High resolution soil moisture mapping using distributed temperature sensing
A data assimilation framework

Dong, Jianzhi

DOI
10.4233/uuid:b6234864-c97b-4efd-91b3-89295b6b5a4a

Publication date
2016

Document Version
Final published version

Citation (APA)
Dong, J. (2016). *High resolution soil moisture mapping using distributed temperature sensing: A data assimilation framework*. [Dissertation (TU Delft), Delft University of Technology].
https://doi.org/10.4233/uuid:b6234864-c97b-4efd-91b3-89295b6b5a4a

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

HIGH RESOLUTION SOIL MOISTURE MAPPING USING DISTRIBUTED TEMPERATURE SENSING

A DATA ASSIMILATION FRAMEWORK

HIGH RESOLUTION SOIL MOISTURE MAPPING USING DISTRIBUTED TEMPERATURE SENSING

A DATA ASSIMILATION FRAMEWORK

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op dinsdag 13 oktober 2016 om 10:00 uur

door

Jianzhi DONG

Master of Science,
Beijing Normal University, Beijing, China,
geboren te Shenyang, China.

Dit proefschrift is goedgekeurd door de

promotor: prof. dr. ir. N. C. van de Giesen
copromotor: dr. ir. S. C. Steele-Dunne

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. N. C. van de Giesen,	Technische Universiteit Delft
Dr. ir. S. C. Steele-Dunne,	Technische Universiteit Delft
Onafhankelijke leden:	
Prof. dr. ir. J.D. Jansen,	Technische Universiteit Delft
Prof. dr. ir. John Selker,	Oregon State University
Prof. dr. D.P. Solomatine,	UNESCO-IHE, Technische Universiteit Delft
Prof. dr. ir. G. De Lannoy,	Katholieke Universiteit Leuven
Dr. rer. nat. C. Montzka,	Forschungszentrum Jülich

Keywords: Soil moisture, Soil temperature, Soil thermal and hydraulic properties, Distributed temperature sensing, Data assimilation

Printed by: Ipskamp Drukkers

Front & Back: Designed by Jianzhi Dong

ISBN 978-94-028-0325-9

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

To my parents

CONTENTS

SUMMARY

Soil moisture is a key variable in controlling land surface and atmospheric interactions. It plays a pivotal role in both the land surface water and energy balance. Large scale soil moisture measurements can be obtained using remote sensing techniques. Calibration/validation of large scale measurements is usually challenging, as it requires accurately representing soil moisture over large areas using point sensors. Capturing the soil moisture spatial variability using point sensors is usually economically infeasible. This thesis provides a way of continuously mapping meter resolution soil moisture over large areas using Distributed Temperature Sensing (DTS).

DTS is an environmental temperature measurement tool, which can provide meter resolution temperature data up to kilometers using fiber optical cables. Soil temperature data at discrete depths can be observed with fine resolution over large areas, when fiber optical cable are installed in the soil. A main focus of this thesis is to extract soil moisture information from soil temperature observations.

Great challenges were shown in estimating soil moisture using soil temperatures in the past decades. A discussion on the difficulties met in the previous studies, and the reasons of why data assimilation is considered to be more suitable for this research are presented in Chapter 1.

The data assimilation approach is more complex than traditional methods, e.g. methods related to thermal inertia, amplitude analysis and inversion of the heat diffusion model. Distinguishing the contributions of different sources of uncertainties and errors in data assimilation is necessary. Synthetic experiments, in which the error sources were known by design, were used to investigate the impacts of each type of error/uncertainty. Multiple-truth synthetic tests were also used to test the robustness of the proposed data assimilation method across a wide range of soil textures.

Motivated by the fact that a sequence of soil temperatures may contain more information of soil moisture, a Particle Batch Smoother (PBS) was proposed to assimilate soil temperature observations within a window to estimate soil moisture estimates of that window. Chapter 3 provides detailed discussions on the performance of the PBS, the impacts of observations intervals, and the window lengths. The PBS was shown to outperform the Particle Filter (PF), which demonstrated the hypothesis that assimilating a sequence of soil temperatures is more suitable for soil moisture estimation.

Updating soil moisture without updating hydraulic soil properties may lead to inconsistent combinations of soil moisture and soil hydraulic properties. Chapter 4 showed the physical links between soil temperatures and soil hydraulic properties. The accuracy of estimating soil hydraulic properties using soil temperatures was tested using a set of synthetic experiments. It was also shown that jointly updating soil moisture and soil hydraulic properties provided significantly improved soil moisture estimates compared to the case that only soil moisture was updated.

To apply the proposed data assimilation method under different vegetation cover, Chapter 5 improved the current state of the art vadose zone model by including a vegetation energy balance scheme. In addition, a data assimilation scheme that can handle the uncertainties of the cable depths was proposed. This is particularly relevant to the DTS applications, since the cable depths of the DTS are rarely known accurately.

Meter resolution soil moisture along a 71 m DTS cable was estimated in Chapter 6. This high resolution soil moisture map was evaluated using field measurements. Furthermore, the statistics of the soil moisture spatial variability presented by this high resolution soil moisture map were shown to be consistent with previous studies. This highlighted the feasibility and the potential of using DTS as a tool for continuous high resolution soil moisture mapping, which may fundamentally improve our knowledge about soil moisture scaling, and large scale soil moisture validation.

In addition to soil moisture mapping, a new data assimilation algorithm, the Adaptive Particle Batch Smoother (APBS) was also proposed in Chapter 6. The APBS algorithm was developed for handling the particle weight degeneracy problem. This APBS algorithm will also be applicable to general hydrological data assimilation as discussed in Chapter 7.

SAMENVATTING

Bodemvocht is een sleutel variabele die interacties tussen het landoppervlak en de atmosfeer controleert. Het speelt een centrale rol in zowel water op het landoppervlak als de energiebalans. Grootschalige bodemvochtmetingen kunnen gedaan worden met remote sensing technieken. Kalibratie/validatie van grootschalige metingen is over het algemeen een uitdaging, omdat bodemvocht accuraat gerepresenteerd moet worden door punt sensoren over grote gebieden. Het vatten van de ruimtelijke variabiliteit van bodemvocht is meestal economisch gezien niet mogelijk. Dit proefschrift voorziet een manier van het continu in kaart brengen van bodemvocht met meter-resolutie over grote gebieden met behulp van gedistribueerde temperatuur waarneming (Distributed Temperature Sensing, DTS).

DTS is een meetinstrument voor omgevingstemperatuur, dat met een lengte van vele kilometers meter-resolutie temperatuurdata kan leveren met behulp van glasvezelkabels. Als glasvezelkabels in de bodem geïnstalleerd zijn, kan de bodemtemperatuur op bepaalde dieptes worden waargenomen met een hoge resolutie over grote gebieden. Een belangrijke focus van dit proefschrift bestaat uit het afleiden van informatie over bodemvocht uit waarnemingen van bodemtemperatuur.

De afgelopen decennia heeft men grote uitdagingen laten zien in het schatten van bodemvocht met behulp van bodemtemperatuur. Een discussie over de moeilijkheden die voorgaande studies zijn tegengekomen en de redenen waarom data-assimilatie voor dit onderzoek meer gepast wordt verondersteld, wordt gepresenteerd in Hoofdstuk 1.

De data-assimilatie benadering is complexer dan traditionele methodes, zoals methodes gerelateerd aan thermische traagheid, amplitude analyse en inversie van het warmtediffusie model. Het onderscheiden van het aandeel van de verschillende bronnen van onzekerheden en fouten in data-assimilatie is noodzakelijk. Kunstmatige experimenten, waar de foutenbronnen bekend zijn door het ontwerp van het experiment, zijn gebruikt om de impact van elk type fout/onzekerheid te onderzoeken. Meerdere-waarheid kunstmatige tests zijn ook gebruikt om de robuustheid van de voorgestelde data assimilatiemethode te testen over een breed scala van bodemtexturen.

Gemotiveerd doordat een serie van bodemtemperaturen mogelijk meer informatie over bodemvocht bevat, is een Particle Batch Smoother (PBS) voorgesteld om waarnemingen van bodemtemperatuur te assimileren binnen een interval om bodemvocht in te schatten van dat interval. Hoofdstuk 3 geeft een gedetailleerde discussie over de prestatie van PBS, de impact van waarnemingsintervals, en de interval lengte. Het wordt aangetoond dat PBS beter presteert dan een Particle Filter (PF), wat de hypothese bevestigt dat het assimileren van bodemtemperaturen geschikter is voor het inschatten van bodemvocht.

Het bijwerken van bodemvocht zonder het bijwerken van hydraulische bodemeigenschappen kan leiden tot inconsistente combinaties van bodemvocht en hydraulische bodemeigenschappen. Hoofdstuk 4 laat de fysieke relaties zien tussen bodemtempera-

turen en hydraulische bodemeigenschappen. De accuraatheid van het schatten van hydraulische bodemeigenschappen met bodemtemperaturen is getest met behulp van een aantal kunstmatige experimenten. Het wordt ook aangetoond dat het gezamenlijk bijwerken van bodemvocht en hydraulische bodemeigenschappen in een significant verbeterde bodemvochtschatting voorziet in vergelijking met het geval waarin alleen bodemvocht wordt bijgewerkt.

Hoofdstuk 5 verbetert het huidige 'state-of-the-art' onverzadigde zone model door het toevoegen van een vegetatie-energiebalans-schema om de voorgestelde data assimilatiemethode toe te passen met een verschillende vegetatiebedekking. Bovendien wordt een data-assimilatieschema dat de onzekerheden van kabeldieptes kan meenemen voorgesteld. Dit is in het bijzonder relevant voor DTS-toepassingen, omdat de kabeldieptes zelden bekend zijn.

Meter-resolutie bodemvocht langs een 71 m DTS-kabel wordt ingeschat in Hoofdstuk 6. Deze hoge resolutie bodemvochtkaart is geëvalueerd met behulp van veldmetingen. Verder wordt aangetoond dat de statistieken van de ruimtelijke variabiliteit van bodemvocht gerepresenteerd door deze hoge resolutie bodemvochtkaart, overeenkomen met voorgaande studies.

Dit benadrukte de haalbaarheid en het potentieel van het gebruik van DTS als een instrument voor het continu in kaart brengen van bodemvocht met een hoge resolutie, wat onze kennis over het schalen van bodemvocht fundamenteel verbetert, en bodemvochtvalidatie op grote schaal.

Als toevoeging op het in kaart brengen van bodemvocht, een nieuw data-assimilatie algoritme, de Adaptive Particle Batch Smoother (APBS) wordt voorgesteld in Hoofdstuk 6. Het APBS algoritme was ontwikkeld voor het omgaan met het probleem van de weegfactor van een deeltje. Dit APBS-algoritme zal ook toepasbaar zijn op algemene hydrologische data-assimilatie, zoals bediscussieerd in Hoofdstuk 7.

PREFACE

Microwave remote sensing might be the most feasible way for us to measure soil moisture globally. The resolution of the microwave remote sensing soil moisture products is usually larger than 10 km^2. It is usually necessary to upscale point measurements to this footprint scale for calibrating and validating the microwave remote sensing products.

The real challenge of scaling soil moisture products is to understand soil moisture variability at different scales. The spatial variability of soil type, vegetation cover, topography and meteorological forcing lead to the spatial variability of soil moisture. We also found that this spatial variability evolves in time. The most reliable way of understanding the soil moisture spatial variability, and the rules of soil moisture scaling, is intensively (both in space and time) measuring soil moisture at point scales. DTS can provide soil temperature measurements with temporal resolution < 1 min, spatial resolution < 1 m, up to kilometers. Hence, DTS might be a suitable tool for continuously measuring high spatial resolution soil moisture over large scales, provided that the soil temperature measurements can be used for soil moisture estimation.

Very limited satisfactory results of estimating soil moisture using soil temperature data were reported in the past decades. In the past 40 years, heat transfer processes were usually simplified by making different types of assumptions. Soil thermal properties are then estimated using the simplified heat conduction equation and the observed soil temperatures. The estimated soil thermal properties can be linked to soil moisture, provided soil thermal property to moisture relationship is known. Structural errors in the simplified soil heat transfer equation and the uncertainties in the soil thermal property to soil moisture sometimes mask the soil moisture information. As a result, it is difficult to accurately quantify either the absolute soil moisture value, or the soil moisture dynamics using these traditional methods.

This thesis uses a different concept of extracting soil moisture information from soil temperature, i.e. data assimilation. In data assimilation, soil moisture is first simulated using the state of the art vadose zone model. Ensemble model runs are used to map the prior distribution of the soil moisture and soil temperature under different sources of uncertainties. Then this prior estimated soil moisture will be constrained using soil temperature. I agree that “the main purpose of science is simplicity" (Edward Teller). However, there are reasons why solving this problem in a more “complex" way is preferred here. By increasing the complexity, more information can be included to constrain the estimated soil moisture. For example, provided the meteorological forcing is known, the forward model can already provide reasonable soil moisture dynamics. This information is not considered in traditional methods. Soil evaporation is a key clue for linking soil moisture and temperature, which is not included in the traditional methods.

The thesis also presents a learning and understanding process of how to extract soil moisture information from such a complex system. Distinguishing the contributions of different sources of uncertainties is challenging. The performance of the forward model

and the robustness of the data assimilation algorithm can both affect the final estimates. In synthetic tests, the uncertainties and the model structural errors are known by design. The “true" values are also perfectly known. This provides a way of accurately evaluating the concept and the data assimilation algorithm. Synthetic tests were frequently used to investigate the conceptual feasibility, and the robustness of the data assimilation algorithm in this thesis. Then, the proposed algorithms were tested using traditional point sensors, which have the best quality controlled soil temperature and moisture measurements. This is to minimize the impacts of observation errors on the algorithm evaluation. Finally, this algorithm is applied to real DTS data, which can continuously map the soil moisture field with high spatial resolution. The estimates are evaluated using field measurements at a nearby site.

I am, or maybe was, (probably) a good fieldworker. At least, I didn't work on a lot of equations before my Ph.D. research. I planned to focus on field experiments, though not messing up all the DTS cables in the field is still a big challenge for me. However, this thesis mainly focused on the algorithm development. This is because robustly estimating soil moisture using soil temperatures is still a big challenge to the soil physics community, and it is more attractive to me. I have to say I am lucky, and I would thank all my colleagues, both in the Netherlands and the USA, who made the field experimental data available for me. This gave me the opportunity of focusing on the equations of the data assimilation, and the physics of soil water and heat transfer processes.

I hope this thesis can contribute to the area of high resolution environmental variable mapping and hydrological data assimilation. In particular, I hope this thesis could also remind myself that I have done something that I am really proud of.

Jianzhi Dong
Delft, June 2016

1

INTRODUCTION

We can't solve problems by using the same kind of thinking we used when we created them.

Albert Einstein

1

1.1. IMPORTANCE OF SOIL MOISTURE

SOIL moisture refers to the water content of a certain volume of soil. Soil moisture plays a central role in the land surface - atmosphere processes [1]. For example, the infiltration capacity of a watershed is largely determined by prevailing soil moisture conditions [2]. Consequently, rainfall - runoff processes strongly depend on surface soil moisture [3]. Soil moisture also determines the availability of the land surface water that can be evaporated into the atmosphere [4], and hence soil moisture also has significant impacts on the land surface energy balance, and land surface atmosphere interactions. [5, 6]. Thus, improved understanding of soil moisture may fundamentally improve flood forecasting [7], drought monitoring [8], and heat wave prediction [6].

1.2. SOIL MOISTURE MEASUREMENT TECHNIQUES

The most feasible way of providing global coverage large scale soil moisture measurements is to use remote sensing techniques. Estimating soil moisture using thermal remote sensing was intensively discussed in the 1970s [e.g. 9–12]. Soil thermal inertia is correlated to the differences of daily maximum and minimum soil temperature. Regression methods can link daily maximum and minimum temperature differences or soil thermal inertia to the soil moisture values, provided the calibration soil moisture data are available. However, it is challenging to retrieve soil moisture using the thermal inertia method at the global scale. First of all, the observed land surface temperature (LST) is a weighted average of soil and vegetation temperatures. Hence, the thermal inertia method is only applicable to bare soil and sparsely vegetated areas [13]. Second, LST cannot be observed during cloudy days, which will limit the availability of the thermal inertia method. Third, the maximum and the minimum soil temperature differences were assumed to be primarily attributed to soil thermal inertia, an assumption which is often violated. For example, low maximum and minimum soil temperature differences can be attributed to solar radiation, rather than soil thermal inertia or soil moisture [9]. Soil evaporation also significantly affects soil temperature evolution, and hence the differences of the maximum and minimum soil temperature. This may lead to substantial errors in soil moisture retrieval using soil thermal inertia methods [14].

Microwave remote sensing is the most popular technique for providing footprint scale ($> 10\text{km}^2$) soil moisture observations. Compared to soil thermal remote sensing methods, microwave remote sensing methods are less affected by the atmospheric conditions. Soil dielectric constant is a strong function of soil moisture, which has a value of approximately 3.5 for dry soil, and approximately 80 for water [15]. In active microwave remote sensing, pulses are sent to the ground, and the backscatter of these pulses are measured. The backscatters depend on soil dielectric constant, which can thus be related to soil moisture [16]. Passive microwave remote sensing measures the soil brightness temperature (T_b), which can be used for solving soil dielectric constant [15]. However, retrieving soil moisture can be complicated by vegetation cover and soil surface roughness. To accurately retrieve soil moisture, different algorithms were proposed [17], which can be primarily categorized into two groups. The first group is to use statistical or machine learning methods to calibrate/train the relationship between ground based soil moisture measurements and brightness temperatures or backscatters [e.g. 18, 19].

Another group of retrieval methods is to invert soil moisture from more physically based models [e.g. 20–23]. Ancillary data, e.g. land cover classification maps, vegetation indices, multiangular microwave remote sensing observations, were usually used to quantify the impacts of vegetation and soil surface roughness in these methods. Since remote sensing measurements are at scales of kilometers, ground based soil moisture measurements at those scales are also required to calibrate or validate the soil moisture retrieval algorithms.

Many ground based methods are available for soil moisture measurement. The only direct soil moisture measurement approach is the gravimetric method, which is usually used for calibrating and validating other soil moisture techniques. In the gravimetric method, soil moisture is derived by oven drying a certain volume of a soil sample. However, the gravimetric method is destructive, laborious and time consuming [24]. Time Domain Reflectometer (TDR) and capacitance probes are alternatives of the gravimetric method for point scale soil moisture measurement. Similar to microwave remote sensing techniques, TDR and capacitance probes also measure soil moisture by measuring the soil dielectric constant [25]. Compared to the gravimetric method, TDR and capacitance probes are more efficient, and good accuracies can be achieved after site specific calibration.

Accurately representing footprint scale soil moisture may require numerous point scale measurements to capture the soil moisture spatial variability. It is challenging since intensively sampling point scale soil moisture across large areas is usually economically infeasible. Innovative devices were proposed to measure intermediate scale soil moisture with the aim of linking point scale and intermediate scale soil moisture measurements. A popular intermediate scale soil moisture measurement technique is Global Positioning System Reflectometry (GPS-R), which was shown to be capable of monitoring the areal mean surface soil moisture of 300 m^2 [26, 27]. The reflection depth depends on land surface soil moisture conditions, which determines the phase offset of the Signal to Noise Ratio (SNR) expression. Hence, the SNR detected by GPS-R receivers could be used for measuring surface soil moisture. Similar to microwave remote sensing, GPS-R technique is sensitive to the soil surface roughness and the vegetation conditions [28], which may be of practical concern for GPS-R applications.

In parallel to the GPS-R technique, the Cosmic-ray probe was proposed to represent the average soil moisture of a large area (~ 300 m radius) [29]. The hydrogen atoms have the greatest capacity of slowing down the fast cosmic-ray neutrons [29]. Therefore, the intensity of fast neutrons is negatively related to total hydrogen atoms of the environment. A reference value of the neutron counting rate over dry soil is required, which can be obtained using field calibration. To estimate soil moisture, other sources of hydrogen atoms (water content), e.g. atmospheric water vapor, soil surface ponding water, biomass and lattice water, have to be separately estimated [30]. Separating the contribution of vegetation water content is usually most challenging, especially when the vegetation water content changes with time, e.g. during the grow seasons. A second key challenge of successfully using cosmic-ray probes is to deal with the variances in the measurement depths, with a measurement depth of 5 cm for wet soils and approximately 70 cm for dry soils [29]. Therefore, the cosmic-ray probe essentially integrates soil moisture values at different depths over the measurement area, instead of one single depth.

This may complicate the interpretation of Cosmic-ray probe measurements. This means calibrating the Cosmic-ray probe may also require high resolution point scale measurements, especially when field spatial variability is high.

1.3. SOIL MOISTURE UPSCALING

As shown in the previous section, measuring soil moisture at large scales is usually complicated by different sources of uncertainties. Different upscaling algorithms have been proposed with the aim of accurately representing the large scale soil moisture using sparse point scale measurements, which may reduce the cost of validating large scale soil moisture measurement techniques.

Temporal stability is a popular upscaling technique. The concept is to only monitor soil moisture at locations that can best represent the areal mean soil moisture [e.g. 31, 32]. Hence, the temporal stability method can reduce the number of point sale sensors required to observe large scale soil moisture. In order to locate the representative locations, high spatial resolution soil moisture measurements with a period of approximately 1 year are usually required [33]. Kriging interpolation methods were also frequently investigated [e.g. 34, 35]. This type of algorithm calculates the areal mean soil moisture by weighted averaging the available point measurements [36]. Similar to the temporal stability method, the upscaling equation has to be parameterized using a sequence of observations. The feasibility of upscaling soil moisture observations using land surface modeling was also explored [37]. Well-parameterized land surface models may be helpful in capturing the spatial variability within a footprint, and improving the accuracy of the upscaled soil moisture estimates. However, it is also shown that point measurements within the footprint are essential in removing the bias of the model estimates. Hence, the accuracy of the algorithm also depends on the availability of the point measurements [37].

Clearly, the first step of successfully applying an upscaling algorithm is to accurately represent the soil moisture spatial variability. Otherwise, applying the upscaling algorithms may even yield larger sampling errors comparing to the cases that no upscaling algorithms are applied [37]. This indicates that though different upscaling algorithms are available, well distributed point scale measurements over large scales are still required.

1.4. SOIL MOISTURE MAPPING USING DTS

As clearly shown in the previous two sections, both intermediate scale soil moisture measurement techniques and upscaling algorithms can be affected by different sources of uncertainties. The most reliable way of bridging the gap between point and intermediate scale soil moisture is to intensively collect point measurements over intermediate scales. Distributed temperature sensing (DTS) has been considered to have the potential of providing cost-effective high resolution soil moisture measurements over large areas [38, 39].

DTS measures environmental temperature using fiber-optic cables. In DTS, a light pulse is first sent into the fiber-optic cable. Due to the Raman scattering, a portion of light will be scattered at a decreased (Stokes) or increased (anti-Stokes) frequency. The backscatter of the Stokes and anti-Stokes will be received by the DTS unit. The inten-

sities of Stokes are linearly correlated with the illumination [40], while the intensities of the anti-Stokes depend on both the illumination and the temperature of fiber-optic cable. Hence, the ratio of the two signals strongly depends on the fiber temperature. The velocity of light in the fiber is a constant, which can be used for locating the measured temperature by recording the time that the backscattered signal is received. DTS can provide high resolution (temporal < 1 min, and spatial < 1 m) environmental soil temperature measurements using cables up to kilometers in length [40].

There are two categories of DTS methods for soil moisture measurement, namely Active DTS and Passive DTS. Active DTS methods apply an electrically generated heat pulse to the fiber optic cables. Due to the large heat capacity and conductivity of water, the electrically generated heat is more likely to dissipate in wetter soils. Consequently, wetter soil will present a smaller temperature increase after heating. Therefore, temperature change during and after this heat pulse can be related to soil moisture using either empirically or physically based equations [39, 41]. Passive DTS, on the other hand, uses soil thermal responses to the net solar radiation to estimate soil moisture. The feasibility of the Passive DTS was demonstrated by an “inversion" method [38]. In the inversion method, the soil thermal diffusivity is estimated by finding the value that gives the best fit between the observed and estimated soil temperature within a certain period. This estimated soil thermal diffusivity can be used for estimating soil moisture, provided soil thermal diffusivity to soil moisture relationship is known.

Active DTS is relatively more accurate, particularly when soil moisture is low. However, the energy consumption of the heat pulses may be a logistic obstacle for field applications. This is particularly true for experiment sites that rely on electric power collected using solar panels.

Passive DTS can overcome the drawbacks of the Active DTS methods. First of all, no energy is required for generating heat pulses. Hence, it is easier to operate the Passive DTS than Active DTS in field applications. Further, Passive DTS uses soil temperature data of a longer time window (e.g. 24 hours), which provides sufficient time for the thermal equilibrium between soil and air gaps. Hence, Passive DTS is insensitive to the gaps between fiber optic cable and soil.

However, the current inversion approach based Passive DTS also has several key challenges in the field applications. First, soil thermal diffusivity is insensitive to soil moisture for a wide range of soil moisture values. Second, physically reasonable estimates cannot be obtained from the inversion approach when solar radiation is low. Third, it is difficult to determine the soil thermal diffusivity to moisture relationship along the cable, when the spatial variability of soil properties is high. Finally, the soil thermal diffusivity/soil moisture estimates are very sensitive to uncertainties in the cable depths.

1.5. DATA ASSIMILATION

Data assimilation may be useful in handling the uncertainties in the Passive DTS method, and improving the accuracy of the soil moisture estimates. Data assimilation refers to a broad range of techniques that can optimally combine model estimates and observations, in which different sources of uncertainties can be considered [42]. The Kalman filter is one of the most popular data assimilation techniques. For linear systems, the

predicted model states (prior) are normally distributed, provided initial states and the model errors are both normally distributed. When observations are available, the prior estimates can be updated (posterior) using Bayes' theorem. By optimally combining two different sources of information, i.e. the prior and the observation, the posterior has smaller uncertainties and is closer to the "truth". The Kalman filter was modified for solving non-linear problems, such as unscented Kalman filter and extended Kalman filter. The Ensemble Kalman Filter (EnKF) algorithm is a popular variant of the Kalman filter ([43]). The EnKF also assumes the model prior is normally distributed and approximates its first two momentums using random samples, since analytical expressions are usually unavailable. Once the prior distribution is estimated, the posterior is calculated using similar procedures as described in the Kalman filter. The EnKF is conceptually simple, and can be easily implemented for complex systems. Hence, the EnKF algorithm is widely used in hydrology related research, e.g in improving footprint scale soil moisture estimates [e.g. 44–48], terrestrial water storage [e.g. 49–51], and stream flow estimates [e.g. 2, 52].

When the Gaussian assumptions are violated and the system is highly non-linear, the EnKF cannot provide optimal solutions. Instead of only preserving the first two momentums of the prior, the Particle Filter (PF) maps the entire prior distribution using Monte Carlo samples and the associated weights [53]. The posterior distribution is derived by updating the weight of each particle using Bayes' theorem. The PF algorithm was intensively tested and widely applied for hydrology related studies [e.g. 53–56]. A recent study also shows PF algorithm outperforms the EnKF, particularly in parameter estimation [57]. However, the PF algorithm is not without drawbacks. The prior distribution is essentially approximated using a discrete form, which means few particles will be sampled at the tails of the prior distribution. As a consequence, when the observations are located at the tails of the prior, the PF algorithm may have severe weight degeneration problems. Hence, the EnKF can usually provide superior estimates under these conditions [56]. To address this problem, Markov Chain Monte Carlo (MCMC) was included to reduce the weight degeneracy problem [58].

Batch smoothing may be more suitable for the estimation problems when a sequence of observations contains more information than observations made at instantaneous time steps. The Ensemble Smoother (ES) is a popular smoothing technique. The ES algorithm is essentially the same as the EnKF algorithm, except that observations are assimilated in a batch to update all states within some window in time. This allows the ES to consider observations beyond the estimation time. The ES is shown to be suitable for estimating land surface soil moisture states by assimilating microwave remote sensing observations [59]. Series of soil temperature measurements contain more information about the soil heating and cooling rate, and therefore may contain more information about land surface evapotranspiration information. Hence, the ES algorithm was also demonstrated to be useful in estimating land surface latent heat fluxes using land surface temperatures [60]. However, the ES may be inaccurate when the system is highly non-linear or the prior distribution is not normally distributed [61].

1.6. DATA ASSIMILATION & DTS

Data assimilation has the potential to improve the current Passive DTS method, tackling the challenges in the inversion method and making DTS a viable tool for high resolution intermediate scale soil moisture measurement.

