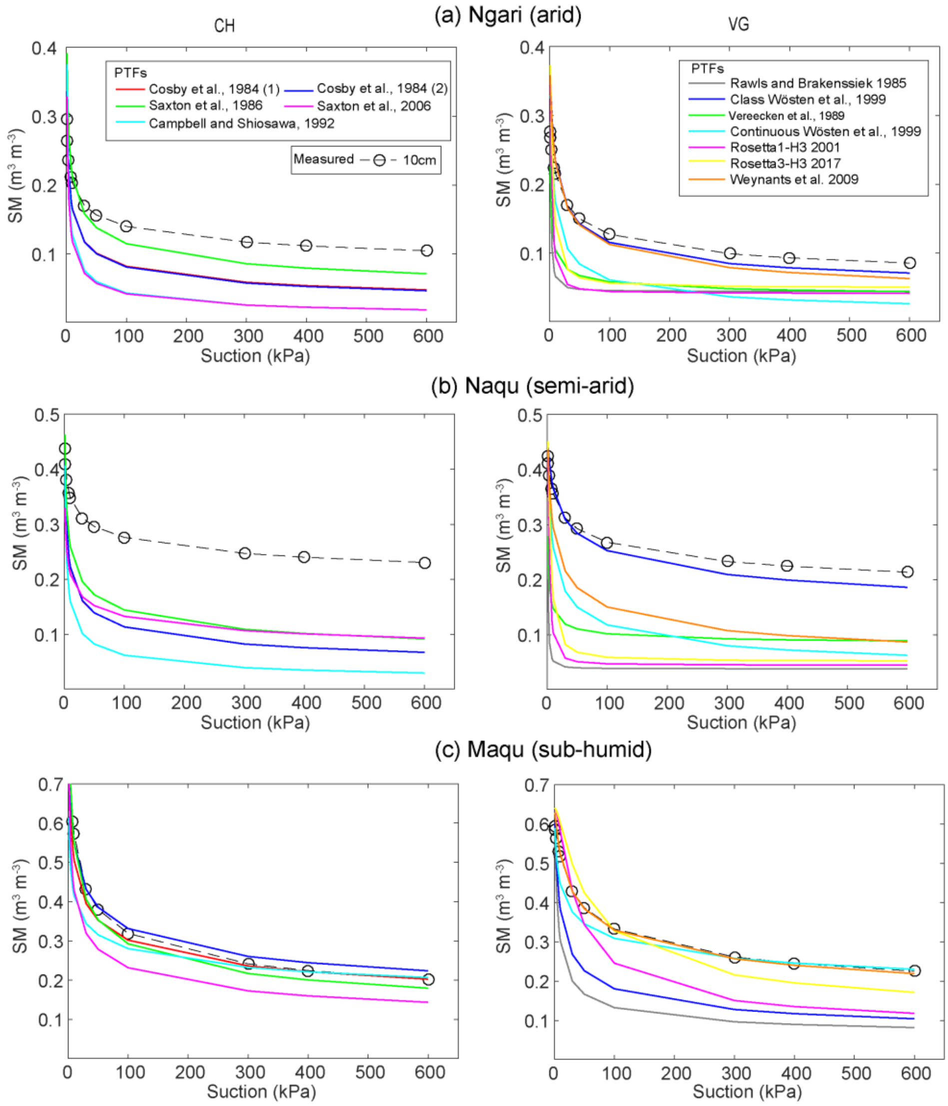


Figure A1.1 Comparisons between estimated SWRCs from PTFs combined with the BD scheme, and the measurement determined SWRCs at 10 cm for three climate zones. It is to note that the SWRC estimated from Vereecken et al. (1989) PTFs was out of range over the sub-humid zone and was removed (right figure in Figure A1.1-C).

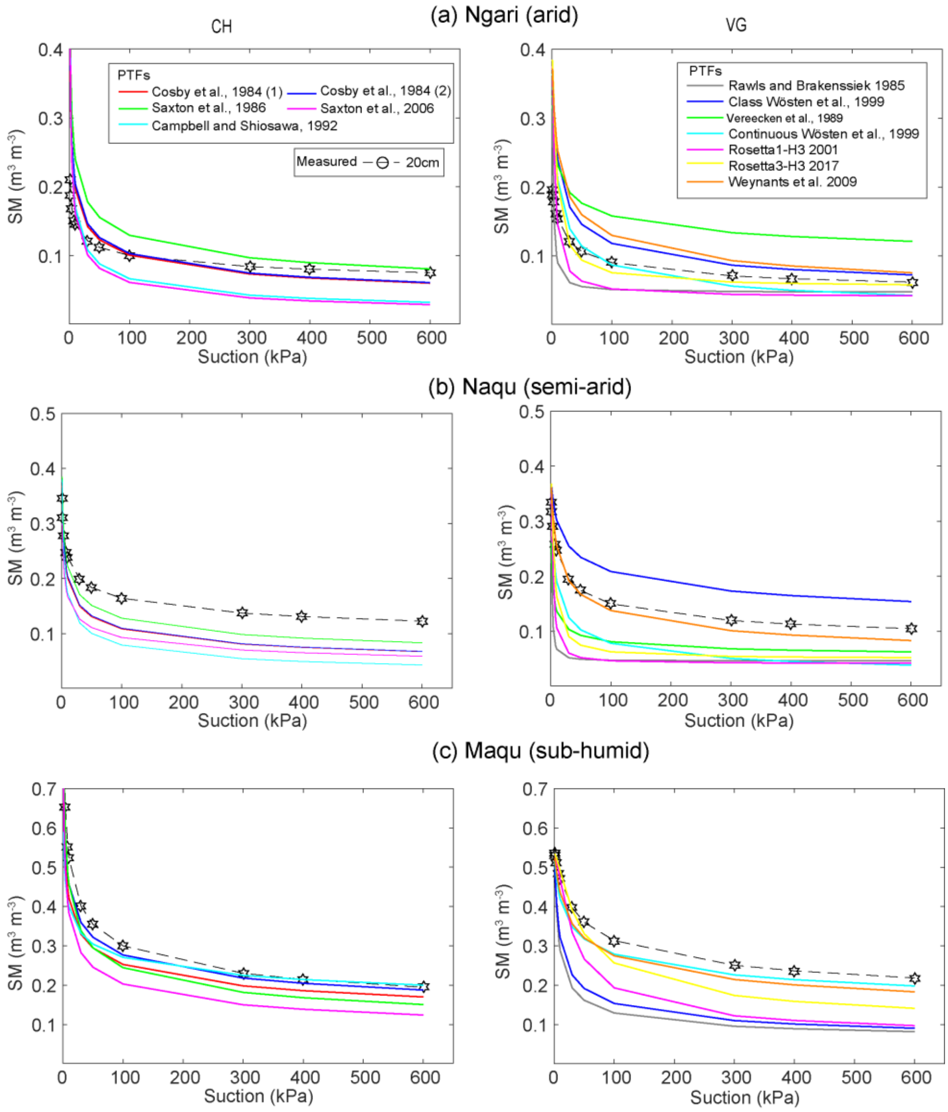


Figure A1.2 Comparisons between the estimated SWRCs from PTFs combined with the BD scheme, and the measurement-determined SWRCs at 20 cm in the three climate zones. Notably, the SWRC estimated from Vereecken et al. (1989) PTFs was beyond the range in the sub-humid zone and not considered (right figure in Figure A1.2-C).

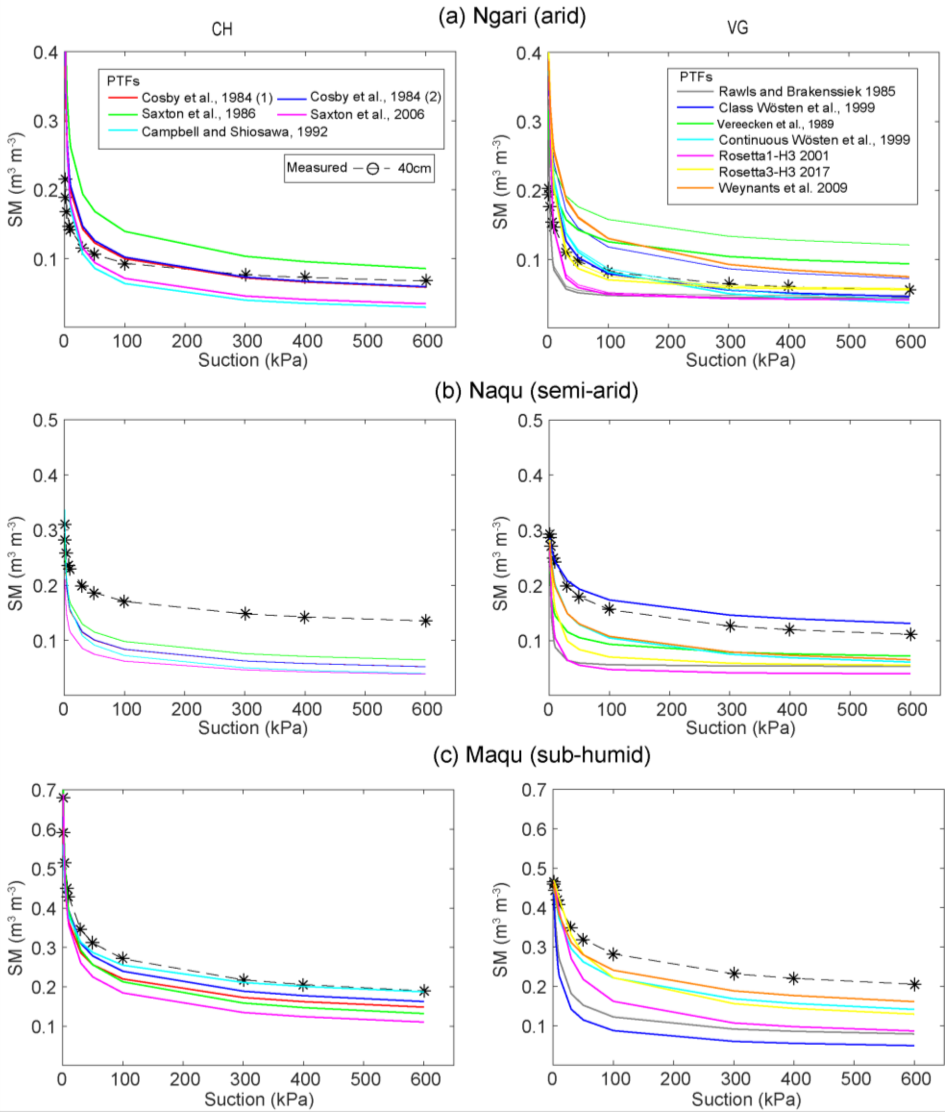


Figure A1.3 Comparisons between the estimated SWRCs from PTFs combined with the BD scheme, and the measurement-determined SWRCs at 40 cm in the three climate zones. Notably, the SWRC estimated from Vereecken et al. (1989) PTFs was beyond the range in the sub-humid zone and not considered (right figure in Figure A1.3-C).

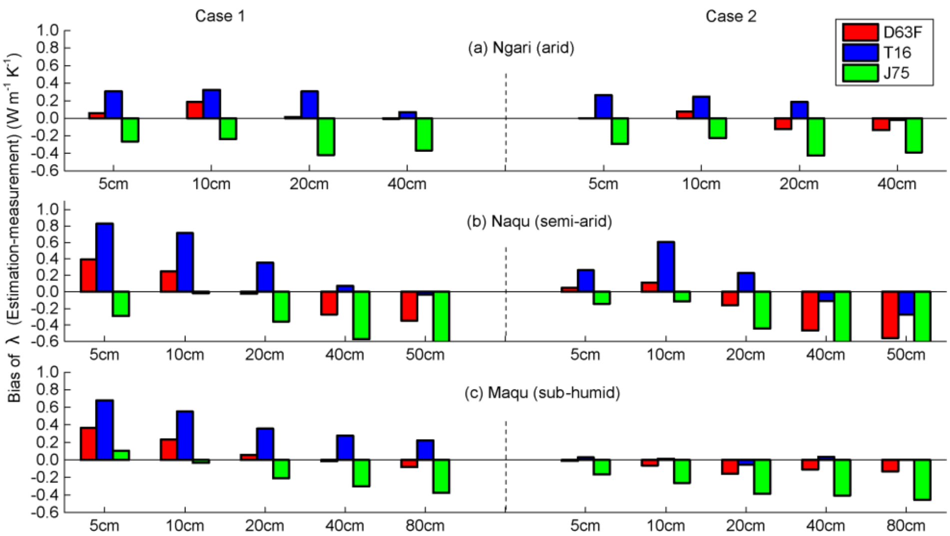


Figure A1.4 Biases of the estimated λ based on the D63F model, the T16 model and the J75 scheme combined with the Cosby-S scheme (Cosby PTFs) in the profile on the three climate zones with the measurements. Case 1 is the bias derived based on schemes not considering gravel impact parameterization on the arid and semi-arid zone and SOC impact parameterization on the semi-humid zone. Case 2 is the bias derived based on schemes considering the foregoing parameterizations.

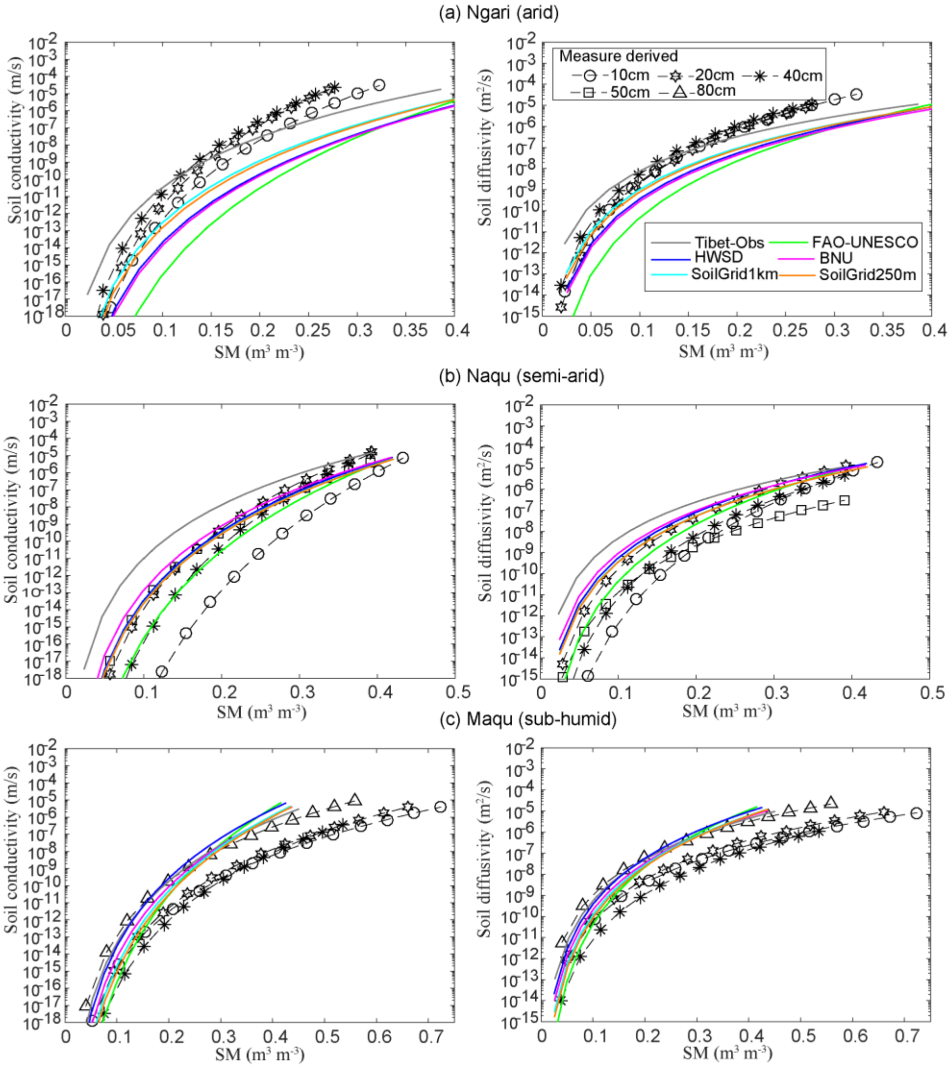


Figure A1.5 Comparisons of derived soil conductivity (K) and soil diffusivity (D) by the CH model based on the six soil datasets, with those derived from the laboratory measurements. Given the relatively homogenous soil profile obtained based on the existing datasets (please refer to Figure 2.12 in the text), the averaged K and D derived from existing datasets over the different depths were illustrated.