In data assimilation, the joint distribution of soil moisture and temperature could be estimated using a fully coupled soil water, heat and vapor transfer model (forward model). Therefore, the prior estimates can already provide reasonable soil moisture estimates in data assimilation. For cases when soil temperature has little soil moisture information (e.g. on low radiation days), data assimilation methods can use the prior guessed soil moisture, instead of converging to physically unreasonable values, as in the inversion approach. Second, the link of soil moisture and soil temperatures are implicitly considered in the coupled processes of soil water, heat and vapor transfer, which is advantageous compared to the inversion method. In the inversion approach, only the impacts of soil moisture on the vertical soil heat transfer are considered. In addition to the soil heat transfer process, soil moisture and soil temperature links were also considered via vapor transfer processes in data assimilation. Hence, soil moisture may be more effectively estimated using soil temperatures in data assimilation. In sequential data assimilation, the joint distribution of soil moisture and temperature is firstly estimated using ensemble runs. Uncertainties in the meteorological forcing, soil properties, model and observation errors were considered in each ensemble run. The prior estimates are then updated using observations to yield posterior estimates, which essentially adjusts the probability (weights) of the prior estimates. Hence, the data assimilation estimates can account for the uncertainties from different sources, and always provide physically reasonable estimates. On the other hand, the inversion approach calculates soil moisture in a deterministic way, i.e. the forward model (soil thermal diffusion model) and soil moisture - diffusivity relationship are all assumed to be perfect. As a result, the uncertainties in the inversion approach will accumulate in the soil moisture, and may result in physically unreasonable estimates.

Although it is promising, the concept of estimating soil moisture by assimilating soil temperatures has never been proposed or tested, and several key challenges have to be addressed. First, it is still unknown whether the coupling between soil water and heat transfer process is strong enough to estimate soil moisture using soil temperatures. If the correlation between soil moisture and temperature is weak, it may be masked by the uncertainties. Second, the optimal data assimilation technique for this estimation problem, and how to implement the data assimilation techniques to yield robust estimates are still yet unknown. DTS can observe soil temperature every minute or less. However, assimilating soil temperature every minute may lead to ensemble collapse and ignore the observations. An optimal assimilation interval,that can provide sufficient soil moisture information and can keep the data assimilation system healthy, is required. Further, a batch smoother, rather than sequential filtering may be more suitable for this estimation problem. A sequence of soil temperatures can better describe the soil heating and cooling rate, and hence it contains more information of soil moisture. However, assimilating soil temperatures within a batch window using ES may yield less accurate soil moisture estimates, since projecting the joint distribution of soil temperature and moisture estimates into Gaussian distributions can be inappropriate. Therefore, a new

batch smoothing algorithm is required to robustly assimilate a sequence of observations. Third, the state of the art of the vadose zone model cannot simulate the surface energy balance under vegetated areas. This will lead to incorrectly estimated soil moisture and temperature relationships, and may result in biased soil moisture estimates. Improvements have to be made in the forward model to account for the vegetation affects on the land surface energy balance. Finally, the observation depths are uncertain due to the difficulties in DTS installation. As shown in the inversion approach, the uncertainties in the observation depths can usually lead to physically unreasonable soil moisture estimates. Dealing with this type of uncertainty is also rarely reported. Thus, a data assimilation algorithm that can account for the observation depth uncertainties is needed. The main objective of this thesis is to design and test such an algorithm.

1.7. RESEARCH OUTLINE

The goal of this thesis is to develop a new Passive DTS framework for large scale soil moisture mapping by assimilation DTS observed soil temperature data. The rest of the thesis is structured as follows:

Chapter 2 tests the hypothesis that soil moisture can be estimated by assimilating soil temperatures. Synthetic experiments under different assumptions of model structural errors are used to test the robustness of the proposed data assimilation method, which demonstrate the strong physical links between the soil heat and water transfer processes. The impacts of assimilating soil temperature at different intervals and at different depths are discussed, which aims to investigate the optimal strategies for applying the data assimilation method in estimating soil moisture using soil temperatures. The popular Ensemble Kalman Filter will be used in this chapter.

Chapter 3 tests the hypothesis that soil temperature evolution within a certain length of time window, rather than an instantaneous soil temperature measurement, contains more soil moisture information. In order to assimilate all the soil temperature observations within a batch window, we proposed a new data assimilation technique, the Particle Batch Smoother (PBS). We will also demonstrate that batch smoothing outperforms sequential filtering.

Chapter 4 tests the hypothesis that soil moisture, thermal properties, and hydraulic properties can be jointly estimated by assimilating soil temperatures. We show that estimating model states and parameters jointly leads to significantly improved estimates compared to the cases of updating state alone, as presented in Chapter 3.

Chapter 5 solves two key challenges in applying the methodology developed in chapter 4 to real DTS data. First, the state of the art vadose zone model is improved by coupling a vegetation energy balance scheme. This allows us to apply the proposed method under various land covers. Second, we provide a method for estimating observation depths. This allows us to estimate soil moisture information when the observation (DTS cable) depths are poorly known. Data collected from an intermediate scale soil moisture network are used for demonstration.

Chapter 6 applies the PBS methods to a section of real DTS data. The theories and algorithms presented in chapter 5 are used in this chapter. The estimated soil moisture, thermal and hydraulic properties are validated using measurements collected at a nearby site. An improved PBS is also presented to provide robust estimates by assimilat-

ing DTS observed soil temperatures.

Chapter 7 synthesizes the key contributions of this thesis in high spatial resolution soil moisture monitoring using Passive DTS. The knowledge generated and the contributions of the thesis are also discussed in a broader context, e.g. intermediate scale soil moisture monitoring, soil - hydrology research, data assimilation.

2

DETERMINING SOIL MOISTURE BY ASSIMILATING SOIL TEMPERATURE MEASUREMENTS USING THE ENSEMBLE KALMAN FILTER

Based on: Dong Jianzhi, Susan C. Steele-Dunne, Tyson E. Ochsner, Nick van de Giesen. Determining soil moisture by assimilating soil temperature measurements using the Ensemble Kalman Filter, 2015, *Advances in Water Resources*, 85, 340-353.

2.1. INTRODUCTION

The previous chapter summarized the importance and the challenges of measuring high spatial resolution soil moisture using Passive DTS over large scales. It also proposed the concept of using data assimilation to handle the uncertainties in Passive DTS. Since the hypothesis that soil moisture can be estimated by assimilating soil temperatures has never been tested, this chapter tests this hypothesis using synthetic tests.

In this chapter, synthetic experiments are used to demonstrate that the ensemble Kalman filter [43, 62] could be used to estimate soil moisture by merging soil moisture and temperature profiles from a fully coupled soil heat and water transport model (Hydrus-1D, [63]) with temperature observations at a limited number of depths (section 2.2). Synthetic experiments are preferred in this chapter is because the true soil temperature and moisture are perfectly known at every time step and depth and the input uncertainties are known by design. Therefore, the performance of the data assimilation algorithm and the concept of this chapter can be explicitly and precisely evaluated. We also demonstrate that synthetic experiments can be a useful tool in designing DTS experiments, e.g. in determining the number of depths at which cables are required and which combinations of cable depths yields the most information on soil moisture (section 2.3).

2.2. METHOD AND MATERIALS

2.2.1. HYDRUS-1D MODEL

In this chapter, the vertical soil water, heat and vapor transport processes in the unsaturated zone are simulated using the Hydrus-1D model [63]. The governing equation for one-dimensional liquid and vapor flow is expressed as:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K_{Th}\frac{\partial h}{\partial z} + K_{Lh} + K_{TT}\frac{\partial T}{\partial z}\right] - S \tag{2.1}$$

where θ is soil water content (m^3m^{-3}) at time t (s), and z is the vertical coordinate (positive upward) (m). K_{Th} and K_{TT} are the isothermal and thermal total hydraulic conductivities, respectively, and K_{Lh} is the isothermal unsaturated hydraulic conductivity. S is a sink term ($m^3m^{-3}s^{-1}$). K_{Lh} and the soil retention curve are determined using van Genuchten's model [64] :

$$K_{Lh} = K_s S_e^l \left[1 - (1 - S_e^{\frac{1}{m}})^m\right]^2 \tag{2.2}$$

$$\theta(h) = \begin{cases} \theta_r + \frac{\theta_s - \theta_r}{[1+|\alpha h|^n]^m} & h < 0 \\ \theta_s & h \geq 0 \end{cases} \tag{2.3}$$

where K_s is the saturated hydraulic conductivity (ms^{-1}), S_e is the effective saturation, l, m, n and α are empirical shape parameters and θ_r and θ_s are the residual and saturated soil water contents (m^3m^{-3}). The parameters in equation (2.2) and (2.3) were estimated using the ROSETTA model [65]. The basic input soil properties for ROSETTA model are soil texture (in terms of percentage sand, silt and clay) and bulk density(ρ_b).

The governing equation for soil heat transport is:

$$\frac{\partial C_p T}{\partial t} + L_0 \frac{\partial \theta_v}{\partial t} = \frac{\partial}{\partial z}\left[\lambda(\theta)\frac{\partial T}{\partial z}\right] - C_w \frac{\partial q_L T}{\partial z} - L_0 \frac{\partial q_v}{\partial z} - C_v \frac{\partial q_v T}{\partial z} - C_w S T \tag{2.4}$$

where T is soil temperature in Kelvin, C_w,C_v and C_p are the volumetric heat capacities of water, vapor and moist soil ($\mathrm{Jm^{-3}K^{-1}}$), L_0 is the volumetric latent heat of vaporization of liquid water ($\mathrm{Jm^{-3}}$), q_L and q_v are the flux densities of liquid water and vapor ($\mathrm{ms^{-1}}$), and $\lambda(\theta)$ is the apparent soil thermal conductivity ($\mathrm{Wm^{-1}K^{-1}}$). $\lambda(\theta)$ is estimated from:

$$\lambda(\theta) = \lambda_0(\theta) + \beta_0 C_w |q_L| \tag{2.5}$$

where β_0 is the thermal dispersivity (m), and thermal conductivity λ_0 can be estimated in Hydrus-1D using either the model of Campbell et al. [66] or Chung and Horton [67]. The Chung and Horton model only provides three parameter sets for sand, loam and clay soil types. The Campbell model can estimate the relationship between soil thermal conductivity and soil moisture for any given soil texture. Therefore, the Campbell model will be used here in the forward model of ensemble open loop and assimilation experiments.

The upper boundary condition in the soil water movement calculations was “Atmospheric boundary condition with surface runoff". Several options for the lower boundary condition for soil water movement are available in Hydrus. In the forward model of ensemble open loop and assimilation experiments, the “free drainage" lower boundary condition will be used. The upper boundary for heat transfer was “Heat flux", which is calculated based on the soil surface energy balance, and the lower boundary is “zero gradient". Model details and a complete description of these boundary conditions are provided by Saito et al. [68]. In this chapter, the model time step is approximately 1min, and the soil column to 1m depth is simulated with a vertical resolution of 1cm.

2.2.2. ENSEMBLE KALMAN FILTER IMPLEMENTATION

In the EnKF([43, 62, 69]), each of the N ensemble members is propagated forward in time using the forward model:

$$\mathbf{x}_t^{i,f} = f\left(\mathbf{x}_{t-1}^{i,a}, \mathbf{u}_t^i, \mathbf{b}^i\right) + \mathbf{w}_t^i \tag{2.6}$$

where f represents the forward model (Hydrus-1D in this case), $\mathbf{x}_t^{i,f}$ is the i^{th} ensemble model forecast (prior) state at time t, and the $\mathbf{x}_{t-1}^{i,a}$ is the updated (posterior) model state at time $t-1$, $\mathbf{u}^i$ is the perturbed forcing data, $\mathbf{b}^i$ is a vector of time invariant model parameters (e.g. soil thermal and hydraulic properties), and $\mathbf{w}^i$ is the model error. In this experiment, the state vector ($\mathbf{x}_t^i$) is defined to include the temperature and soil moisture in each layer of the soil column, and we use 100 ensemble members . The state update equation is as follows:

$$\mathbf{x}_t^{i,a} = \mathbf{x}_t^{i,f} + \mathbf{K}(\mathbf{d}_t^i - \mathbf{H}\mathbf{x}_t^{i,f}) \tag{2.7}$$

where $\mathbf{d}_t^i$ is the vector of perturbed observations, $\mathbf{K}$ is the Kalman gain term and $\mathbf{H}$ is the measurement operator that relates the model states to the observations. In this chapter, the observations are temperatures at a limited number of depths.

The Kalman gain term (**K**) is given by:

$$\mathbf{K} = \mathbf{P}_f \mathbf{H}^T (\mathbf{H}\mathbf{P}_f \mathbf{H}^T + \mathbf{R})^{-1} \tag{2.8}$$

where **R** is the observation error covariance, and $\mathbf{P}_f$ is the forecast error covariance. In EnKF, $\mathbf{P}_f$ is estimated from the ensemble as follows:

$$\mathbf{P}_f = \frac{1}{N-1} \sum_{i=1}^{N} (\mathbf{x}_t^{i,f} - \bar{\mathbf{x}}_t^f)(\mathbf{x}_t^{i,f} - \bar{\mathbf{x}}_t^f)^T \tag{2.9}$$

$$\bar{\mathbf{x}}_t^f = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_t^{i,f} \tag{2.10}$$

2.2.3. INVERSION METHOD

Steele-Dunne et al. [38] estimated thermal diffusivity from DTS temperature measurements at three depths. The temperatures measured in the top and bottom cables provide boundary conditions for a diffusion model:

$$\frac{\partial T}{\partial t} = \frac{\lambda(\theta)}{C(\theta)} \frac{\partial^2 T}{\partial z^2} = D(\theta) \frac{\partial^2 T}{\partial z^2} \tag{2.11}$$

where D is soil thermal diffusivity ($m^2 s^{-1}$), which is the ratio of soil thermal conductivity to heat capacity. Matlab's "fminsearch" function (a Nelder-Mead simplex direct search) was used to find the diffusivity value that minimizes the RMSE between simulated and observed temperature in the middle cable. The optimal diffusivity is found for a 24 hour window, which is moved in 3 hourly increments. The initial value for each window is the best estimate from the previous window. This moving window strategy ensures that enough data is available to estimate diffusivity, while yielding estimates at an acceptable temporal resolution. The soil moisture can be inferred from the estimated soil thermal diffusivity if the relationship between soil moisture and thermal diffusivity is known or can be assumed. Note that a single value of D is estimated for the soil column between the top and bottom cables, i.e. it is assumed that D is uniform between the top and bottom cables. Additional details on the implementation of this inversion method are provided in [38].

2.2.4. DATA ASSIMILATION EXPERIMENTS

The meteorological data and soil properties used in this chapter were collected at the Soil Moisture Active Passive (SMAP) Marena Oklahoma In-Situ Sensor Testbed (MOISST) site (refer to A.1). Uncertainty is included in the meteorological forcing data, soil texture and the initial conditions. For each ensemble member, these are perturbed using the assumptions summarized in Table 2.1. Precipitation and solar radiation were perturbed assuming multiplicative errors to ensure that the values remain positive. The assumed standard deviations reflect instrument error as well as spatial representativeness - weather data is collected at a point, while DTS observations are distributed in space. Figure 2.1 (a) and (b) show a histogram of the ensemble of precipitation values and solar radiation values at some model time step.

Table 2.1: Generation of perturbed inputs (soil property and forcing) for each ensemble member.

Variable	Error Distribution	Mean	Standard deviation	Bound
Sand (%)	Gaussian, Additive	0 (%)	17 (%)	15, 55
Silt (%)	Gaussian, Additive	0 (%)	5 (%)	-, -
$\rho_b (g/cm^3)$	Gaussian, Additive	0 (g/cm^3)	0.05 (g/cm^3)	-, -
Precipitation (mm)	Gaussian, Multiplicative	1	× 0.2	-, -
Radiation (W/m^2)	Gaussian, Multiplicative	1	× 0.075	-, 1350
Air temperature (^{o}C)	Gaussian, Multiplicative	1	× 0.05	-, -
Relative humidity (%)	Gaussian, Multiplicative	1	× 0.05	-, 100
Wind speed (Km/h)	Gaussian, Multiplicative	1	× 0.2	-, -

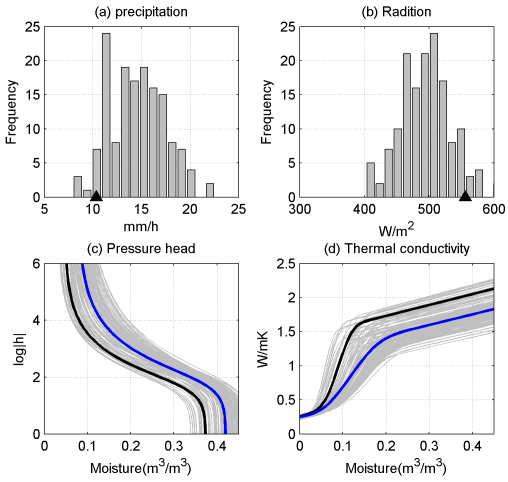

Figure 2.1: Distribution of the model forcing for the ensemble members (histogram) and the truth (triangle) at one model step, taking precipitation (a) and radiation (b) as an example. Soil water retention (c) and soil thermal conductivity (d) as a function of soil moisture. Each thin gray line represents the soil property of one ensemble member, and the ensemble mean is shown as the blue line. The soil properties used for truth generation are shown as the black lines.

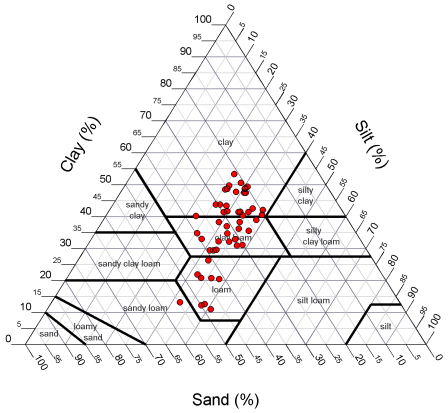

Figure 2.2: The "true" soil texture for each of the 50 synthetic truths as being presented in USDA textural triangle. The silt and sand content are sampled from the distribution decided in Table 2.1, with the clay content calculated as (100-Silt-Sand)%.

The ensemble mean values of the sand and silt contents and bulk density (ρ_b) are set to the values measured at site A. The sand and silt content are perturbed assuming additive Gaussian error with zero mean and a standard deviation based on the observed variability in the field ($\approx$ 17% for sand, and $\approx$ %5 for silt). The clay content for each ensemble member is calculated as the residual (100 – Sand – Silt)%. We perturbed bulk density using additive white Gaussian error, with standard deviation of 0.05 gcm^{-3}, which is larger than calculated using field measurements. Soil properties (texture and bulk density) are assumed to be uniform over the entire profile. For each ensemble member, soil hydraulic properties were calculated from this perturbed soil texture and bulk density using ROSETTA. The Campbell model was used to generate soil thermal conductivity curves for each ensemble member. Figure 2.1 (c) and (d) show the soil water retention and thermal conductivity curves for each ensemble member. The initial soil moisture and temperature profiles are assumed to be homogeneous with depth. The initial soil moisture and temperature values are sampled from uniform probability distributions with a range of 0.2 to 0.35 m^3/m^3 and 17 to 25 ^{o}C respectively.

The synthetic truth was generated as an additional ensemble member from the model assuming the error distributions described in Table 2.1. This means that the forcing data used to generate the “true” soil temperature and moisture is different from the individual ensemble members and the ensemble mean. In addition, the truth is generated using a randomly sampled soil texture and bulk density. As a result, the derived soil hydraulic and thermal properties also differ from the individual ensemble members and the ensemble mean (see Figure 2.1 c and d). This is to represent the errors/uncertainties in soil hydraulic and thermal properties caused by spatial variability and measurement error.

Similarly, the “true” soil temperature and moisture are initialized using the same distribution as the ensemble. Hence, the truth used here considered the error from model forcing data, initial conditions and the error caused by incorrect soil parameters (i.e. soil hydraulic and thermal properties). Synthetic DTS observations were generated by adding Gaussian error to the “true” soil temperature from a limited number of depths. DTS temperature resolution can approach 0.03K over integration times of 60 to 300s [70]. With good calibration, even less expensive DTS systems can achieve 0.1K accuracies for integration times of 60s [71]. Therefore, as a conservative estimate, an observation error of an observation error of 0.1K is assumed for the synthetic “DTS" temperature observations.

Using this set-up, the following experiments were performed:

1. **Illustrative example** For a single “truth”, temperature observations at three depths (5,10 and 20cm) are assimilated into Hydrus-1D every hour. The estimates after data assimilation (EnKF) are compared to the synthetic truth (Truth), and an ensemble open loop run of Hydrus-1D without any assimilation (EnOL). In this illustrative case, the soil water retention and thermal conductivity curves used for the truth are shown in black in Figure 2.1. The data assimilation approach will also be compared with the inversion method as employed by Steele-Dunne et al [38]. Note that in the inversion method, the relationship between soil moisture and thermal diffusivity is assumed to be known and is based on the soil texture from the truth. This requirement, that the relationship between soil moisture and thermal diffusivity is known, is one of the major limitations of the inversion approach [38].

2. **Robustness test** To investigate the robustness of our data assimilation algorithm across a wide range of soil properties, the first experiment is repeated for fifty different “truths”. Each truth has different perturbed inputs (e.g. soil properties, initial condition and forcing). Figure 2.2 shows the fifty soil texture values considered. This range of values was chosen to be consistent with the range of soil texture values observed at the SMAP MOISST site. This gives an indication of how uncertain soil texture would be along a Passive DTS installation at this site, with the soil type varying from clay to loam and one instance of sandy loam.

3. **Impacts of data assimilation interval**. For ten different “truths”, the soil moisture and temperature profiles were estimated by assimilating temperature data at 5, 10 and 20cm at intervals of 1, 3, 6, 12, 24 and 72 hours.

4. **Comparing observation strategies** Observing System Simulation Experiments (OSSEs) are used in remote sensing to evaluate and develop the optimal observation strategy and to analyze the influence of observation strategy and errors on retrievals. Here, an OSSE is conducted to determine how many cables are really needed, and at which depths they should be installed. Installing cables at fewer depths allows us to measure over a larger area with the same amount of cable.

5. **Impact of model structural errors - lower boundary condition** This set of experiment investigates the impact of an incorrectly defined lower boundary condition. The truth is generated in a similar manner to the previous experiments, except

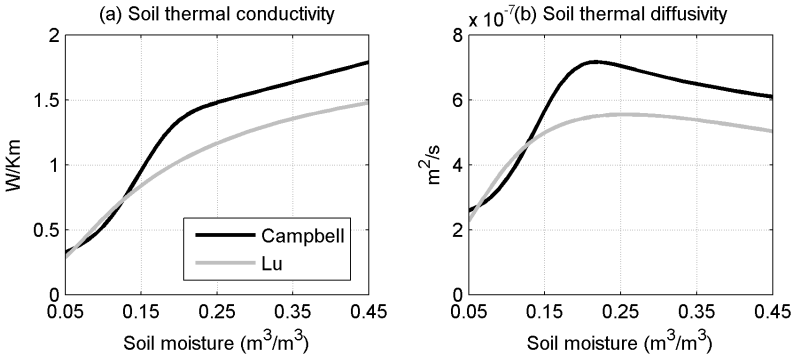

Figure 2.3: An illustrative example of comparing soil thermal conductivity curve (a) and diffusivity curve estimated using the Campbell and Lu model. Soil properties for the two models are identical, with a sand content of 32%, and bulk density of 1.45 g/cm^3

that the hydraulic lower boundary condition is set to be “zero flux". The EnKF and the EnOL used here are the same as previous experiments, i.e. the hydraulic lower boundary condition was assumed to be “free drainage".

6. **Impact of model structural errors - thermal conductivity model** In previous experiments, the Campbell model was used in generating the true relationship between soil moisture and thermal conductivity, as well as the relationship in the forward model of the EnOL and the EnKF. Here, we consider the case where the “true" relationship is generated using the model of Lu et al. [72] and the EnOL and the EnKF use the Campbell model. A comparison of the modeled soil thermal conductivity and diffusivity using the two models are shown in Figure 2.3.

7. **Impact of model structural errors - heterogeneity soil property profile** All of the previous experiments were based on the assumption that the true soil properties are homogeneous with depth. However, this assumption may be violated in reality. This set of experiments is to evaluate the impact of vertical heterogeneity in the true soil thermal and hydraulic property profile on the performance of the proposed data assimilation method. A vertically heterogeneous soil texture and bulk density profile is generated using the field measurements of soil texture at discrete depths from site C. As shown in Figure 2.4, the soil texture is quite variable at shallow depths, and the profile mean is significantly different from that used in the EnKF. Soil thermal and hydraulic properties for the truth are generated using the heterogeneous soil texture and bulk density profiles, and hence the true soil thermal and hydraulic properties are vertically heterogeneous. The EnOL and EnKF are the same as in the previous experiments, i.e. the soil texture and bulk density profiles, and therefore the soil thermal and hydraulic properties, are vertically homogeneous (Figure 2.4).

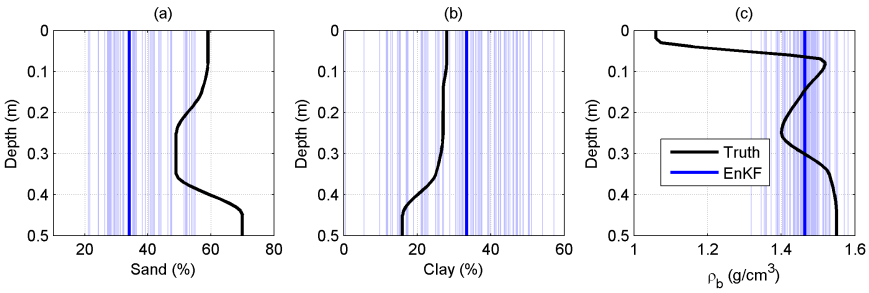

Figure 2.4: Sand (a), Clay (b) and bulk density (c) profiles used for generating soil thermal and hydraulic properties in the truth and the EnKF. Only the top 0.5m of the simulation soil column is presented. Each thin blue line represents one ensemble member.

2.3. RESULTS AND DISCUSSION

2.3.1. AN ILLUSTRATIVE EXAMPLE

Figure 2.5 (a) to (c) show the true temperature profile, and the difference between the true profile and that obtained from the EnOL and EnKF. It is clear that this particular realization of the "truth" is much warmer than the EnOL, with errors up to 5K at depth. Assimilation of temperature observations at 5, 10 and 20cm reduces this simulation error to close to zero throughout the soil column. Figures 2.5 (d) to (f) show time series of the error in the estimated temperature at 1, 5 and 10cm. They show that solar radiation, and uncertainty in solar radiation result in a significant daily cycle in the error in estimated temperature from the EnOL. The maximum errors coincide with the daily maximum values of solar radiation at noon. The EnKF reduces the error of the temperature estimates considerably, and the daily cycle in temperature error is no longer apparent. Soil temperature close to the surface is more likely affected by the uncertainties of the forcing (e.g. radiation, air temperature, and wind speed). Hence, the temperature estimates have larger errors than those at depth. The smallest errors in estimated temperature from the EnOL are during large precipitation events (e.g. 05/19, 06/15 and 07/04) when temperatures in the top 20 cm are primarily influenced by water fluxes in the soil column.

It is noticed that the innovations are biased in Figure 2.5 (d) to (f). This may be avoided when a bias correction algorithm is used [e.g. 73–76]. However, the bias of the innovations is attributed to the biased model parameters (Figure 2.3). Hence, a more physically based way of removing this bias is to remove the bias in the prior guessed model parameters. Providing unbiased model parameter along each meter of DTS cable is challenging, if possible. Hence, we are interested in investigating the accuracy of the soil moisture estimates, when the prior guessed model parameters are biased, in this chapter.

In Figure 2.6, the soil moisture truth is compared to the ensemble mean from EnOL and EnKF for the entire profile. By comparing Figures 2.6 (d) and (e), one can see that, assimilating temperature observations also leads to an improved soil moisture estimate. Comparing Figures 2.6 (a) and (b), it is clear that the EnOL overestimates the soil moisture throughout the profile. The extended periods of elevated soil moisture after pre-

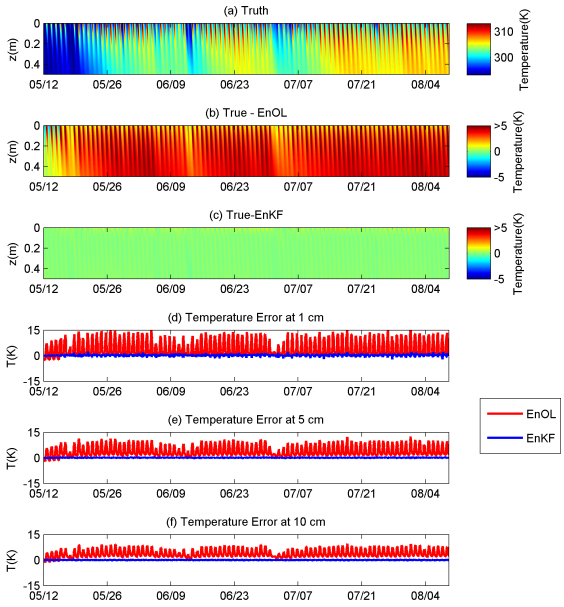

Figure 2.5: Comparison of soil temperature profile estimated by EnOL and EnKF. a: the synthetic true soil temperature profile, b: the difference between the True and the EnOL estimated soil temperature profile, and d to f: the error of the EnOL and EnKF estimates at 1, 5 and 10 cm.

cipitation events are particularly striking. The bias is particularly noticeable at depth in the second half of the experiment. From Figure 2.6 (d), the most significant improvement due to assimilation is in drying the surface layers after precipitation events. In the EnKF, the soil moisture at depth is updated through the update equation directly, but also through the model physics. If the surface moisture near the surface is updated to a moister condition, this extra moisture ultimately increases the soil moisture at depth. Similarly, the physics of the model ensure that the "extraction" of moisture from the surface layers in the EnKF contribute to a drier soil at depth between precipitation events. Figure 2.6 (e) shows that the largest discrepancy between the estimated (EnKF) and true soil moisture occurs during precipitation (e.g. 05/19, 06/15 and 07/04).

Figure 2.7 provides insight into how the soil temperature and moisture states are updated by the temperature observations. It shows the forecast error covariance between instantaneous temperature and temperature ($\mathbf{P}_{TT}$) and temperature and moisture ($\mathbf{P}_{T\theta}$) states across the three observation depths (5, 10 and 20cm). The forecast error covariance was determined at each time step, and averaged over the experiment duration to yield this daily cycle of values.

When the forecast error covariance is low, the corresponding element of the Kalman gain matrix will be close to zero and the states will barely be updated. When the forecast error covariance is high, the EnKF updates towards the observations. Any highly correlated states (including unobserved states) will also be updated.

From the left column, $\mathbf{P}_{TT}$ clearly has a significant daily pattern driven primarily by

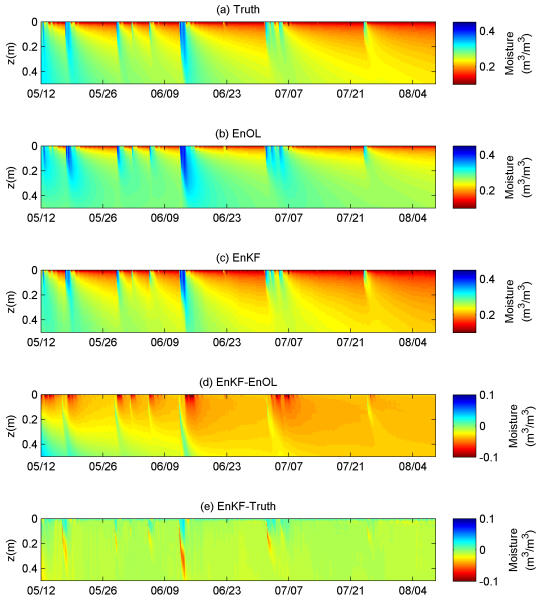

Figure 2.6: Comparison of soil moisture profile estimated by EnOL and EnKF.

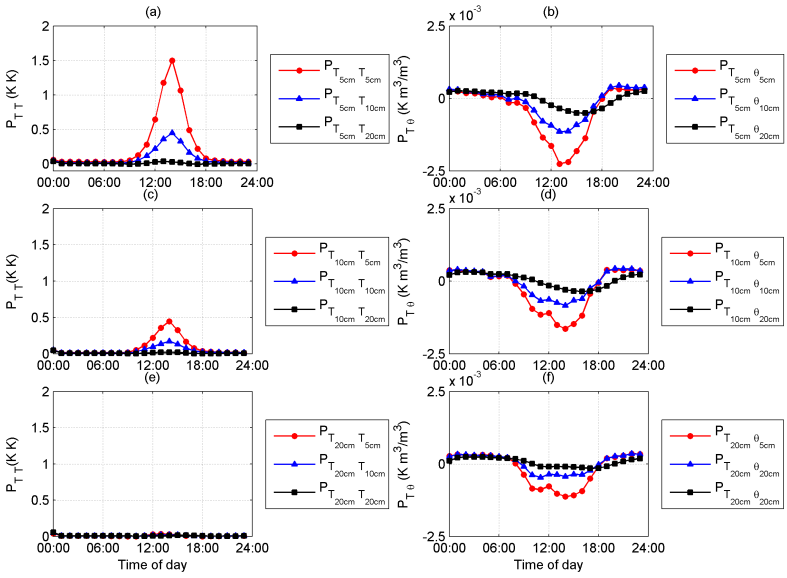

Figure 2.7: Averaged daily cycle of temperature and soil moisture error covariance at three depths. The left column presents the error covariance between modeled temperature, and the right column shows the error covariance between temperature and moisture. The error covariances presented are vectors in $\mathbf{P}_f$, e.g. $\mathbf{P}_{T_{5cm}\theta_{5cm}} = \frac{1}{N-1}\sum_{i=1}^{N}(T_{5cm} - \overline{T_{5cm}})(\theta_{5cm} - \overline{\theta_{5cm}})^T$.

the uncertainty in solar radiation. The largest forecast error covariance occurs around the noon, while smallest values occur when solar radiation is low (e.g. 0:00 to 6:00 and 18:00 to 24:00). The temperature forecast error covariance decreases with depth due to the exponential decrease in soil temperature amplitude with depth [77].

The right column in Figure 2.7 shows the error covariances between instantaneous soil moisture and temperature ($\mathbf{P}_{T\theta}$). This strong negative covariance between temperature and soil moisture explains why the soil moisture can be updated using observations of temperature alone. $\mathbf{P}_{T\theta}$ also has a significant daily cycle due to the temperature errors, with a maximum absolute value around noon. This suggests that assimilating temperature observations around noon will be most effective in updating soil temperature and moisture estimates. $\mathbf{P}_{T\theta}$ also decreases with depth, but to a lesser degree than $\mathbf{P}_{TT}$ does. Consequently, temperature observations at 20cm can still contribute to soil moisture estimates. It is worth noting that soil moisture at different depths is highly correlated, so any update to the shallow soil moisture will also influence values at depth.

Recall from Figure 2.6 (e) that the largest discrepancy between the estimated (EnKF) and true soil moisture occurs during precipitation. At these times, heat convection due to water flux during precipitation cools the soil column at all layers (Figure 2.5 (d) and (e)). The ensemble of temperature values collapses, reducing the covariance with soil moisture and prevents the soil moisture from being updated with temperature during and just after precipitation events.

The EnKF approach is compared with the inversion method in Figure 2.8. The inversion method can only provide the average of the soil thermal diffusivity between the boundaries (5 and 20cm in this case), as shown in Figure 2.8 (a). Therefore, the true values are also averaged for the purpose of comparison. The spikes in the estimated diffusivity are due to the fact that heat convection processes are neglected in the inversion method (e.g. May 9th and June 13th). In the dry period (e.g. July 7th to August 4th), the estimated diffusivity is biased, which might be explained by ignoring the heat transported by vapor fluxes. Diffusivity is a non-monotonic function of soil moisture. We can select the right estimates by selecting the soil moisture value that ensures the soil in drying down, except during precipitation events. Provided the soil diffusivity curve is perfectly known, the inversion method can provide reasonable soil moisture estimates. In practice, it is difficult to have this relationship for every meter of cable in the field. The EnOL deviates from the truth due to the incorrect parameterization and model forcing errors. When soil temperature data are assimilated, soil moisture estimates are significantly improved. The soil moisture estimated using the EnKF is closer to the truth than the estimates from the inversion method. It is worth noting that, in contrast to the inversion method, the data assimilation approach does not require a known relationship between soil thermal diffusivity and soil moisture. This will greatly facilitate the field implementation of this method.

2.3.2. ROBUSTNESS OF THE DATA ASSIMILATION ALGORITHM

Figure 2.9 compares the RMSE of soil moisture estimated by EnOL and EnKF at different depths for each of the fifty "truths". Generally, the EnKF provides better soil moisture estimates. The improvement is greatest close to the surface where the covariance between temperature and soil moisture is largest. The EnOL estimate is unconstrained by

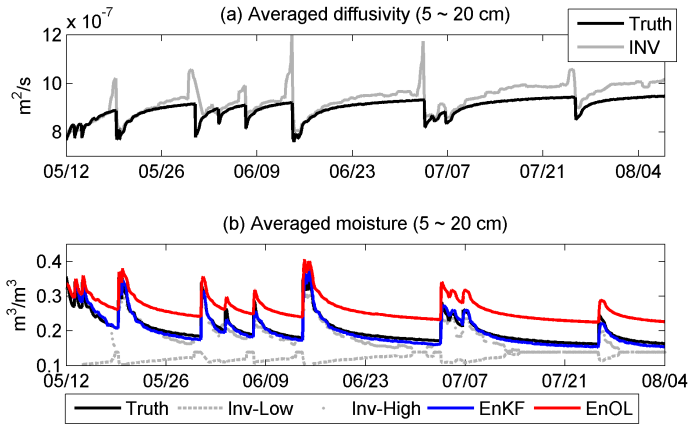

Figure 2.8: Comparison of soil thermal diffusivity estimated using inversion method with the truth (a), and the soil moisture estimated using inversion method (Inv-High and Inv-low), EnKF, EnOL with the truth (b).

observations, so the estimate is always the same. Sometimes, the "truth" coincides with this mean, yielding a very low RMSE for the EnOL. In this case, the EnKF yields little improvement. Conversely, if the synthetic truth is far from this EnOL mean, the EnKF yields a significant improvement by drawing the ensemble mean towards the observations.

Recall that for the fifty tests, each truth was generated with a different "true" soil texture, which can be significantly different from the ensemble mean (see also Figure 2.1). This means that the DA approach has a significant advantage over the inversion approach. Steele-Dunne et al. [38] demonstrated that the need to know the relationship between soil thermal conductivity and soil moisture for each meter of the cable was a serious limitation of the inversion approach. The DA approach used here can estimate soil moisture in the presence of uncertainty in soil texture and therefore overcome this limitation.

2.3.3. IMPACT OF ASSIMILATION INTERVAL

Figure 2.10 shows the impact of assimilation interval on the RMSE in estimated temperature and soil moisture. For assimilation intervals of 12, 24 or 72 hours, the states were updated at noon to maximize the covariance between temperature and soil moisture.

Soil temperature at the surface is influenced by high frequency variations in solar radiation. As thermal energy is propagated through the soil column, evidence of higher frequency fluctuations is damped and the temperature is dominated by longer (e.g. daily, seasonal and annual) cycles of temperature variation. Figure 2.10 (a) shows that at depth (e.g. at 50cm), an assimilation interval of 72 hours or 24 hours yields a substantial improvement over the EnOL. Reducing the assimilation interval further has a limited benefit at depth because fluctuations on this time scale are less relevant deep in the soil column. Closer to the surface (e.g. at 10cm), infrequent assimilation (72 hours) has limited benefit, but reducing the assimilation interval continues to yield a reduction in RMSE. Therefore, to minimize the RMSE in estimated temperature, temperature observations

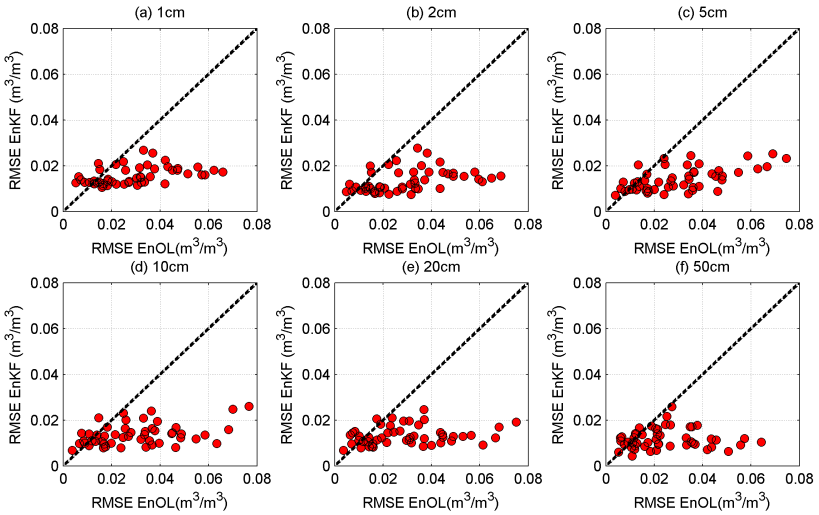

Figure 2.9: Comparison of RMSE in soil moisture estimated using EnKF and EnOL. Results are presented at six depths. Each cycle represents the RMSE from a single synthetic experiment.

should be assimilated as often as possible.

From Figure 2.10 (b), assimilating temperature observations every 72 hours reduces the RMSE in soil moisture by almost 50% compared to the EnOL. Results are similar throughout the profile. The profile in RMSE is almost uniform apart from a slight increase between the location of the first cable and the surface where the estimate is primarily updated through the state update and not through model physics (heat and moisture were primarily propagated downward during this simulation period). Generally, more frequent assimilation reduces the RMSE, though there is little difference between hourly and three-hourly assimilation. Given that hourly assimilation yields an almost uniform profile in RMSE in both temperature and soil moisture, one hour is considered a reasonable assimilation interval. However, reducing the assimilation interval further to five minutes leads to ensemble collapse and an increase in RMSE. Hourly assimilation ensures that the ensemble has a chance to grow between updates.

2.3.4. OBSERVATION STRATEGY

In Figure 2.11, different observation strategies (number of cables and depths) are compared to the EnOL (no observations), and EnKF using cables at all three depths. Figure 2.11 (a) compares the RMSE in temperature for these two benchmarks against the EnKF assimilating observations at a single depth which is varied from 5cm to 50cm. As expected, all single depth EnKF assimilation results lie between the two benchmarks and the largest RMSEs are closest to the surface. The minimum RMSE is at the depth of the observation because the model and observations are most correlated at this depth. The RMSE increases with increasing distance from the observations. From Figure 2.11 (c), it is clear that using observations at two depths greatly improves the temperature estimates

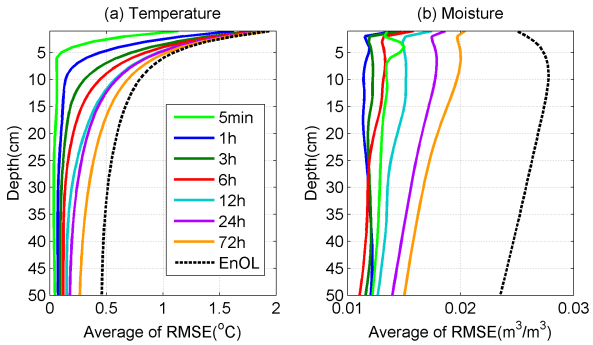

Figure 2.10: RMSE of estimated temperature and soil moisture as a function of data assimilation intervals. Each line represents the averaged RMSE from 10 synthetic experiments. Temperature observations at noon were assimilated in the cases with assimilation interval of 12, 24 and 72h.

of the entire profile. This is because observations at two or more depths constrain the temperature gradient, yielding improved estimates of the entire soil temperature profile. As the greatest errors occur close to the surface, Figure 2.11 (c) shows that locating the uppermost cable close to the surface yields the best results.

Results in Figure 2.11 (b) and (d) indicate that installing cables at two depths might also be sufficient for soil moisture estimation. If only one cable is to be installed, it appears from Figure 2.11 (b) that this should be installed at 10 or 20cm. While the temperature at 5cm layer is dominated by solar radiation, the temperature at 20cm is influenced by the thermal properties (and hence moisture content) of the overlying soil. Installing any deeper (e.g. at 50cm) means observing close to or deeper than the damping depth, where the temperature variations are probably too low to yield useful information.

From Figure 2.11 (d), there is no clear optimal pair of depths at which temperatures should be observed to estimate soil moisture. Though the RMSE values are lower than those obtained with temperature observations at one depth, there is little difference in RMSE between depth combinations. It appears that capturing the temperature gradient somewhere in the soil column is enough to constrain the soil moisture estimate. Observing temperatures at 10 and 20cm yields just a marginal improvement over the other options.

Using the proposed data assimilation approach to estimate soil moisture from DTS observations would allow us to measure at two, rather than three, depths. This makes installation in the field much easier, and can allow us to measure over a larger area with the same amount of cable. The data assimilation approach allows us to determine, for a given soil type, the depths at which the cables should be installed. Hence, by integrating DTS and data assimilation, we can design better DTS installations.

2.3.5. IMPACT OF MODEL STRUCTURE ERRORS

The impact of incorrectly defining the lower boundary condition is shown in Figure 2.12. Based on 50 truths, the RMSE of soil moisture estimates in the surface layers (above 10cm) are significantly reduced using the EnKF. At 50cm, the soil moisture estimates can be worse after data assimilation. This is because the impacts of the lower bound-

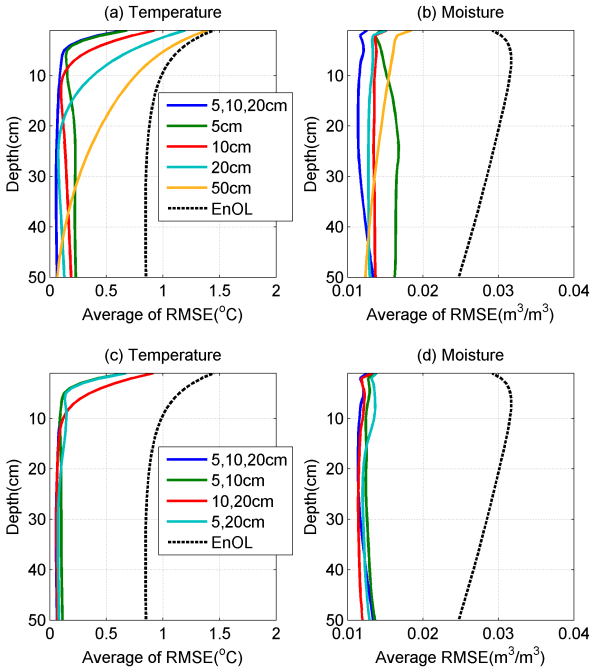

Figure 2.11: RMSE of estimated temperature (a and c) and soil moisture (b and d) using observations at different depths in the EnKF. In the top row, temperature observations at a single depth are assimilated. In the bottom row, two observation depths were assimilated (5 and 10cm, 5 and 20cm, and 10 and 20cm respectively). For each case, 10 synthetic truths were used for testing. Each line represents the averaged RMSE of the 10 tests. Results are compared to EnOL and EnKF results in which observations from all 3 depths were assimilated.

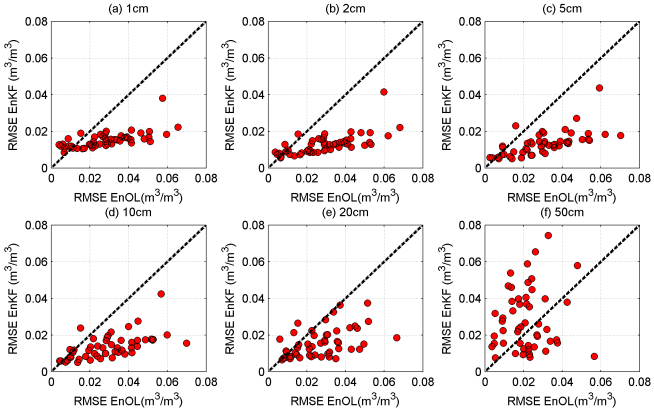

Figure 2.12: Similar with Figure 2.9, but the truth is generated using a lower boundary of "zero flux".

ary condition mainly affects the soil moisture at deep layers. Furthermore, only surface soil temperatures (5 to 20cm) were assimilated and their correlation with the deep soil moisture is limited.

Figure 2.13 shows the impact of assuming an incorrect model structure for the soil thermal conductivity as a function of soil moisture. Recall that in this experiment, the truth was generated using the Lu Model, while the EnKF and the EnOL use the Campbell model. At 2cm, the soil moisture estimates are greatly improved by EnKF. With increasing depth, the improvement decreases gradually (the red circles are getting closer to the 1:1 line). Solar radiation propagates downward from the surface to the lower boundary, so the soil heat transfer process is from the surface downwards. As a result, the error due to assuming an incorrect soil thermal conductivity model accumulates from the surface to the lower boundary. Hence, soil moisture estimates at deeper layers are more likely to be affected.

Table 2.2 shows the RMSE of soil moisture estimated using the EnOL and EnKF when the assumption of vertically homogeneous soil thermal and hydraulic property profile is violated. The soil moisture estimated using EnOL has a RMSE above 0.05 m^3/m^3 at all depths. Compared to the EnOL estimates, the EnKF provides improved soil moisture estimates across the entire profile. The largest RMSE of the EnKF estimates are at 5cm, where there is a sharp change of soil bulk density in the vertical profile (Figure 4). Identical to the EnOL, the EnKF assumes the soil thermal and hydraulic properties are uniformly distributed in the profile. As a result, the error covariance between the soil temperature and moisture predicted by the forward model will be less accurate, particularly at depths where there are sharp changes of soil thermal/hydraulic properties. Nevertheless, this set of experiments demonstrates soil temperature at shallow surfaces can be used to improve and constrain the soil moisture profile estimates, even when soil thermal/hydraulic properties are heterogeneously distributed in the profile. However, the performance is obviously better when the soil properties are relatively homogeneous with depth, i.e. when the model is closer to the truth. Hence, prior knowledge

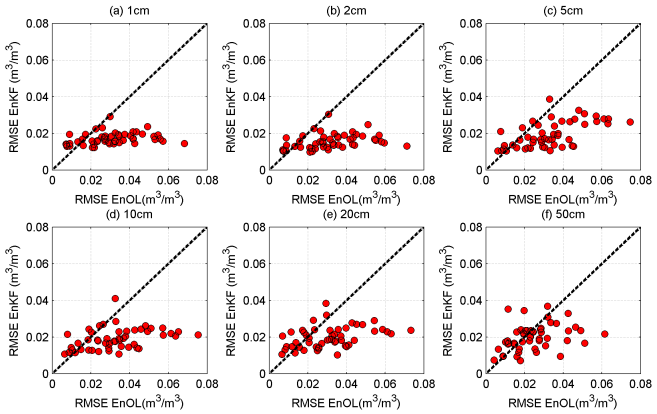

Figure 2.13: Similar with Figure 2.9, but the thermal conductivity of the truth is estimated using Lu model.

Table 2.2: The RMSE of soil moisture (m^3/m^3) estimated using the EnOL and the EnKF at 6 depths with truth generated using a heterogeneity soil thermal and hdyraulic property profile, as shown in Figure 2.4.

Depths	**1cm**	**2cm**	**5cm**	**10cm**	**20cm**	**50cm**
EnOL	0.0566	0.0589	0.0632	0.0813	0.0529	0.0675
EnKF	0.0398	0.0371	0.0430	0.0317	0.0175	0.0315

or methodologies that can constrain soil thermal and hydraulic property profiles in the forward model may further improve the proposed method.

2.4. CONCLUSIONS

This synthetic study demonstrates that the profile of soil moisture and temperature can be estimated more accurately by assimilating soil temperature observations into a coupled heat and moisture transport model. The soil temperature is observed directly at a limited number of depths, but the strong correlation in temperature across the profile ensures that temperature at all depths is updated by each observation. The strong negative error covariance between instantaneous soil moisture and temperature is enough to ensure that information from the temperature observations also updates the soil moisture throughout the profile. In addition to the increments from the update equation itself, the physics of the model also play an important role in updating the states at depth e.g. if the surface soil moisture is reduced, that ultimately leads to drier soil at depth.

The data assimilation approach presented here offers a promising new way to estimate soil moisture from Passive DTS. It offers several advantages over an amplitude analysis or the previously used inversion method: 1) It provides a detailed vertical distribution of soil moisture and temperature in the soil column. 2) It avoids the need to infer soil moisture from thermal diffusivity, which can be problematic due to the non-monotonic nature of the relationship between soil moisture and thermal diffusivity. 3) The estimate is constrained by the model as well as the observations. If the net radiation

is low, the inversion method fails because there is no daily temperature signal to propagate into the soil column. In this data assimilation approach, the model keeps the estimate on track in this case. 4) Data assimilation approaches, particularly the EnKF used here, are appealing because they provide a means to handle uncertainty. This is particularly relevant due to the importance of uncertainty in soil properties in the application of DTS to estimate soil moisture in the field. In this chapter, it has been shown that providing the model with a distribution of soil texture that reflects the expected range of values, soil moisture can be estimated regardless of the "true" soil texture. 5) While the inversion method required observations at three depths, assimilating temperature at two depths is enough to constrain the temperature gradient and yield equally good results as three depths. For DTS, this means that cables need to be installed at two, rather than three, depths. This allows us to measure over a larger area with the same amount of cable.

The model structural error impacts were also investigated by generating the truth using different assumptions to those made in the EnOL and EnKF simulations. Results suggest that soil moisture estimates down to 20cm can be robustly estimated even when the lower boundary condition or the soil thermal properties are poorly parameterized. The performance of the proposed method might be affected when true soil thermal and hydraulic properties have sharp changes in the vertical profile. The results presented here are limited to synthetic experiments. It is important to note that synthetic experiments are based on a "truth" that is infinitely simpler than reality. Applications of data assimilation in real world situations are therefore always more complicated and less accurate than in synthetic cases. In this chapter, some of the main sources of uncertainty are considered (i.e. soil thermal and hydraulic properties caused by the uncertainties in the soil texture, soil moisture-thermal conductivity relationship, vertical heterogeneity), albeit in a simplistic way. In practice, there are many other potential sources of uncertainty including the structural error in modeling water retention curve (i.e. van Genuchten's model) and preferential flow. Consequently, the performance of this approach in a real world case is expected to be poorer than the synthetic study case here. The obvious next step is to evaluate this approach in the field. Application of the methodology for a real-world case with additional types of errors is still a further, challenging and also interesting step. The primary goal will be to see how well this approach works when the physics of the true soil column are more complicated than that simulated by Hydrus-1D.

The model forecasts are shown to be biased in some cases, which is primarily due to the differences between the ensemble mean model parameters and the true parameters. Assuming the model parameters are unbiased everywhere along the DTS cable may not be valid for the DTS applications. A bias correction algorithm [e.g. 73–76] might be useful in removing the forecast minus observation biases, and further improve the estimates. When a bias correction scheme is included, the data assimilation performance in terms of unbiased metrics will change and the findings in the results would be different. However, since the bias of the forecasts are primarily attributed to the model parameters, an alternative approach of removing the forecast biases is to jointly estimating the model parameters with the states [69]. Provided the model parameters can converge to the "truth", the forecast biases will be removed automatically. This may also allow us to extract soil property information along the DTS cable.

Future research will also explore the potential of using smoothing rather than fil-

tering for assimilation. This is often beneficial in soil moisture estimation due to the strong correlation in time [78]. In this application, it is particularly relevant because soil moisture influences the amplitude of the diurnal temperature wave as it is propagated through the soil column. So, there may be additional information in a sequence of temperature observations compared to an instantaneous temperature measurement [79].

3

A PARTICLE BATCH SMOOTHER FOR SOIL MOISTURE ESTIMATION USING SOIL TEMPERATURES

Based on: Dong Jianzhi, Susan C. Steele-Dunne, Jasmeet Judge, Nick van de Giesen. A particle batch smoother for soil moisture estimation using soil temperature observations, 2015, *Advances in Water Resources*, 83, 111-222.

3.1. INTRODUCTION

The previous chapter demonstrated that soil moisture can be estimated by assimilating soil temperatures sequentially using the Ensemble Kalman Filter (EnKF). Soil moisture determines the portion of energy that used for evaporative cooling. Consequently, a sequence of soil temperature observations contains more information about soil moisture. Hence, assimilating a sequence of soil temperatures might lead to superior soil moisture estimates than sequentially assimilating soil temperatures at instantaneous time steps. This chapter tests this hypothesis.

The Ensemble Smoother (ES) can assimilate a sequence of observations. However, the ES may lead to incorrect estimates when the system is highly nonlinear or the Gaussian assumptions are not satisfied. Apparently, projecting the joint distribution of a sequence of soil temperature and soil moisture estimates into a multivariate Gaussian distribution can be inappropriate and risky. Particle approaches, e.g. Particle Filter (PF), approximate the entire model posterior distribution using Monte Carlo sampling. Thus, instead of preserving just the first two moments of the distribution, they also track the higher moments. It may be possible to extend this PF algorithm to a batch smoother algorithm, in which soil temperatures within a batch window can be assimilated, and robustly estimate the soil moisture within that window.

In this chapter, a new data assimilation method, the Particle Batch Smoother (PBS), is proposed in section 3.2. Section 3.4 provides detailed discussions and comparisons of the PF and the PBS algorithm, which shows that show the PBS outperforms the PF algorithm in soil moisture estimation. This also demonstrates that the evolution of soil temperature, rather than instantaneous points, contains more soil moisture information.