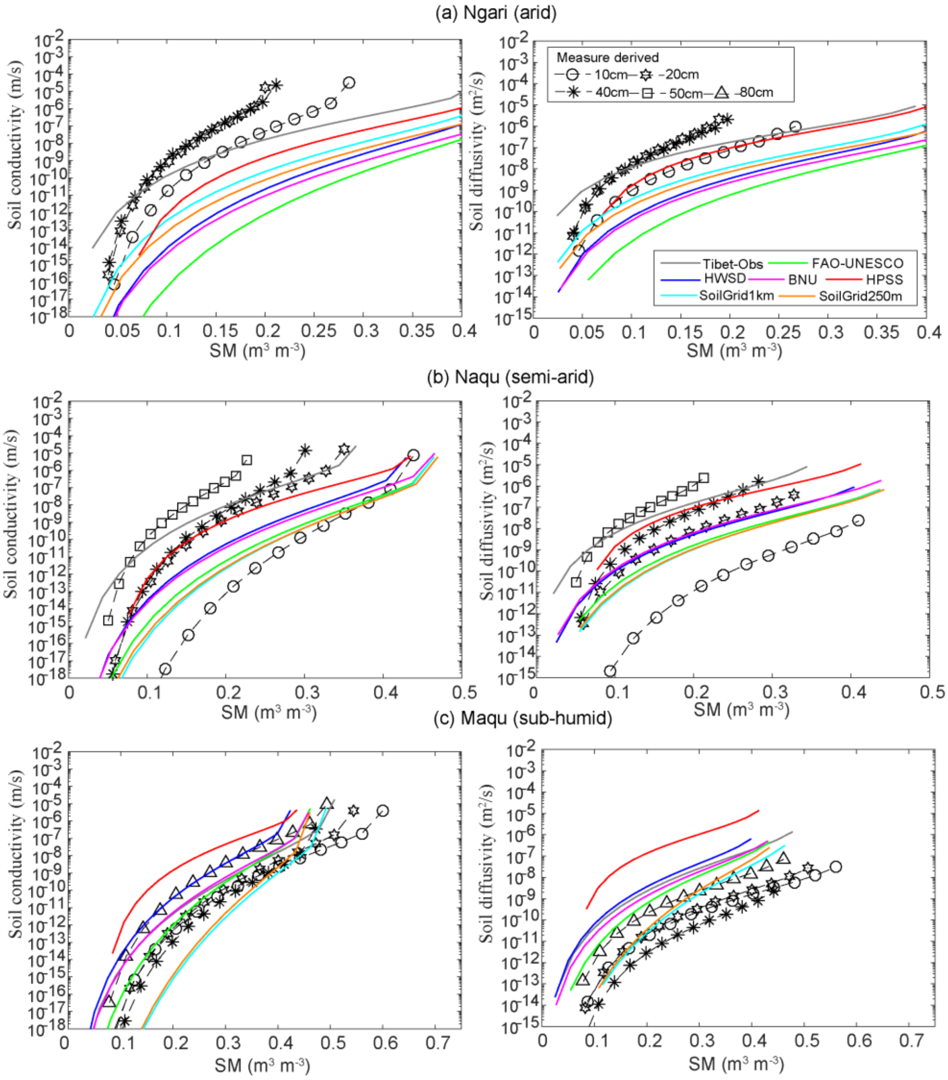


Figure A1.6 Comparisons of the derived soil conductivity (K) and soil diffusivity (D) by the VG model based on the seven soil datasets, with those derived from the laboratory measurements. Given the relatively homogenous soil profile obtained based on the existing datasets (please refer to Figure 2.12 in the text), the averaged K and D derived from existing datasets over the different depths were illustrated. The HPSS only provides values of the hydraulic parameters of the VG model.

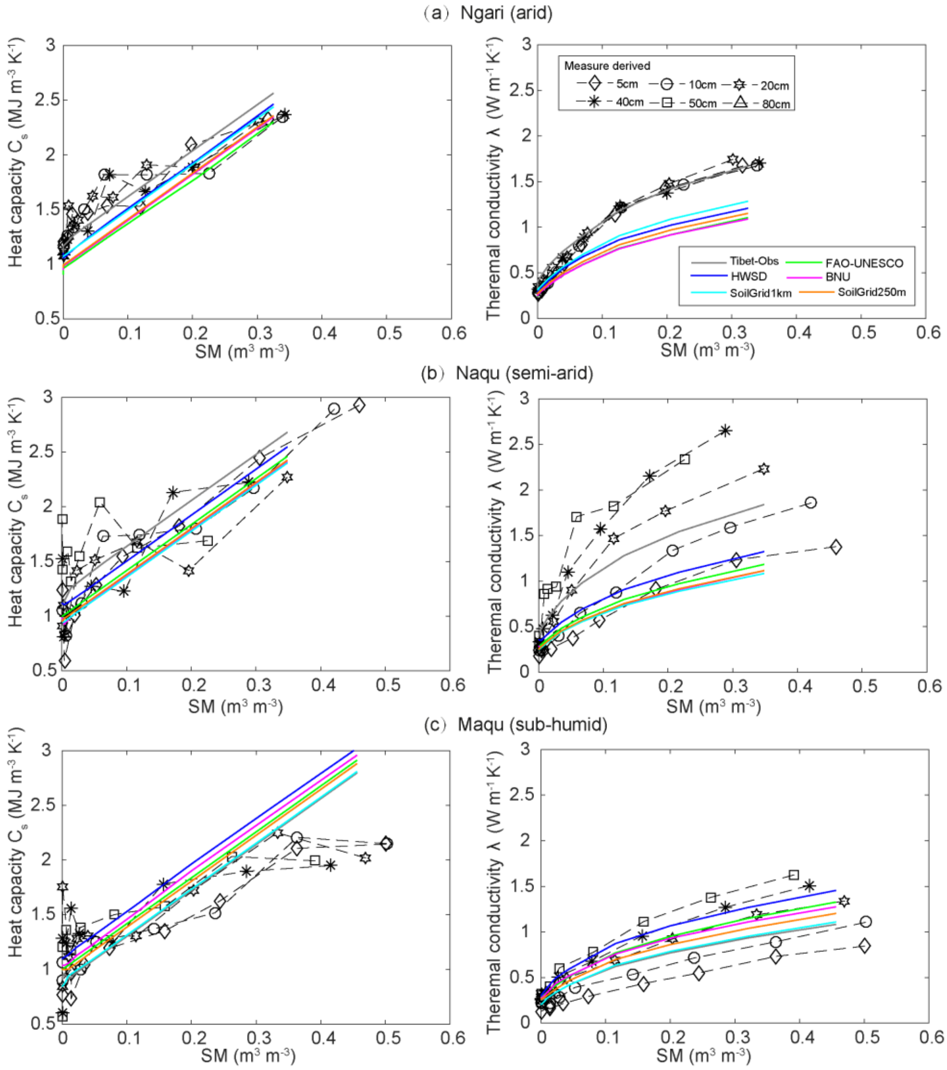


Figure A1.7 Comparisons of the derived C_s -SM (left panel) and λ-SM (right-panel) by the D63F model based on the various datasets, with the measurements. Given the relatively homogenous soil profile obtained based on the existing datasets (please refer to Figure 2.12 in the text), the averaged C_s-SM and λ-SM derived from the existing datasets over the different depths were illustrated.

Appendix B. Inputs of the ATS+AIEM+TVG Model

Table A2.1 Input parameters of the integrated ATS+AIEM+TVG model.

	Parameter name	Value	Reference
TVG: Vegetation part	Leaf area index (LAI)	MCD15A2H, Figure 3.2 in Chapter 3	
	Plant moisture content (kg kg^{-1})	0.59	Wang et al., 2018
	Leaves: disc radius (cm)	1.4	Dente et al., 2014
	Leaves: disc thickness (cm)	0.02	Dente et al., 2014
	Leaves: disc angular distribution	Random	Dente et al., 2014
ATS+AIEM: Soil part	Volumetric soil moisture and soil temperature	*In situ* measurements at 2.5, 5, 10, 20, 35 and 60 cm	Su et al., 2020
	Soil texture	*In situ* measurements	Zhao et al., 2018
	The standard deviation of the surface height (cm)	0.9	Dente et al., 2014
	Correlation length (cm)	9	Dente et al., 2014
	Autocorrelation function	Exponential	Dente et al., 2014
Sensor configuration	Incidence angle (°)	40	
	Frequency	1.4 GHz	
	Polarization	H and V	

Figure A2.1 Snapshots of the Maqu site (left, taken in October) and Naqu site (right, taken in Augusts) on the Tibetan Plateau, where the grassland is grazed by yarks.

B.1 Comparisons of T_B^p simulations with considering the impact of only h_{SS} and h_{SS}+h_{SV}

To quantify the change in simulated T_B^p due to the consideration of the roughness resulting from only topsoil structures (h_{SS}) and from both the topsoil structure and soil volume ($h_{SS} + h_{SV}$, current case in Chapter 3), we extended T_B^p simulations by only considering the former (i.e., h_{SS}) and compared results to current simulation results shown in Chapter 3. The conclusions in Chapter 3 show that the ATS model using the *in situ* SM at 2.5 cm (ATS-Lv2.5) can be applied during the whole study period except for the soil freeze-thaw transition period. Accordingly, the configuration of the ATS-Lv2.5 model was used here for investigations.

To only consider the impact of h_{SS} on T_B^p simulations, we set h_{SV} to zero. As described in Chapter 3, there is an exponential dependence of the roughness thickness on soil moisture and the dielectric constant of air as the upper boundary and that of the bulk soil as the lower boundary. To obtain the dielectric profile, we utilized the Fermi-function, which describes the probability of energy (dielectric in this case) distribution. The parameterization of k_{AS} was unchanged. With air being the upper boundary, namely $\varepsilon(z^* = 0) \approx \varepsilon_{air}$, we arbitrarily chose a constant set for z_{AS}, which replaced the h_{SV} in equation (3.7) in Chapter 3, and it referred to at which $z^* = 0$ for the case with only considering h_{SS}. The 5 cm of z_{AS} was chosen in the simulation experiments. Correspondingly, the depth at which the averaging procedure (please refer to equations (3.9, 3.10) in Chapter 3) applied was adjusted to $z_{avg}^* = z_{AS} + h_{SS}/2.0$, and the variation of layer thickness around the z_{avg}^{*} was assumed due to the physical geometric roughness (s) at the top, and that was $z_{ub}^{*} = z_{avg}^{*} - s$.

Figures A2.2 and A2.3 show that T_B^p simulated by the ATS-Lv2.5 model only considering h_{SS} are higher than those from baseline simulations, especially when the soil is wet (please refer to soil moisture in Figure 3.5 in Chapter 3). However, without considering the impact of h_{SV}, T_B^p especially under H polarization is still underestimated, and the underestimation is compensated by simulations considering the impacts of both h_{SS}+h_{SV}. This indicates again the necessity of the ATS model in L-band modeling, which incorporates the parameterization of dielectric roughness resulting from both the topsoil structure and soil volume in this study. The error metrics including R and RMSEs are adopted and the corresponding values are listed in Table A2.2 for comparisons.

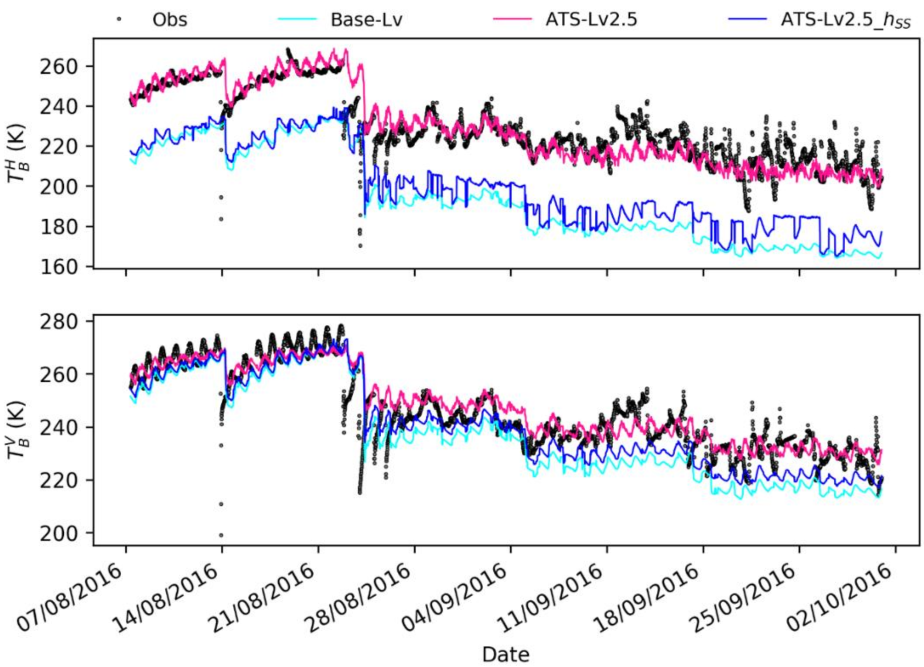


Figure A2.2 Comparisons of simulated T_B^p by the different models to the ELBARA-III T_B^p observations during the late-monsoon period. p denotes H or V polarization mode. Base-Lv refers to the experiment using the AIEM+TVG model in combination with the Lv model. ATS-Lv2.5 denotes the experiments considering $h_{SS}+h_{SV}$, and the ATS-Lv2.5_h_{SS} model considering h_{SS}. (As a matter of fact, the Base-Lv simulation results give the lower limit, which excludes the effect of h_{SV} in the ATS-Lv2.5 experiment simulations and is demonstrated by the ATS-Lv2.5_h_{SS} experiment simulations).

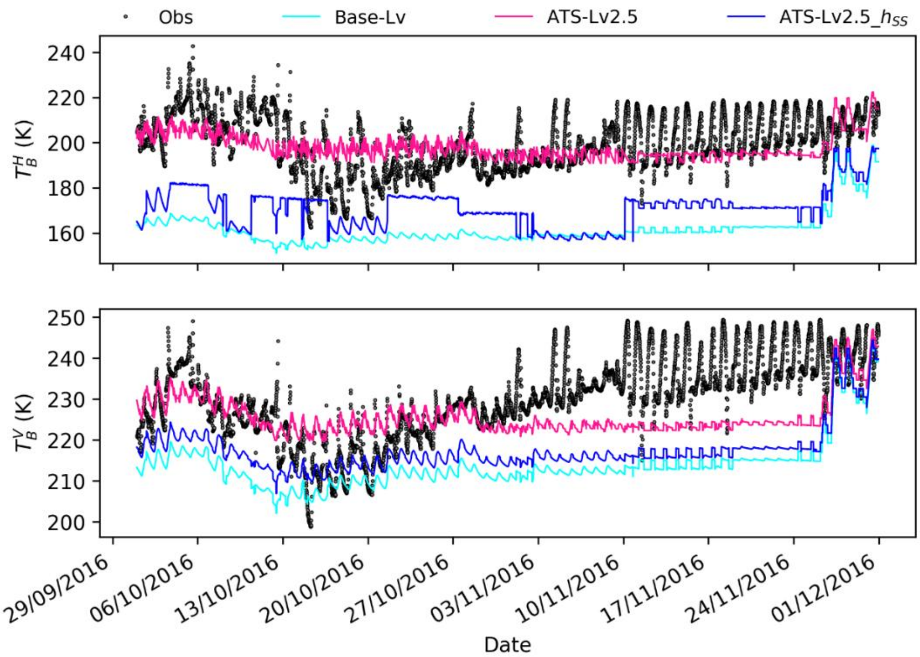


Figure A2.3 Same as Figure A2.2 but for the post-monsoon period.

Table A2.2 Comparisons of simulated T_B^p by the different models to the ELBARA-III observations. R is the Pearson correlation coefficient and the unit of RMSE is K.

Polarization	Period Experiments	Late-monsoon R	Late-monsoon RMSE (K)	Post-monsoon R	Post-monsoon RMSE (K)
H	Base-Lv	0.90	37.6	0.36	39.4
	ATS-Lv2.5	0.89	9.1	0.29	11.9
	ATS-Lv2.5_h_{SS}	0.88	31.4	0.3	30.9
V	Base-Lv	0.88	12.1	0.54	18.1
	ATS-Lv2.5	0.87	7.5	0.26	10.5
	ATS-Lv2.5_h_{SS}	0.88	9.4	0.48	15.1

B.2 T_B^p simulations under incidence angles of 50° and 60°

We consider the mechanism of the surface-radiation interaction has less to do with the footprint. The bigger footprint may average out certain surface irregularities, however, it does not alter the wavelength scale interactions between the waves and the surface features. Nevertheless, we tried to demonstrate the impact of the footprint size on T_B^p modeling using the measured multi-incidence angular data by comparisons to the baseline and ATS-based simulation results.

The configuration of the ATS-Lv2.5 model was used, please refer to the reason described in section B.1. Accordingly, we extended simulations using the Base-LV model and ATS-Lv2.5 model, respectively, at incidence angles of 50° and 60° as an example. Here we only focused on model comparisons during the late-monsoon period because of the limitation of the current ATS model used during the post-monsoon period (please refer to discussions in section 3.5.1 in Chapter 3), in which the soil surface starts to experience freeze-thaw processes.

Figure 3.7 in Chapter 3 and Figures A2.4-1 and A2.5-1 below show that the underestimation of T_B^H (≈ 30-50 K) by the baseline model at incidence angles of 40°, 50° and 60° respectively is obviously compensated by integrating the ATS model. Accordingly, it is observed that the integration of the ATS model reduces the impact of the surface roughness on emission modeling at different incidence angles under H polarization. Regarding the imperfect capture of the explicitly measured diurnal variations, this consideration is beyond the scope of this study, because, generally, apart from the soil part, which is mainly focused in this chapter, the vegetation part such as the vegetation water content, vegetation growing/senescing, vegetation intercepted water also contribute to T_B^p variations as well as the surface runoff after rainfall events.

Figure 3.7 in Chapter 3 shows that the consideration of the ATS model improves $T_B^V(40°)$ simulations during the late-monsoon period. Figure A2.4-2 below also shows a slight improvement of $T_B^V(50°)$ simulation with the ATS-based model compared to the baseline simulation results. Although Figure A2.5-2 below shows $T_B^V(60°)$ overestimated by both the baseline and ATS-based models, the ATS-based model attempts to decrease the emissivity and led to the simulated $T_B^V(60°)$ closer to the observations than those simulated by the baseline model. The effect of the Brewster angle might account for the aforementioned description (Shi et al., 2002). The T_B^V simulations are observed exhibiting less sensitivity to the surface roughness especially under large incidence angles (e.g., 50°-60°). As such, factors rather than the surface roughness resulting in the overestimations of $T_B^V(60°)$ (Figure A2.5-2) is also beyond the scope of this study.