3.2. METHOD AND MATERIALS

3.2.1. SEQUENTIAL DATA ASSIMILATION

Same as Chapter 2, the model states of interest are soil moisture and temperature from the surface (0cm) to 1 meter, which are propagated forward using the Hydrus-1D model (Eq.2.6). The model estimates are related to the observations by:

$$\hat{\mathbf{y}}_t^i = h(\mathbf{x}_t^i) + \mathbf{v}_t^i \tag{3.1}$$

where $\hat{\mathbf{y}}_t^i$ is the simulated observation vector, h is a nonlinear operator relating the prior estimated states ($\mathbf{x}_t^i$) to the measured variable and $\mathbf{v}_t^i$ is the observation error [58]. The observation error is set to be 0.1K, according to the field calibration of the temperature sensors. In this study, the observations are the temperatures at 4 and 8 cm.

3.2.2. PARTICLE FILTER AND BATCH SMOOTHER

In the PF, the posterior distribution of the model state can be written in recursive form as follows [80]:

$$p(\mathbf{x}_{1:t}^i|\mathbf{y}_{1:t}) = p(\mathbf{x}_{1:t-1}^i|\mathbf{y}_{1:t-1}) \frac{p(\mathbf{x}_t^i|\mathbf{x}_{1:t-1}^i) p(\mathbf{y}_t|\mathbf{x}_t^i)}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})} \tag{3.2}$$

Integrating out $\mathbf{x}^i_{1:t-1}$, gives the marginal distribution [80, 81]:

$$p(\mathbf{x}^i_t|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t|\mathbf{x}^i_t)\,p(\mathbf{x}^i_t|\mathbf{y}_{1:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})} \tag{3.3}$$

where $\mathbf{y}_{1:t}$ is the observation vector. This equation is called the updating step in which the likelihood function, $p(\mathbf{y}_t|\mathbf{x}^i_t)$, is used to update the prior estimates. The prior estimates at time step t are described as [81]:

$$p(\mathbf{x}^i_t|\mathbf{y}_{1:t-1}) = \int p(\mathbf{x}^i_t|\mathbf{x}^i_{t-1})\,p(\mathbf{x}^i_{t-1}|\mathbf{y}_{1:t-1})\,d\mathbf{x}^i_{t-1} \tag{3.4}$$

In the PF, the model state posterior density is approximated as:

$$p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) = \sum_{i=1}^{N} w^i_t\,\delta(\mathbf{x}_{1:t} - \mathbf{x}^i_{1:t}) \tag{3.5}$$

where $\delta()$ is the Direc delta function [82]. Because the posterior density is difficult to sample directly, importance sampling is often used to draw particles from a known function $q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ and assign the weights according to:

$$w^{i*}_t = \frac{p(\mathbf{x}^i_{1:t}|\mathbf{y}_{1:t})}{q(\mathbf{x}^i_{1:t}|\mathbf{y}_{1:t})} \tag{3.6}$$

where w^{i*}_t is the importance weight. Equation (3.6) can be expressed as [53]:

$$w^{i*}_t \propto w^{i*}_{t-1}\frac{p(\mathbf{y}_t|\mathbf{x}^i_t)\,p(\mathbf{x}^i_t|\mathbf{x}^i_{t-1})}{q(\mathbf{x}^i_t|\mathbf{x}^i_{t-1},\mathbf{y}_t)} \tag{3.7}$$

where $p(\mathbf{y_t}|\mathbf{x}^i_t)$ is the likelihood, $p(\mathbf{x}^i_t|\mathbf{x}^i_{t-1})$ is the transition prior, i.e. the probability of moving to $\mathbf{x}^i_t$ from $\mathbf{x}^i_{t-1}$, and $q(\mathbf{x}^i_{k+1}|\mathbf{x}^i_k,\mathbf{y}_t)$ is the proposal distribution in the importance sampling. Usually, the transition prior is used for the proposal distribution [53]. Hence, Equation (3.7) can be simplified as:

$$w^{i*}_t \propto w^{i*}_{t-1}\,p(\mathbf{y}_t|\mathbf{x}^i_t) \tag{3.8}$$

The likelihood function is expressed as:

$$p(\mathbf{y_t}|\mathbf{x}^i_t) = \frac{1}{(2\pi)^{n/2}det(\mathbf{R})^{1/2}}e^{[-0.5(\mathbf{y}_t-\hat{\mathbf{y}}^i_t)^T\mathbf{R}^{-1}(\mathbf{y}_t-\hat{\mathbf{y}}^i_t)]} \tag{3.9}$$

where $\mathbf{R}$ is the error covariance of observations, and n is the number of the observations, e.g. $n = 2$ in this case (temperatures at 4 and 8 cm were assimilated). The normalized weight (w^i_t) is calculated as:

$$w^i_t = \frac{w^{i*}_t}{\sum_{i=1}^{N} w^{i*}_t} \tag{3.10}$$

Figure 3.1 (top panel) illustrates the PF algorithm at an update step. Initially, the uniformly distributed weights are assigned to the particles. When observations are available, the PF will update the weights of each particle according to the prior distribution

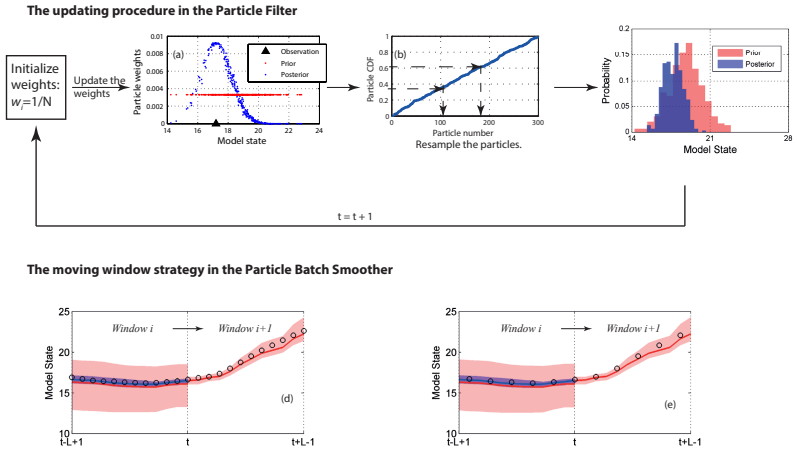

Figure 3.1: A diagram for illustrating the PF at one updating step (upper panel) and the Particle Batch Smoother (PBS) moving window strategy (lower panel). In the PBS moving window strategy plot, the solid lines are the mean of the prior and posterior, and the distributions of the prior and posterior are shown as the shaded red and blue area. Black circles represent the observations, and the strategy of using different observation interval in the PBS is shown in (d) and (e) respectively.

and likelihood (i.e. Equation (3.9)). As shown in Figure 3.1(a), the PF adds weights to the particles closer to the observation. Resampling is usually required to prevent weight degeneration, i.e. a situation where most of the particles have negligible weights. Moradkhni et al. [53] demonstrated that resampling the posterior after each update will avoid the degeneracy problem, and result in a significantly improved performance. A detailed description of resampling is given by Moradkhni et al. [53]. Therefore, we also perform resampling when the particle weights are updated (Figure 3.1b), which results in updated states (Figure 3.1 c).

Figure 3.1 (lower panel) illustrates the implementation of the PBS. While the PF assimilates observations sequentially, the PBS assimilates all of the observations within a window in a single batch. Therefore, the posterior density is calculated for a series of model states in a window (L) i.e. $p(\mathbf{x}_{t-L+1:t}|\mathbf{y}_{1:t})$.

Similar to the PF algorithm, this marginal distribution is used to derive the weights updating equation:

$$w^{i*}_{t,s} \propto w^{i*}_{t-L,s} p(\mathbf{y}_{t-L+1:t}|\mathbf{x}^{i}_{t-L+1:t}) \tag{3.11}$$

$$w^{i}_{t,s} = \frac{w^{i*}_{t,s}}{\sum_{i=1}^{N} w^{i*}_{t,s}} \tag{3.12}$$

where the subscript s denotes smoother. The calculated weights will be assigned to the entire window, i.e. $w^{i}_{t-L+1:t,s} = w^{i}_{t,s}$. The likelihood function is calculated based on the

Table 3.1: Generation of perturbed inputs (soil property and forcing) for each particle member.

Variable	Error Distribution	Mean	Standard deviation	Bound
Sand (%)	Uniform	-	-	55, 95
Silt (%)	Uniform	-	-	0, 20
$\rho_b(g/cm^3)$	Gaussian, Additive	0	0.05	-, -
Precipitation (mm)	Gaussian, Multiplicative	1	$\times$ 0.2	-, -
Radiation (W/m^2)	Gaussian, Multiplicative	1	$\times$ 0.075	-, 1350
Air temperature (^{o}C)	Gaussian, Multiplicative	1	$\times$ 0.05	-, -
Relative humidity (%)	Gaussian, Multiplicative	1	$\times$ 0.05	-, 100
Wind speed (Km/h)	Gaussian, Multiplicative	1	$\times$ 0.2	-, -

observations of the entire window:

$$p(\mathbf{y}_{t-L+1:t}|\mathbf{x}^i_{t-L+1:t}) = \prod_{j=t-L+1}^{t} \frac{1}{(2\pi)^{n_o/2} det(\mathbf{R})^{1/2}} e^{\left[-0.5(\mathbf{y}_j-\hat{\mathbf{y}}^i_j)^T\mathbf{R}^{-1}(\mathbf{y}_j-\hat{\mathbf{y}}^i_j)\right]} \tag{3.13}$$

where n_o is the number of observation depths. Clearly, the PF algorithm can be considered as a special case of the PBS algorithm where $L = 1$. The update procedure is the same as illustrated in Figure 3.1, except that the state vector includes all states within that window. After resampling, the algorithm will move to the next window with all the weights initialized to $1/N$. Different observations intervals can be used in the PBS algorithm, as shown in Figure 3.1 (d) and (e). When a longer observation interval is used, the PBS will discard the observations collected between the observation intervals.

3.2.3. DATA ASSIMILATION SET-UP

Data from the Microwex-2 experiment (A.2) are used to test and compare the performance of the PF and PBS. The distribution assumed for each perturbed model input is shown in Table 3.1. We are interested in investigating whether the proposed algorithms can handle the biases/ uncertainties in the soil properties, when only rough soil texture classification information is available. This is often the case in distributed temperature sensing applications where soil texture can vary considerably over the scale of the installation [38]. Sand and silt content were sampled uniformly within the USDA classes of sand and sandy loam, i.e. sand from 55% to 95% and silt from 0% to 20%. This is to represent the spatial variability of the soil texture. Similar to the previous chapter, meteorological forcing data were also perturbed to represent the instrument error and the spatial variability.

The objectives of this chapter are to demonstrate that particle-based data assimilation can determine soil moisture by assimilating soil temperature observations in real world data, and that particle batch smoothing yields better results than filtering.

First, we will demonstrate that the PF can be used to assimilate soil temperature observations at two depths into the Hydrus-1D model to estimate soil moisture estimates. In this first test, an observation interval of 45 minutes will be assumed, and the number of the particles will be set to be 100. Results from the PF will be compared to a model open loop (OL) run, which is identical to running the 100 particles without performing

any assimilations. Results from the PF and OL are also compared to in-situ observations of soil moisture throughout the profile.

Second, we will determine the optimal observation interval in the PF algorithm. With DTS, observations can be made every minute. However, assimilating these observations too frequently can limit the range of particle values. If this becomes too small, the observation may fall outside the particle range, resulting in a poor update. To examine the impact of increasing observation intervals on the range of particle values, we will vary this interval from 15 minutes to 3 hours. Due to the limited duration of the experiment, the initial condition has a significant influence on the estimated soil moisture, particularly at depth. To account for this, experiments are repeated 20 times with different initial conditions, and the median and range of RMSE will be used to determined the optimal observation interval for the PF.

Next, we will focus on the PBS algorithm. We will examine the impact of observation interval on the PBS algorithm. As in the PF, the observation interval may affect the performance. The observation interval will be varied from 15 minutes to 1.5 hours. In this case, a window length of 3 hours will be assumed. Then we will investigate the impact of assumed window length on the PBS estimate. As shown in Section 2.2, if the window length (L) is one (i.e. the window length is equal to the observation interval), the PBS and the PF are equivalent. Our hypothesis is that the evolution of temperature in time contains more information on soil moisture than the relationship between instantaneous soil moisture and temperature. Hence, increasing the window length should lead to an improved estimate. However, increasing the length of the window increases the dimension of the distributions to be estimated. This may increase the number of particles required, and hence the computational expense. Maintaining the observation interval of 15 minutes, the window length will be varied from 1 to 12 hours. Then we will compare the optimal PF and the optimal PBS.

Finally, we consider the number of particles required by the two optimal approaches. Though we expect the PBS to yield an improved estimate, this approach may require more particles and therefore greater computational expense. A trade-off is necessary between the reduction in RMSE and the increase in computational demand between the two approaches. The PF and the PBS will be compared as the number of particles is increased from 10 to 300.

3.3. DETERMINISTIC AND PROBABILISTIC PERFORMANCE ASSESSMENT

The RMSE is calculated as following:

$$RMSE = \sqrt{\frac{1}{N_t}\sum_{t=1}^{N_t}\left(\overline{\theta}_t - \theta_{t,obs}\right)^2} \tag{3.14}$$

where N_t is the number of total time steps that has observations, $\overline{\theta_t}$ is the mean of the particle estimates at time step t, and $\theta_{t,obs}$ is the observed soil moisture at t. The abso-

lute bias of the estimates will also be calculated as:

$$Bias = \left| \frac{1}{N_t} \sum_{t=1}^{N_t} \left(\overline{\theta}_t - \theta_{t,obs} \right) \right| \tag{3.15}$$

Probabilistic verification tools Quantile - Quantile (Q-Q) plot and reliability (α_r) will be used for assessing the performance of the proposed algorithms. Probabilistic metrics are mainly concerning whether the uncertainty of the estimates (particle range or ensemble spread) is appropriate, instead of the accuracy of the estimates. The quantile of the predictive distribution is calculated at each time step [83]:

$$z_t = \frac{1}{N} \sum_{i=1}^{N} k_i \tag{3.16}$$

where z_t is the quantile of the predictive distribution calculated at time t, $k_i = 1$ when $\theta_{t,obs} > \theta_{t,i}$, and $k_i = 0$, otherwise [84]. In the perfect case, the cumulative distribution of z_t should be the same as the cumulative uniform distribution ($U[0,1]$). If z_t clustered at the middle range, it indicates the uncertainty is overestimated. The uncertainty is underestimated when the z_t clustered around the tails. In the case that z_t is constantly lower/higher than $U[0,1]$, it indicates the estimates are biased [85]. Based on the Q-Q plot, the reliability (α_r) of the estimates can be computed as:

$$\alpha_r = 1 - \frac{2}{N_t} \sum_{t=1}^{Nt} \left| z_t - U[0,1] \right| \tag{3.17}$$

The reliability (α_r) varies from 0 (zero reliability) to 1 (perfect reliability).

3.4. RESULTS AND DISCUSSION

3.4.1. THE PARTICLE FILTER

Figure 3.2 shows an example of the soil moisture profile estimated using the PF when temperature observations at 4 cm and 8 cm are assimilated every 45 minutes. Note that the case with a significantly biased initial soil moisture condition is presented here. In the DTS implementation, the soil moisture is supposed to be measured every meter up to kilometers by assimilating soil temperatures at shallow soil depths. Hence, it is impossible to provide correct initial soil moisture profiles everywhere along the DTS installation. As a result, we are interested in the ability of the proposed data assimilation approach to correct for errors in the initial condition. The first thing to note is that the estimated soil moisture from the PF is generally closer to the observed soil moisture than that from the open loop (OL). The OL generally overestimates soil moisture at all depths. The range of particle values in the OL (indicated in pink) is due to the sources of uncertainty described in Table 3.1. Because the OL estimate is never constrained by observations, the impact of uncertainty in the initial condition persists through the study interval. The range of soil moisture values from the OL generally even exceeds the dynamic range of the observed soil moisture.

The greatest improvements due to the PF are observed at 4 cm and 8cm, which are the depths at which soil temperatures were assimilated. At 4 cm, the RMSE is reduced

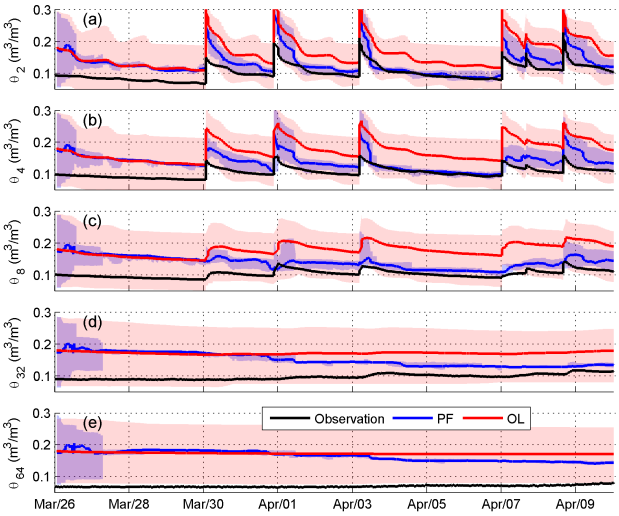

Figure 3.2: Comparison of soil moisture estimated by the open loop (OL) and the Particle Filter (PF) with observations at 5 depths (a to e represent depths 2, 4, 8, 32 and 64 cm, respectively). The PF algorithm, in this case, updates soil moisture using temperature observations every 45min. The Solid lines for the OL and the PF are the mean of the particles. The shaded areas are the range of the particles for the OL (red) and the PF (blue).

from 0.069 to 0.039m^{3}m^{-3} and at 8 cm it is reduced from 0.074 to 0.045m^{3}m^{-3}. The soil moisture estimates are improved when the PF resamples the particles with larger likelihoods. The soil moisture estimates at 2 cm are also greatly improved (from 0.060 to 0.037m^{3}m^{-3}), primarily by resampling the particles that provide larger likelihoods at 4 and 8 cm. This resampling also has a significant impact on the range of particles. It generally takes about a day for the PF to shed the influence of the uncertain initial condition. Due to uncertainty in precipitation, the range in soil moisture between 2 and 8 cm from the PF increases after precipitation. Assimilation with the PF has limited impact at 32 cm and 64 cm (Figure 3.2 (d) and (e)). Because of the sandy soil and high evaporative demand, there is little variability in soil moisture in response to precipitation at these depths. The estimated soil moisture is largely determined by the prescribed initial condition. The lack of correlation between soil moisture and temperature at this depth means that assimilation is ineffective. The only manner in which soil moisture at these depths is corrected is through the eventual impact of the updated surface (2 to 8 cm) soil moisture on the lower layers through the model physics. The lower soil moisture values in the PF compared to the OL between 2 cm and 8 cm eventually lead to drier soil at 32 cm (after April 1) and 64 cm (after April 3). Figure 3.3 gives some additional insight into how and when the PF is most effective in updating soil moisture. Figure 3.3 (a) and (b) show the influence of uncertainty in shortwave radiation and soil texture on the soil temperatures at 4 and 8 cm. The spread of particle values is largest during sunlight hours and at a maximum in the afternoon. Figure 3.3 (c) shows the spread of particle values of soil moisture after a precipitation event. Note that the PF is ineffective until just before noon when the spread of temperature values is large enough for the PF to

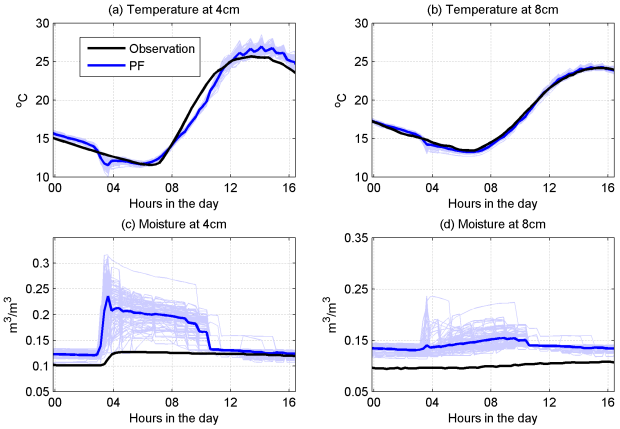

Figure 3.3: Soil temperature and moisture estimates at 4 and 8 cm on 2nd April. The thin blue lines represent the ensemble members.

update temperature and hence soil moisture. This is also apparent at 8 cm (Figure 3.3 (d)). This explains why the soil moisture estimated in the PF is not updated immediately after precipitation in Figure 3.2.

Figure 3.4 shows the impact of observation interval on the performance of the PF. From Figure 3.2, it is clear that the initial condition can influence the estimate, particularly at depth. Therefore, results for this experiment are presented in terms of box plots constructed using 20 cases, each of which has a different initial condition. Hence, Figure 3.4 also provides insight into the robustness of the PF to different initial conditions. From 2 to 8 cm (Figure 3.4 (a) to (c)), the median RMSE from the PF is always lower than that of the OL. At greater depth (Figure 3.4 d and e), the median RMSE is close to that of the OL regardless of the observation interval. The range of RMSE values is also larger than that closer to the surface. Recall that the particle range of soil moisture at depths reduces dramatically after a few updates. The limited influence of precipitation and temperature at depth mean that the particle range at depths will not grow again. Therefore, if the initial distribution of particles leads to the PF updating soil moisture towards an incorrect value, it is difficult to correct the estimates towards the observations. At 32 cm and 64 cm, Figure 3.4 shows that the estimated soil moisture can even be worse than that from the OL.

With less frequent assimilation, the median PF RMSE approaches that of the OL. Assimilating more often than every 45 minutes yields little improvement in terms of median RMSE, but the interquartile range (IQR) and the full range of values is often higher at lower observation intervals. The PF can only adjust the weights of particles, so the range of particle values must be wide enough to include the true value if the estimate is to be correct. Due to the observation error is very small, assimilating too frequently prevents the range of values from growing and can therefore yield a poorer estimate. Though no distinct RMSE minimum is shown in Figure 3.4 between assimilation interval of 30 to 120 minutes, the IQR seems to be the minimum when assimilation interval is

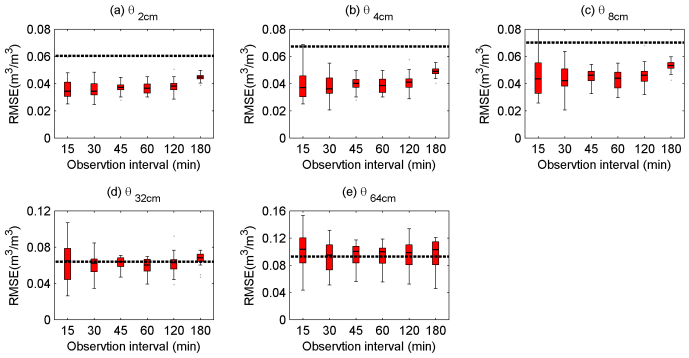

Figure 3.4: Comparison of RMSE of soil moisture estimated using the PF with different data observation intervals. The PF was tested using different initialized model inputs, and 20 runs were used for each observation interval. The black dashed line represents the RMSE of the open loop. In the box plot, the middle black line denotes the median value, the edges of the box are the interquartile range (IQR), the maximum length of the whiskers is set to be the 1.5 times of the IQR, and values larger/smaller than the maximum/minimum the whiskers are considered as outliers (black dots). The legends are the same for the following boxplots.

45 minutes. For the purpose of comparison, 45 minutes is considered to be the optimal observation interval for the PF, which will be compared with the PBS estimates.

3.4.2. THE PARTICLE BATCH SMOOTHER

First, we will compare the PF and the PBS algorithms assuming they both use an observation interval of 45 minutes (PF-45min and PBS-45min in Table 3.2). The window length for the PBS is 3 hours, i.e. the PBS will update temperature and soil moisture every 45 minutes in a 3 hour window using all temperature observations (at 4 cm and 8 cm) within that window. Table 3.2 shows the RMSE of soil moisture estimated using the PF and the PBS. The PBS results in a statistically significant reduction in RMSE compared to the PF (using a two-tailed paired T-test, $p < 0.05$). This is analogous to the improvement observed by using an ensemble batch smoother compared to an ensemble Kalman filter (e.g. [59]). Furthermore, by assimilating a series of temperature observations, the PBS exploits the influence of soil moisture in propagating a thermal wave from the surface into the soil [79]. This is a stronger relationship than that between instantaneous temperature and soil moisture values.

The advantage of using the PBS algorithm is also shown in the reduction of the bias, and the increased reliability. At depth above 32cm, the reliability of the PBS is nearly twice as high as that in the PF.

Despite resampling, the PF reduces the range of particle values every time there is an update as it proceeds sequentially through the 3 hour period. The PBS, on the other hand, allows the range of particle values to grow over the 3 hour window. Across all model time steps in the whole study period, the range of the particle values is on average 17% larger than that of the PF algorithm. Consequently, the observations are more likely to fall within the range of values considered by the PBS.

Table 3.2: Comparison of soil moisture RMSE (m^3m^{-3}), Bias (m^3m^{-3}) and Reliability estimated using the PF and PBS at five layers. The value presented are the averaged RMSE from 20 tests. The observation interval being considered are 15 minutes (PBS-15min) and 45 minutes (PF-45min and PBS-45min). The window length for the PBS is 3 hours.

Metric	Algorithm	Depths				
		2cm	4 cm	8 cm	32cm	64 cm
RMSE	PF-45min	0.037	0.039	0.045	0.062	0.094
	PBS-45min	0.035	0.037	0.042	0.059	0.094
	PBS-15min	0.029	0.029	0.030	0.039	0.074
Bias	PF-45min	0.029	0.035	0.042	0.059	0.098
	PBS-45min	0.027	0.032	0.036	0.052	0.090
	PBS-15min	0.019	0.020	0.023	0.033	0.081
Reliability	PF-45min	0.087	0.031	0.021	0.002	0.000
	PBS-45min	0.143	0.074	0.053	0.004	0.000
	PBS-15min	0.346	0.314	0.155	0.004	0.000

IMPACT OF OBSERVATION INTERVAL

Figure 3.5 shows the influence of the observation interval on the PBS algorithm. The total number of the observations within the assumed 3 hour window is varied from 2 (1.5 hour interval) to 12 (15 minute assimilation interval). If the observation interval is equal to or greater than 45 minutes, the PBS yields little if any improvement over the PF with an observation interval of 45 minutes (shown in black dashed line for reference). However, unlike the PF algorithm (Figure 3.4), the median RMSE consistently increases with increased assimilation interval. This is particularly noticeable between 2cm to 8 cm. Therefore, the best results are obtained when all available observations are assimilated. This may prove particularly useful in the context of DTS as observations can be made every minute or less.

IMPACT OF WINDOW LENGTH

The impact of window length in the PBS algorithm is explored in Figure 3.6, where an observation interval of 15 minutes, and 100 particles are used. From 2cm to 32cm, increasing the window length from 1 to 3 hours results in a reduction in the median RMSE. The improvement is greatest at the assimilation depths of 4 cm and 8 cm. This confirms that assimilating a series of temperature observations may contain more soil moisture information than sequentially assimilation of instantaneous observations. However, increasing the window length further to 6 or 12 hours leads to an increase in both median RMSE and the IQR. There are two contributors to this degradation in performance. First, when precipitation occurs within a window, the estimated soil moisture prior to the precipitation event may be drawn to a moister condition by the subsequent observations. This was observed when brightness temperature was assimilated to estimate soil moisture using an ensemble batch smoother [59]. Second, the dimension of the distributions to be estimated increases with window length, and additional particles may be needed to accurately capture the posterior distribution.

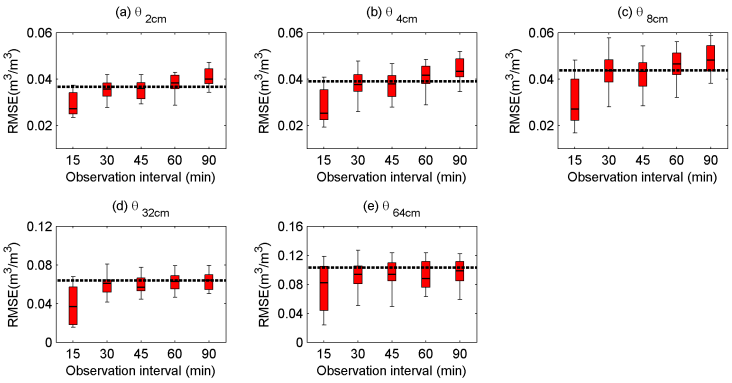

Figure 3.5: Impacts of soil temperature observation interval on the PBS algorithm. The PBS algorithm uses a 3 hour window length with 100 particles. The black dashed line represents the averaged RMSE of the estimates from the PF using an observation interval of 45 minutes. Similar to Figure 3.4, 20 tests with different model initialization are used.