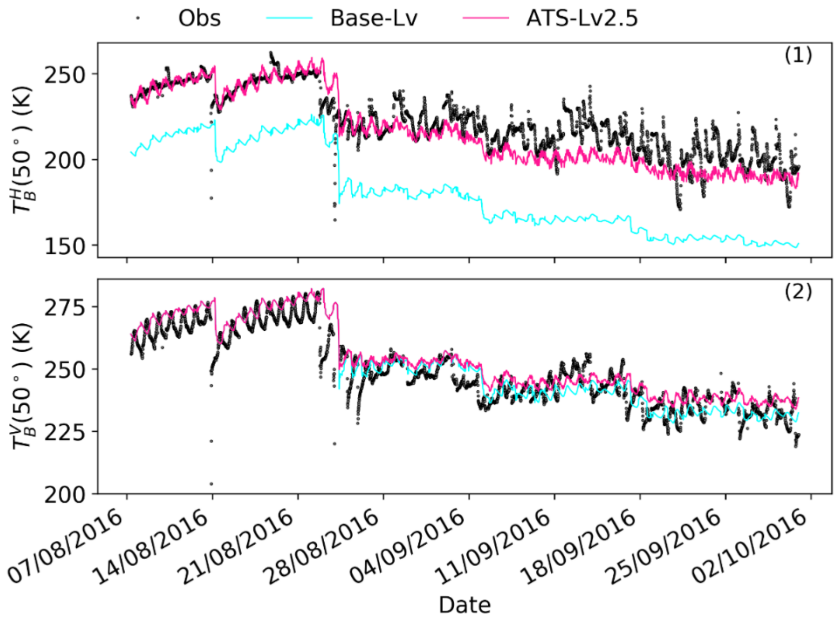


Figure A2.4 Comparisons of simulated $T_B^p(50^\circ)$ by the different models to the ELBARA-III $T_B^p(50^\circ)$ observations during the late-monsoon period. p denotes H or V polarization mode. Base-Lv refers to the experiment using the AIEM+TVG model in combination with the Lv model. ATS-Lv2.5 denotes the experiments using the ATS+AIEM+TVG model in combination with the Lv model and soil moisture at 2.5 cm for the calculation of the dielectric constant of bulk soil.

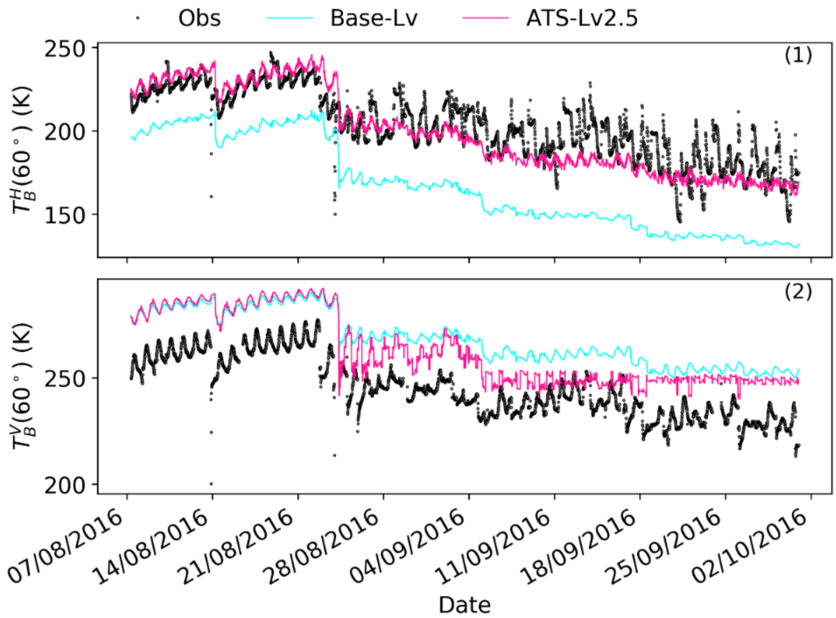


Figure A2.5 Same as Figure A2.4 but for $T_B^p(60^\circ)$.

Appendix C. Community Land Model and Retrieval Results

The modified Richards equation is formulated in equation (A2.1), which maintains the hydrostatic equilibrium soil moisture distribution,

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[k\left(\frac{\partial(\psi - \psi_E)}{\partial z}\right)\right] - Q \tag{A2.1}$$

where θ is the volumetric soil water content. z is the height in the soil column (mm). k is hydraulic conductivity (mm s^{-1}). ψ is the soil matrix potential and ψ_E is the equilibrium soil matric potential (mm). Q is the soil sink term. ψ_E is linked to a constant hydraulic potential (C) above the water table z_∇ (unit: m),

$$C = \psi_E + z = \psi_{sat} + z_\nabla \quad \text{(A2.2)}$$

The hydraulic properties of soil are assumed to be a weighted combination of the mineral properties, which are determined according to the percentages of the sand and the clay (i.e., %sand, %clay) investigated by Campbell (1974) and Cosby et al. (1984), and the soil organic matter investigated by Lawrence and Slater (2008).

The hydraulic conductivity k_i is defined at the depth of the interface of the two adjacent layers $z_{h,i}$ and is a function of the saturated hydraulic conductivity $k_{sat}[z_{h,i}]$, the liquid volumetric soil moisture of the two layers θ_i and θ_{i+1} and an ice impedance factor Θ_{ice}.

$$k[z_{h,i}] = \begin{cases} \Theta_{ice} k_{sat}[z_{h,i}][\frac{0.5(\theta_i + \theta_{i+1})}{0.5(\theta_{sat,i} + \theta_{sat,i+1})}]^{2B_i+3} & 1 \\ \Theta_{ice} k_{sat}[z_{h,i}](\frac{\theta_i}{\theta_{sat,i}})^{2B_i+3} & i = N_{levoi} \end{cases} \quad \text{(A2.3)}$$

$$\leq i \leq N_{levsoi}$$

The ice impedance factor is a function of the ice content, which is used to quantify the increased tortuosity of the water flow when the part of the pore space is filled with ice. The power law of $\Theta_{ice} = 10^{-\Omega F_{ice}}$ where $\Omega = 6$ and $F_{ice} = \frac{\theta_{ice}}{\theta_{sat}}$ is the ice-filled fraction of the pore space. θ_{sat} is the saturated soil moisture (usually regarded as soil porosity in hydrology). Considering organic matter impact, θ_{sat} is calculated with equation (A2.4),

$$\theta_{sat,i} = \left(1 - f_{om,i}\theta_{sat,min,i}\right) + f_{om,i}\theta_{sat,om} \quad \text{(A2.4)}$$

where $f_{om,i}$ is the soil organic matter fraction using $f_{om} = \rho_{om}/130.0$, in which ρ_{om} is soil organic matter density. $\theta_{sat,om}$ is the porosity of organic matter and set to 0.88 (Letts et al., 2000). The porosity of mineral soils $\theta_{sat,min,i}$ is computed by

$$\theta_{sat,min,i} = 0.489 - 0.001268(\%\text{sand})_i \qquad \text{(A2.5)}$$

The soil matrix potential ψ (mm) is defined as,

$$\psi_i = \psi_{sat,i}(\theta_i / \theta_{sat,i})^{-B_i} \geq -1 \times 10^8 \qquad \text{(A2.6)}$$

where ψ_s is the saturated matric potential (mm). B_i is the pore size distribution index (dimensionless).

$$\psi_{sat,i} = (1 - f_{om,i})\psi_{sat,min,i} + f_{om,i}\psi_{sat,om} \qquad \text{(A2.7)}$$

$$B_i = (1 - f_{om,i})B_{min,i} + f_{om,i}B_{om}$$

where $\psi_{sat,om}$ is the saturated organic matter matric potential and valued at $-10.2\ mm$ (Letts et al., 2000). $B_{om} = 6.1$ (Letts et al., 2000). The saturated mineral soil matric potential $\psi_{sat,min,i}$ and $B_{min,i}$ are calculated by equation (A2.8),

$$\psi_{sat,min,i} = -10.0 \times 10^{1.88 - 0.0131(\%sand)_i} \qquad \text{(A2.8)}$$

$$B_{min,i} = 2.91 + 0.159(\%sand)_i$$

Due to the organic matter fraction less than the threshold of 50%, no connected flow pathways consisting of organic materials are assumed to exist and span the soil space. Instead, unconnected flow pathways that pass through the mineral and organic components in series are assumed and the corresponding saturated hydraulic conductivity is calculated by equation (A2.9),

$$k_{sat,uncon} = f_{uncon}\left(\frac{1 - f_{om}}{k_{sat,mim}} + \frac{f_{om}}{k_{sat,om}}\right) \qquad \text{(A2.9)}$$

where $f_{uncon,i}$ is one when organic matter density is less than 50%. $k_{sat,om}$ is a saturated hydraulic conductivity of organic matter and valued at 0.002 mm/s (Letts et al., 2000). Saturated hydraulic conductivity of mineral soils $k_{sat,mim}$ (mm/s) depends on soil texture (Cosby et al., 1984) as

$$k_{sat,mim}[z_{h,i}] = 0.0070556 \times 10^{-0.884+0.0153(\%sand)} \quad \text{(A2.10)}$$

The bulk soil layer saturated hydraulic conductivity is then computed as

$$k_{sat}[z_{h,i}] = f_{uncon,i}k_{sat,uncon}[z_{h,i}] + (1 - f_{uncon,i})k_{sat,om}[z_{h,i}] \quad \text{(A2.11)}$$

Parameterization of soil heat transfer and thermal properties can be found in section A.5.

Table A3.1 Values of the retrieved soil properties in the third layer (i.e., 11.89 cm) in the two data assimilation experiments with the assimilation of SMAP T_B^H and the ensemble size of 30, and the derived standard deviation values.

Soil property	True	$\boldsymbol{z}^b$	$\boldsymbol{z}^b$_std	Only_Para		Joint_Updt	
				$\boldsymbol{z}^a$	$\boldsymbol{z}^a$_std	$\boldsymbol{z}^a$	$\boldsymbol{z}^a$_std
Sand fraction (%)	43.71	46.06	1.01	51.57	1.07	47.72	0.99
Clay fraction (%)	9.30	17.62	1.21	12.96	0.97	13.56	0.97
Organic matter density (kg/m^3)	9.61	17.22	0.47	7.86	0.26	7.18	0.24

Table A3.2 Values of RMSE of the retrieve soil properties in the third layer (i.e., 11.89 cm) in the two data assimilation experiments with the assimilation of SMAP T_B^H and the ensemble size of 30, and the derived standard deviation values.

Soil property	$\boldsymbol{z}^b$ _RMSE	Only_Para		Joint_Updt	
		$\boldsymbol{z}^a$ _RMSE	Reduction	za_RMSE	Reduction

Sand fraction (%)	7.96	7.93	0.4%	4.13	48.1%
Clay fraction (%)	8.4	3.79	54.8%	4.37	48.0%
Organic matter density (kg/m^3)	1.87	1.77	5.3%	2.44	Ineffective

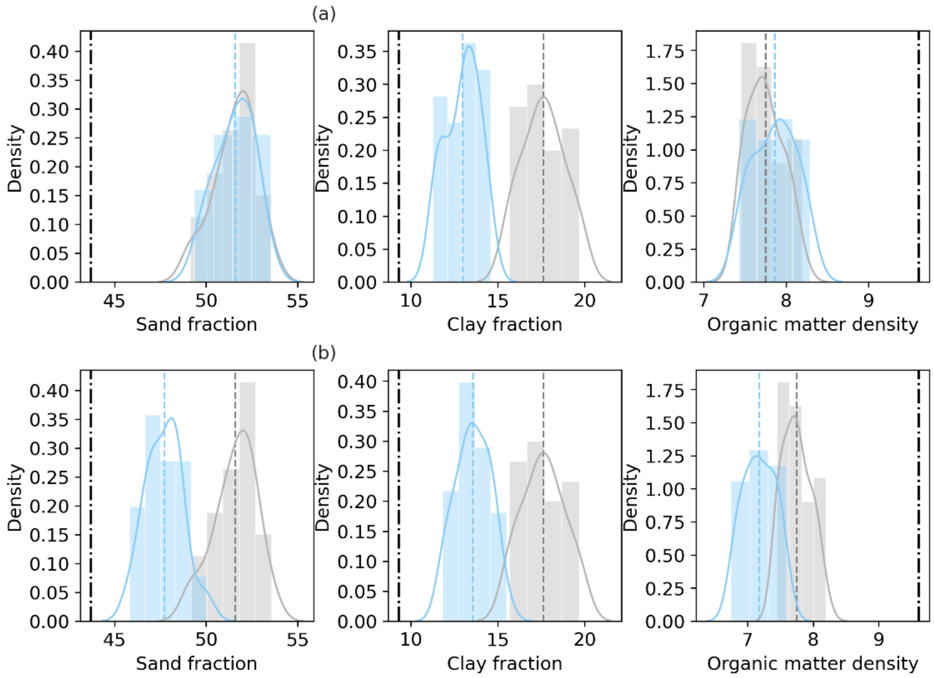


Figure A3.1 Prior and posterior distributions of the sand fraction (%), clay fraction (%) and organic matter density (kg/m^3) in the third layer (i.e., 11.89 cm). Grey denotes the prior and light blue denotes the posterior, and the black dashed-dotted line refers to the laboratory measurements. (a) is for Only_Para experiment and (b) is for Joint_Updt experiment by assimilating SMAP T_B^H with an ensemble size of 30.

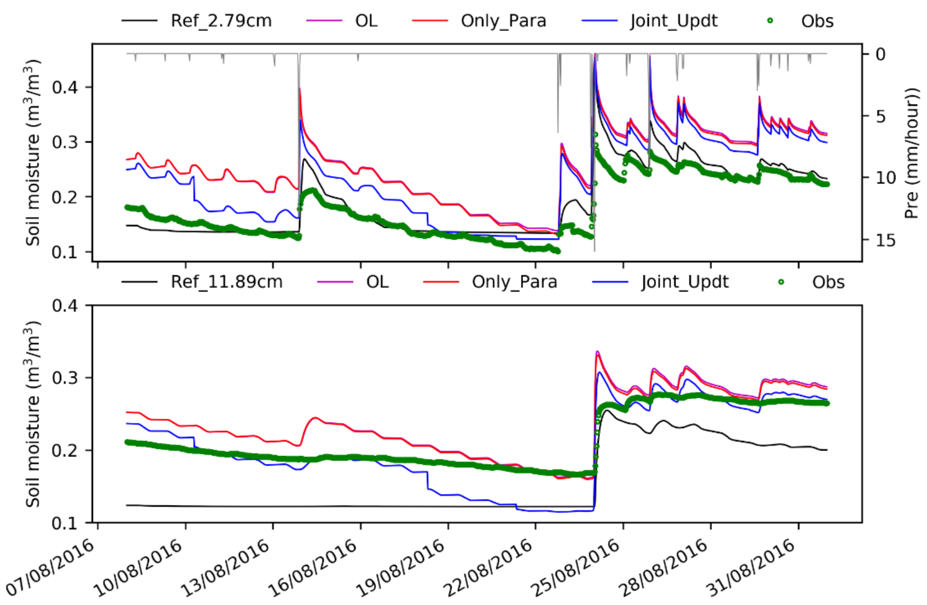


Figure A3.2 Soil moisture time series of the second (2.79 cm) and fourth (11.89 cm) layers in the reference (Ref) experiment, open loop (OL) experiment, the experiment with updating only the soil properties (Only_Para) and the experiment with updating both the soil properties and soil moisture (Joint_Updt) based on the assimilation of SMAP T_B^V during the period from 07/08/2016 to 08/31/2016.

List of figures

List of tables

List of symbols

Symbol	Name	Unit
θ	Volumetric soil moisture content	$m^3\ m^{-3}$
θ_s	Saturated soil moisture content	$m^3\ m^{-3}$
θ_r	Residual soil moisture content	$m^3\ m^{-3}$
φ	Soil matric potential	cm or mm
φ_s	Soil matric potential at saturation	cm or mm
B	The pore size distribution index	-
b	The pore size distribution index	-
n	A shape parameter	-
α	A parameter corresponding approximately to the inverse of the air-entry value	1/cm
K	Soil hydraulic conductivity	m/s
K_s	Saturated hydraulic conductivity	mm/s or m/s
D	Soil diffusivity	m^2/s
C_s	Soil heat capacity	$MJ\ m^{-3}\ K^{-1}$
λ	Soil thermal conductivity	$W\ m^{-1}\ K^{-1}$
T_B^p	Brightness temperature, with p = H or V polarization	K
T_{eff}	Soil effective temperature	K
λ_0	Wavelength	cm
h	The roughness thickness	cm
s	The standard deviation of the surface height	cm
L	The correlation length of the surface height	cm
θ_i	Incidence angle	°
e_p	Emissivity	-
$S_h(z^*)$	The probability density function of the dielectric roughness height	-
$V_h(z^*)$	The cumulative probability of density of $S_h(z^*)$	-
h_{SS}	The dielectric roughness thickness in the topsoil structures.	cm

h_{SV}	The dielectric roughness thickness in the soil volume.	cm
ε_{soil}	Soil dielectric constant	-
ε'_{soil}	The real part of the dielectric constant of the bulk soil	-
ε_{soil}''	The imaginary part of the soil dielectric constant	-
$\alpha(z^*)$	A rate parameter, namely, the power attenuation coefficient	1/m
Np	A polarization modulation parameter	-
k_{AS}	The steepness parameter	cm
$\varepsilon(z^*)$	The dielectric depth profile	-
z^*_{avg}	The average dielectric thickness	cm
z^*_{lb}	The lower boundary of the average dielectric surface	cm
f_i^p	The fraction of absorption	-
T_i	The temperature of the ith layer	K
R_s^p	The reflectivity of the smooth air-soil interface	-
δ_T	The thermal sampling depth derived from the Wilheit (1978) model	cm
δ_{PD}	The thermal sampling depth derived from the Lv (2014) model	cm
H_R	An empirical soil roughness parameter adopted in the zeroth-order Radiative Transfer Model	-
$\boldsymbol{x}$	Model state vector	
$\boldsymbol{M}$	A nonlinear model	
$\boldsymbol{y}_t^o$	A vector of observations at time t	
$\boldsymbol{H}$	The observation operator maps the state vector to the observations	
$\boldsymbol{\epsilon}$	Model errors	
$\boldsymbol{\varepsilon}$	Observation errors	
$\boldsymbol{R}$	Observation covariance matrix	
$\overline{\boldsymbol{x}}^b$	Mean of the background model state vector	
$\overline{\boldsymbol{y}}^b$	Mean of the observation vector	
$\boldsymbol{P}^b$	Background covariance matrix	
$\boldsymbol{P}^a$	Analysis covariance matrix	
$\boldsymbol{J}(\boldsymbol{x})$	The Kalman filter cost function	
$\boldsymbol{X}^b$	Background model state perturbation matrix	

S	Column space
$\tilde{S}$	Column space
$\boldsymbol{w}$	A vector in $\tilde{S}$
$\boldsymbol{I}$	Identity matrix
$\boldsymbol{Y}^b$	Background observation perturbation matrix

List of abbreviations

AIEM	Advanced integral equation model
ATS	Air-to-soil transition
BD	Dry soil bulk density
BNU dataset	Soil dataset from Shangguan et al. (2013) released by the Beijing Normal University
CMEM	Community microwave emission model
CH	Clapp and Hornberger (1978) model
CLM	Community land model
DasPy	The open-source multivariate land data assimilation framework
DMM	Soil dielectric mixing model
ECMWF	European Center for Medium-Range Weather Forecasts
ECV	Essential Climate Variable
FAO	Food and Agriculture Organisation
FC	Field capacity
FD	The particle diameter of fine minerals
GD	The particle diameter of gravel particles
GGF	Gravimetric gravel fraction
H	Sensible heat fluxes
HANTS	Harmonic analysis of the time series
H-TESSEL	Hydrology-tiled scheme for surface exchanges over land
HWSD	Harmonized World Soil Database
LAI	Leaf area index
LE	Latent heat fluxes
LETKF	Local ensemble transform Kalman filter
LSMs	Land surface models
MODIS	Moderate-resolution imaging spectroradiometer
NCAR	The National Center for Atmospheric Research
OL	Open loop
R	Pearson correlation coefficient
RMSE	Root mean square error
PTFs	Pedotransfer functions
RTM	Radiative transfer model
PWP	Permanent wilting point
SHP	Soil hydraulic properties

SM	Soil moisture
SMAP	Soil Moisture Active Passive
SMOS	Soil Moisture and Ocean Salinity
SMST	Soil moisture and soil temperature
ST	Soil temperature
STP	Soil thermal properties
SWRC	Soil water retention curve
Tair	Air temperature
Tibet-Obs	The Tibetan plateau observatory for soil moisture and soil temperature
TG	Ground surface temperature
TP	Tibetan Plateau
TVG	Tor Vergata discrete scattering model
USDA	The United States Department of Agriculture
VG	van Genuchten (1980) - Mualem (1976)
VGF	Volumetric gravel fraction
WoSIS	World Soil Information
4DEnVar	Four-dimensional ensemble variational

Bibliography

Angulo-Jaramillo, R., Vandervaere, J.-P., Roulier, S., Thony, J.-L., Gaudet, J.-P., & Vauclin, M. (2000). Field measurement of soil surface hydraulic properties by disc and ring infiltrometers: A review and recent developments. *Soil and Tillage Research*, *55*(1), 1-29. https://doi/org/10.1016/S0167-1987(00)00098-2

Assouline, S., & Or, D. (2013). Conceptual and Parametric Representation of Soil Hydraulic Properties: A Review. *Vadose Zone Journal*, *12*(4), 0. https://doi/org/10.2136/vzj2013.07.0121

Babaeian, E., Sadeghi, M., Jones, S. B., Montzka, C., Vereecken, H., & Tuller, M. (2019). Ground, Proximal and Satellite Remote Sensing of Soil Moisture. *Reviews of Geophysics*, *57*(2), 530-616. https://doi/org/10.1029/2018RG000618

Bai, X., Zeng, J., Chen, K. S., Li, Z., Zeng, Y., Wen, J., et al. (2019). Parameter Optimization of a Discrete Scattering Model by Integration of Global Sensitivity Analysis Using SMAP Active and Passive Observations. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(2), 1084-1099. https://doi/org/10.1109/TGRS.2018.2864689

Balland, V., & Arp, P. A. (2005). Modeling soil thermal conductivities over a wide range of conditions. *Journal of Environmental Engineering and Science*, *4*(6), 549-558. https://doi/org/10.1139/s05-007.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., & Betts, A. K. (2009). A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System. *Journal of Hydrometeorology*, *10*(3), 623-643. https://doi/org/10.1175/2008jhm1068.1

Bandara, R., Walker, J. P., & Rüdiger, C. (2014). Towards soil property retrieval from space: Proof of concept using in situ observations. *Journal of Hydrology*, *512*, 27-38. https://doi/org/10.1016/j.jhydrol.2014.02.031

Bandara, R., Walker, J. P., Rüdiger, C., & Merlin, O. (2015). Towards soil property retrieval from space: An application with disaggregated satellite observations. *Journal of Hydrology*, *522*, 582-593. https://doi/org/10.1016/j.jhydrol.2015.01.018

Bannister, R. N. (2017). A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *143*(703), 607-633. https://doi/org/10.1002/qj.2982

Bass, M., Van Stryland, E., William, R., & Wolfe, W. (1995). Optical properties of films and coatings. In *Handbook of Optics* (Vol. 11, pp. 42.49–42.14). NewYork opticalMcGraw-Hill.

Batjes, N. H., Ribeiro, E., van Oostrum, A., Leenaars, J., Heng, T., & Mendes de Jesus, J. (2017). WoSIS: Serving standardised soil profile data for the world. *Earth System Science Data, 9*(1), 1. https://doi/org/10.5194/essd-9-1-2017

Baver, L. D., Gardner, W. H., & Gardner, W. R. (1956). *Soil physics* (Vol. 489): Wiley New York.

Ben-Dor, E. (2019). *Passive remote sensing of soil throughout the years sensor platforms and data analyses*. World soils user consultation meeting, ESA-ESRIN, Frascati, Italy.

Benninga, H.-J. F., van der Velde, R., & Su, Z. (2018). *Uncertainty of Effective Roughness Parameters Calibrated on Bare Agricultural Land Using Sentinel-L SAR.* Paper presented at the IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium.

Bircher, S., Skou, N., & Kerr, Y. H. (2013). Validation of SMOS L1C and L2 Products and Important Parameters of the Retrieval Algorithm in the Skjern River Catchment, Western Denmark. *IEEE Transactions on Geoscience and Remote Sensing, 51*(5), 2969-2985. https://doi/org/10.1109/tgrs.2012.2215041

Bittelli, M., Campbell, G. S., & Tomei, F. (2015). *Soil Physics with Python: Transport in the Soil-Plant-Atmosphere System*: OUP Oxford.

Bouma, J. (1989). Using soil survey data for quantitative land evaluation. In *Advances in Soil Science* (pp. 177-213): Springer.

Bracaglia, M., Ferrazzoli, P., & Guerriero, L. (1995). A fully polarimetric multiple scattering model for crops. *Remote Sensing of Environment, 54*(3), 170-179. https://doi/org/10.1016/0034-4257(95)00151-4

Brooks, R. H., & Corey, A. T. (1964). Hydraulic properties of porous media and their relation to drainage design. *Transactions of the ASAE, 7*(1), 26-0028.

Brunke, M. A., Broxton, P., Pelletier, J., Gochis, D., Hazenberg, P., Lawrence, D. M., et al. (2016). Implementing and Evaluating Variable Soil Thickness in the Community Land Model, Version 4.5 (CLM4.5). *Journal of Climate, 29*(9), 3441-3461. https://doi/org/10.1175/jcli-d-15-0307.1

Brutsaert, W. (1966). Probability laws for pore-size distributions. *Soil science, 101*(2), 85-92.

Buckingham, E. (1907). Studies on the movement of soil moisture.

Buol, S. W., Southard, R. J., Graham, R. C., & McDaniel, P. A. (2011). *Soil genesis and classification*: John Wiley & Sons.

Burdine, N. (1953). Relative permeability calculations from pore size distribution data. *Journal of Petroleum Technology, 5*(03), 71-78.

Burke, E. J., Gurney, R. J., Simmonds, L. P., & Jackson, T. J. (1997). Calibrating a soil water and energy budget model with remotely sensed data to obtain quantitative information about the soil. *Water Resources Research, 33*(7), 1689-1697. https://doi/org/10.1029/97WR01000.

Camillo, P. J., O'Neill, P. E., & Gurney, R. J. (1986). Estimating soil hydraulic parameters using passive microwave data. *IEEE Transactions on Geoscience and Remote Sensing, GE-24*(6), 930-936. https://doi/org/10.1109/TGRS.2014.2368831.

Campbell, G., & Shiozawa, S. (1992). Prediction of hydraulic properties of soils using particle-size distribution and bulk density data. *Indirect methods for estimating the hydraulic properties of unsaturated soils. University of California, Riverside*, 317-328.

Campbell, G. S. (1974). A simple method for determining unsaturated conductivity from moisture retention data. *Soil science, 117*(6), 311-314. https://doi/org/10.1097/00010694-197406000-00001

Campbell, G. S. (1985). *Soil physics with BASIC: transport models for soil-plant systems*: Elsevier.

Carrassi, A., Bocquet, M., Bertino, L., & Evensen, G. (2018). Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*. https://doi/org/10.1002/wcc.535

Charoenhirunyingyos, S., Honda, K., Kamthonkiat, D., & Ines, A. V. M. (2011). Soil hydraulic parameters estimated from satellite information through data assimilation. *International Journal of Remote Sensing, 32*(23), 8033-8051. https://doi/org/10.1080/01431161.2010.532170

Chen, F., & Avissar, R. (1994). Impact of land-surface moisture variability on local shallow convective cumulus and precipitation in large-scale models. *Journal of Applied Meteorology, 33*(12), 1382-1401.

Chen, F., & Dudhia, J. (2001). Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly Weather Review, 129*(4), 569-585. https://doi/org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2

Chen, K.-S., Wu, T.-D., Tsang, L., Li, Q., Shi, J., & Fung, A. K. (2003a). Emission of rough surfaces calculated by the integral equation method with comparison to three-dimensional moment method simulations. *Geoscience and Remote Sensing, IEEE Transactions on, 41*(1), 90-101.

Chen, K. S., Tzong-Dar, W., Leung, T., Qin, L., Jiancheng, S., & Fung, A. K. (2003b). Emission of rough surfaces calculated by the integral equation method with comparison to three-dimensional moment method simulations. *IEEE Transactions on Geoscience and Remote Sensing, 41*(1), 90-101. https://doi/org/10.1109/tgrs.2002.807587

Chen, Y., Yang, K., Tang, W., Qin, J., & Zhao, L. (2012). Parameterizing soil organic carbon’s impacts on soil porosity and thermal parameters for Eastern Tibet grasslands. *Science China Earth Sciences, 55*(6), 1001-1011. https://doi/org/10.1007/s11430-012-4433-0

Choudhury, B., Schmugge, T., & Mo, T. (1982). A parameterization of effective soil temperature for microwave emission. *Journal of*

Geophysical Research: Oceans, *87*(C2), 1301-1304. https://doi/org/10.1029/JC087iC02p01301

Choudhury, B., Schmugge, T. J., Chang, A., & Newton, R. (1979). Effect of surface roughness on the microwave emission from soils. *Journal of Geophysical Research: Oceans*, *84*(C9), 5699-5706. https://doi/org/org/10.1029/JC084iC09p05699

Clapp, R. B., & Hornberger, G. M. (1978). Empirical equations for some soil hydraulic properties. *Water Resources Research*, *14*(4), 601-604. https://doi/org/10.1029/WR014i004p00601

Cooper, E., Blyth, E., Cooper, H., Ellis, R., Pinnington, E., & Dadson, S. J. (2020). Using data assimilation to optimize pedotransfer functions using large-scale in-situ soil moisture observations. *Hydrology and Earth System Sciences Discussion*. https://doi/org/10.5194/hess-2020-359

Corbari, C., Mancini, M., Li, J., & Su, Z. (2015). Can satellite land surface temperature data be used similarly to river discharge measurements for distributed hydrological model calibration? *Hydrological Sciences Journal*, *60*(2), 202-217. https://doi/org/10.1080/02626667.2013.866709

Cosby, B., Hornberger, G., Clapp, R., & Ginn, T. (1984). A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resources Research*, *20*(6), 682-690. https://doi/org/10.1029/WR020i006p00682

Côté, J., & Konrad, J.-M. (2005). A generalized thermal conductivity model for soils and construction materials. *Canadian Geotechnical Journal*, *42*(2), 443-458. https://doi/org/10.1139/t04-106

Dai, Y., Shangguan, W., Duan, Q., Liu, B., Fu, S., & Niu, G. (2013). Development of a China Dataset of Soil Hydraulic Parameters Using Pedotransfer Functions for Land Surface Modeling. *Journal of Hydrometeorology*, *14*(3), 869-887. https://doi/org/10.1175/JHM-D-12-0149.1

Dai, Y., Shangguan, W., Wei, N., Xin, Q., Yuan, H., Zhang, S., et al. (2019). A review of the global soil property maps for Earth system models. *Soil*, *5*(2), 137-158. https://doi/org/10.5194/soil-5-137-2019

Das, N. N., Entekhabi, D., & Njoku, E. G. (2010). An algorithm for merging SMAP radiometer and radar data for high-resolution soil-moisture retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(5), 1504-1512. https://doi/org/10.1109/TGRS.2010.2089526

De Lannoy, G. J. M., & Reichle, R. H. (2016). Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model. *Hydrology and Earth System Sciences*, *20*(12), 4895-4911. https://doi/org/10.5194/hess-20-4895-2016

de Rosnay, P., Drusch, M., & Sabater, J. ı. M. n. (2009). *Milestone 1 Tech Note - Part 1: SMOS Global Surface Emission Model* (0143-1161).

Retrieved from Shinfield Park, Reading: https://www.ecmwf.int/node/8923

de Rosnay, P., Muñoz-Sabater, J., Albergel, C., Isaksen, L., English, S., Drusch, M., & Wigneron, J.-P. (2020). SMOS brightness temperature forward modelling and long term monitoring at ECMWF. *Remote Sensing of Environment, 237*. https://doi/org/10.1016/j.rse.2019.111424

De Vries, D. A. (1963). Thermal properties of soils. *Physics of plant environment*, 211-234.

Della Vecchia, A. (2006). *Advances in modeling microwave interactions with vegetation for active and passive remote sensing.* PhD thesis, University of Rome Tor Vergata, Geoinformation PhD School,

Dente, L., Ferrazzoli, P., Su, Z., van der Velde, R., & Guerriero, L. (2014). Combined use of active and passive microwave satellite data to constrain a discrete scattering model. *Remote Sensing of Environment, 155*, 222-238. https://doi/org/10.1016/j.rse.2014.08.031

Dente, L., Vekerdy, Z., Wen, J., & Su, Z. (2012). Maqu network for validation of satellite-derived soil moisture products. *International Journal of Applied Earth Observation and Geoinformation, 17*, 55-65. https://doi/org/10.1016/j.jag.2011.11.004

Desroziers, G., Camino, J.-T., & Berre, L. (2014). 4DEnVar: link with 4D state formulation of variational assimilation and different possible implementations. *Quarterly Journal of the Royal Meteorological Society, 140*(684), 2097-2110. https://doi/org/10.1002/qj.2325

Dimitrov, M., Vanderborght, J., Kostov, K., Debecker, B., Schulze Lammers, P., Damerow, L., & Vereecken, H. (2015). Soil Hydraulic Parameters of Bare Soil Plots with Different Soil Structure Inversely Derived from L-Band Brightness Temperatures. *Vadose Zone Journal, 14*(8). https://doi/org/10.2136/vzj2014.09.0133

Dimitrov, M., Vanderborght, J., Kostov, K. G., Jadoon, K. Z., Weihermüller, L., Jackson, T. J., et al. (2014). Soil Hydraulic Parameters and Surface Soil Moisture of a Tilled Bare Soil Plot Inversely Derived from L-Band Brightness Temperatures. *Vadose Zone Journal, 13*(1), 1-18. https://doi/org/10.2136/vzj2013.04.0075

Dirmeyer, P. A. (2000). Using a global soil wetness dataset to improve seasonal climate simulation. *Journal of Climate, 13*(16), 2900-2922.

Dong, J., Steele-Dunne, S. C., Ochsner, T. E., & van de Giesen, N. (2016). Estimating soil moisture and soil thermal and hydraulic properties by assimilating soil temperatures using a particle batch smoother. *Advances in Water Resources*.

Douville, H., Chauvin, F., & Broqua, H. (2001). Influence of soil moisture on the Asian and African monsoons. Part I: Mean monsoon and daily precipitation. *Journal of Climate, 14*(11), 2381-2403.