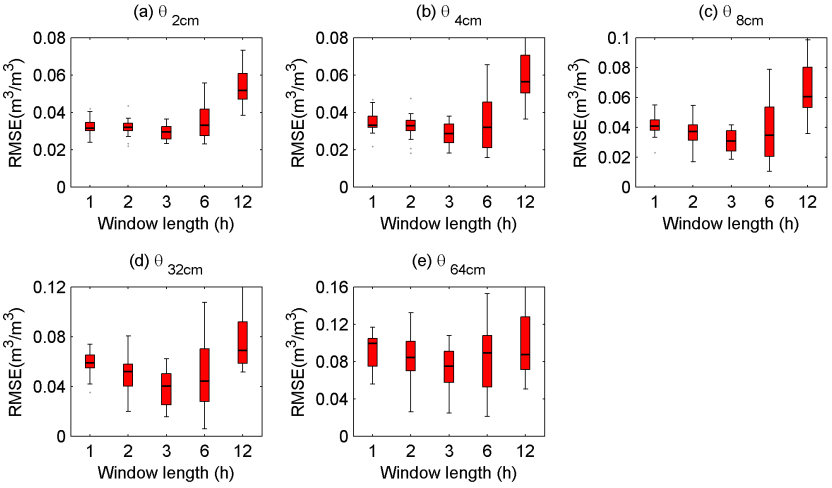

Figure 3.6: Comparison of RMSE of soil moisture estimated using the PBS with different assimilation window lengths at 5 depths. For each window length, 20 tests with different model initialization is used. The observation interval is 15 minutes, with 100 particles.

COMPARISON OF THE "OPTIMAL" PF AND THE PBS APPROACHES

Based on the results in Figure 3.5 and Figure 3.6, the optimal assimilation strategy for the PBS combines a 15 minutes observation interval with a window length of 3 hours. Figure 3.7 shows an example of soil moisture estimated using this optimal PBS and the optimal PF (45 minute observation interval). The initial particle settings are the same for both algorithms. The soil moisture at all depths (Figure 3.7) benefit from the significant improvement in the correction from the assumed initial conditions. The PBS draws the estimate to the truth within the first day, while the PF adjusts it slightly before the particle range collapses inhibiting further improvement. At 32cm and 64 cm, the PBS estimate continues to track the truth as the impact of improved soil moisture at 2cm to 8 cm is propagated downwards through the model physics; while the PF estimate is drying out at 8 cm to 64 cm, the PBS increases at a similar rate to the true observed soil moisture. At 2cm and 4 cm, the PBS also draws the moisture estimates closer towards observations after precipitation (e.g. 30 Mar, 01, 03 and 07 Apr). The Q-Q plot in Figure 3.8 compares the PF and the PBS from a probabilistic point of view. In general, the PF estimates are biased compared to the observed soil moisture, since the predicted quantiles are consistently lower than the uniform distribution. As discussed above, this is caused by the errors in the initial condition, and the uncertainties in the parameters. Compared with the PF, the PBS reduces the bias of the estimates, which is most significant at depth of 32cm. Both the PF and the PBS are overconfident in the estimates, i.e. the particle range is too narrow to encompass the observations. This is partly because soil moisture was not directly assimilated, and partly because the observation accuracy of soil temperature is very small, which leads to particle weight degeneracy.

Results in Figure 3.7 and Figure 3.8 are illustrative and based on one assumed initial condition. Table 3.2 shows the averaged RMSE from 20 tests, each with a different assumed initial condition. It is clear that the optimal PBS (PBS-15min) yields a significant reduction in RMSE compared to the optimal PF (PF-45min). Similar to the results in Figure 3.7, the greatest improvements are at 32cm and 64 cm where the RMSE is reduced by 0.023 and 0.020m^3m^{-3} respectively. The bias in the PBS-15min is significantly smaller than that of PF-45min estimated soil moisture. This may indicate that the PBS algorithm is more suitable in correcting the errors in the initial conditions. The reliability of the PBS-15min is approximately 4 to 7 times higher than that of PF-45min at depths above 8 cm. At 32 and 64 cm, the differences in reliability are insignificant. This indicates that both algorithms are overconfident in the estimates, which is consistent with Figure 3.8. The PBS may be further improved by including a MCMC algorithm [e.g. 58, 80], or by modifying the likelihood function to a distribution with heavy tails [86]. Both techniques may help to identify the location with larger posterior probability, and reduce the overconfidence of the estimates.

3.4.3. COMPUTATIONAL BURDEN OF THE PBS AND THE PF

Here, we quantify the potential increase in computational burden associated with using a smoothing approach (PBS) rather than a filter (PF). There are two factors to consider. The first is that the posterior distribution has a larger dimension in the PBS as so additional particles may be required to capture it. Figure 3.9 (a) to (e) show the impact of the number of particles on the RMSE in soil moisture at each depth. For any given number

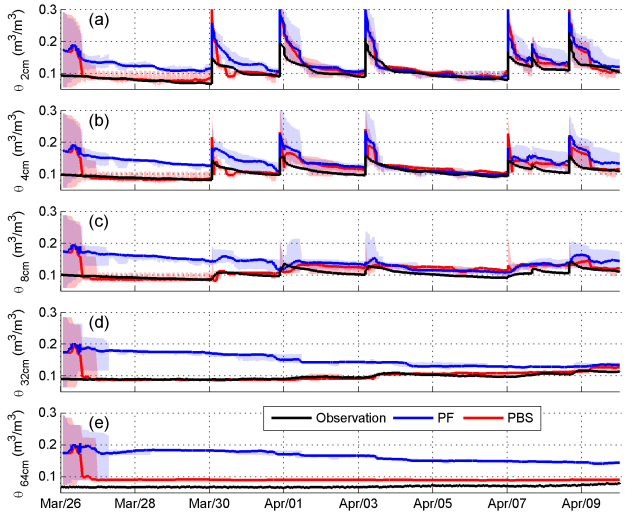

Figure 3.7: An illustrative example of comparing the soil moisture estimates at 5 depths using the optimal PF (observation interval of 45 minutes), and the optimal PBS (observation interval of 15 minutes and window length of 3 hours). The blue and red solid lines represent the mean of the PF and the PBS estimates. The shade areas are the particle range of the PF (blue) and the PBS (red). The initial conditions for the PBS and PF are the same, and identical to these used in Figure 3.2.

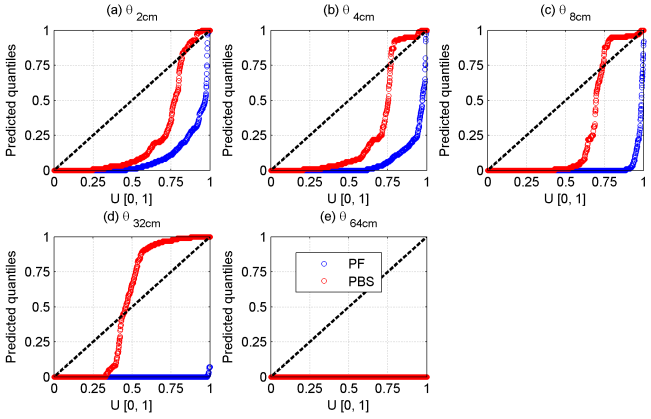

Figure 3.8: The QQ plot for the PF and the PBS estimated soil moisture at five depths. The estimates are from the results presented in Figure 3.7.

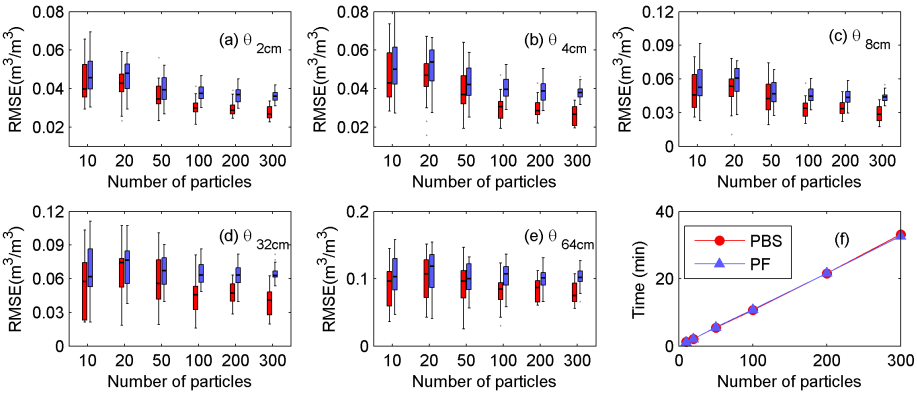

Figure 3.9: Comparison of RMSE of soil moisture estimated using the PBS and the PF with different number of particle sizes at 5 depths (a to e), and the computation time for different number of particles (f). A window length of 3 hours and 15 minutes observation interval is used for the PBS. The observation interval for the PF is 45 minutes. For each window length, 20 different initialized model inputs were used for testing.

of the particles, the PBS outperforms the PF at each depth. For both algorithms, the median and the range of the RMSE is generally lower for a larger number of particles. The greatest reduction in RMSE is observed between 2cm and 8 cm as the number of particles is increased up to 100. Any further increase leads to a marginal, if any, reduction in median RMSE for both algorithms.

The second is the additional cost of performing the sequential importance sampling calculations for the larger state vector. Furthermore, the states at previous times are stored in memory which adds input/output costs. Figure 3.9 (f) shows that the difference in clock time required by the PBS and PF is negligible (< 3%). The dominant control is the number of particles, so the computational burden is determined by the Hydrus-1D simulations rather than the PF or PBS update steps.

3.5. CONCLUSION

In this chapter, we investigated the potential to use particle approaches to estimate soil moisture from temperature observations. Two particle-based approaches (i.e. the Particle Filter, PF and the Particle Batch Smoother, PBS) were tested by assimilating temperature observations and validating the estimated soil moisture profile against soil moisture observed using conventional Hydra probes. The PBS uses the evolution of soil temperature within a window, instead of instantaneous measurements. Therefore, the PBS may be more suitable for capturing the temperature heating/cooling rate, hence more suitable for soil moisture estimation. We considered sources of uncertainty comparable to those which would be encountered in a DTS application, i.e uncertain meteorological forcing, soil texture parameters etc.

Results demonstrate that assimilation using a PF yields a significant improvement over an open loop (no assimilation) run. The best estimates were obtained between the surface and the depth of the deepest temperature observation. Estimates at greater

depth were particularly sensitive to the prescribed initial condition as the range of particle values collapsed soon after the first updates preventing any additional update. The PF updates the states by updating the weights of the particles, giving those that agree with the observations a larger weight. It can only adjust within the range of predicted values, so it is essential that this range includes the observation. Results indicate that the PF performs best when observations are assimilated at an interval that allows this predicted range to grow. Further research will investigate whether including MCMC and/or alternative proposal distributions could alleviate this problem.

In addition, we implemented a PBS algorithm, in which a series of temperature observations within a window are assimilated to update the trajectory of soil moisture in that window. Results demonstrated that this smoothing approach yielded a statistically significant reduction in RMSE compared to the PF. Furthermore, this improvement was achieved with a negligible increase in computational cost. The PBS uses a trajectory of temperature observations within a window, which contains more moisture information than one instantaneous observation. This results in a reduction of RMSE up to 0.023 m^3m^{-3}, or a RMSE reduction of 33%, compared with PF algorithm. The PBS updates the prior moisture estimates once per window, which allows us to use all available observations without particle degeneration. As the PBS performs assimilation on the entire window at once, the dimension of the joint distribution of the states is larger than in the PF. For soil moisture estimates, a window length of 3 hour with observation frequency of 15 minutes was shown to yield the best results. For both the PF and PBS, the number of particles was found to be about 100. Increasing the number any further resulted in only a marginal improvement in RMSE, while the computational burden increased linearly with the number of particles.

The approaches studied here were developed to use Distributed Temperature Sensing to estimate soil moisture. Applying the particle approaches with DTS would allow us to estimate soil moisture every 1 m or less along fiber-optic cables that can be several kilometers in length. Though these first results are already promising, we will consider the additional benefit of performing dual state-parameter estimation, e.g. [53, 69], in the next chapter. Improving the soil parameters, would improve the performance of the model itself and potentially provide a means to monitor soil heat flux using DTS.

4

ESTIMATING SOIL MOISTURE AND SOIL PROPERTIES IN BARE SOIL AREAS BY ASSIMILATING SOIL TEMPERATURES

Based on: Dong Jianzhi, Susan C. Steele-Dunne, Tyson E. Ochsner, Nick van de Giesen. Estimating soil moisture and soil thermal and hydraulic properties by assimilating soil temperatures using a particle batch smoother, 2016, *Advances in Water Resources*, 91, 104-116.

4.1. INTRODUCTION

The previous chapter showed the importance of assimilating a sequence of soil temperature observations for soil moisture estimation. Improved model parameters can significantly improve the performance of the forward model, which means the prior estimated model states can better encompass the observations. Thus, updating model states together with model parameters can yield significantly improved estimates, compared with the cases that only model states are updated [e.g. 54, 87, 88]. This joint model state-parameter estimation scheme may also improve the data assimilation approach presented in the previous chapter, provided the soil temperature observations contain information of soil thermal and hydraulic properties. This joint model state-parameter estimation scheme can keep the consistency of the states and the parameters, and hence provide physically reasonable soil state and parameter combinations.

Soil thermal properties can be estimated using soil temperatures is already shown in several previous studies [e.g. 12, 38, 89]. As shown in the previous chapters, different soil moisture dynamics will eventually lead to different soil temperature evolution patterns. Soil hydraulic properties are key factors that control soil moisture dynamics. Hence, soil hydraulic properties may also be inferred from soil temperature observations.

This chapter tests the hypothesis that soil thermal and hydraulic properties can be estimated using soil temperatures. The benefit of jointly estimating soil properties with soil states will also be investigated. This chapter will first present a sensitivity analysis between soil temperature and soil hydraulic properties, which aims to explore the physical links between the two quantities (section 4.3.1). The algorithm of jointly estimating model states and parameters using the particle batch smoother to extract soil moisture and soil property information from soil temperatures is presented in section 4.2. Finally, the robustness and the benefit of this joint model state-parameter estimation scheme will be tested using a series of synthetic experiments 4.4.

4.2. JOINT PARAMETER AND STATE ESTIMATION

The forward model and the implementation of the PBS are essentially the same with the algorithms described in section 3.2.1 and 3.2.2, except that the state vector ($\mathbf{x}_t^i$) is augmented as:

$$\mathbf{X}_t^i = \left[\mathbf{x}_t^i, \mathbf{b}_t^i\right]^T \tag{4.1}$$

Consequently, the estimated posterior distribution is expressed as:

$$\begin{aligned} p(\mathbf{X}_{t-L+1:t}|\mathbf{y}_{1:t}) &= p(\mathbf{x}_{t-L+1:t}, \mathbf{b}_{t-L+1:t}|\mathbf{y}_{1:t}) \\ &= \sum_{i=1}^{N} w_t^i \delta\left(\mathbf{x}_{t-L+1:t} - \mathbf{x}_{t-L+1:t}^i, \mathbf{b}_t - \mathbf{b}_t^i\right) \end{aligned} \tag{4.2}$$

Note that the model parameters are assumed to be constant within each batch window (i.e. $\mathbf{b}_{t-L+1:t}^i = \mathbf{b}_t^i$). Perturbing the estimated parameter set is usually required to avoid parameter impoverishment [53, 54]:

$$\mathbf{b}_t^i = \mathbf{b}_t^i + \epsilon_b \tag{4.3}$$

where ϵ_b is normally distributed noise with zero mean, and standard deviation (Std) of s. In this study, s is determined as:

$$s = max\,(0, Q_0 - \mathrm{Std}(\mathbf{b}_t)) \tag{4.4}$$

where Q_0 is a prescribed threshold [90], which is set to be 10% of the initial parameter standard deviation. A discussion of the choice of Q_0 can be found in [90] and [91]. This perturbation technique guarantees the parameter distribution has a minimum spread, and this spread cannot grow uncontrollably through perturbation [91].

TUNING FACTOR FOR THE PBS

In case observation is near perfect (i.e. $\mathbf{R}$ is small) or all the particles have large errors in the prior estimated model states, resampling cannot entirely avoid the weight degeneracy problem [86]. As a result, the variance of the weights is high, and estimated posterior relies on only a few particles. This will lead to unreliable estimates. Stordal et al. [86] suggest to avoid this problem by approximating a posterior that has heavier tails, rather than the true posterior. Though biases might be introduced, the final estimates will almost surely converge to the true posterior [86, 92]. Similar to their study, we introduce a tuning factor of β which modifies the likelihood function (Equation 3.13) in the form of:

$$p(\mathbf{y}_{t-L+1:t}|\mathbf{x}^i_{t-L+1:t}) = \prod_{j=t-L+1}^{t} \frac{1}{(2\pi)^{n_o/2}\det(\mathbf{R})^{1/2}} e^{\left[-0.5\beta^2(\mathbf{y}_j-\hat{\mathbf{y}}^i_j)^T\mathbf{R}^{-1}(\mathbf{y}_j-\hat{\mathbf{y}}^i_j)\right]} \tag{4.5}$$

The tuning factor of β ranges from zero to one. The modified likelihood function is equivalent to the original likelihood function if β is set to be one. For the cases $\beta < 1.0$, the variance of the estimated weights will be significantly reduced, which may efficiently avoid the weight degeneracy problem. The β value smaller than 1.0 is essential to compensate for cases in which the PBS is overly confident with respect to estimating parameter values.

4.3. EXPERIMENT SET-UP

4.3.1. SENSITIVITY OF MODEL STATES TO SOIL HYDRAULIC PROPERTIES

Data collected from Marena Oklahoma In-Situ Sensor Testbed (MOISST) site (A.1) are used in this Chapter. Prior to the data assimilation experiments, we will first demonstrate that different soil hydraulic properties will lead to different soil temperatures. A Monte Carlo simulation with 300 samples (runs) using identical model inputs, except for the soil hydraulic properties, will be conducted. The forcing data is the nominal value of the field measurements, and the soil hydraulic properties were generated using ROSETTA [65] with randomly sampled soil texture and bulk density drawn from the distributions described in Table 4.1. The variance of the soil moisture and temperature of the 300 simulations can shed light on the extent to which the soil temperature variations can be explained by the variance of soil hydraulic properties.

Then, we will provide a more quantitative sensitivity test between the change of soil temperatures to the change of the soil hydraulic properties. A reference run will be first

Table 4.1: Generation of perturbed inputs (soil property and forcing) for each particle.

Variable	Error Distribution	Mean	Standard deviation	Bound
Sand (%)	Uniform	-	-	0, 55
Silt (%)	Uniform	-	-	20, 30
$\rho_b(g/cm^3)$	Uniform	-	-	1.05, 1.85
Precipitation (mm)	Gaussian, Multiplicative	1	× 0.2	-, -
Radiation (W/m^2)	Gaussian, Multiplicative	1	× 0.075	-, 1350
Air temperature (^{o}C)	Gaussian, Multiplicative	1	× 0.05	-, -
Relative humidity (%)	Gaussian, Multiplicative	1	× 0.05	-, 100
Wind speed (Km/h)	Gaussian, Multiplicative	1	× 0.2	-, -

performed using Hydrus-1D, in which soil texture and bulk density is set to be the nominal value of the field measurements. Soil thermal and hydraulic properties were then calculated based on this set of soil texture and bulk density for the reference run. We will then fix all the inputs but only vary one of the soil hydraulic properties within the range of [90% 110%] of the nominal value. Five soil hydraulic properties (θ_r, θ_s, α, n, K_s) will be considered.

4.3.2. DATA ASSIMILATION EXPERIMENTS

Five data assimilation algorithms are compared in this study: (1) the open-loop (OL), which is identical to running the particles in parallel without any data assimilation. (2) PBS-State, in which soil temperature is used to update soil moisture and temperature states. (3) PBS-DT, where the state vector is augmented with the soil thermal properties. In the Campbell model, the sand, clay and bulk density are input parameters to determine the relationship between thermal conductivity and soil moisture. This study takes the three quantities as three free soil thermal property parameters. (4) PBS-DH, in which state vector is augmented with the soil hydraulic parameters (θ_r, θ_s, α, n, K_s), and (5) PBS-DTH, in which both the soil thermal and hydraulic properties are jointly updated with soil moisture and temperature.

This study is limited to synthetic experiments. The truth is generated using perturbed forcing data and parameters from the distributions summarized in Table 4.1. Precipitation and solar radiation were perturbed using multiplicative errors to ensure that the values remain positive. Sand, clay content and bulk density (ρ_b) are randomly sampled from a uniform distribution as described in Table 4.1, which covers the soil texture spatial variability of the study area. The clay content for each particle is calculated as the residual, i.e. (100 - Sand -Silt)%. Soil texture and bulk density are assumed to be uniform over the entire profile. The soil texture and bulk density for each particle were sampled from the distribution described in Table 4.1. This randomly sampled soil texture and bulk will be used to generate the prior guessed soil hydraulic and thermal properties using ROSETTA [65] and Campbell model. Initial soil temperature and moisture profiles are assumed to be homogeneous in depth. The initial soil temperature value is drawn from a uniform distribution with the range 17 to 25 ^{o}C, and soil moisture is drawn from a uniform distribution with the range 0.2 to 0.35 m^3/m^3. The assumed errors in the forcing data, initial soil moisture and temperature profile, and soil properties reflect the

measurement error as well as spatial representativeness - weather and soil property data are collected at a point, while DTS observations are distributed in space. Synthetic DTS temperature observations are generated by adding Gaussian error with zero mean and standard deviation of 0.1 ^{o}C to the “true” soil temperature at these depths.

A detailed discussion on the impacts of the window length and assimilation interval is provided in the previous chapter. They demonstrated that soil temperature within a short window length (e.g. 1 hour) may not contain enough soil temperature information. On the other hand, using overly long window lengths will increase the dimension of the state vector and hence of all the probability distributions. The previous chapter showed that a window length of 3 hours with an observation interval of 15 minutes was the most suitable. Because the experiment design of the current study is similar to the previous chapter, the same window length and assimilation interval are used here. The PBS requires sufficient particles to map the prior and the posterior distributions. The previous chapter shows using 100 particles provides similar estimates to the cases using 300 particles, when only states were estimated. In this study, model parameters were estimated in addition to the states. We found that 300 particles were sufficient to ensure that the PBS could map the prior and the posterior distribution.

The first experiment is an illustrative example for a single “truth" to compare the four assimilation approaches. The soil thermal and hydraulic properties estimated from the four assimilation approaches will be compared, and the benefit of estimating the soil properties on the soil moisture and temperature estimates will be investigated.

To provide a more robust comparison of the different data assimilation strategies, the experiment described above will then be repeated for 10 randomly selected truths. The initial conditions and the prior guess of soil thermal and hydraulic properties for the PBS data assimilation strategies will be the same as that described in the previous experiment.

Finally, we will explore the impact of the tuning factor (β) on the performance of the PBS. As shown in Equation 4.5, the estimated posterior distribution will have heavier tails when β is small. This might be helpful in avoiding severe weight degeneracy problem, and ensuring that the particle range is large enough to encompass the observations. However, an extremely small β will lead to very uncertain estimates, and the observations may be ignored. For example, the PBS is equivalent to the OL if $\beta = 0$. Therefore, in the final experiment, the impact of the tuning factor will be investigated by varying β from 0.25 to 1. Note that in all other experiments, the tuning factor β is assumed to be 0.75 for the PBS-State, and 0.25 for PBS-DT, PBS-DH and PBS-DTH.

4.4. RESULTS AND DISCUSSION

4.4.1. SENSITIVITY OF MODEL STATES TO SOIL HYDRAULIC PROPERTIES

Figure 4.1 shows the mean and the standard deviation (Std) of the soil temperature and moisture simulated in the 300 Monte Carlo runs. Only soil hydraulic properties were varied for the simulations. Hence, in Figure 4.1 (b), a larger standard deviation in the simulated soil moisture indicates a higher sensitivity of soil moisture to the soil hydraulic properties. The soil hydraulic properties determine the liquid water that infiltrates into the soil column during the precipitation events. As a result, the highest standard devia-

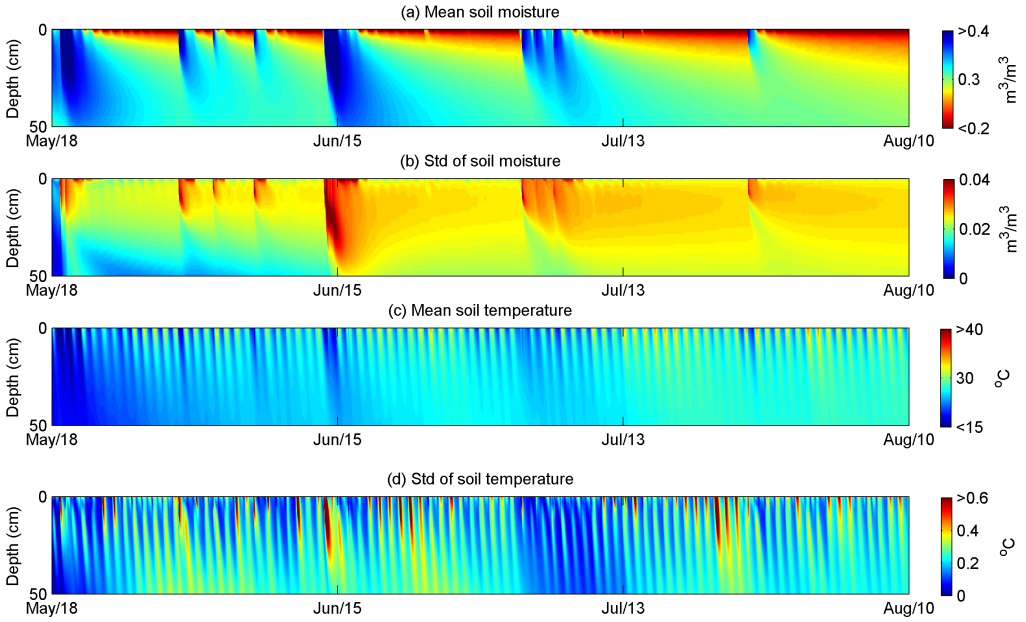

Figure 4.1: The mean and standard deviation (Std) of soil moisture (a and b) and soil temperature (c and d) using 300 Monte Carlo runs which differ in soil hydraulic properties. For each run, the soil texture and bulk density are sampled from the distribution described in Table 4.1, and ROSETTA [65] is used to determine the soil hydraulic properties.

tion in soil moisture occurs around precipitation events (e.g. June 15 in Figure 4.1 b), and decreases gradually as the soil dries down. Solar radiation forcing is the main driver for the diurnal variations observed in both the mean (Figure 4.1 c) and standard deviation (Figure 1d) of soil temperature. However, soil moisture affects soil thermal properties which in turn influences the temperature. Hence, variations in soil temperature also occur due to variations in soil moisture which lead to different rates of heat diffusion and evaporative cooling. The standard deviation of soil temperature can exceed 0.6 ^{o}C, which is significantly larger than the accuracy of most temperature observations.

Figure 4.2 shows the temporal mean of soil moisture and temperature at 5 cm as a function of soil hydraulic properties. Of the five soil hydraulic properties, soil moisture and soil temperature are most sensitive to n. Increasing/Decreasing n by 10% can result in 0.2 ^{o}C change in the temporal mean of soil temperature. θ_s, and to a lesser degree θ_r, also have a small impact on the temporal mean of soil temperature. The soil temperature is shown to be less sensitive to α and K_s. Note that the range within which the mean temperature varies is less than 0.5 ^{o}C. From Figure 4.1, it is clear that variations are higher at certain times of day and in response to precipitation events.

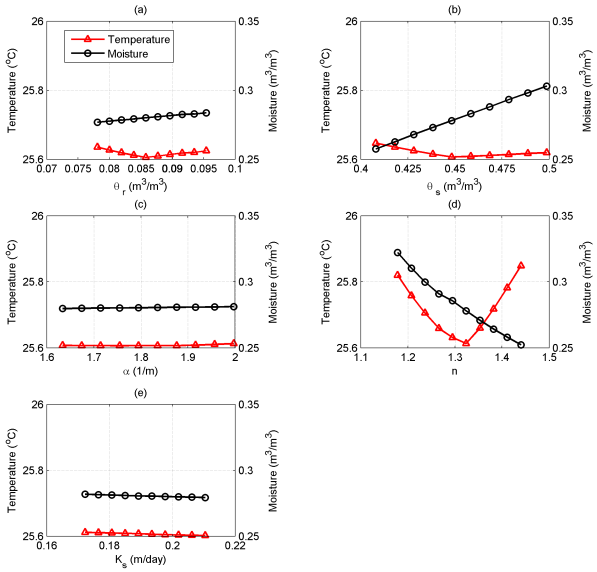

Figure 4.2: The temporal mean (from 18 May to 10 August) of soil temperature and moisture at 5 cm using different soil hydraulic properties. Each symbol represents the results from one simulation.