Drusch, M. (2007). Initializing numerical weather prediction models with satellite - derived surface soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and the TMI soil moisture data set. *Journal of Geophysical Research: Atmospheres, 112*(D3).

Ellerbrock, R., Gerke, H., Bachmann, J., & Goebel, M. (2005). Composition of organic matter fractions for explaining wettability of three forest soils. *Soil Science Society of America Journal, 69*(1), 57-66.

Elrick, D. E., W. D. Reynolds, and K. A. Tan. (1989). Hydraulic Conductivity Measurements in the Unsaturated Zone Using Improved Well Analyses. *Groundwater Monitoring & Remediation, 8*(3), 193. https://doi/org/10.1111/j.1745-6592.1989.tb01162.x

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., et al. (2010). The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE, 98*(5), 704-716. https://doi/org/10.1109/JPROC.2010.2043918

Entekhabi, D., Yueh, S., O'Neill, P. E., Kellogg, K. H., Allen, A., Bindlish, R., et al. (2014). SMAP handbook–soil moisture active passive: Mapping soil moisture and freeze/thaw from space.

Eom, H. J., & Fung, A. K. (1984). A scatter model for vegetation up to Ku-band. *Remote Sensing of Environment, 15*(3), 185-200. https://doi/org/10.1016/0034-4257(84)90030-0

Escorihuela, M. J., Chanzy, A., Wigneron, J. P., & Kerr, Y. H. (2010). Effective soil moisture sampling depth of L-band radiometry: A case study. *Remote Sensing of Environment, 114*(5), 995-1001. https://doi/org/10.1016/j.rse.2009.12.011

Escorihuela, M. J., Kerr, Y. H., de Rosnay, P., Wigneron, J. P., Calvet, J. C., & Lemaitre, F. (2007). A Simple Model of the Bare Soil Microwave Emission at L-Band. *IEEE Transactions on Geoscience and Remote Sensing, 45*(7), 1978-1987. https://doi/org/10.1109/tgrs.2007.894935

FAO/IIASA/ISRIC/ISSCAS/JR. (2012). Harmonized World Soil Database (version 1.2). *Food and Agric. Organ., Rome.*

FAO/UNESCO. (2007). The FAO-UNESCO soil map of the World.

Farouki, O. T. (1981). The thermal properties of soils in cold regions. *Cold Regions Science and Technology, 5*(1), 67-75. https://doi/org/10.1016/0165-232X(81)90041-0

Farouki, O. T. (1986). Thermal properties of soils.

Fatichi, S., Or, D., Walko, R., Vereecken, H., Young, M. H., Ghezzehei, T. A., et al. (2020). Soil structure is an important omission in Earth System Models. *Nat Commun, 11*(1), 522. https://doi/org/10.1038/s41467-020-14411-z

Fayer, M. J., & Simmons, C. S. (1995). Modified soil water retention functions for all matric suctions. *Water Resources Research, 31*(5), 1233-1238. https://doi/org/10.1029/95WR00173

Feddes, R., Menenti, M., Kabat, P., & Bastiaanssen, W. (1993). Is large-scale inverse modelling of unsaturated flow with areal average evaporation and surface soil moisture as estimated from remote sensing feasible? *Journal of Hydrology, 143*(1), 125-152.

Fernandez-Moran, R., Al-Yaari, A., Mialon, A., Mahmoodi, A., Al Bitar, A., De Lannoy, G., et al. (2017). SMOS-IC: An Alternative SMOS Soil Moisture and Vegetation Optical Depth Product. *Remote Sensing, 9*(5). https://doi/org/10.3390/rs9050457

Ferrazzoli, P., & Guerriero, L. (1996). Emissivity of vegetation: theory and computational aspects. *Journal of Electromagnetic Waves and Applications, 10*(5), 609-628. https://doi/org/10.1163/156939396X00559

Fredlund, D. G., & Xing, A. (1994). Equations for the soil-water characteristic curve. *Canadian Geotechnical Journal, 31*(4), 521-532. https://doi/org/10.1139/t94-061

Frings, R. M., Schüttrumpf, H., & Vollmer, S. (2011). Verification of porosity predictors for fluvial sand - gravel deposits. *Water Resources Research, 47*(7). https://doi/org/10.1029/2010WR009690

Fung, A. K. (1994). *Microwave scattering and emission models and their applications*. Artech House: Boston London.

Garcia Gonzalez, R., Verhoef, A., Luigi Vidale, P., & Braud, I. (2012). Incorporation of water vapor transfer in the JULES land surface model: Implications for key soil variables and land surface fluxes. *Water Resources Research, 48*(5). https://doi/org/10.1029/2011WR011811

GCOS. (2006). *Systematic Observation Requirements for Satellite-based Products for Climate. Supplemental details to the satellite-based component of the Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC*. Retrieved from

Gelb, A. (1974). *Applied optimal estimation*: MIT press.

GlobalSoilMap. (2009). Specifications for GlobalSoilMap.net products. In. http://www.globalsoilmap.net/Rome/May20_2010_Final_Specifications_for_Output_Products_GSM_Nov_2009_V8.pdf.

Grant, J. P., Saleh-Contell, K., Wigneron, J. P., Guglielmetti, M., Kerr, Y. H., Schwank, M., et al. (2008). Calibration of the L-MEB Model Over a Coniferous and a Deciduous Forest. *IEEE Transactions on Geoscience and Remote Sensing, 46*(3), 808-818. https://doi/org/10.1109/tgrs.2007.914801

Gutmann, E., & Small, E. (2005). The effect of soil hydraulic properties vs. soil texture in land surface models. *Geophysical Research Letters, 32*(2), 1-4. https://doi/org/10.1029/2004GL021843

Gutmann, E. D., & Small, E. E. (2010). A method for the determination of the hydraulic properties of soil from MODIS surface temperature for use in land-surface models. *Water Resources Research, 46*(6), n/a-n/a. https://doi/org/10.1029/2009wr008203

Han, X., Franssen, H.-J. H., Montzka, C., & Vereecken, H. (2014a). Soil moisture and soil properties estimation in the Community Land Model with synthetic brightness temperature observations. *Water Resources Research, 50*(7), 6081-6105. https://doi/org/10.1002/2013wr014586
Han, X., Hendricks Franssen, H., Rosolem, R., Jin, R., Li, X., & Vereecken, H. (2014b). Correction of systematic model forcing bias of CLM using assimilation of cosmic-ray neutrons and land surface temperature: a study in the Heihe catchment, China. *Hydrol. Earth Syst. Sci. Discuss, 11*, 9027-9066.
Han, X., Li, X., He, G., Kumbhar, P., Montzka, C., Kollet, S., et al. (2015). DasPy 1.0–the Open Source Multivariate Land Data Assimilation Framework in combination with the Community Land Model 4.5. *Geoscientific Model Development Discussions, 8*(8). https://doi/org/10.5194/gmdd-8-7395-2015
Han, X., Li, X., Hendricks Franssen, H., Vereecken, H., & Montzka, C. (2012). Spatial horizontal correlation characteristics in the land data assimilation of soil moisture. *Hydrology and Earth System Sciences, 16*(5), 1349-1363. https://doi/org/10.5194/hessd-8-8831-2011
He, J., Yang, K., Tang, W., Lu, H., Qin, J., Chen, Y., & Li, X. (2020). The first high-resolution meteorological forcing dataset for land process studies over China. *Sci Data, 7*(1), 25. https://doi/org/10.1038/s41597-020-0369-y
Hengl, T. (2017). SoilGrids250m: Global Gridded Soil Information Based on Machine Learning. *PloS one*. https://doi/org/10.1371/journal.pone.0169748
Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B., Ribeiro, E., et al. (2014a). SoilGrids1km—global soil information based on automated mapping. *PloS one, 9*(8), e105992. https://doi/org/10.1371/journal.pone.0105992
Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B., Ribeiro, E., et al. (2014b). SoilGrids1km--global soil information based on automated mapping. *PloS one, 9*(8), e105992. https://doi/org/10.1371/journal.pone.0105992
Hillel, D. (1998). *Environmental soil physics: Fundamentals, applications, and environmental considerations*: Academic press.
Hillel, D. (2003). *Introduction to environmental soil physics*: Academic press.
Hollmann, R., Merchant, C. J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., et al. (2013). The ESA Climate Change Initiative: Satellite Data Records for Essential Climate Variables. *Bulletin of the American Meteorological Society, 94*(10), 1541-1552. https://doi/org/10.1175/bams-d-11-00254.1
Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter.

Physica D: Nonlinear Phenomena, *230*(1-2), 112-126. https://doi/org/10.1016/j.physd.2006.11.008

Ines, A. V., & Mohanty, B. P. (2008a). Near - surface soil moisture assimilation for quantifying effective soil hydraulic properties using genetic algorithm: 1. Conceptual modeling. *Water Resources Research, 44*(6). https://doi/org/10.1029/2007WR005990

Ines, A. V., & Mohanty, B. P. (2009). Near - surface soil moisture assimilation for quantifying effective soil hydraulic properties using genetic algorithms: 2. Using airborne remote sensing during SGP97 and SMEX02. *Water Resources Research, 45*(1). https://doi/org/10.1029/2007WR005990

Ines, A. V. M., & Mohanty, B. P. (2008b). Near-Surface Soil Moisture Assimilation for Quantifying Effective Soil Hydraulic Properties under Different Hydroclimatic Conditions. *Vadose Zone Journal*, *7*(1), 39. https://doi/org/10.2136/vzj2007.0048

Jackson, T. J. (1993). III. Measuring surface soil moisture using passive microwave remote sensing. *Hydrological Processes*, *7*(2), 139-152.

Johansen, O. (1975). *Thermal conductivity of soils*. Retrieved from https://archive.org/details/DTIC_ADA044002

Kamann, P. J., Ritzi, R. W., Dominic, D. F., & Conrad, C. M. (2007). Porosity and permeability in sediment mixtures. *Ground Water, 45*(4), 429-438. https://doi/org/10.1111/j.1745-6584.2007.00313.x

Kang, S., Xu, Y., You, Q., Flügel, W.-A., Pepin, N., & Yao, T. (2010). Review of climate and cryospheric change in the Tibetan Plateau. *Environmental Research Letters, 5*(1), 015101. https://doi/org/10.1088/1748-9326/5/1/015101

Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., Lola da Costa, A. C., & Gentine, P. (2019). Implementing Plant Hydraulics in the Community Land Model, Version 5. *Journal of Advances in Modeling Earth Systems, 11*(2), 485-513. https://doi/org/10.1029/2018ms001500

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., et al. (2012). The SMOS soil moisture retrieval algorithm. *IEEE Transactions on Geoscience and Remote Sensing, 50*(5), 1384-1403. https://doi/org/10.1109/TGRS.2012.2184548

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., et al. (2010). The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proceedings of the IEEE, 98*(5), 666-687. https://doi/org/10.1109/JPROC.2010.2043032

Kersten, M. S. (1949). *Laboratory Research for The Determination of The Thermal Properties of Soils*. Retrieved from

Khlosi, M., Cornelis, W. M., Douaik, A., van Genuchten, M. T., & Gabriels, D. (2008). Performance Evaluation of Models That Describe the Soil

Water Retention Curve between Saturation and Oven Dryness. *Vadose Zone Journal, 7*(1), 87. https://doi/org/10.2136/vzj2007.0099

Khlosi, M., Cornelis, W. M., Gabriels, D., & Sin, G. (2006). Simple modification to describe the soil water retention curve between saturation and oven dryness. *Water Resources Research, 42*(11). https://doi/org/10.1029/2005WR004699

Kishné, A. S., Yimam, Y. T., Morgan, C. L., & Dornblaser, B. C. (2017). Evaluation and improvement of the default soil hydraulic parameters for the Noah Land Surface Model. *Geoderma, 285*, 247-259. https://doi/org/10.1016/j.geoderma.2016.09.022

Kittel, C. (1976). *Introduction to solid state physics* (Vol. 8): Wiley New York.

Koltermann, C. E., and Steven M. Gorelick. (1995). Fractional packing model for hydraulic conductivity derived from sediment mixtures. *Water Resources Research, 31*(12), 3297. https://doi/org/10.1029/95WR02020

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., et al. (2004). Regions of strong coupling between soil moisture and precipitation. *Science, 305*(5687), 1138-1140. https://doi/org/10.1126/science.1100217

Koster, R. D., Liu, Q., Mahanama, S. P. P., & Reichle, R. H. (2018). Improved Hydrological Simulation Using SMAP Data: Relative Impacts of Model Calibration and Data Assimilation. *J Hydrometeorol, 19*(4), 727-741. https://doi/org/10.1175/JHM-D-17-0228.1

Kosugi, K. i. (1994). Three - parameter lognormal distribution model for soil water retention. *Water Resources Research, 30*(4), 891-901. https://doi/org/10.1029/93WR02931

Kreith, F., & Black, W. Z. (1980). *Basic heat transfer*: Harper & Row New York.

Kumar, S. V., Reichle, R. H., Koster, R. D., Crow, W. T., & Peters-Lidard, C. D. (2009). Role of subsurface physics in the assimilation of surface soil moisture observations. *Journal of Hydrometeorology, 10*(6), 1534-1547. https://doi/org/10.1175/2009JHM1134.1

Laliberte, G. (1969). A mathematical function for describing capillary pressure-desaturation data. *Hydrological Sciences Journal, 14*(2), 131-149.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., et al. (2019). The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty. *Journal of Advances in Modeling Earth Systems, 11*(12), 4245-4287. https://doi/org/10.1029/2018ms001583

Lawrence, D. M., & Slater, A. G. (2008). Incorporating organic soil into a global climate model. *Climate Dynamics, 30*(2-3), 145-160. https://doi/org/10.1007/s00382-007-0278-1

Lawrence, H., Wigneron, J.-P., Demontoux, F., Mialon, A., & Kerr, Y. H. (2013). Evaluating the semiempirical H-Q model used to calculate the L-band emissivity of a rough bare soil. *IEEE Transactions on*

Geoscience and Remote Sensing, *51*(7), 4075-4084. https://doi/org/10.1109/TGRS.2012.2226995
Lee, J. H. (2014). Spatial-Scale Prediction of the SVAT Soil Hydraulic Variables Characterizing Stratified Soils on the Tibetan Plateau from an EnKF Analysis of SAR Soil Moisture. *Vadose Zone Journal*, *13*(11), 0. https://doi/org/10.2136/vzj2014.06.0060
Lee, J. H., Pellarin, T., & Kerr, Y. H. (2014). Inversion of soil hydraulic properties from the DEnKF analysis of SMOS soil moisture over West Africa. *Agricultural and Forest Meteorology*, *188*, 76-88. https://doi/org/10.1016/j.agrformet.2013.12.009
Leij, F. J., Russell, W. B., & Lesch, S. M. (1997). Closed‐form expressions for water retention and conductivity data. *Ground Water*, *35*(5), 848-858. https://doi/org/10.1111/j.1745-6584.1997.tb00153.x.
Letts, M. G., Roulet, N. T., Comer, N. T., Skarupa, M. R., & Verseghy, D. L. (2000). Parametrization of peatland hydraulic properties for the Canadian Land Surface Scheme. *Atmosphere-Ocean*, *38*(1), 141-160. https://doi/org/10.4324/9781315170206-8
Lievens, H., Verhoest, N. E. C., De Keyser, E., Vernieuwe, H., Matgen, P., Álvarez-Mozos, J., & De Baets, B. (2011). Effective roughness modelling as a tool for soil moisture retrieval from C- and L-band SAR. *Hydrology and Earth System Sciences*, *15*(1), 151-162. https://doi/org/10.5194/hess-15-151-2011
Liu, C., Xiao, Q., & Wang, B. (2008). An Ensemble-Based Four-Dimensional Variational Data Assimilation Scheme. Part I: Technical Formulation and Preliminary Test. *Monthly Weather Review*, *136*(9), 3363-3373. https://doi/org/10.1175/2008mwr2312.1
Liu, Y., & Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, *43*(7). https://doi/org/10.1029/2006wr005756
Livneh, B., Kumar, R., & Samaniego, L. (2015). Influence of soil textural properties on hydrologic fluxes in the Mississippi river basin. *Hydrological Processes*, *29*(21), 4638-4655. https://doi/org/10.1002/hyp.10601
Lu, N., & Likos, W. J. (2004). *Unsaturated soil mechanics*: Wiley.
Lu, S., Ren, T., Gong, Y., & Horton, R. (2007). An Improved Model for Predicting Soil Thermal Conductivity from Water Content at Room Temperature. *Soil Science Society of America Journal*, *71*(1), 8. https://doi/org/10.2136/sssaj2006.0041
Lu, Y., Dong, J., Steele-Dunne, S. C., & van de Giesen, N. (2016). Estimating surface turbulent heat fluxes from land surface temperature and soil moisture observations using the particle batch smoother. *Water Resources Research*, *52*(11), 9086-9108. https://doi/org/10.1002/2016wr018943

Lv, S., Wen, J., Zeng, Y., Tian, H., & Su, Z. (2014). An improved two-layer algorithm for estimating effective soil temperature in microwave radiometry using in situ temperature and soil moisture measurements. *Remote Sensing of Environment*, *152*, 356-363. https://doi/org/10.1016/j.rse.2014.07.007

Lv, S., Zeng, Y., Wen, J., & Su, Z. (2016a). A reappraisal of global soil effective temperature schemes. *Remote Sensing of Environment, 183*, 144-153. https://doi/org/10.1016/j.rse.2016.05.012

Lv, S., Zeng, Y., Wen, J., Zhao, H., & Su, Z. (2018). Estimation of Penetration Depth from Soil Effective Temperature in Microwave Radiometry. *Remote Sensing*, *10*(4), 1-19. https://doi/org/10.3390/rs10040519

Lv, S., Zeng, Y., Wen, J., Zheng, D., & Su, Z. (2016b). Determination of the Optimal Mounting Depth for Calculating Effective Soil Temperature at L-Band: Maqu Case. *Remote Sensing*, *8*(6), 1-19. https://doi/org/10.3390/rs8060476

Ma, J., Zhu, Y., Wobus, R., & Wang, P. (2012). An effective configuration of ensemble size and horizontal resolution for the NCEP GEFS. *Advances in Atmospheric sciences*, *29*(4), 782-794. https://doi/org/10.1007/s00376-012-1249-y

Ma, Y., Ma, W., Zhong, L., Hu, Z., Li, M., Zhu, Z., et al. (2017). Monitoring and Modeling the Tibetan Plateau's climate system and its impact on East Asia. *Sci Rep*, *7*, 44574. https://doi/org/10.1038/srep44574

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, *11*(2), 431-441. https://doi/org/10.1137/0111030

Mattia, F., Satalino, G., Dente, L., & Pasquariello, G. (2006). Using a priori information to improve soil moisture retrieval from ENVISAT ASAR AP data in semiarid regions. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(4), 900-912. https://doi/org/10.1109/TGRS.2005.863483

Mätzler, C. (1994). Microwave (1-100 GHz) dielectric model of leaves. *IEEE Transactions on Geoscience and remote sensing.*, *32*(4), 947-949. https://doi/org/10.1109/36.298024

Mätzler, C. (2006). Surface emission. In E. V. J. P. J. B. Clarricoats (Ed.), *Thermal microwave radiation: applications for remote sensing* (Electromagnetic Waves Series ed., Vol. 52, pp. 225-425). London, UK: The Institution of Engineering and Technology.

Milly, P. (1987). Estimation of Brooks - Corey Parameters from water retention data. *Water Resources Research*, *23*(6), 1085-1089.

Mironov, V. L., Dobson, M. C., Kaupp, V. H., Komarov, S. A., & Kleshchenko, V. N. (2004). Generalized refractive mixing dielectric model for moist soils. *Geoscience and Remote Sensing, IEEE Transactions on*, *42*(4), 773-785.

Mironov, V. L., Kerr, Y. H., Kosolapova, L. G., Savin, I. V., & Muzalevskiy, K. V. (2015). A Temperature-Dependent Dielectric Model for Thawed and Frozen Organic Soil at 1.4 GHz. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*(9), 4470-4477. https://doi/org/10.1109/JSTARS.2015.2442295

Mo, T., Schmugge, T. J., & Wang, J. R. (1987). Calculations of the microwave brightness temperature of rough soil surfaces: Bare field. *IEEE Transactions on Geoscience and Remote Sensing*(1), 47-54. https://doi/org/10.1109/TGRS.1987.289780

Mohanty, B. P. (2013). Soil hydraulic property estimation using remote sensing: A review. *Vadose Zone Journal, 12*(4). https://doi/org/10.2136/vzj2013.06.0100

Montzka, C., Herbst, M., Weihermüller, L., Verhoef, A., & Vereecken, H. (2017). A global data set of soil hydraulic properties and sub-grid variability of soil water retention and hydraulic conductivity curves. *Earth System Science Data, 9*(2), 529. https://doi/org/10.5194/essd-9-529-2017

Montzka, C., Moradkhani, H., Weihermüller, L., Franssen, H.-J. H., Canty, M., & Vereecken, H. (2011). Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter. *Journal of Hydrology, 399*(3-4), 410-421. https://doi/org/10.1016/j.jhydrol.2011.01.020

Moran, M. S., Peters-Lidard, C. D., Watts, J. M., & McElroy, S. (2004). Estimating soil moisture at the watershed scale with satellite-based radar and land surface models. *Canadian journal of remote sensing, 30*(5), 805-826. https://doi/org/10.5589/m04-043

Morel‐Seytoux, H. J., & Nimmo, J. R. (1999). Soil water retention and maximum capillary drive from saturation to oven dryness. *Water Resources Research, 35*(7), 2031-2041. https://doi/org/10.1029/1999WR900121

Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research, 12*(3), 513-522. https://doi/org/10.1029/WR012i003p00513

Mwangi, S., Zeng, Y., Montzka, C., Yu, L., & Su, Z. (2020). Assimilation of cosmic‐ray neutron counts for the estimation of soil ice content on the eastern Tibetan Plateau. *Journal of Geophysical Research: Atmospheres, 125*(3), 1-23. https://doi/org/10.1029/2019JD031529

Nemes, A., Rawls, W. J., & Pachepsky, Y. A. (2005). Influence of Organic Matter on the Estimation of Saturated Hydraulic Conductivity. *Soil Science Society of America Journal, 69*(4), 1330. https://doi/org/10.2136/sssaj2004.0055

Nimmo, J. R. (1991). Comment on the treatment of residual water content in “A consistent set of parametric models for the two‐phase flow of

immiscible fluids in the subsurface" by L. Luckner et al. *Water Resources Research*, *27*(4), 661-662.
Njoku, E. G., & Kong, J.-A. (1977). Theory for passive microwave remote sensing of near-surface soil moisture. *Journal of Geophysical Research*, *82*(20), 3108-3118. https://doi/org/10.1029/JB082i020p03108
Njoku, E. G., Wilson, W. J., Yueh, S. H., Dinardo, S. J., Li, F. K., Jackson, T. J., et al. (2002). Observations of soil moisture using a passive and active low-frequency microwave airborne sensor during SGP99. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(12), 2659-2673. https://doi/org/10.1109/tgrs.2002.807008
O'Neill, P., Bindlish, R., Chan, S., Njoku, E., & Jackson, T. (2018). Algorithm Theoretical Basis Document. Level 2 & 3 Soil Moisture (Passive) Data Products.
O'Neill, P., Bindlish, R., Chan, S., Njoku, E., & Jackson, T. (2020). *Algorithm Theoretical Basis Document. Level 2 & 3 Soil Moisture (Passive) Data Products*. Retrieved from
O'Neill, P., Njoku, E., Jackson, T., Chan, S., & Bindlish, R. (2015). SMAP algorithm theoretical basis document: Level 2 & 3 soil moisture (passive) data products. *Jet Propulsion Lab., California Inst. Technol., Pasadena, CA, USA, JPL D-66480*.
Oleson, K., Lawrence, D., Bonan, G., Drewniak, B., Huang, M., Koven, C., et al. (2013). *Technical description of version 4.5 of the community land model (CLM), NCAR Tech. Note*. Retrieved from
Oleson, K., Lawrence, D., Bonan, G., Flanner, M., Kluzek, E., Lawrence, P., et al. (2010). Technical description of version 4.5 of the Community Land Model (CLM), NCAR Tech. *Notes (NCAR/TN-478+ STR)*.
Oleson, K., Niu, G. Y., Yang, Z. L., Lawrence, D., Thornton, P., Lawrence, P., et al. (2008). Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal of Geophysical Research: Biogeosciences*, *113*(G1). https://doi/org/10.1029/2007JG000563
Owe, M., de Jeu, R., & Holmes, T. (2008). Multisensor historical climatology of satellite-derived global land surface moisture. *Journal of Geophysical Research*, *113*(F1). https://doi/org/10.1029/2007jf000769
Pachepsky, Y., & Rawls, W. J. (2004). *Development of pedotransfer functions in soil hydrology* (Vol. 30): Elsevier.
Park, C.-H., Behrendt, A., LeDrew, E., & Wulfmeyer, V. (2017). New Approach for Calculating the Effective Dielectric Constant of the Moist Soil for Microwaves. *Remote Sensing*, *9*(7). https://doi/org/10.3390/rs9070732
Park, C. H., Montzka, C., Jagdhuber, T., Jonard, F., De Lannoy, G., Hong, J., et al. (2019). A Dielectric Mixing Model Accounting for Soil Organic Matter. *Vadose Zone Journal*, *18*(1). https://doi/org/10.2136/vzj2019.04.0036

Parrens, M., Calvet, J.-C., de Rosnay, P., & Decharme, B. (2014). Benchmarking of L-band soil microwave emission models. *Remote Sensing of Environment*, *140*, 407-419. https://doi/org/10.1016/j.rse.2013.09.017

Parrens, M., Wigneron, J.-P., Richaume, P., Mialon, A., Al Bitar, A., Fernandez-Moran, R., et al. (2016). Global-scale surface roughness effects at L-band as estimated from SMOS observations. *Remote Sensing of Environment*, *181*, 122-136. https://doi/org/10.1016/j.rse.2016.04.006

Paustian, K., Lehmann, J., Ogle, S., Reay, D., Robertson, G. P., & Smith, P. (2016). Climate-smart soils. *Nature*, *532*(7597), 49-57. https://doi/org/10.1038/nature17174

Peake, W. (1959). Interaction of electromagnetic waves with some natural surfaces. *IEEE Transactions on Antennas and Propagation*, *7*(5), 324-329. https://doi/org/10.1109/TAP.1959.1144736

Peck, A., & Watson, J. (1979). *Hydraulic conductivity and flow in non-uniform soil.* Paper presented at the Workshop on soil physics and field heterogeneity. CSIRO Division of Environmental Mechanics, Canberra, Australia.

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification.

Peters-Lidard, C. D., Mocko, D. M., Garcia, M., Santanello, J. A., Tischler, M. A., Moran, M. S., & Wu, Y. (2008). Role of precipitation uncertainty in the estimation of hydrologic soil properties using remotely sensed soil moisture in a semiarid environment. *Water Resources Research*, *44*(5), n/a-n/a. https://doi/org/10.1029/2007wr005884

Pinnington, E., Quaife, T., & Black, E. (2018). Impact of remotely sensed soil moisture and precipitation on soil moisture prediction in a data assimilation system with the JULES land surface model. *Hydrology and Earth System Sciences*, *22*(4), 2575-2588. https://doi/org/10.5194/hess-22-2575-2018

Pinnington, E., Quaife, T., Lawless, A., Williams, K., Arkebauer, T., & Scoby, D. (2020). The Land Variational Ensemble Data Assimilation Framework: LAVENDAR v1.0.0. *Geoscientific Model Development*, *13*(1), 55-69. https://doi/org/10.5194/gmd-13-55-2020

Pitman, A. (2003). The evolution of, and revolution in, land surface schemes designed for climate models. *International Journal of Climatology*, *23*(5), 479-510. https://doi/org/10.1002/joc.893

Prentice, I. C., Liang, X., Medlyn, B. E., & Wang, Y.-P. (2015). Reliable, robust and realistic: the three R's of next-generation land-surface modelling. *Atmospheric Chemistry and Physics*, *15*(10), 5987-6005. https://doi/org/10.5194/acp-15-5987-2015

Qin, J., Liang, S., Yang, K., Kaihotsu, I., Liu, R., & Koike, T. (2009). Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal. *Journal of Geophysical Research, 114*(D15). https://doi/org/10.1029/2008jd011358

Qiu, J. (2008). China: the third pole. *Nature News, 454*(7203), 393-396. https://doi/org/10.1038/454393a

Raju, S., Chanzy, A., Wigneron, J.-P., Calvet, J.-C., Kerr, Y., & Laguerre, L. (1995). Soil moisture and temperature profile effects on microwave emission at low frequencies. *Remote Sensing of Environment, 54*(2), 85-97. https://doi/org/10.1016/0034-4257(95)00133-L

Rawls, W. J., & Brakensiek, D. (1985). *Prediction of soil water properties for hydrologic modeling.* Paper presented at the Proceedings of the Symposium on Watershed Management in the Eighties, New York, NY, USA.

Reichle, R. H., Crow, W. T., & Keppenne, C. L. (2008). An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water Resources Research, 44*(3). https://doi/org/10.1029/2007wr006357

Reichle, R. H., De Lannoy, G. J., Liu, Q., Ardizzone, J. V., Colliander, A., Conaty, A., et al. (2017). Assessment of the SMAP level-4 surface and root-zone soil moisture product using in situ measurements. *Journal of Hydrometeorology, 18*(10), 2621-2645. https://doi/org/10.1175/JHM-D-17-0063.1

Reichle, R. H., De Lannoy, G. J. M., Forman, B. A., Draper, C. S., & Liu, Q. (2013). Connecting Satellite Observations with Water Cycle Variables Through Land Data Assimilation: Examples Using the NASA GEOS-5 LDAS. *Surveys in Geophysics, 35*(3), 577-606. https://doi/org/10.1007/s10712-013-9220-8

Reichle, R. H., Koster, R. D., Liu, P., Mahanama, S. P., Njoku, E. G., & Owe, M. (2007). Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR‐E) and the Scanning Multichannel Microwave Radiometer (SMMR). *Journal of Geophysical Research: Atmospheres, 112*(D9). https://doi/org/10.1029/2006JD008033

Ribeiro, E., Batjes, N. H., Leenaars, J. G., Van Oostrum, A., & de Mendes, J. (2015). Towards the standardization and harmonization of world soil data: Procedures Manual ISRIC World Soil Information Service (WoSIS version 2.0). *ISRIC-World Soil Information: Wageningen, The Netherlands*.

Richards, L. A. (1931). Capillary conduction of liquids through porous mediums. *Journal of Applied Physics, 1*(5), 318-333.

Robinson, D., Campbell, C., Hopmans, J., Hornbuckle, B., Jones, S. B., Knight, R., et al. (2008). Soil moisture measurement for ecological and hydrological watershed-scale observatories: A review. *Vadose Zone Journal, 7*(1), 358-389.

Rodell, M., Houser, P., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., et al. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, *85*(3), 381-394. https://doi/org/10.1175/BAMS-85-3-381

Ross, P., Williams, J., & Bristow, K. L. (1991). Equation for extending water‐retention curves to dryness. *Soil Science Society of America Journal*, *55*(4), 923-927. https://doi/org/10.2136/sssaj1991.03615995005500040004x

Ross, P. J., & Smettem, K. R. (1993). Describing soil hydraulic properties with sums of simple functions. *Soil Science Society of America Journal*, *57*(1), 26-29.

Rossi, C., & Nimmo, J. R. (1994). Modeling of soil water retention from saturation to oven dryness. *Water Resources Research*, *30*(3), 701-708. https://doi/org/10.1029/93WR03238

Sakaki, T., & Smits, K. M. (2015). Water Retention Characteristics and Pore Structure of Binary Mixtures. *Vadose Zone Journal*, *14*(2), 0. https://doi/org/10.2136/vzj2014.06.0065

Saleh, K., Kerr, Y. H., Richaume, P., Escorihuela, M. J., Panciera, R., Delwart, S., et al. (2009). Soil moisture retrievals at L-band using a two-step inversion approach (COSMOS/NAFE'05 Experiment). *Remote Sensing of Environment*, *113*(6), 1304-1312. https://doi/org/10.1016/j.rse.2009.02.013

Saleh, K., Wigneron, J.-P., de Rosnay, P., Calvet, J.-C., Escorihuela, M. J., Kerr, Y., & Waldteufel, P. (2006). Impact of rain interception by vegetation and mulch on the L-band emission of natural grass. *Remote Sensing of Environment*, *101*(1), 127-139. https://doi/org/10.1016/j.rse.2005.12.004

Santanello, J., Joseph, A., & Carlson, T. (2001). Mesoscale simulation of rapid soil drying and its implications for predicting daytime temperature. *Journal of Hydrometeorology*, *2*(1), 71-88. https://doi/org/10.1175/1525-7541(2001)002<0071:MSORSD>2.0.CO;2

Santanello, J. A., Peters-Lidard, C. D., Garcia, M. E., Mocko, D. M., Tischler, M. A., Moran, M. S., & Thoma, D. P. (2007). Using remotely-sensed estimates of soil moisture to infer soil texture and hydraulic properties across a semi-arid watershed. *Remote Sensing of Environment*, *110*(1), 79-97. https://doi/org/10.1016/j.rse.2007.02.007

Saxton, K., & Rawls, W. (2006). Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Science Society of America Journal*, *70*(5), 1569-1578. https://doi/org/10.2136/sssaj2005.0117

Saxton, K., Rawls, W. J., Romberger, J., & Papendick, R. (1986). Estimating generalized soil-water characteristics from texture. *Soil Science Society* *of America Journal*, *50*(4), 1031-1036. https://doi/org/10.2136/sssaj1986.03615995005000040039x

Schaap, M. G., Leij, F. J., & van Genuchten, M. T. (1998). Neural network analysis for hierarchical prediction of soil hydraulic properties. *Soil Science Society of America Journal*, *62*(4), 847-855.