4.4.2. AN ILLUSTRATIVE EXAMPLE

The soil moisture profiles estimated using the OL, PBS-State, PBS-DT, PBS-DH and PBS-DTH are shown in Figure 4.3. A case from multiple truth tests, in which PBS-State estimated soil moisture significantly deviates from the truth is presented here to illustrate the benefit of estimating the soil properties. Due to the errors in the prior estimated soil hydraulic properties, the entire profile in the OL is generally much drier than the truth (Figure 4.3 a and b). Comparing the four data assimilation strategies, it demonstrates updating soil hydraulic properties is the key to provide accurate soil moisture estimates (Figure 4.3 c to f).

In this illustrative case, the estimated bulk density (Figure 4.4 a) and sand (Figure 4.4 b) and clay contents (Figure 4.4 c) from the PBS-DT and PBS-DTH converge quickly to the true values. This results in almost perfect estimates of the soil thermal conductivity and diffusivity curves (Figure 4.4 d and e).

The soil hydraulic properties estimated with the PBS-DH and PBS-DTH are shown in Figure 4.5. Recall from Figure 4.2 that θ_s and n were found to be the most sensitive. When only the hydraulic properties are updated with the states, θ_s does not converge to the truth (Figure 4.5 b). Because the thermal properties are not updated in the PBS-DH, it seems that the PBS has overfitted the hydraulic parameters to compensate for the errors in the soil thermal properties. When both the thermal and hydraulic properties are estimated (PBS-DTH), all parameters converge to the truth (Figure 4.5 a to d). The temporal evolution of α estimated using both methods is noisier than that of other soil

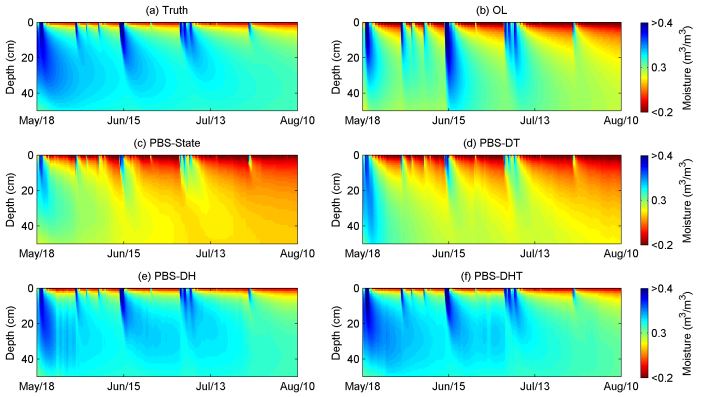

Figure 4.3: The synthetic truth (a) in compared to the open loop (b), PBS with state estimation only (c), PBS for state and thermal properties estimation (d), PBS for state and hydraulic properties (e), and PBS for state, and thermal and hydraulic properties estimation (f).

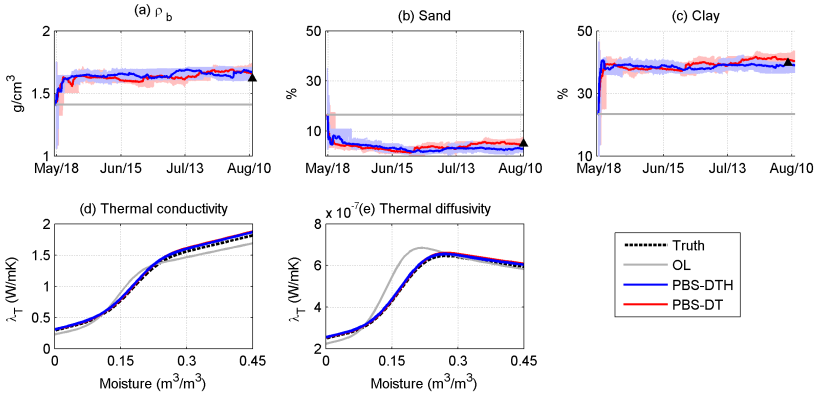

Figure 4.4: The convergence of the soil thermal properties (a to c). The “true" values are from one single realization, and comparison of estimated and true soil thermal conductivity (d) and diffusivity curve (e), and. The shaded area represents the range of the particles. The value of the true parameter is shown as the black triangle at the final time step.

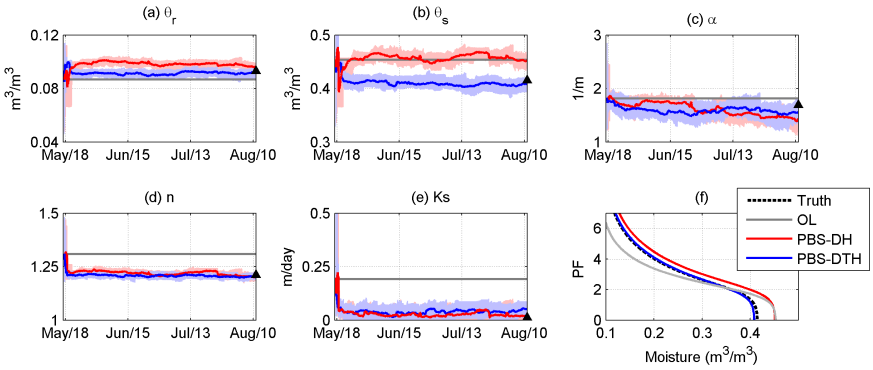

Figure 4.5: An illustrative example of the convergence of the soil hydraulic properties. The “true" values are from one single realization (a to e), and comparison of the true and estimated soil water retention curve (f). The shaded area represents the range of the particles. The value of the true parameter is shown as the black triangle at the final time step.

4

hydraulic properties (Figure 4.5 c), which is due to the low sensitivity to the soil temperature (Figure 4.2 c). Both the PBS-DH and PBS-DTH provide a good estimate of the soil water retention curve, with the PBS-DTH performing better at higher moisture contents (Figure 4.5 f).

Results from the illustrative case presented in Figure 4.3 to Figure 4.5 confirm what we expect from Figure 4.2. First of all, the PBS-State may correct the prior estimated soil moisture to a significantly biased value when only states were updated. A dry soil moisture can provide a similar thermal response to that of the wet soil moisture ([38], also see Figure 4.2). In this illustrative example, the soil hydraulic property is biased, which leads to a faster dry down compared to the truth (Figure 4.2 b). Hence, the PBS-State may converge to an incorrect dry soil moisture value. The PBS-DT does not significantly benefit the soil moisture estimates, since it cannot remove the biases in the prior estimated soil moisture. Updating soil hydraulic properties contributes the most to the soil moisture estimation (Figure 4.3 e and f). The PBS-DTH provides the best performance, as it ensures consistency between the model parameters and model states for both soil heat and water transportation.

4.4.3. ROBUSTNESS TEST USING MULTIPLE TRUTHS

The experiment discussed in the previous section was repeated 10 times for each of 10 different synthetic truths, which will provide insights about the robustness of the different PBS schemes across a wide range of soil textures. The estimated soil thermal properties from these tests are compared to the true values in Figure 4.6, and performance metrics are reported in Table 4.2. The RMSE and bias are lower and the correlation is higher when the PBS-DTH is used rather than the PBS-DT. The reduction in bias is particularly noteworthy. These results are consistent with those in the illustrative case. When the soil hydraulic properties are not updated, the soil moisture estimates may converge to an incorrect value resulting in incorrect or biased soil thermal property estimates.

The soil hydraulic properties estimated using the PBS-DH and the PBS-DTH are com-

Table 4.2: The RMSE, absolute bias and the correlation between the true and estimated soil thermal properties using the PBS-DT and PBS-DTH. The values presented are calculated based on the estimates from 10 tests

Metric	**Approach**	$\rho_{\mathbf{b}}(g/cm^3)$	**Sand** (%)	**Clay** (%)
RMSE	PBS-DT	0.052	8.147	12.324
	PBS-DTH	0.014	3.266	2.111
Bias	PBS-DT	0.028	3.693	4.666
	PBS-DTH	0.002	0.538	0.044
Correlation	PBS-DT	0.881	0.915	0.797
	PBS-DTH	0.984	0.984	0.993

Table 4.3: The RMSE, absolute bias and the correlation between the true and estimated soil hydraulic properties using the PBS-DH and PBS-DTH. The values presented are calculated based on the estimates from 10 tests

Metric	**Approach**	$\theta_r(m^3/m^3)$	$\theta_s(m^3/m^3)$	$\alpha(1/m)$	n	$K_s(m/day)$
RMSE	PBS-DH	0.006	0.024	0.320	0.033	0.039
	PBS-DTH	0.001	0.008	0.170	0.013	0.026
Bias	PBS-DH	0.002	0.002	0.076	0.002	0.023
	PBS-DTH	0.000	0.001	0.102	0.004	0.015
Correlation	PBS-DH	0.926	0.885	0.566	0.930	0.991
	PBS-DTH	0.996	0.989	0.703	0.990	0.996

pared in Table 4.3 and Figure 4.7. The bias of the estimated soil hydraulic properties using the two approaches are similar. However, the PBS-DTH shows significantly reduced RMSEs and improved correlations between the true and estimated parameters. Since only soil temperatures are assimilated, improved soil thermal properties also benefit the estimated soil moisture and hydraulic properties.

The RMSE in soil temperature estimated using OL, PBS-State, PBS-DT, PBS-DH and the PBS-DTH tested by the 10 truths are shown in Figure 4.8. Compared to the OL estimates, all PBS strategies (i.e. the PBS-State, PBS-DT, PBS-DH and PBS-DTH) significantly improve the soil temperature estimates across the entire profile. At depths where soil temperatures are observed (i.e. 5 and 10 cm), the 4 PBS strategies provide similar results. However, at depths without temperature observations, updating soil thermal properties (i.e. PBS-DT and PBS-DTH) can reduce the median and IQR of the RMSE of

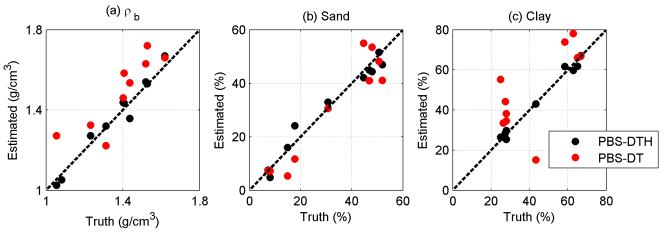

Figure 4.6: Comparison of the estimated and the true soil thermal properties tested using ten different truths.

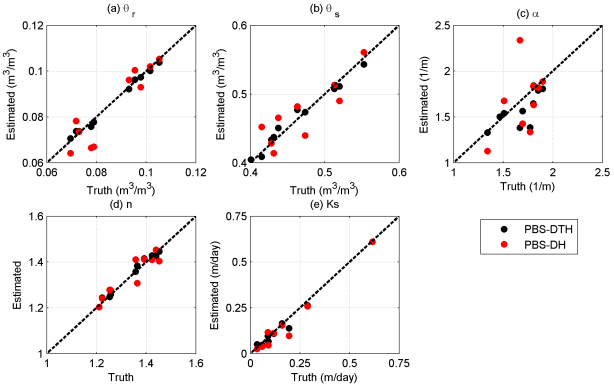

Figure 4.7: Comparison of the estimated and the true soil hydraulic properties tested using ten different truths.

the soil temperature estimates, compared to PBS-State and PBS-DH. This is particularly noticeable at 20 and 50 cm (Figure 4.8 d and e). Because heat transfer in the soil is primarily downwards, better soil thermal properties improve the model's ability to simulate temperature at depth. Soil moisture and hydraulic properties determine the soil evaporation processes, which essentially control the cooling rate of the soil surface. Hence, a better knowledge of the soil surface moisture and hydraulic properties can significantly improve the soil temperature estimates, and remove the bias of the estimated soil temperatures at shallow soil depths.

The RMSE and the bias in the soil moisture estimates using OL, PBS-State, PBS-DT, PBS-DH and the PBS-DTH across the 10 different truths are shown in Figure 4.9. Consistent with the results presented in the illustrative case, Figure 4.9 shows that estimating soil hydraulic properties leads to improved soil moisture estimates. If the soil hydraulic properties are not updated, the PBS may converge to an incorrect dry soil moisture value, as occurred in Figure 4.3. Hence, the IQR of the RMSE using the PBS-State and PBS-DT are both larger than the OL, though the median of the RMSE is reduced. Compared to the PBS-State and PBS-DT, both the median value and the IQR of the RMSE of the soil moisture estimates are significantly reduced by PBS-DH.

This set of experiments demonstrates the importance of removing the errors/biases in the soil moisture estimates by improving the soil hydraulic properties. The further reduction in the median and IQR of RMSE when PBS-DTH is used rather than PBS-DH, indicates the benefit of updating soil thermal properties as well as hydraulic properties. This confirms the importance of maintaining consistent model parameters and states.

4.4.4. IMPACTS OF TUNING FACTOR

The results presented so far are from experiments in which a tuning factor β is set to 0.75 in the state estimation case, and 0.25 in the parameter estimation. In this section, this choice of values will be justified.

Figure 4.10 shows an illustrative example of the evolution of the estimated soil temperature, moisture and porosity (θ_s) when different β values were used in the PBS-DTH.

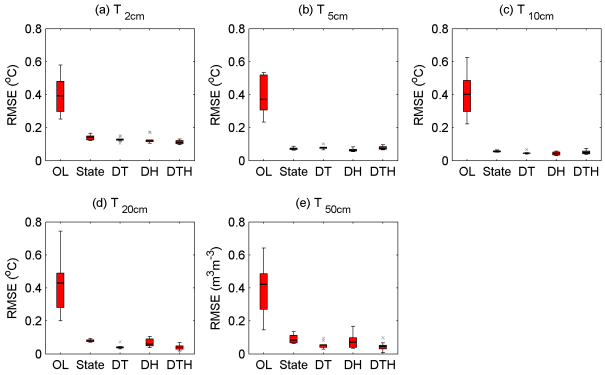

Figure 4.8: Comparison of soil temperature estimates using OL, PBS-State, PBS-DT, PBS-DH, and PBS-DTH at 5 depths. Each box plot presents the results from 10 tests using different truths. In the box plot, the middle black line denotes the median value, the edges of the box are the interquartile range (IQR), the maximum length of the whiskers is set to be the 1.5 times the IQR, and values larger/smaller than the maximum/minimum the whiskers are considered as outliers (black crosses). The legends are the same for the following box plots.

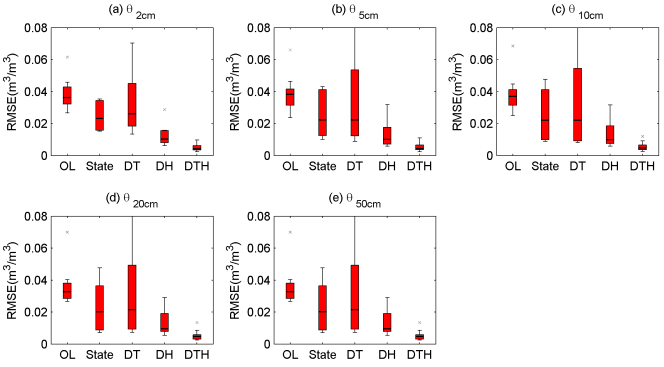

Figure 4.9: Comparison of soil moisture estimates using OL, PBS-State, PBS-DT, PBS-DH, and PBS-DTH at 5 depths. Each box plot presents the results from 10 tests using different truths.

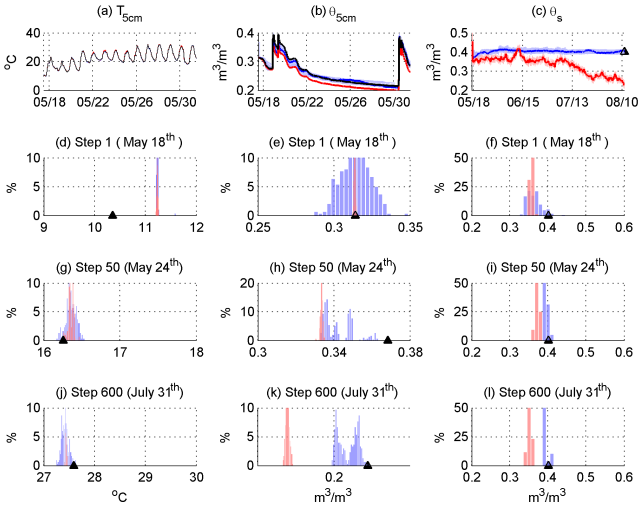

Figure 4.10: The PBS-DTH estimated the soil temperature (first column), moisture (second column) at 5 cm, and the convergence of soil hydraulic properties (third column) , using saturated water content (θ_s) as an example. The first row shows the temporal evolution of the estimates. The estimated posterior at some model steps are shown in row 2 to 4, and the black triangle denotes the value of the truth. The shaded blue area is the estimated posterior using $\beta = 0.25$, and shaded red is for $\beta = 1.00$.

Only the first 14 days are shown in the soil temperature and moisture estimates. When $\beta = 0.25$, the estimates are closer to the synthetic true values. However, the estimates are significantly biased when β is set to be 1.00 (Figure 4.10 b and c). The second to fourth rows provide some insight into the impact of β. After the first update step, using different β values provide similar estimated soil temperature posteriors. The main differences are shown in the soil moisture and θ_s posteriors. Using β of 0.25 results in a heavier tailed distribution, which can better encompass the truth (Figure 4.10 f). After about 10 update steps (Figure 4.10 g-i), the true soil moisture and θ_s are almost always beyond the range of the posterior for $\beta = 1.0$, even though soil temperature can still be reasonably estimated (Figure 4.10 j). When $\beta = 0.25$, the estimated posterior is wide enough to encompass the truth. This is critical for successfully implementing PBS, since PBS only adjusts the weights of the particles. Hence, the optimal value cannot be estimated using PBS if the truth falls out of the particle range. Different soil moisture and hydraulic properties can provide similar soil thermal responses (Figure 4.2). If the truth is not within the particle range, the PBS estimated soil moisture and hydraulic properties can converge to an incorrect value. This is evident when β of 1.0 is used. The incorrectly estimated soil moisture and soil properties, and the narrow particle range in turn affect the soil temperature estimates (Figure 4.10 g and j).

The impact of β on the error of the estimated soil thermal and hydraulic properties tested on 10 randomly selected truths is shown in Figure 4.11. As expected, the error of the estimated parameters generally increases with increased β. The IQR of the parameter estimation error with $\beta = 1.0$ is generally 5 to 10 times higher than that using β of

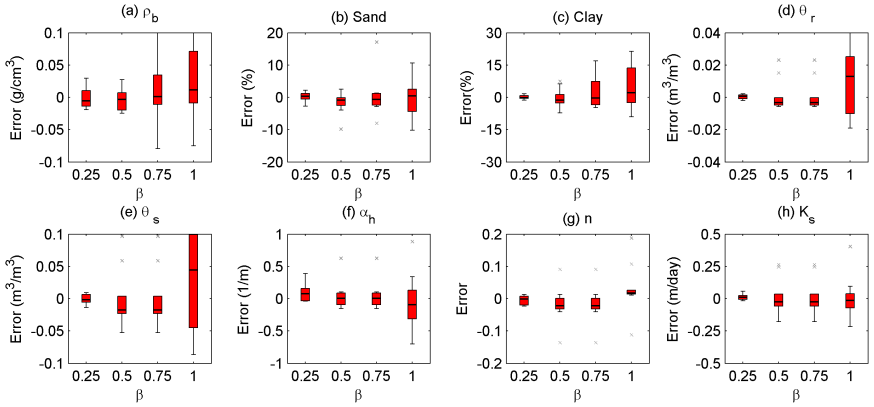

Figure 4.11: The error of the estimated soil thermal and hydraulic properties as a function of the tuning factor β using the PBS-DTH tested using 10 different truths.

0.25.

The impact of β on soil temperature estimates are shown in the left column of Figure 4.12. In contrast to the PBS-State, the IQR of the RMSE of the PBS-DTH estimated soil temperature increases with increased β. It is quite striking that PBS-DTH generally has a larger IQR of soil temperature estimates compared to the PBS-State, when β is larger than 0.25. There are two contributors to this degraded soil temperature estimate. First, the estimated soil properties may converge to an incorrect value in the PBS-DTH, when a large β value is used (Figure 4.10 to 4.11). Second, because the variance of the soil thermal parameters is reduced by the PBS-DTH, the particle range of soil temperature estimates will also be significantly reduced. As a result, the particle range is less likely to encompass the truth. Hence, using a large β value may lead to worse temperature estimates, compared with the PBS-State.

The right column of Figure 4.12 shows the impact of β on the RMSE of the estimated soil moisture. As in the temperature case, the PBS-State is not sensitive to β. Using a β of 0.75 seems to produce slightly lower median RMSE at the surface. However, PBS-DTH is more sensitive to the choice of β. The median and the IQR of the RMSE of the PBS-DTH estimated soil moisture show a clear minimum at $\beta = 0.25$ at all depths. Similar to the soil temperature estimates, the dependency of the PBS-DTH estimated soil moisture on β can also be partly explained by the accuracy of the soil hydraulic parameters as shown in Figure 4.11. Soil hydraulic parameters are most accurately estimated when β is set to 0.25. The error of the estimated soil hydraulic properties increases with increased β, and hence the RMSE of soil moisture estimates also increases.

Figure 4.10 to 4.12 also show that once a reasonable β value is selected, soil moisture, temperature, thermal and hydraulic properties can be robustly estimated, regardless of the soil types. This indicates that β is not a site dependent parameter. Rather, the purpose of β is to avoid severe particle weight degeneration in the PBS. Hence, a reasonable value for β can be determined by examining the parameter estimates. If they are highly variable in time and show little convergence, this suggests that the PBS is overfitting the

parameters and a smaller β is needed.

4.5. CONCLUSIONS

In this chapter, we investigated the potential of estimating soil states (temperature and moisture profile) and soil properties simultaneously by assimilating soil temperatures into Hydrus-1D using a Particle Batch Smoother (PBS). Different data assimilation strategies, specifically state versus joint state-parameter estimation, were compared to investigate the impact of estimating soil thermal and hydraulic properties on the soil moisture estimates.

It was shown that soil hydraulic properties determine the temporal evolution of soil moisture, which results in different soil temperatures. Hence, the soil hydraulic properties can be inferred from soil temperature observations. Correcting the soil hydraulic properties improves the performance of the forward model. As a result, the estimated prior distribution is more likely to encompass the truth, which results in significantly improved soil temperature and moisture estimates. However, an improperly tuned PBS algorithm may have a severe weight degeneracy problem when parameters are jointly estimated. The potential for weight degeneracy is associated with using very small observation errors, and an insufficient parameter range, i.e. in particle methods, the weights are updated, not the particles themselves. So the prior must encompass the truth at each time, otherwise there will be a significant bias. There are a few possible solutions, e.g. using tuning factors to modify the distributions with heavier tails [86, 92], using more sophisticated parameter perturbation techniques e.g. Variable Variance Multiplier (VVM) [93] to increase the variability of the parameter distribution, and using the MCMC algorithm [88] to iteratively provide better proposal distributions. Alternatively, Yang and DelSole [94], and Su et al. [90] have handled similar problems in ensemble data assimilation using so-called “smoothing" factors. This chapter primarily focused on how to use a tuning factor to provide robust and computational effective estimated model states and parameters. Combining the tuning factor with VVM and MCMC algorithms may also have the potential of improving the model parameter estimates.

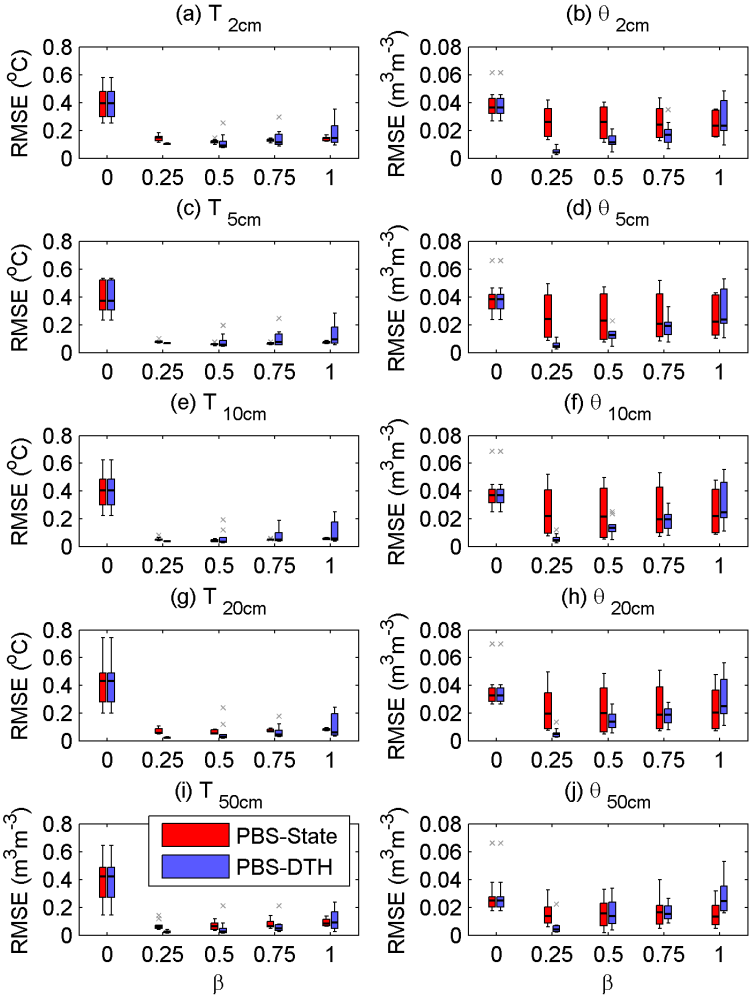

Figure 4.12: Impact of β on the RMSE of estimated soil temperature (left column) and moisture (right column) at 5 depths using the PBS-State and PBS-DTH.

5

DETERMINING SOIL MOISTURE AND SOIL PROPERTIES IN VEGETATED AREAS BY ASSIMILATING SOIL TEMPERATURES

Based on: Dong Jianzhi, Susan C. Steele-Dunne, Tyson E. Ochsner, Nick van de Giesen. Determining soil moisture and soil properties in vegetated areas by assimilating soil temperatures, *Water Resources Research*, 52, 4280-4300.

5.1. INTRODUCTION

It was demonstrated that PBS can robustly estimate the soil moisture and soil thermal and hydraulic properties in Chapter 4. There are still two key challenges in applying this joint model state-parameter estimation approach presented in real DTS data. The first challenge is that the forward model is limited to bare soil and the data assimilation scheme presented in Chapter 4 assumes that the cable depths are perfectly known. In order for Passive DTS to be more broadly applicable as a soil hydrology research tool, and as a remote sensing soil moisture product validation tool, it must be applicable in vegetated areas. Further, the uncertainties in the DTS cable depths has to be handled, since the cable depths are rarely perfectly known in the DTS installations.

This chapter solves the first challenge by including a canopy energy balance scheme into Hydrus-1D model as presented in Section 5.2.1. The second challenge is solved in section 5.2.2, in which the cable depths is estimated as a free parameter. In Chapter 5.3, synthetic tests are used to demonstrate that without the canopy energy balance scheme, the PBS estimated soil moisture could be even worse than the open loop case (no assimilation). When the improved Hydrus-1D model is used as the forward model in the PBS, vegetation impacts on the soil heat and water transfer were well accounted for. This led to accurate and robust estimates of soil moisture and soil properties. Synthetic tests are also used to demonstrate that observation depths can be jointly estimated with other model states and parameters. The state and parameter results are only slightly poorer than those obtained when the cable depths were perfectly known. Finally, in-situ temperature data collected from the SMAP MOISST site (refer to A.1) are used to test the proposed approach. Results show good agreement between the observed and estimated soil moisture, hydraulic properties, thermal properties and observation depths at all locations. The proposed method resulted in soil moisture estimates in the top 10 cm with RMSE values typically < 0.04 m^3/m^3.

5.2. METHOD AND MATERIALS

5.2.1. SOIL WATER, HEAT AND VAPOR TRANSFER UNDER VEGETATED AREA

In this chapter, the state of the art vadose zone model, Hydrus-1D is coupled with a canopy energy balance scheme originally presented by Oleson et al. [95]. The surface energy balance equations for the vegetation and ground are as follows:

$$\overrightarrow{S}_v - \overrightarrow{L}_v = H_v + \lambda E_v \tag{5.1}$$

$$\overrightarrow{S}_g - \overrightarrow{L}_g = H_g + \lambda E_g + G \tag{5.2}$$

where $\overrightarrow{S}_v$ is the solar radiation absorbed by the vegetation canopy, $\overrightarrow{L}_v$ is the net longwave radiation for the vegetation canopy (positive upwards from the canopy), and $\overrightarrow{S}_g$ and $\overrightarrow{L}_g$ are defined similarly such that $\overrightarrow{S}_g - \overrightarrow{L}_g$ is the net radiation at the ground surface. The radiation terms were estimated using the radiative transfer model proposed by [96]. H_v and λE_v are the sensible and latent heat fluxes from the canopy, H_g and λE_g are the sensible and latent heat fluxes from the ground, $\lambda \approx 2.501 \times 10^6$Jkg^{-1}, and G is the ground heat flux [95]. All the energy fluxes have a unit of Wm^{-2}. The water vapor

flux from canopy (i.e. E_v, in $\mathrm{kgs^{-1}m^{-2}}$) includes evaporation of intercepted water and transpiration.