Schaap, M. G., Leij, F. J., & van Genuchten, M. T. (2001). ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of Hydrology*, *251*(3), 163-176. https://doi/org/10.1016/S0022-1694(01)00466-8

Schmugge, T., & Choudhury, B. (1981). A comparison of radiative transfer models for predicting the microwave emission from soils. *Radio Science*, *16*(5), 927-938. https://doi/org/10.1029/RS016i005p00927

Schneeberger, K., Schwank, M., Stamm, C., De Rosnay, P., Mätzler, C., & Flühler, H. (2004). Topsoil structure influencing soil water retrieval by microwave radiometry. *Vadose Zone Journal*, *3*(4), 1169-1179. https://doi/org/10.2113/3.4.1169

Schwank, M., Guglielmetti, M., Matzler, C., & Fluhler, H. (2008). Testing a new model for the L-band radiation of moist leaf litter. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(7), 1982-1994. https://doi/org/10.1109/TGRS.2008.916983

Schwank, M., Matzler, C., Guglielmetti, M., & Fluhler, H. (2005). L-band radiometer measurements of soil water under growing clover grass. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(10), 2225-2237. https://doi/org/10.1109/tgrs.2005.855135

Schwank, M., Stahli, M., Wydler, H., Leuenberger, J., Matzler, C., & Fluhler, H. (2004). Microwave L-band emission of freezing soil. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(6), 1252-1261. https://doi/org/10.1109/tgrs.2004.825592

Schwank, M., Wiesmann, A., Werner, C., Mätzler, C., Weber, D., Murk, A., et al. (2010). ELBARA II, an L-band radiometer system for soil moisture research. *Sensors*, *10*(1), 584-612. https://doi/org/10.3390/s100100584

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., et al. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, *99*(3), 125-161. https://doi/org/10.1016/j.earscirev.2010.02.004

Shangguan, W., Dai, Y., Liu, B., Ye, A., & Yuan, H. (2012). A soil particle-size distribution dataset for regional land and climate modelling in China. *Geoderma*, *171*, 85-91. https://doi/org/10.1016/j.geoderma.2011.01.013

Shangguan, W., Dai, Y., Liu, B., Zhu, A., Duan, Q., Wu, L., et al. (2013). A China data set of soil properties for land surface modeling. *Journal of Advances in Modeling Earth Systems*, *5*(2), 212-224. https://doi/org/10.1002/jame.20026

Shellito, P. J., Small, E. E., & Cosh, M. H. (2016). Calibration of Noah Soil Hydraulic Property Parameters Using Surface Soil Moisture from

SMOS and Basinwide In Situ Observations. *Journal of Hydrometeorology*, *17*(8), 2275-2292. https://doi/org/10.1175/jhm-d-15-0153.1

Shi, J., Chen, K.-S., Li, Q., Jackson, T. J., O'Neill, P. E., & Tsang, L. (2002). A parameterized surface reflectivity model and estimation of bare-surface soil moisture with L-band radiometer. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(12), 2674-2686. https://doi/org/10.1109/TGRS.2002.807003

Soet, M., & Stricker, J. (2003). Functional behaviour of pedotransfer functions in soil water flow simulation. *Hydrological Processes*, *17*(8), 1659-1670. https://doi/org/10.1002/hyp.1207

Straiton, A. W., Tolbert, C. W., & Britt, C. O. (1958). Apparent Temperatures of Some Terrestrial Materials and the Sun at 4.3‐Millimeter Wavelengths. *Journal of Applied Physics*, *29*(5), 776-782. https://doi/org/10.1063/1.1723283

Su, Z., de Rosnay, P., Wen, J., Wang, L., & Zeng, Y. (2013). Evaluation of ECMWF's soil moisture analyses using observations on the Tibetan Plateau. *Journal of Geophysical Research: Atmospheres*, *118*(11), 5304-5318. https://doi/org/10.1002/jgrd.50468

Su, Z., Troch, P. A., & De Troch, F. P. (1997). Remote sensing of bare surface soil moisture using EMAC/ESAR data. *International Journal of Remote Sensing*, *18*(10), 2105-2124. https://doi/org/10.1080/014311697217783

Su, Z., Wen, J., Dente, L., van der Velde, R., Wang, L., Ma, Y., et al. (2011). The Tibetan Plateau observatory of plateau scale soil moisture and soil temperature (Tibet-Obs) for quantifying uncertainties in coarse resolution satellite and model products. *Hydrology and Earth System Sciences*, *15*(7), 2303-2316. https://doi/org/10.5194/hess-15-2303-2011

Su, Z., Wen, J., Zeng, Y., Zhao, H., Lv, S., van der Velde, R., et al. (2020a). Multiyear in-situ L-band microwave radiometry of land surface processes on the Tibetan Plateau. *Scientific Data*, *7*(1), 317. https://doi/org/10.1038/s41597-020-00657-1

Su, Z., Zeng, Y., Romano, N., Manfreda, S., Francés, F., Ben Dor, E., et al. (2020b). An Integrative Information Aqueduct to Close the Gaps between Satellite Observation of Water Cycle and Local Sustainable Management of Water Resources. *Water*, *12*(5). https://doi/org/10.3390/w12051495

Su, Z., Zeng, Y., Wen, J., & Zhao, H. (2019). *A multiyear ELBARA-III L-band radiometric observation dataset of the Third Pole Environment*.

Swenson, S. C., & Lawrence, D. M. (2014). Assessing a dry surface layer-based soil resistance parameterization for the Community Land Model using GRACE and FLUXNET-MTE data. *Journal of Geophysical Research: Atmospheres*, *119*(17), 10,299-210,312. https://doi/org/10.1002/2014jd022314

Swenson, S. C., & Lawrence, D. M. (2015). A GRACE-based assessment of interannual groundwater dynamics in the Community Land Model. *Water Resources Research, 51*(11), 8817-8833. https://doi/org/10.1002/2015wr017582

Tarnawski, V. R., & Leong, W. H. (2012). A series-parallel model for estimating the thermal conductivity of unsaturated soils. *International Journal of Thermophysics, 33*(7), 1191-1218.

Tarnawski, V. R., & Wagner, B. (1992). A new computerized approach to estimating the thermal properties of unfrozen soils. *Canadian Geotechnical Journal, 29*(4), 714-720. https://doi/org/10.1139/t92-079

Taylor, C. M., de Jeu, R. A., Guichard, F., Harris, P. P., & Dorigo, W. A. (2012). Afternoon rain more likely over drier soils. *Nature, 489*(7416), 423. https://doi/org/10.1038/nature11377

Tian, Z., Lu, Y., Horton, R., & Ren, T. (2016). A simplified de Vries-based model to estimate thermal conductivity of unfrozen and frozen soil. *European journal of soil science, 67*(5), 564-572. https://doi/org/10.1111/ejss.12366

Topp, G. C., Davis, J., & Annan, A. P. (1980). Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. *Water Resources Research*, *16*(3), 574-582. https://doi/org/10.1029/WR016i003p00574

Ulaby, F. T., Long, D. G., Blackwell, W. J., Elachi, C., Fung, A. K., Ruf, C., et al. (2014). Electromagnetic Wave Propagation and Reflection. In *Microwave radar and radiometric remote sensing* (Vol. 4, pp. 44-50). Ann Arbor, US: University of Michigan Press.

van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal, 44*(5), 892-898. https://doi/org/10.2136/sssaj1980.03615995004400050002x

van Genuchten, M. T., & Nielsen, D. (1985). On describing and predicting the hydraulic properties of unsaturated soils. *Ann. Geophys, 3*(5), 615-628.

Van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., et al. (2017). Pedotransfer Functions in Earth System Science: Challenges and Perspectives. *Reviews of Geophysics*, *55*(4), 1199-1256. https://doi/org/10.1002/2017rg000581

Vereecken, H., Kasteel, R., Vanderborght, J., & Harter, T. (2007). Upscaling hydraulic properties and soil water flow processes in heterogeneous soils. *Vadose Zone Journal*, *6*(1), 1-28. https://doi/org/10.2136/vzj2006.0055

Vereecken, H., Maes, J., Feyen, J., & Darius, P. (1989). Estimating the soil moisture retention characteristic from texture, bulk density, and carbon content. *Soil science*, *148*(6), 389-403. https://doi/org/10.1097/00010694-198912000-00001

Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M., & Genuchten, M. T. (2010). Using pedotransfer functions to estimate the van Genuchten–Mualem soil hydraulic properties: a review. *Vadose Zone Journal*, *9*(4), 795-820. https://doi/org/10.2136/vzj2010.0045

Verhoef, W. (1996). *Application of harmonic analysis of NDVI time series (HANTS)*. Retrieved from Wageningen:

Wang, J., & Choudhury, B. (1981). Remote sensing of soil moisture content, over bare field at 1.4 GHz frequency. *Journal of Geophysical Research: Oceans*, *86*(C6), 5277-5282. https://doi/org/10.1029/JC086iC06p05277

Wang, J. R., & Schmugge, T. J. (1980). An empirical model for the complex dielectric permittivity of soils as a function of water content. *IEEE Transactions on Geoscience and Remote Sensing*(4), 288-295. https://doi/org/10.1109/TGRS.1980.350304

Wang, Q., van der Velde, R., & Su, Z. (2018). Use of a discrete electromagnetic model for simulating Aquarius L-band active/passive observations and soil moisture retrieval. *Remote Sensing of Environment*, *205*, 434-452. https://doi/org/10.1016/j.rse.2017.10.044

Wang, T., Wedin, D., & Zlotnik, V. A. (2009). Field evidence of a negative correlation between saturated hydraulic conductivity and soil carbon in a sandy soil. *Water Resources Research*, *45*(7). https://doi/org/10.1029/2008wr006865

Webb, S. W. (2000). A simple extension of two ‐ phase characteristic curves to include the dry region. *Water Resources Research*, *36*(6), 1425-1430. https://doi/org/10.1029/2000WR900057

Weynants, M., Vereecken, H., & Javaux, M. (2009). Revisiting Vereecken pedotransfer functions: Introducing a closed-form hydraulic model. *Vadose Zone Journal*, *8*(1), 86-95. https://doi/org/10.2136/vzj2008.0062

Whitaker, J. S., & Hamill, T. M. (2012). Evaluating Methods to Account for System Errors in Ensemble Data Assimilation. *Monthly Weather Review*, *140*(9), 3078-3089. https://doi/org/10.1175/mwr-d-11-00276.1

Wigneron, J.-P., Chanzy, A., Kerr, Y. H., Lawrence, H., Shi, J., Escorihuela, M. J., et al. (2011). Evaluating an Improved Parameterization of the Soil Emission in L-MEB. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(4), 1177-1189. https://doi/org/10.1109/tgrs.2010.2075935

Wigneron, J.-P., Kerr, Y., Waldteufel, P., Saleh, K., Escorihuela, M.-J., Richaume, P., et al. (2007). L-band Microwave Emission of the Biosphere (L-MEB) Model: Description and calibration against experimental data sets over crop fields. *Remote Sensing of Environment*, *107*(4), 639-655.

Wigneron, J.-P., Laguerre, L., & Kerr, Y. H. (2001). A simple parameterization of the L-band microwave emission from rough agricultural soils. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(8), 1697-1707. https://doi/org/10.1109/36.942548

Wigneron, J. P., Jackson, T. J., O'Neill, P., De Lannoy, G., de Rosnay, P., Walker, J. P., et al. (2017). Modelling the passive microwave signature from land surfaces: A review of recent results and application to the L-band SMOS & SMAP soil moisture retrieval algorithms. *Remote Sensing of Environment*, *192*, 238-262. https://doi/org/10.1016/j.rse.2017.01.024

Wilheit, T. T. (1978). Radiative transfer in a plane stratified dielectric. *IEEE Transactions on Geoscience Electronics*, *16*(2), 138-143. https://doi/org/10.1109/TGE.1978.294577

Wösten, J., Lilly, A., Nemes, A., & Le Bas, C. (1999). Development and use of a database of hydraulic properties of European soils. *Geoderma*, *90*(3), 169-185. https://doi/org/10.1016/S0016-7061(98)00132-3

Wösten, J., Schuren, C., Bouma, J., & Stein, A. (1990). Functional sensitivity analysis of four methods to generate soil hydraulic functions. *Soil Science Society of America Journal*, *54*(3), 832-836.

Wu, W., & Wang, S. S. (2006). Formulas for sediment porosity and settling velocity. *Journal of Hydraulic Engineering*, *132*(8), 858-862. https://doi/org/10.1061/ASCE0733-9429

Xie, S., Qu, J., Lai, Y., Zhou, Z., & Xu, X. (2015). Effects of freeze-thaw cycles on soil mechanical and physical properties in the Qinghai-Tibet Plateau. *Journal of Mountain Science*, *12*(4), 999-1009. https://doi/org/10.1061/(ASCE)0733-9410(1992)118:7(1083)

Yang, K., Chen, Y., He, J., Zhao, L., Lu, H., Qin, J., et al. (2020). Development of a daily soil moisture product for the period of 2002–2011 in Mainland China. *Science China Earth Sciences*, 1-13. https://doi/org/10.2136/vzj2012.0072

Yang, K., Chen, Y., & Qin, J. (2009a). Some practical notes on the land surface modeling in the Tibetan Plateau. *Hydrology and Earth System Sciences*, *13*(5), 687-701. https://doi/org/10.5194/hess-13-687-2009

Yang, K., Koike, T., Ishikawa, H., Kim, J., Li, X., Liu, H., et al. (2008). Turbulent Flux Transfer over Bare-Soil Surfaces: Characteristics and Parameterization. *Journal of Applied Meteorology and Climatology*, *47*(1), 276-290. https://doi/org/10.1175/2007jamc1547.1

Yang, K., Koike, T., Kaihotsu, I., & Qin, J. (2009b). Validation of a dual-pass microwave land data assimilation system for estimating surface soil moisture in semiarid regions. *Journal of Hydrometeorology*, *10*(3), 780-793. https://doi/org/10.1175/2008JHM1065.1

Yang, K., Qin, J., Zhao, L., Chen, Y., Tang, W., Han, M., et al. (2013). A multiscale soil moisture and freeze-thaw monitoring network on the third pole. *Bulletin of the American Meteorological Society*, *94*(12), 1907-1916.

Yang, K., Watanabe, T., Koike, T., Li, X., FUJII, H., Tamagawa, K., et al. (2007). Auto-calibration system developed to assimilate AMSR-E data into a land surface model for estimating soil moisture and the surface

energy budget. *Journal of the Meteorological Society of Japan. Ser. II, 85*, 229-242. https://doi/org/10.2151/jmsj.85A.229
Yang, K., Wu, H., Qin, J., Lin, C., Tang, W., & Chen, Y. (2014). Recent climate changes over the Tibetan Plateau and their impacts on energy and water cycle: A review. *Global and Planetary Change*, *112*, 79-91.
Yang, K., Ye, B., Zhou, D., Wu, B., Foken, T., Qin, J., & Zhou, Z. (2011). Response of hydrological cycle to recent climate changes in the Tibetan Plateau. *Climatic change*, *109*(3-4), 517-534.
Yang, K., Zhu, L., Chen, Y., Zhao, L., Qin, J., Lu, H., et al. (2016). Land surface model calibration through microwave data assimilation for improving soil moisture simulations. *Journal of Hydrology*, *533*, 266-276. https://doi/org/10.1016/j.jhydrol.2015.12.018
Yao, T., Thompson, L. G., Mosbrugger, V., Zhang, F., Ma, Y., Luo, T., et al. (2012). Third pole environment (TPE). *Environmental Development*, *3*, 52-64. https://doi/org/10.1016/j.envdev.2012.04.002
Yu, L., Zeng, Y., & Su, Z. (2020). Understanding the mass, momentum, and energy transfer in the frozen soil with three levels of model complexities. *Hydrology and Earth System Sciences*, *24*(10), 4813-4830. https://doi/org/10.5194/hess-24-4813-2020.
Yu, L., Zeng, Y., Wen, J., & Su, Z. (2018). Liquid-Vapor-Air Flow in the Frozen Soil. *Journal of Geophysical Research: Atmospheres*, 7393-7415. https://doi/org/10.1029/2018jd028502
Zeng, J., Chen, K.-S., Bi, H., Zhao, T., & Yang, X. (2017). A Comprehensive Analysis of Rough Soil Surface Scattering and Emission Predicted by AIEM With Comparison to Numerical Simulations and Experimental Measurements. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(3), 1696-1708. https://doi/org/10.1109/tgrs.2016.2629759
Zeng, J., Li, Z., Chen, Q., Bi, H., Qiu, J., & Zou, P. (2015). Evaluation of remotely sensed and reanalysis soil moisture products over the Tibetan Plateau using in-situ observations. *Remote Sensing of Environment*, *163*, 91-110.
Zeng, X., & Decker, M. (2009). Improving the Numerical Solution of Soil Moisture–Based Richards Equation for Land Models with a Deep or Shallow Water Table. *Journal of Hydrometeorology*, *10*(1), 308-319. https://doi/org/10.1175/2008jhm1011.1
Zeng, Y., & Su, Z. (2013). Reply to comment by Binayak P. Mohanty and Zhenlei Yang on "A simulation analysis of the advective effect on evaporation using a two‐phase heat and mass flow model". *Water Resources Research*, *49*(11), 7836-7840;. https://doi/org/10.1002/2013WR013764
Zeng, Y., Su, Z., van der Velde, R., Wang, L., Xu, K., Wang, X., & Wen, J. (2016). Blending Satellite Observed, Model Simulated, and in Situ Measured Soil Moisture over Tibetan Plateau. *Remote Sensing*, *8*(3), 1-22. https://doi/org/10.3390/rs8030268

Zeng, Y., Su, Z., Wan, L., & Wen, J. (2011a). Numerical analysis of air-water-heat flow in unsaturated soil: Is it necessary to consider airflow in land surface models? *Journal of Geophysical Research*, *116*(D20). https://doi/org/10.1029/2011jd015835

Zeng, Y., Su, Z., Wan, L., & Wen, J. (2011b). Numerical analysis of air - water - heat flow in unsaturated soil: Is it necessary to consider airflow in land surface models? *Journal of Geophysical Research: Atmospheres*, *116*(D20). https://doi/org/10.1029/2011JD015835

Zeng, Y., Su, Z., Wan, L., & Wen, J. (2011c). A simulation analysis of the advective effect on evaporation using a two - phase heat and mass flow model. *Water Resources Research*, *47*(10). https://doi/org/10.1029/2011WR010701

Zeng, Y., Su, Z., Wan, L., Yang, Z., Zhang, T., Tian, H., et al. (2009a). Diurnal pattern of the drying front in desert and its application for determining the effective infiltration. *Hydrology and Earth System Sciences*, *13*(6), 703. https://doi/org/10.5194/hess-13-703-2009

Zeng, Y., Wan, L., Su, Z., Saito, H., Huang, K., & Wang, X. (2009b). Diurnal soil water dynamics in the shallow vadose zone (field site of China University of Geosciences, China). *Environmental geology*, *58*(1), 11-23. https://doi/org/10.1007/s00254-008-1485-8

Zhang, P., Zheng, D., van der Velde, R., Wen, J., Zeng, Y., Wang, X., et al. (2020). Status of the Tibetan Plateau observatory (Tibet-Obs) and a 10-year (2009-2019) surface soil moisture dataset. *Earth Syst. Sci. Data Discuss.* https://doi/org/10.5194/essd-2020-209

Zhang, Y., & Schaap, M. G. (2017). Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3). *Journal of Hydrology*, *547*, 39-53. https://doi/org/10.1016/j.jhydrol.2017.01.004

Zhang, Z., Ma, W., Feng, W., Xiao, D., & Hou, X. (2016). Reconstruction of Soil Particle Composition During Freeze-Thaw Cycling: A Review. *Pedosphere*, *26*(2), 167-179. https://doi/org/10.1016/S1002-0160(15)60033-9

Zhang, Z. F., Ward, A. L., & Keller, J. m. (2011). Determining the Porosity and Saturated Hydraulic Conductivity of Binary Mixtures. *Vadose Zone Journal*, *10*(1), 313. https://doi/org/10.2136/vzj2009.0138

Zhao, H., Zeng, Y., Lv, S., & Su, Z. (2018a). Analysis of soil hydraulic and thermal properties for land surface modeling over the Tibetan Plateau. *Earth System Science Data*, *10*(2), 1031-1061. https://doi/org/10.5194/essd-10-1031-2018

Zhao, H., Zeng, Y., Su, B., & Lv, S. (2018b). *Soil Hydraulic and Thermal Properties for Land Surface Modelling over the Tibetan Plateau. University of Twente. 10.4121/uuid:c712717c-6ac0-47ff-9d58-97f88082ddc0.*

Zhao, H., Zeng, Y., Wen, J., Wang, X., Wang, Z., Meng, X., & Su, Z. (2021). An Air-to-Soil Transition Model for Discrete Scattering-Emission Modelling at L-Band. *Journal of Remote Sensing*. https://doi/org/10.34133/2021/3962350

Zheng, D., Li, X., Wang, X., Wang, Z., Wen, J., van der Velde, R., et al. (2019). Sampling depth of L-band radiometer measurements of soil moisture and freeze-thaw dynamics on the Tibetan Plateau. *Remote Sensing of Environment, 226*, 16-25. https://doi/org/10.1016/j.rse.2019.03.029

Zheng, D., van der Velde, R., Su, Z., Wang, X., Wen, J., Booij, M. J., et al. (2015a). Augmentations to the Noah model physics for application to the Yellow River source area. Part I: Soil water flow. *Journal of Hydrometeorology, 16*(6), 2659-2676.

Zheng, D., van der Velde, R., Su, Z., Wang, X., Wen, J., Booij, M. J., et al. (2015b). Augmentations to the Noah model physics for application to the Yellow River source area. Part II: Turbulent heat fluxes and soil heat transport. *Journal of Hydrometeorology, 16*(6), 2677-2694. https://doi/org/10.1175/JHM-D-14-0199.1

Zheng, D., van der Velde, R., Wen, J., Wang, X., Ferrazzoli, P., Schwank, M., et al. (2018a). Assessment of the SMAP Soil Emission Model and Soil Moisture Retrieval Algorithms for a Tibetan Desert Ecosystem. *IEEE Transactions on Geoscience and Remote Sensing, 56*(7), 3786 - 3799. https://doi/org/10.1109/tgrs.2018.2811318

Zheng, D., Wang, X., van der Velde, R., Ferrazzoli, P., Wen, J., Wang, Z., et al. (2018b). Impact of surface roughness, vegetation opacity and soil permittivity on L-band microwave emission and soil moisture retrieval in the third pole environment. *Remote Sensing of Environment, 209*, 633-647. https://doi/org/10.1016/j.rse.2018.03.011

Zheng, D., Wang, X., van der Velde, R., Zeng, Y., Wen, J., Wang, Z., et al. (2017). L-Band Microwave Emission of Soil Freezesc-Thaw Process in the Third Pole Environment. *IEEE Transactions on Geoscience and Remote Sensing, 55*(9), 5324 - 5338. https://doi/org/10.1109/tgrs.2017.2705248

Zhuang, R., Zeng, Y., Manfreda, S., & Su, Z. (2020). Quantifying long-term land surface and root zone soil moisture over Tibetan Plateau. *Remote Sensing, 12*(3), 509. https://doi/org/10.3390/rs12030509

Samenvatting

In weer- en klimaat voorspellingen speelt het grondvochtgehalte van de bodem een belangrijke rol als zijnde de onderste randvoorwaarde van atmosfeermodellen. Basale grondeigenschappen, i.e. grondtextuur en het gehalte organisch materiaal, en daarbij behorende hydraulische grondeigenschappen (HGE), i.e. grondvocht retentiecurve en hydraulische geleiding, en thermische grondeigenschappen (TGE), i.e. de warmte-inhoud van de grond en thermische geleiding, spelen een essentiële rol bij het schatten van het grondvochtgehalte en de grondtemperatuur middels aardoppervlakmodellen (AOMn). Vanwege het gebrek aan gedetailleerde grondeigenschapkaarten kunnen grondparametrisatiemethodes welke gebruikt worden in AOMn mogelijk niet representatief zijn, wat resulteert in onzekerheden bij het schatten van grondoppervlakcondities en warmtestromen. Door gebruik te maken van het fysische verband tussen de grondfysica-eigenschappen, het grondvochtgehalte, de grondtemperatuur enerzijds en de effectieve diëlektrische constante van de grond anderzijds, kunnen met behulp van een microgolfemissie-observatiemodel gekoppeld aan een AOM - binnen een data-assimilatie kader - de grondfysica-eigenschappen afgeleid worden.

Remote sensing met behulp van L-band microgolfemissie (ook wel passieve remote sensing genoemd) is de meest veelbelovende techniek om het grondvochtgehalte direct aan het aardoppervlak te achterhalen. Dit gebeurt middels de gemeten helderheidstemperatuur (Engels: brightness temperature) (T_B^p, p =H, V). Deze trend versnelt sinds de lancering van twee innovatieve satellietmissies: SMOS (Soil Moisture and Ocean Salinity) en SMAP (Soil Moisture Active and Passive), welke beide uitgerust zijn met L-band radiometers. Sinds de lancering hebben beide satellieten T_B^p- en grondvochtproducten voor het gehele aardoppervlak aangeleverd met een bijna dagelijkse meetfrequentie. Het

afleiden van grondfysica-eigenschappen middels de assimilatie van T_B^p-metingen in het gekoppelde AOM en L-band emissiemodel kan worden beschouwd as zijnde een grondobservatiesysteem dat gebruikmaakt van aardobservatie-data uit de ruimte, lokaal gemeten data en modelleren. Deze analogie gaat vooral op als men afgelegen gebieden beschouwd.

Dit onderzoek heeft als doel het begrip te verbeteren van het afleiden van grondfysica-eigenschappen middels het hiervoor genoemde assimilatie-systeem. De twee belangrijkste componenten binnen dit assimilatie-systeem zijn het AOM als zijnde het systeem-model en het microgolfemissiemodel als observatiefunctie. In AOMn worden de schattingen van de grondvocht- en grondtemperatuurprofielen voor het grootste gedeelte bepaald door de basale grondeigenschappen en de hierbij behorende HGE en TGE. Bij het modelleren van L-band microgolfemissie is juist een belangrijke rol weggelegd voor de oppervlakteruwheid, vooral de emissie met horizontale (H) polarisatie is hier erg gevoelig voor. Meetonzekerheden in de resulterende emissie zullen daarom doorwerken in de hieruit afgeleide grondfysica-eigenschappen. Voordat het afleiden van grondeigenschappen aan de orde komt, zal daarom in dit proefschrift eerst afzonderlijk worden ingaan op de grondfysica-eigenschappen die gebruikt worden voor AOMn (hoofdstuk 2) en vervolgens op het effect van oppervlakteruwheid op L-band T_B^p schattingen (hoofdstuk 3). Als onderzoekslocatie voor het helpen uitvoeren van dit onderzoek is gekozen voor het Maqu station (33.91°N, 102.16°E) dat zich op het oostelijk Tibetaans hoogplateau (TH) bevind en waarvoor een uitgebreid scala aan veldobservaties voorhanden is.

Om de afgeleide fysische grondeigenschappen en de daaropvolgende AOM-analyse te kunnen valideren zijn veldmetingen en bijbehorende laboratoriummetingen van basale grondeigenschappen en SHP & STP uitgevoerd in verschillende klimaatzones binnen het TH. Met deze verzamelde grondprofielen van droge (Ngari), semi-droge (Naqu) tot sub-vochtige (Maqu) klimaatzones stelden wij de zogenaamde Tibet-Obs dataset van grondeigenschappen samen. Met deze dataset wordt/worden in hoofdstuk 2 vervolgens: 1) een analyse gedaan naar de verschillen in basale grondeigenschappen en SHP & STP over de hierboven genoemde klimaatzones; 2) verschillende methodes onderzocht om de poreusheid en SHP & STP van de grond van het TP te bepalen; en 3) de onzekerheid van bestaande basale grondeigenschappen en de daarvan afgeleide SHP & SP over het TP gekwantificeerd. Wij vonden hierbij dat: 1) de basale grondeigenschappen en SHP & STP voor elke klimaatzone verschillend waren en tevens dat deze varieerden over de bodemprofielen; 2) de pedotransferfuncties (PTFs) van Cosby e.a. (1984) beter van toepassing waren voor het afschatten van HGE bij het model van Clapp en Hornberber (CH) (1978), en dat de continue PTFs van Wösten e.a. (1999) meer van toepassing waren voor het van Genuchten (1980) – Mualen (1976) (VG) model. Het semi-empirische model van De Vries (1963) bleek superieur voor het schatten van STP. 3) SoilGrids1km aan te bevelen is voor gebruik in droge en sub-vochtige zones, een combinatie van FAO-UNESCO voor ondiepe lagen en HWSD voor diepe lagen in semi-droge zones op het TP.

De dataset is gepubliceerd bij 4TU.Center for Research Data (https://data.4tu.nl/articles/Soil_Hydraulic_and_Thermal_Properties_for_Land_Surface_Modelling_over_the_Tibetan_Plateau_version_1_/12721418/2). Wij hopen dat de dataset en onze bevindingen kunnen bijdragen aan het modelleren van, en het onderzoek doen naar het derde poolgebied door de hydro-klimatologische gemeenschap en dat deze de geografische gaten zal kunnen dichten tussen de reeds gepubliceerde globale grond-database van de grondonderzoeksgemeenschap.

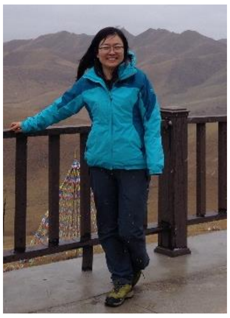
Om rekening te houden met de effecten van structuren in de bovenlaag van de bodem en de inhomogene verdeling van bodemvocht op L-band emissie, ontwikkelen wij in hoofdstuk 3 een lucht-naar-grond transitie model (LNG) waarin wij het begruik van diëlektrische ruwheidsparametrisatie voorstellen. Het Tor Vergata discrete scattering model (TVG), geïntegreerd met het advanced integral equation model (AIEM), wordt gebruikt als referentiemodel voor het simuleren van T_B^p (p =H, V) voor L-band. Dit gecombineerde model wordt vervolgens uitgebreid met het ATS model. We tonen aan dat deze uitbreiding met het ATS model noodzakelijk is om de te laag geschatte T_B^p waarde van het referentiemodel (≈ 20-50 K) omhoog te halen. Wij suggereren om de vaste ruwheidsparameter H_R, welke in de huidige grondvocht-afleidingsprocessen van SMAP en SMOS gebruikt wordt, te vervangen door onze diëlektrische ruwheid om zo hydrometeorologische veranderingen in de bodem, gerelateerd aan de oppervlakteruwheid te kunnen vastleggen. Echter, het verschil tussen de gemodelleerde- en de gemeten T_B^p tijdens het bevriezings- en dooiproces van de bodem suggereert dat het ATS model nog verbeterd kan worden door tevens de effecten van de oppervlaktetemperatuur, oppervlaktewaterfractie en het mengsel van water en ijs mee te nemen in het berekenen van de diëlektrische ruwheid.

In hoofdstuk 4 koppelen we het verbeterde, op fysica-gebaseerde, discrete scattering- and emissie model, genaamd ATS+AIEM+TVG (hoofdstuk 3), met het Community Land Model 4.5 (CLM) om zo grondfysica-eigenschappen af te leiden middels het Local Ensemble Tranform Kalman Filter (LETKF) algoritme door het afzonderlijk assimileren van SMAP Level-1C (L1C) T_B^H en T_B^V. Om te bepalen of enkel het afleiden van grondeigenschappen voldoende is om de schattingen van grondvocht, en daaruit volgend, de hittestromen boven land te verbeteren, zijn twee data-assimilatie experimenten uitgevoerd en vergeleken.

Bij het eerste werden enkel de grondeigenschappen in het assimilatieproces vernieuwd (Only_Para), terwijl dit bij het tweede experiment zowel de grondeigenschappen als het grondvocht waren (Joint_Updt). De resultaten van beide experimenten lieten potentiële verbeteringen zien bij de schattingen van grondeigenschappen van de toplaag door SMAP T_B^p ($p = $H, V) assimilatie, en verbeteringen bij het beschrijven van grondeigenschappen met een vooraf bekende diepteverhouding. Bij het Only_Para experiment bleek het gebruik van T_B^H meer gevoelig voor het afleiden van de kleifractie en de hoeveelheid organisch materiaal, en bleek het gebruik van T_B^V meer gevoelig voor het afleiden van de zandfractie. Aan de andere kant konden bij het Joint_Updt experiment zand- én klei-fractie worden afgeleid door assimilatie van zowel T_B^H als T_B^V. Verder gaf het Jonit_Updt experiment betere schattingen van grondvocht, grondtemperatuur en hittestromen over land gedurende de geassimileerde periode vergeleken met het Only_Para experiment. Echter, voor het gunstig beïnvloeden van grondvochtschattingen bleken deze nauwkeuriger afgeleide grondeigenschappen toch weinig relevant. In plaats daarvan zullen bij toekomstig onderzoek de onzekerheden van de modelstructuren van het CLM worden beschouwd, zoals de vaste PTFs, de hydraulische functie welke de retentiecurve van het grond vocht beschrijft, de waterstresfunctie welke de wateropname door plantwortels bepaald, en ook de parameters voor hittestromen over land.

Als laatste opmerking willen we benadrukken dat het in dit proefschrift ontwikkelde data-assimilatiesysteem voor het afleiden van grondeigenschappen in potentie regionale of zelfs globale grondparametersets kan leveren die zowel fysisch als op meerdere schaalgroottes consistent zijn. Tegelijkertijd merken we op dat verder onderzoek naar de onzekerheden in AOMn en L-band emissiemodellen nodig zijn, zodat zowel het afleiden van grondeigenschappen als het schatten van hittestromen over land verbeterd kan worden.

Biography



Hong Zhao was born on 06 November 1988 in Huazhuang Village, Lanzhou City, Gansu province, China. In June 2012 and 2015 respectively, she received her B.S. and M.Sc. in Cartography and Geographic Information System from Lanzhou University, China. In the same 2015, she was awarded a four-year doctorial scholarship from the China Scholarship Council (CSC), and started to pursue her Ph.D. at the Department of Water Resources, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, the Netherlands. Her research interest involves microwave scattering-emission modeling, soil physical property retrieval, data assimilation, land surface modeling, and soil physical property measurements.

Publications

Zhao, H., Zeng, Y., Wen, J., Wang, X., Wang, Z., Meng, X., & Su, Z. (2021). "An Air-to-Soil Transition Model for Discrete Scattering-Emission Modelling at L-Band", Journal of Remote Sensing, vol. 2021, Article ID 3962350, 20 pages.

Zhao, H., Zeng Y., Lv S., and Su Z. "Analysis of Soil Hydraulic and Thermal Properties for Land Surface Modeling over the Tibetan Plateau." Earth System Science Data 10, no. 2 (2018): 1031-61.

Zhao H, Huang W, Xie T, Wu X, Xie Y, Feng S, Chen F. Optimization and evaluation of a monthly air temperature and precipitation gridded dataset with a 0.025° spatial resolution in China during 1951–2011. Theoretical and Applied Climatology. 2019 Oct 1;138(1-2):491-507.

Zhao, H., Zeng, Y., Han, X., & Su, Z. (2021). Retrieving Soil Physical Properties via Assimilating SMAP Brightness Temperature Observations in the Community Land Model. (Under review in Journal of Advances in Modelling Earth System).

Su, Z., Wen J., Zeng Y., **Zhao H.**, Lv S., van der Velde R., et al. "Multiyear in-Situ L-Band Microwave Radiometry of Land Surface Processes on the Tibetan Plateau." Scientific Data 7, no. 1 (2020): 317.

Lv, S., Zeng, Y., Wen, J., **Zhao, H.** and Su, Z., 2018. Estimation of penetration depth from soil effective temperature in microwave radiometry. Remote sensing, 10(4), p.519.

Xie, Y., **Zhao, H.** & Wang, G., 2015. Spatio-temporal changes in oases in the Heihe River Basin of China: 1963–2013. Écoscience, 22(1), pp.33-46.