The sensible heat fluxes are calculated as follows:

$$H_v = -\rho_{atm} C_p \frac{(T_s - T_v)}{r_b} (L_{leaf} + L_{stem}) \tag{5.3}$$

$$H_g = -\rho_{atm} C_p \frac{\left(T_s - T_g\right)}{r'_{ah}} \tag{5.4}$$

where ρ_{atm} is the density of air ($\mathrm{kgm^{-3}}$), C_p is the specific heat capacity of air at constant pressure ($\mathrm{Jkg^{-1}K^{-1}}$), T_s, T_v and T_g are the temperatures of canopy air (K), vegetation and ground surface respectively, L_{leaf} and L_{stem} are the exposed leaf and stem area indices, r_b is the leaf boundary layer resistance ($\mathrm{ms^{-1}}$), and r'_{ah} is the aerodynamic resistance to the sensible heat transfer between the ground surface and canopy air. Assuming the air within the vegetation canopy does not store heat, it gives:

$$H = H_v + H_g \tag{5.5}$$

$$H = -\rho_{atm} C_p \frac{(\theta_{atm} - T_s)}{r_{ah}} \tag{5.6}$$

where θ_{atm} is the potential air temperature (K), and r_{ah} is the aerodynamic resistance to the sensible heat transfer from the atmosphere to the vegetated land surface. Combining Eq.(5.3) to Eq.(5.6), the temperature of the canopy air (T_s) can be estimated as:

$$T_s = \frac{c_a^h \theta_{atm} + c_v^h T_v + c_g^h T_g}{c_v^h + c_a^h + c_g^h} \tag{5.7}$$

where $c_a^h = 1/r_{ah}$, $c_v^h = \left(L_{leaf} + L_{stem}\right)/r_b$ and $c_g^h = 1/r'_{ah}$. Water vapor fluxes from the vegetation and soil surface are calculated using:

$$E_v = -\rho_{atm} \frac{\left(q_s - q_{sat}^{T_v}\right)}{r_{total}} \tag{5.8}$$

$$E_g = -\rho_{atm} \frac{(q_s - q_g)}{r_{gw}} \tag{5.9}$$

where q_s is the specific humidity of the canopy air, q_g is the specific humidity at the ground surface, $q_{sat}^{T_v}$ is the saturation water vapor specific humidity at the vegetation temperature ($\mathrm{kgkg^{-1}}$), r_{total} and r_{gw} are the resistances of the vapor transfer from the canopy and the ground surface to the canopy air, in $\mathrm{ms^{-1}}$. Similar to the sensible heat flux, the water vapor flux is given by:

$$E = -\rho_{atm} \frac{(q_{atm} - q_s)}{r_{aw}} \tag{5.10}$$

$$E = E_g + E_v \tag{5.11}$$

where q_{atm} is the specific humidity of the atmosphere, and r_{aw} is the aerodynamic resistance to vapor transfer in ms^{-1}. Combining Eq.(5.8) to Eq.(5.11), the specific humidity of the canopy air q_s is solved as:

$$q_s = \frac{c_a^w q_{atm} + c_v^w q_{sat}^{T_v} + c_g^w q_g}{c_a^w + c_v^w + c_g^w} \tag{5.12}$$

where $c_a^w = 1/r_{aw}$, $c_g^w = 1/r_{gw}$, and $c_v^w = 1/r_{total}$. All the resistance and conductance terms were estimated using measured wind speed, relative humidity, air temperatures and vegetation properties as outlined by [97]. The initial vegetation parameters (e.g. vegetation optical and structural properties) were guessed to be the default values of "C3 grass" as prescribed in [97].

To solve the energy fluxes (i.e. H_v, H_g, λE_v, λE_g and G), iterations between the surface energy balance and the soil water, heat and vapor transport equation were required. The iteration procedure will provide the estimated T_v, resistances and the energy fluxes that satisfy the energy balance and the water balance at each time step. The details of the numerical implementation of the algorithm are provided by Oleson et al. [95].

5

5.2.2. ONLINE ESTIMATION OF OBSERVATION DEPTH USING THE PBS

In previous chapters, it was assumed that the depths of the temperature measurements were perfectly known. While this is a reasonable assumption for carefully installed point sensors, cable depths in DTS installations can vary considerably due to ground roughness, stones etc..

Here, a version of the PBS algorithm (denoted PBS-D) includes the observation depths as free parameters. The observation depths are randomly sampled from a prior distribution for each particle.

The observation depths are updated jointly with other parameters. Once the observation depths are updated, the measurement operator (h in Eq.(3.1)) will be modified to ensure that the simulated observations are at depths consistent with the estimated cable depths.

5.2.3. DATA ASSIMILATION EXPERIMENTS

First, we quantified the improvement due to the inclusion of the canopy energy balance scheme in the forward model using synthetic tests. Two open loop runs were performed, i.e. the particles were run in parallel without any data assimilation. The open loop for vegetated areas (OL-V) uses Hydrus-V, the version of Hydrus-1D that includes the new canopy energy balance scheme. The open loop for bare soil (OL-B) uses the original Hydrus-1D model. Results from both open loop runs will be compared to a truth generated using Hydrus-V with perturbed parameters and model forcing. Similarly, the PBS scheme will be implemented using both versions of the forward model. PBS-V and PBS-B denote the cases where Hydrus-V and Hydrus-1D are used. Both PBS schemes assimilate synthetic soil temperature observations generated by adding white noise with a standard deviation of 0.5 ^{o}C to the synthetic true temperatures at 5 and 10 cm. In both PBS schemes the soil temperature and moisture profiles, soil hydraulic properties (θ_r, θ_s, α, n, K_s) and soil thermal property (λ_{sat}) are jointly updated. A priori soil hydraulic

properties were generated using ROSETTA [65] with randomly sampled soil texture and bulk density as described in Table 5.1. The randomly sampled soil texture and bulk density were also used to provide the initial guesses of the λ_{sat} [72] for each particle. In PBS-V, LAI is also estimated. To test the robustness of the approach, this experiment will be repeated 10 times, where each of the 10 synthetic truths is generated using randomly sampled parameters covering a wide range of soil and vegetation properties. Hence, this multiple truths tests will provide more comprehensive comparisons of the performances of the PBS-V and PBS-B.

Table 5.1: Generation of perturbed inputs (soil and vegetation property and forcing) for each particle.

Variable	Error Distribution	Mean	Std.	Bound
Sand (%)	Uniform	-	-	15, 75
Silt (%)	Uniform	-	-	0, 100 - Sand
$\rho_b(g/cm^3)$	Uniform	-	-	1.1, 1.7
Air temperature (^{o}C)	Gaussian, Additive	0	+ 0.5	-, -
Precipitation (mm)	Gaussian, Multiplicative	1	× 0.2	-, -
Radiation (W/m^2)	Gaussian, Multiplicative	1	× 0.075	-, 1350
Relative humidity (%)	Gaussian, Multiplicative	1	× 0.05	-, 100
Wind speed (Km/h)	Gaussian, Multiplicative	1	× 0.2	-, -
Vegetation parameters	Gaussian, Multiplicative	1	× 0.2	-, -

Next, we will consider the case where the observation depths are not perfectly known. In DTS installations, cable depths can have an uncertainty of a few centimeters [38]. Hence, it is more realistic to assume that the true cable depth is unknown which means that the a priori mean cable depth may be biased with respect to the truth. In other words, we are interested in whether the PBS can draw the estimated cable depths closer to the truth, even when the initial guesses are biased. In our synthetic experiment, the "true" observation depths are known to be 5 and 10 cm. However, to simulate a scenario where the actual depth is uncertain, the mean of the initial guess for the observation depths was drawn from a uniform distribution between 3 and 7 cm, and 8 and 12 cm respectively. The standard deviation of the initial guess is assumed to be 1 cm. This extension of the PBS-V, in which the observation depths are also estimated, will be denoted "PBS-D". In synthetic tests, using the same 10 truths as before, the PBS-D algorithm will be benchmarked against the PBS-V. Finally, the PBS-D algorithm will be tested using real world data from four observed soil temperature and moisture profiles at the SMAP MOISST site.

5.3. RESULTS AND DISCUSSION

5.3.1. THE CANOPY ENERGY BALANCE SCHEME

Figure 5.1 shows an illustrative case that compares the estimated soil moisture using the OL-V, PBS-V and PBS-B. From 2.5 to 20 cm deep, the estimated soil moisture from the PBS-B is (dry) biased compared to the truth. In the PBS-V estimates, the large errors in initial conditions persist for a few days, but the PBS-V tracks the truth accurately from July 1st to the end of the simulation period. The soil moisture at 50 cm from the PBS-V

is significantly wetter than the truth, and is worse than both the OL-V and PBS-B. This is because only surface soil temperatures (5 and 10 cm) were assimilated, and these temperatures contain little information on soil moisture at 50 cm [98]. Furthermore, as the soil moisture at depth has little response to meteorological forcing during this time period (Figure 5.1 e), the soil moisture at depth is not correlated with shallow soil moisture. Under these conditions, the primary impact of assimilating surface soil temperatures on the soil moisture at depths is through model physics [98], i.e. by adding or extracting water at surface layers the PBS will draw the deeper layers towards a wetter or drier condition. In this specific case, the true soil moisture is wetter than the OL-V at the surface and drier than the OL-V at 50 cm. The PBS-V corrects the estimates from 2.5 to 20 cm towards a wetter condition, which leads to a higher (incorrect) soil moisture at 50 cm. The bias at depth (e.g. Figure 5.1 e) is largely an uncorrected bias in the prior initial condition. In Figure 5.1 a to e, the OL-V, PBS-B and PBS-V all start with the same prior initial condition, i.e. initial soil moisture and temperature are assumed to be homogeneous with depth, and the values randomly sampled from U[0.15 m^3/m^3, 0.4 m^3/m^3] and U[20 ^oC, 40 ^oC] for soil moisture and temperature respectively. It is clear from Figure 5.1 a that assimilation of temperature observations, and the impact of precipitation in "re-initializing" surface soil moisture are both effective in removing this initial bias at the surface. The degree to which this occurs decreases with depth (see Figures 5.1 b to e). The root zone soil moisture has little correlation to the surface soil moisture, and little response to meteorological forcing. Hence, as shown by Dong et al. [98] the error in the prior guessed root zone soil moisture is difficult to correct using the soil temperatures from the shallow subsurface. As shown in Figure 5.1 d, the bias persists for approximately 4 to 5 weeks at 20 cm, until the soil moisture is gradually drawn to the truth at the end of the simulation. At 50cm, it persists beyond the simulation period. The estimated soil moisture, particularly at depth, could be improved by employing bias correction methods analogous to those employed by Ryu et al. [99], De Lannoy et al. [75], Dee and Da Silva [73], Monsivais-Huertero et al. [100].

The soil hydraulic properties estimated using the PBS-B and PBS-V in this illustrative case are shown in Figure 5.2. The θ_r, and α estimates in the PBS-B converge to the truth during the simulation period. However, the PBS-B draws the estimated n and K_s towards a wrong value, which is significantly worse than the prior guess. As a result, the estimated soil water retention curve is significantly biased compared to the truth (Figure 5.2 h). The estimated soil hydraulic parameters from the PBS-V all converge to the true values after a few updates. The estimates are nearly constant during the entire simulation period, which means the estimates are accurate and robust. Because the parameters are correct, the soil water retention curve from the PBS-V fits the truth accurately (Figure 5.2 h). In this specific case, the true LAI is quite close to the mean of the prior guess. Hence, little improvement is shown in the mean of the PBS-V estimates compared to the prior guess. For the thermal properties, the estimated λ_{sat} is shown in Figure 5.2 f. Similar to the soil hydraulic properties, the PBS-B provides an even worse estimate than the prior guess. The PBS-V quickly draws the initial guess closer to the truth, which results in an improved soil thermal conductivity curve (Figure 5.2 i). Consistent with a previous study [101], the soil thermal conductivity curve is more difficult to estimate with this approach than the water retention curve. This is because the soil hydraulic properties,

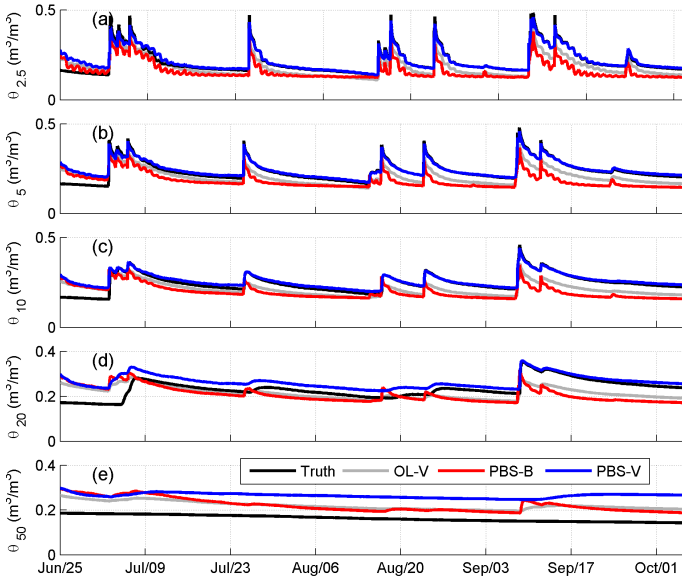

Figure 5.1: Comparison of soil moisture estimated using different approaches at 5 depths.

through their impact on variations in soil moisture, are the dominant control on soil temperature dynamics.

Box plots comparing the open loop and PBS soil moisture estimates for the 10 synthetic truths are given in Figure 5.3. The median RMSE of soil moisture estimates using OL-V constantly is smaller than that estimated using OL-B at all depths. In general, the differences between OL-B and OL-V increases with depth. This is because the OL-B ignores the root water uptake processes in the root zone. The increased interquartile range (IQR) of the RMSE of the PBS-B estimated soil moisture indicates that the PBS-B can be even worse than the OL-B for some cases.

When the canopy energy balance scheme is included in the forward model (PBS-V), the soil moisture can be robustly estimated across the entire profile, with RMSE values often < 0.02 m^3m^{-3}. It is shown that the performance of the PBS-B significantly depends on the temporal mean of the truth. Figure 5.3 f shows the difference between the RMSE of soil moisture estimated using PBS-B and PBS-V at 5 cm as a function of true soil moisture. In general, when the temporal mean of the true soil moisture is low, the performance of the PBS-B and PBS-V are comparable. However, when the temporal mean of the truth is high, using PBS-B yields significantly degraded soil moisture estimates compared to the PBS-V.

This can be explained by the fact that soil heat and water transfer processes are more tightly coupled under dry soils. Soil heat transfer strongly depends on soil moisture when soil moisture is low [38]. Hence, heat transfer processes between two depths (i.e. 5 cm and 10 cm) contain more information about soil moisture when the soil is dry. Furthermore, the energy lost through evaporative cooling is primarily controlled by the availability of soil moisture under dry soils. This means the cooling/heating rate of the

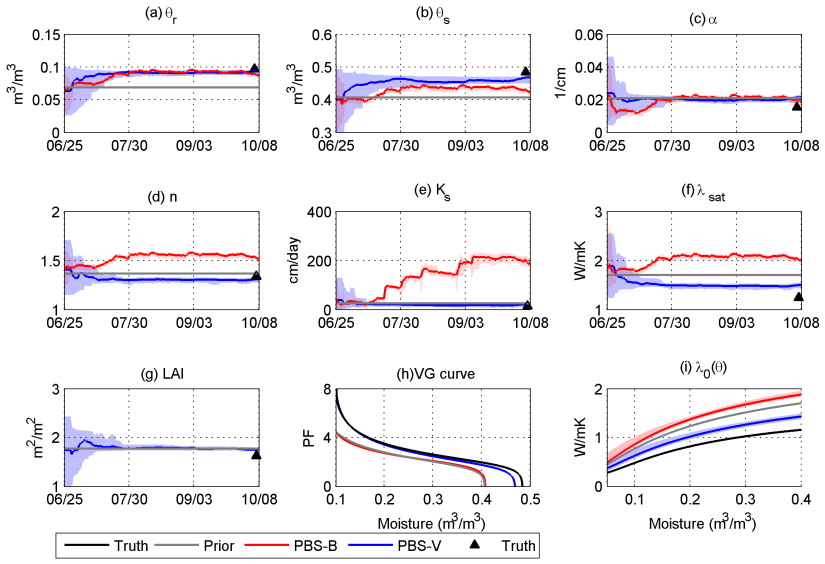

Figure 5.2: The convergence of the estimated soil hydraulic properties (a to e), λ_{sat} (f), LAI (g) and the soil water retention curve (h) and the thermal conductivity curve (i) using the estimated soil hydraulic and thermal parameters

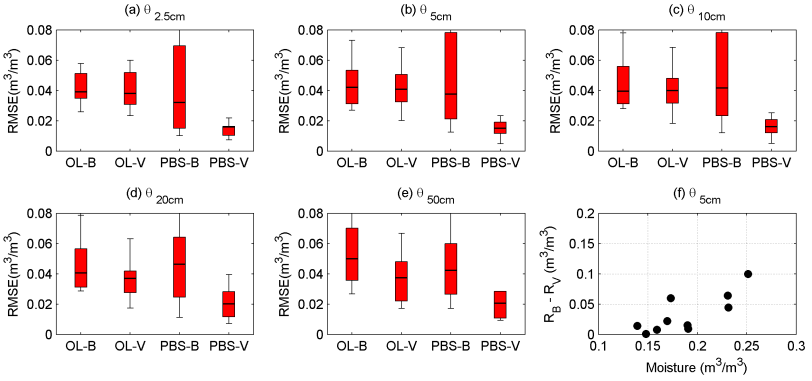

Figure 5.3: The RMSE of soil moisture estimates using 4 different approaches at 5 depths (a to e), and the differences of the RMSE of the PBS-B and PBS-V estimated 5 cm soil moisture as a function of the temporal mean of the true soil moisture at 5 cm. Each boxplot contains results derived from synthetic tests using 10 randomly selected truths.

soil is a strong function of soil moisture. When the truth is dry, soil moisture can be relatively well estimated using soil temperatures, even though the forward model has large structural errors (e.g. PBS-B).

The estimated model parameters are compared to the synthetic “true" values in Figure 5.4 and Table 5.2. As expected, the estimated parameters from the PBS-V are closer to the 1:1 line than those estimated using PBS-B. Several parameters estimated using the PBS-B are even negatively correlated to the truth (Table 5.2). The large errors in estimated θ_{33} (Figure 5.4 g) from the PBS-B imply a poor agreement between the estimated and true soil water retention curves.

In the PBS-V, α and K_s are relatively poorly estimated. However, the other parameters suffice to ensure that the θ_{33} is well-estimated (Figure 5.4 g and Table 5.2), and hence that the estimated and true water retention curves are similar.

LAI determines canopy extinction of the net radiation, which influences the amplitude of the simulated soil temperatures. Hence, LAI can be accurately inferred from the soil temperature observations using PBS-V (Figure 5.4 h).

5.3.2. STATE - PARAMETER ESTIMATION WITH UNKNOWN OBSERVATION DEPTHS

The illustrative case is revisited in Figure 5.5 (a) and (b) which show the estimated sensor (observation) depths using PBS-D. In this case, the prior guess for the depths are 4.3 and 10.7 cm. For the top sensor, the PBS-D converges to the true sensor depth after a few updates. Though the estimated depth lower sensor is closer to the true value than the prior, it is still 0.5 cm from of the true value by the final time step. Once the estimated depths converge from the highly uncertain initial guess, the estimates of the depths remain constant. The reduction of the uncertainties in the sensor depths is also quite noticeable.

Figure 5.5 c shows box plots of the errors in estimated cable depths for 10 different truths. The median error of the estimated cable depth is approximately 0.3 cm for the top sensor depth, and 0.5 cm for the lower sensor depth. The amplitude of the temperature wave induced by the diurnal cycle of solar radiation is damped with increasing depth [38, 77]. As a result, the error in the estimated lower sensor depth may reach up to 1.2 cm, which is significantly larger than that of the top cable. It is more challenging to correct the initial errors of the lower cable (Figure 5.5 d). This is because the solar signals decrease exponentially with depth. Hence, little information is available to estimate the cable depth of the lower cable. As for the upper cable, the depth can be well estimated regardless of whether the prior guess is positively or negatively biased with respect to the truth.

The parameters estimated using PBS-D can be compared to those from PBS-V in Figure 5.4 and Table 5.2. In general, the PBS-D provides very similar results to the PBS-V. Consequently, the accuracy of the soil moisture estimates using the PBS-V and the PBS-D are also quite similar (Table 5.3). At 2.5 to 20 cm, the maximum difference in RMSE is just 0.03 m^3/m^3. The similar performance of PBS-D and PBS-V is remarkable given that the sensor depths are perfectly known in PBS-V, and unknown but estimated in PBS-D. This means that in DTS applications, the PBS-D can be used to estimate soil moisture even when there is significant uncertainty in the cable depths.

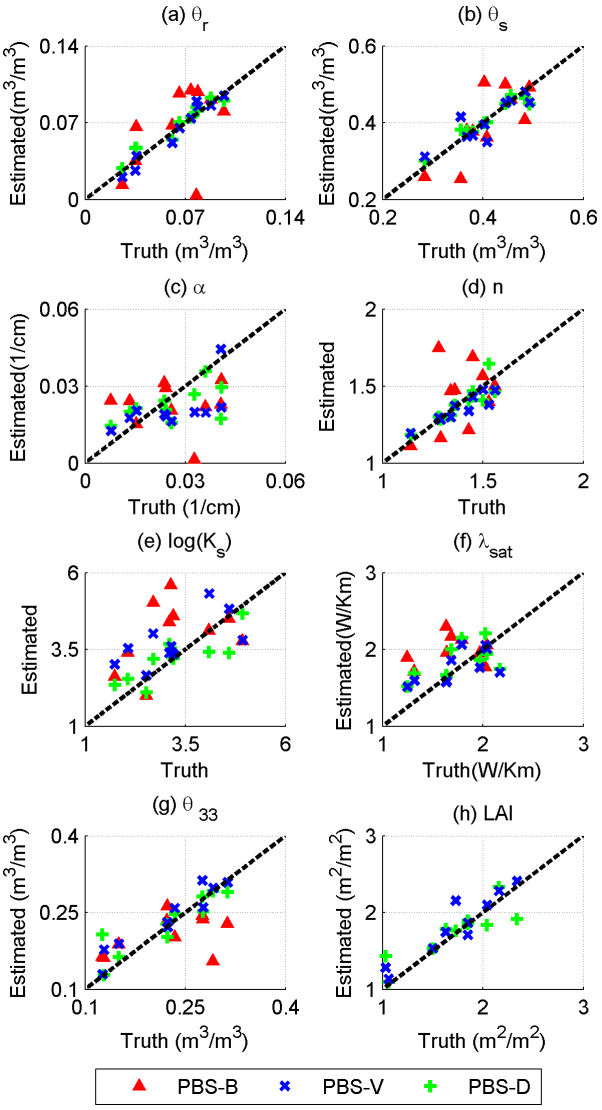

Figure 5.4: Comparison of the estimated soil hydraulic properties (a to e), λ_{sat} (f), θ_{33} (g), and LAI (h) using 3 different PBS strategies.

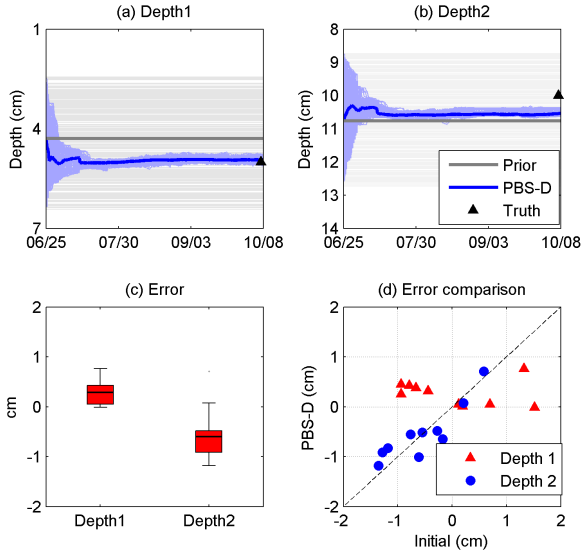

Figure 5.5: The convergence of the two estimated observation depths in one illustrative example (a and b), the error of the estimated observation depths tested by 10 randomly selected truths (c), and a comparison of the PBS-D and the prior guessed cable depths error (d) . In a and b, each thin line represents one particle. The true observation depth is shown at the final estimation step.

Table 5.2: Correlation coefficient of the estimated and the true parameters using three different data assimilation approaches. The values are calculated using the estimates from 10 randomly selected truths

Par.	**PBS-B**	**PBS-V**	**PBS-D**
θ_r	0.51	0.97	0.97
θ_s	0.77	0.86	0.96
α	-0.01	0.61	0.55
n	0.39	0.91	0.88
K_s	-0.09	0.42	0.82
λ_{sat}	-0.38	0.63	0.57
θ_{33}	0.49	0.95	0.90
LAI	-	0.87	0.94

Table 5.3: The RMSE of soil moisture estimates (m^3/m^3) at 5 depths using the PBS-V and the PBS-D. The RMSEs are averaged from 10 randomly selected truths

Depth	**2.5 cm**	**5 cm**	**10 cm**	**20 cm**	**50 cm**
PBS-V	0.013	0.015	0.016	0.018	0.025
PBS-D	0.016	0.017	0.018	0.021	0.031

Table 5.4: The absolute error (cm) of the estimated sensor depths at the four sites

	Prior	Site A	Site B	Site C	Site D
Depth 1	2.117	0.153	0.086	0.659	0.413
Depth 2	1.954	0.698	0.027	1.578	0.255

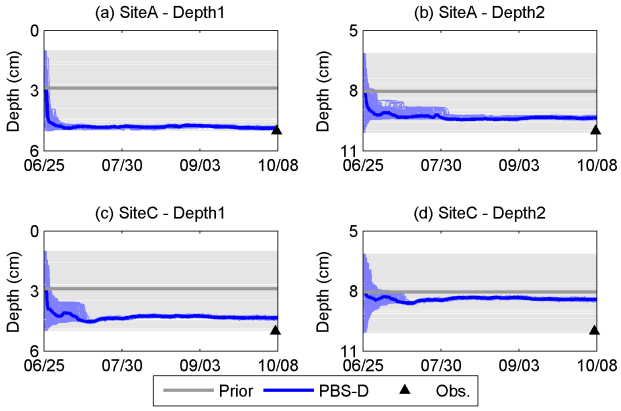

Figure 5.6: The convergence of the estimated sensor depths at Site A (a and b) and Site C (c and d). The observed sensor depth are shown at the final time step of the simulation.

5.3.3. REAL DATA APPLICATION

The PBS-D algorithm was applied to four real, observed soil temperature profiles at the SMAP MOISST site. First, the error of the estimated sensor depths for the four sites are shown in Table 5.4, and the convergence of the estimated sensor depths at site A and C are shown in Figure 5.6 . The initial/prior sensor depths at site A are approximately 2 cm biased for both the upper and the lower sensors (Figure 5.6 a and b). The PBS-D draws the estimates to the observed sensor depths, and the error of the final estimates are 0.15 and 0.69 cm for the upper and lower sensor, respectively. The uncertainties of the sensor depths are considerably reduced. The same prior sensor depths were assumed for site C, and again they are biased with respect to the measured values. It takes longer for the PBS-D estimate of the top sensor depth to converge. When it does, the estimate is closer to the truth but still has an error of 0.66 and 1.58 cm for the upper and lower sensor (Table 5.4). Similar to the synthetic tests, estimates for the lower sensor are less robust (Table 5.4). As discussed in the previous section, the signal from solar radiation is damped exponentially with depth. In addition, the time lag in the propagated soil temperature fluctuations also increases with depth. Hence, soil temperature contains less information of the observation depths for the lower sensor.

Figure 5.7 illustrates the convergence of the estimated model parameters using site A and C for illustration. The estimated θ_r, θ_s and α converged to similar values at the two sites. However, the PBS-D estimates show that site C has a significantly larger n and K_s. The sand content of site C is approximately 22% higher than that of site A on average (Figure A.2). Saturated soil hydraulic conductivity and the van Genuchten parameter n

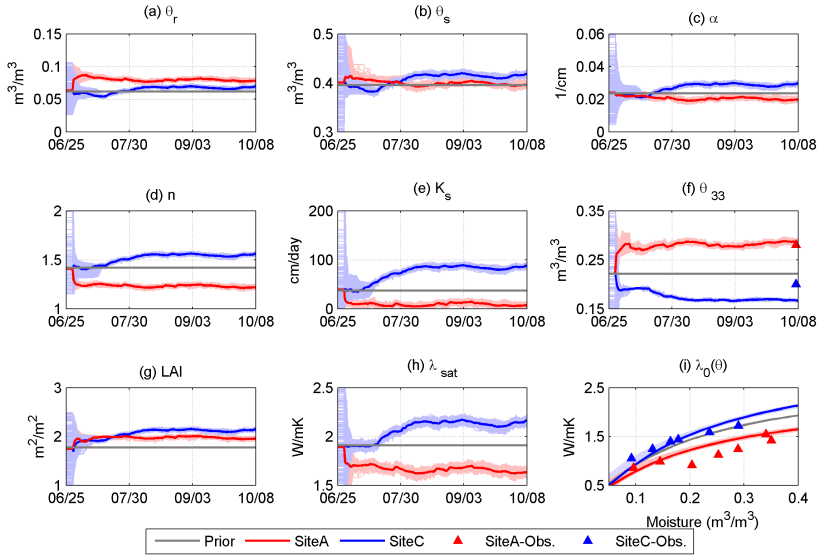

Figure 5.7: The convergence of the soil hydraulic properties at Site A and Site C (a to e), LAI (d), a comparison of the PBS-D estimated and the observed θ_{33} (f) and the soil thermal conductivity curve (i).

typically increases with sand content [102]. This suggests that the differences in estimated K_s and n between the two sites are reasonable. The estimated θ_{33} based on the estimated soil hydraulic properties at the two sites is shown in Figure 5.7 f. The PBS-D draws the estimated θ_{33} closer to the observed value at site A. At site C, the θ_{33} seems to be slightly overfitted, i.e. the value from PBS-D is further from the observed value than the prior guess. Another parameter that clearly shows the difference between the two sites is λ_{sat}. This leads to different estimated soil thermal conductivity curves at the two sites, both of which are closer to the observations (Figure 5.7 i) than the prior. The evolution of the estimated parameters is slightly more variable than in the synthetic tests (Figure 5.2). This may be due to increased model structural errors in the real world application. For example, the forward model assumes the soil property profile is uniform, which is not true in reality (Figure A.2).

In Figures 5.8 and 5.9, results are compared from the OL-V and PBS-D to show the benefit of assimilation in a real DTS application. The forward model is the same in both OL-V and PBS-D. The PBS-V and PBS-B are not included here as they require that the cable depths are perfectly known. The estimated soil moisture at site A is shown in Figure 5.8. Compared with the OL-V, the main benefit of assimilating soil temperatures is between the surface and 10 cm deep, i.e. the depth of the deepest soil temperature observation. The reduction in RMSE at shallow depths is significant (Figure 5.8 a and b and Table 5.5). By adjusting the shallow soil moisture to a wetter condition, the soil moisture estimates at 20 and 50 cm were also improved through the model.

At site C, the OL-V provides reasonable estimates of soil moisture at all depths, despite the large particle ranges (Figure 5.9). The PBS-D provides further improvement at

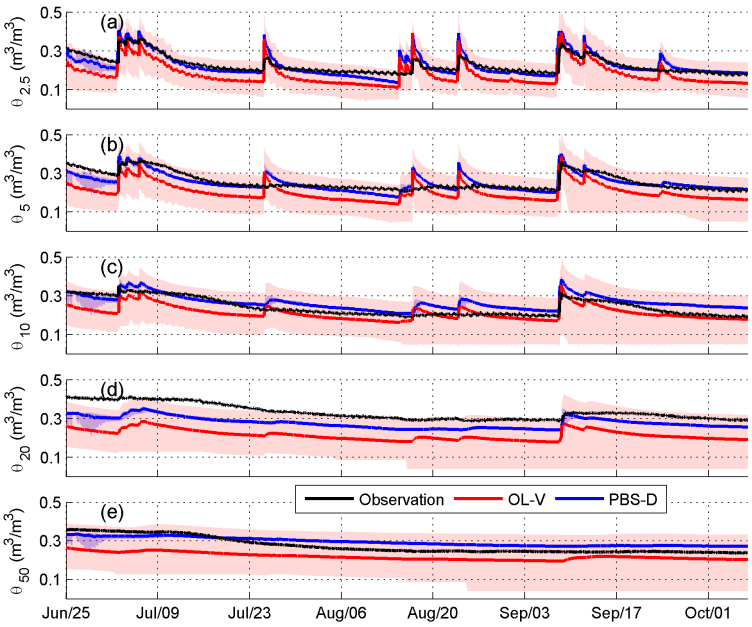

Figure 5.8: Comparison of estimated soil moisture using PBS-D and OL-V at 5 depths at site A. The shaded area represents the range of the particles.

depths from 5 to 40 cm. In general, the largest errors in PBS-D estimated soil moisture occur in the first week of the simulation (i.e. before July 9). This is because soil temperature and not soil moisture is assimilated, and so the PBS-D adjusts the soil moisture to compensate for errors in the initial guessed model states and parameters.

In general, the largest RMSE of the soil moisture estimates at depths above 10 cm is approximately 0.04 m^3/m^3 (Table 5.5). Improvements below 20 cm are primarily through model physics. Improving soil moisture estimates at shallow depths will eventually lead to improved root zone soil moisture estimates. However, this means errors in the initial deep soil moisture persist for a few months, since the correlation between the surface and deep soil moisture is relatively low and there are limited dynamics below 20 cm. Table 5.5 shows the RMSE in the estimated soil moisture at all depths, at sites A to D. From the surface to a depth of 10 cm, most RMSE values are around 0.02 to 0.04 m^3/m^3, which is comparable in magnitude to the measurement error associated with the “ground truth" soil moisture probe. Consistent with the synthetic tests, the best assimilation results are obtained in the top 10 cm, i.e. at and above the cable depths. At depths greater than 20 cm, the impact of solar radiation is reduced so there is limited variation in temperature and therefore little correlation with the temperatures observed and simulated closer to the surface. Assimilation therefore leads to a limited improvement at depth due to the persistence of bias from the initial condition. With the exception of Site C, the RMSE from PBS-V is lower than that from OL-V, confirming that assimilation generally leads to an improvement over the open loop. At Site C, the OL-V

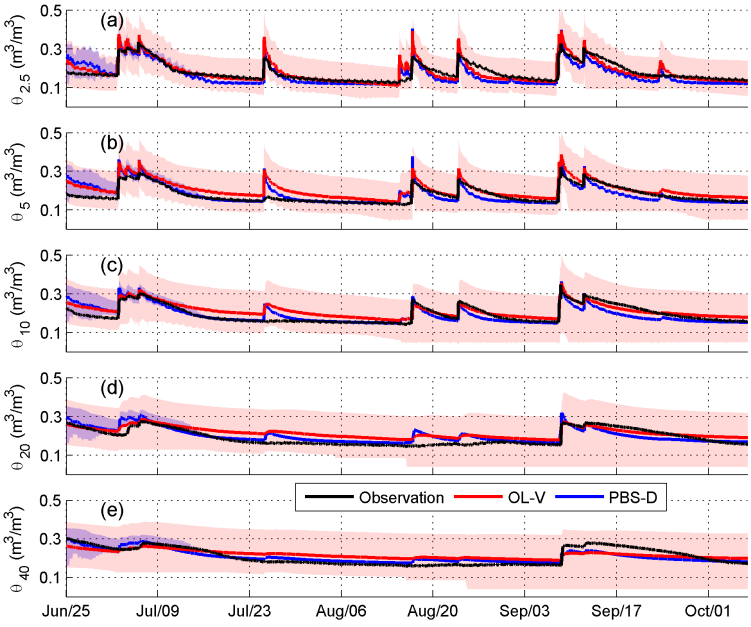

Figure 5.9: Similar to Figure 5.8, but for site C

yields RMSE values that are comparable to the measurement error of the soil moisture probe at Site C. Though assimilation appears to worsen the estimate, the magnitude of these errors means that the estimates have essentially the same accuracies. The value of including the vegetation scheme is clear from the reduction in RMSE when PBS-V is used instead of PBS-B. The reduction of RMSE is typically around 20 to 30%. The most important result, in the context of applying this approach to a real DTS application, is that the PBS-D performance is very similar to that of PBS-V. Recall that PBS-V and PBS-D are identical in terms of model physics. The only difference is that the cable depths are unknown in PBS-D and are included with the parameters to be estimated. Though the performance of the PBS-V is excellent, its value is limited by the need to know the cable depths. The fact that the PBS can estimate these depths and yield comparable performance when the cable depths are unknown removes the final barrier to applying the PBS approach to real DTS data. Finally, it is noteworthy that the relative performance of the PBS implementations is consistent with that observed in the synthetic experiments.

Figure 5.10 further demonstrates the potential of using soil temperature observations to detect the spatial variability of soil moisture and soil properties. The OL-V uses the same a priori model parameters and forcing data, and hence it provides the same estimates across all the four sites, i.e. no soil moisture or soil property difference can be detected by the OL-V. The temporal mean soil moisture is most accurately estimated by PBS-D at the depth of 2.5 cm. It correctly shows that site C is drier than the other three sites. Also, sites A, B and D have similar temporal mean soil moisture. The accuracy of the PBS-D estimates decreases with increasing depth, and is significantly biased at sites

Table 5.5: The RMSE of soil moisture estimates (m^3/m^3) at 5 depths using the OL-V and the PBS-D at 4 sites. The deepest measurement and estimated soil moisture depth being compared is 50 cm for Site A, B and D, and 40 cm for site C.

Sites	Appro.	2.5 cm	5 cm	10 cm	20 cm	40/50 cm
A	OL-V	0.053	0.060	0.039	0.074	0.037
	PBS-V	0.034	0.040	0.040	0.057	0.030
	PBS-D	0.027	0.026	0.034	0.057	0.030
	PBS-B	0.046	0.051	0.058	0.106	0.092
B	OL-V	0.055	0.026	0.056	0.115	0.159
	PBS-V	0.041	0.035	0.036	0.095	0.127
	PBS-D	0.041	0.030	0.041	0.114	0.140
	PBS-B	0.050	0.050	0.046	0.093	0.138
C	OL-V	0.023	0.031	0.027	0.032	0.035
	PBS-V	0.028	0.034	0.029	0.040	0.038
	PBS-D	0.030	0.027	0.025	0.025	0.023
	PBS-B	0.041	0.024	0.024	0.026	0.027
D	OL-V	0.056	0.058	0.066	0.092	0.153
	PBS-V	0.026	0.029	0.030	0.044	0.093
	PBS-D	0.032	0.031	0.034	0.061	0.111
	PBS-B	0.047	0.046	0.045	0.081	0.184

A, B and D at 20 cm. The variability of the temporal mean soil moisture across the four sites is still well captured. Similar to the soil moisture estimates, the PBS-D correctly shows the variability of the θ_{33} across the 4 sites using soil temperature dynamics. The estimated θ_{33} at site B to D are slightly biased (0.037 m^3/m^3 at sites B and D, and 0.044 m^3/m^3 for site C). However, given the accuracy of the field measured θ_{33}, the PBS-D estimates are quite acceptable. PBS-V generally provides similar results to PBS-D, which is consistent with the synthetic test. The PBS-B can also reasonably estimated the mean soil moisture and the θ_{33} across the four sites. This may demonstrate the feasibility and robustness of using soil temperatures from DTS for high resolution soil moisture mapping.

5.4. CONCLUSIONS

In this chapter, we solved the two key remaining barriers to the application of Passive DTS for large-scale high-resolution soil moisture monitoring. First, the inclusion of a new surface energy balance scheme in the Hydrus-1D model extends the applicability of the PBS data assimilation approach to vegetated areas. This is an essential development to enable the use of Passive DTS for soil hydrology research, and the validation of remote sensing observations under agricultural and natural vegetation. Second, an algorithm was proposed to estimate observation depths jointly with other model states and parameters. This is particularly relevant for DTS applications, since measuring the cable depths every meter along the DTS cable is logistically impractical. Requiring the cable depths to be measured everywhere would render the data assimilation scheme proposed

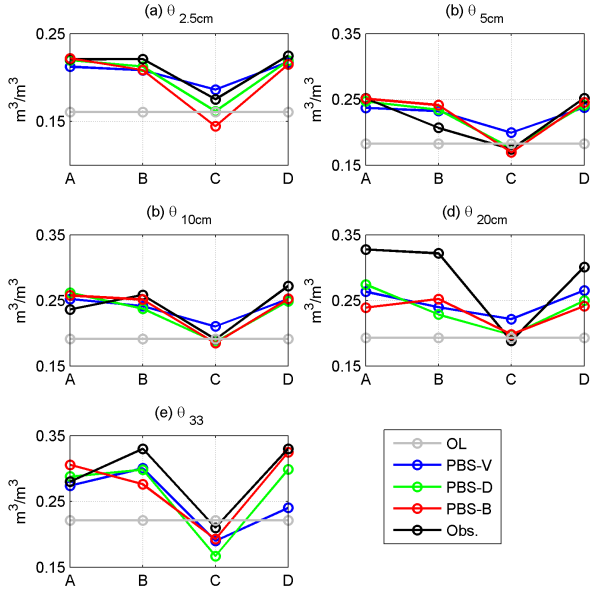

Figure 5.10: Comparison of the estimated and observed temporal mean soil moisture (a to d) and the θ_{33} (e) at the four sites.

by [98, 101, 103] useless in DTS applications. Therefore, estimating the cable depths is essential.

The new approach (PBS-D) was tested using both synthetic and real world data. The synthetic experiments were necessary because in real DTS installations the cable depth is difficult to measure and therefore often unknown. Synthetic tests, in which the cable depth is perfectly known, were used to demonstrate that the proposed approach worked and to illustrate the value of being able to estimate the observation depth. The “real" data were from four observation profiles from traditional sensors rather than DTS observations. These profiles had very distinct soil texture profiles. Though the prior guesses for sensor depths, soil moisture and properties were identical for all profiles, the PBS-D approach was able to reproduce observed differences in soil properties and soil moisture between the four sites. Hence, the results demonstrate that the PBS-D can be used to assimilate in-situ soil temperature observations to simultaneously estimate soil moisture dynamics along with soil hydraulic and thermal properties. It is worth nothing that the cable depths could be calibrated off-line if they are time-invariant parameters. This would reduce the degree of freedom of the estimation problem, potentially improving the estimates of the remaining states and parameters. However, the ability to estimate them is limited by the uncertainty and/or biases in the soil properties in the prior guess. This would require an iterative procedure to allow both to converge. Furthermore, cable depth is not necessarily time-invariant. The cables may shift in the soil due to soil compaction, animal or vehicle traffic or sedimentation processes. Estimating them in-line provides the flexibility to handle these temporal variations.

6

MAPPING HIGH RESOLUTION SOIL MOISTURE AND PROPERTIES USING DISTRIBUTED TEMPERATURE SENSING DATA

Based on: Dong Jianzhi, Susan C. Steele-Dunne Tyson E. Ochsner, Christine E. Hatch, John Selker, Scott Tyler, Michael H. Cosh, and Nick van de Giesen. Mapping high resolution soil moisture and properties using distributed temperature sensing data and an adaptive particle batch smoother, *Water Resources Research*, accepted.

6.1. INTRODUCTION

Data assimilation methods were proposed and investigated for estimating soil moisture and soil properties using soil temperatures in the previous chapters. All of the data assimilation developments to date have been tested and developed using synthetic data and observations from point in-situ temperature sensors. In this chapter, the Particle Batch Smoother will be applied to real Distributed Temperature Sensing data from the Soil Moisture Active Passive (SMAP) Marena Oklahoma MOISST site (refer to A.1).

As shown in the previous chapters, a tuning factor is necessary to avoid severe particle weight degeneracy and to provide accurate estimates, which is particularly true when model parameters are jointly estimated with model states. Due to the limited scale of the problem, it was possible to determine a temporally constant tuning factor by trial and error [101]. To apply the PBS to real DTS data, an objective, automatic procedure is needed to determine an appropriate tuning factor.

This chapter first proposes an adaptive Particle Batch Smoother (APBS), in which a tuning factor will be automatically determined to avoid severe weight degeneration. Since the truth and the uncertainties are known by design, synthetic tests were then used to test the accuracy and the robustness of this APBS algorithm. Finally, We applied this APBS algorithm to real DTS data, which yields high spatial soil moisture and soil property maps along this section of DTS cable. The estimated soil moisture and properties will be evaluated using data collected at a nearby site.

6.2. METHOD AND MATERIALS

6.2.1. ADAPTIVE PARTICLE BATCH SMOOTHER

The adaptive particle batch smoother is essentially the same as the PBS algorithm described above, except that the tuning factor for each batch window is determined by maximizing the reliability of the soil temperature estimates. The probabilistic metric of reliability is calculated using the Quantile - Quantile (Q-Q) plot, which indicates whether the estimated uncertainty (particle range or ensemble spread) is appropriate. For each batch window, the quantile of the predictive distribution is calculated at each time step at each observation depth within the batch window [58]:

$$z_{t_j,j} = \frac{1}{N}\sum_{i=1}^{N} k_i \tag{6.1}$$

where N is the number of the particles, $z_{t_j,j}$ is the quantile of the predictive distribution calculated at time t_j depth j, $k_i = 1$ when observed soil temperature is larger than the ith particle simulated soil temperature at time t_j, and $k_i = 0$, otherwise [84]. In the perfect case, the distribution of $z_{t_j,j}$ should follow the uniform distribution ($U[0,1]$). If $z_{t_j,j}$ are clustered at the middle range, it indicates that the uncertainty is overestimated. The uncertainty is underestimated when the $z_{t_j,j}$ are clustered around the tails. In the case where $z_{t_j,j}$ is constantly lower/higher than $U[0,1]$, it indicates the estimates are biased [85]. The differences between the $z_{t_j,j}$ and $U[0,1]$ are measured by reliability (α_r):

$$\alpha_r = 1 - \frac{2}{n_0 \times n_t}\sum_{j=1}^{n_0}\sum_{t_j=1}^{n_t}\left|z_{t_j,j} - U[0,1]\right| \tag{6.2}$$

where n_t denotes the total number of time steps within the batch window. For example, when hourly soil temperatures were observed at two depths and a window length of 12 h was used, n_0 was two and n_t was 12. Consequently, the reliability of this batch window will be estimated using 24 observations. Reliability (α_r) varies from 0 (zero reliability) to 1 (perfect reliability). It was demonstrated that the PBS can provide robust estimates, once this β value is within a reasonable range [101]. Hence, β is varied from 0 to 1 in increments of 0.05 and the optimal β is that which yields the largest reliability in the APBS. This adaptive approach may be particularly suitable for the cases when particle weight degeneracy is severe, or the prior estimates have large errors. Under such circumstances, the prior distribution may be less capable of encompassing the truth, which also means little model parameter and state information can be inferred from the soil temperatures. Hence, updating the particles using the original PBS (i.e. $\beta = 1$) will unavoidably lead to overconfident estimates and reduced α_r. Further, the parameter estimates will also be over-fitted to model or observation errors. While in the APBS method, a small tuning factor will be used to maximize the α_r. Using a small tuning factor will allow the particle range to grow to encompass the observations in the following steps. Further, since a small tuning factor is used, the model parameters are less likely to be overfitted.

6.2.2. DATA ASSIMILATION EXPERIMENTS

First, we will show the necessity of using the adaptive Particle Batch Smoother (APBS) to avoid severe particle weight degeneracy and to improve the estimates. A single illustrative synthetic test is used to compare the original Particle Batch Smoother (i.e. $\beta = 1.0$) to the APBS. In this synthetic test, the truth was generated using the forward model with perturbed forcing data (Table 5.1). Soil textures and soil bulk density drawn from the distributions described in Table 5.1 were used to generate the true parameters. To account for uncertainties in the cable depths in real DTS applications, the observation depths were randomly drawn from the uniform distributions U[3 cm, 7 cm] and U[8 cm, 12 cm] for the upper and lower cable, respectively. Synthetic DTS observations were generated by adding zero mean Gaussian distributed noise to the synthetic true soil temperatures. As a conservative assumption, the observation error used in this study is 0.5 ^{o}C. An open-loop run (OL), in which the particles are run in parallel without performing any data assimilation, was used to evaluate the improvement made by the different PBS algorithms. For the OL run, the model parameters were generated using randomly sampled soil properties as described in Table 5.1. In the PBS algorithms, the initial guessed cable depths for each particle were drawn from a Gaussian distribution with standard deviation of 1 cm, and mean of 5 cm for the upper cable, and 10 cm for the lower cable. The cable depths were jointly estimated with soil temperature and moisture profiles, soil hydraulic properties (θ_r, θ_s, α, n, K_s), soil thermal property (λ_{sat}) and LAI in both PBS schemes.

Next, we will investigate the robustness of the APBS algorithm using a multiple truth test. We repeated the experiment described above 15 times, where using 15 different truths, generated using randomly sampled model parameters. Chapter 4 and 5 demonstrated that setting $\beta = 0.25$ in the original PBS algorithm (note as PBS$_\beta$) yields significantly improved estimates. A comparison of the PBS$_\beta$ and the APBS may provide insight into the benefits of using an adaptive tuning factor in the PBS, rather than simply assum-

ing some arbitrary value. Hence, in this multiple truth test, three PBS schemes, i.e. PBS, PBS$_\beta$ and APBS, were compared.

Finally, the APBS algorithm was tested using the real DTS data from the SMAP MOISST. The APBS estimated soil moisture and soil properties were evaluated using the data collected from an observation site approximately 70 m away. First, we examined the variability in the estimated cable depths. Then, we validated the estimated thermal conductivity and field capacity against observed values, and compared estimated soil moisture along the cable to that observed with in-situ Hydra probes at the nearest enclosure (Site B). Finally, we used the estimated soil moisture to quantify spatial variability in moisture along the cable and consider possible sources of this variability.

6.3. RESULTS AND DISCUSSION

6.3.1. AN ILLUSTRATIVE CASE

Figure 6.1 illustrates how the value of β influences the Q-Q plot, and hence how the Q-Q plot can be used to determine β adaptively. In the top panel, $\beta = 0.0$, which means that the state and parameters were not updated. This is equivalent to an open loop simulation. Figure 6.1 a and b show that the posterior estimate is biased. As a result, Figure 6.1 c shows that the predicted quantile is constantly above the cumulative uniform distribution. At the other extreme, Figures 6.1 e to h show what happens when $\beta = 1.0$. Because the observation error was so small with respect to the a priori estimate, only a few particles were preserved after resampling (Figure 6.1 e and f). Consequently, the Q-Q plot shows that the predicted quantiles are clustered in the middle range, which indicates that the estimated posterior is overconfident (Figure 6.1 g). This also indicates that the particle weights are significantly degenerated. As a result, the posterior of the model parameter, e.g. θ_s, is concentrated on a few particles, and therefore cannot encompass the true value (Figure 6.1 h). Figure 6.1 i and j show the impact of selecting a tuning factor that maximizes the reliability of the soil temperature estimates. More particles are accepted in the posterior distribution which also benefits the model parameter estimation. Though the mode of the posterior is biased, the range of the posterior is still wide enough to encompass the truth (Figure 6.1 l). In this case, the Q-Q plot shows that predicted quantiles are almost perfectly aligned with those from a uniform distribution (Figure 6.1 k).

Figure 6.2 compares the convergence of the model parameter estimates when β is set to 1.0 (PBS) to the case when β is estimated using the APBS. In this specific illustrative case, the PBS fails to converge to the true parameters (Figure 6.2 a to g). The evolution of the parameter estimates is also quite unstable, even at the end of the experiment, which suggests that the PBS overfits the model parameter to the model error. The APBS provides significantly improved soil hydraulic properties, compared to the PBS approach (Figure 6.2 a to e). This yields a nearly perfectly estimated soil water retention curve (Figure 6.2 h). Among the five soil hydraulic properties, the α and K_s parameters provide the poorest agreement with the true values. This is because these two parameters are less sensitive to the soil temperature evolution than the other parameters [101]. The prior guessed thermal conductivity, λ_{sat}, is very close to the truth in this specific case. Both PBS and APBS provide degraded estimates of λ_{sat} and that thermal conductivity curve

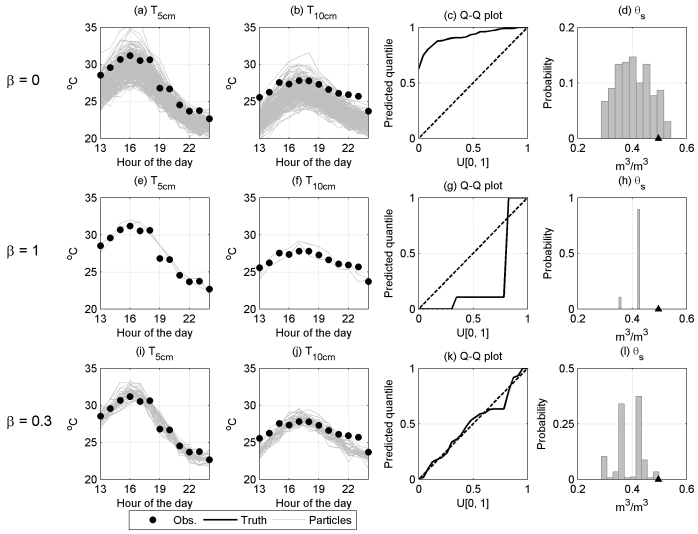

Figure 6.1: Soil temperature estimates at 5 (first column) and 10 cm (second column), the Q-Q plot (third column), and the posterior distribution of the estimated model parameter using θ_s as an example (fourth column) in one batch window. The top row is the case when $\beta = 0$, which is equivalent to open loop. The second row shows the estimates using $\beta = 1.0$, i.e. no tuning is used. The last row uses an automatically determined β, which is 0.3 in this specific window.

(Figure 6.2 f and i). Because the soil moisture drives dynamics in soil thermal properties, the soil hydraulic properties are easier to estimate than the soil thermal properties [101]. The prior guessed LAI is already very close to the truth (Figure 6.2 g). Both the PBS and the APBS have converged to the true LAI at the final time step. However, it is noticeable that the estimate from the PBS varies considerably during the simulation period.

The estimated upper and lower cable depths using PBS and APBS for this synthetic, illustrative case are shown in Figure 6.3. The prior guess for the cable depth of the upper cable was biased by approximately 1.5 cm. Both the PBS and the APBS draw the prior estimate closer to the truth, and result in an error of less than 0.5 cm for both methods. The estimated lower cable depth is less accurate compared to the upper cable in this specific case (Figure 6.3 b). The solar signal is damped exponentially with depth, and hence it is more difficult to estimate the depth of the lower cable as shown in the previous chapter.

Figure 6.4 shows the soil moisture estimates using the two PBS schemes. Because the PBS yielded little improvement in the soil hydraulic properties compared with the prior (Figure 6.2), the estimated soil moisture from the PBS is very similar to the OL estimates at all depths. On the other hand, the APBS benefits from the near perfect soil property estimates from approximately May 10 onward (Figure 6.2). Hence, it provides very accurate surface soil moisture estimates. Soil moisture at depth has little response to the climatic forcing at the upper boundary during the simulation period. Furthermore, only 5 and 10 cm soil temperatures were assimilated in this study, which has limited corre-

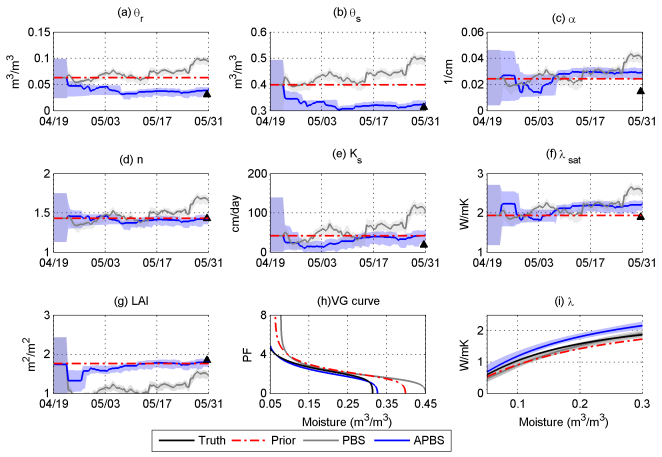

Figure 6.2: The estimated soil hydraulic and thermal properties (a to f) and LAI (g). The soil water retention curve and the soil thermal conductivity curve calculated using the estimated soil properties are shown in (h) and (i). The true model parameters are shown at the final time step. The shaded area represents the range of the particles.

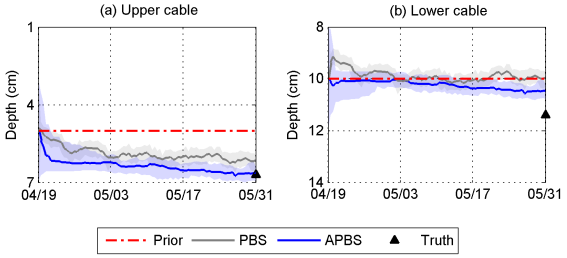

Figure 6.3: Estimated upper (a) and lower (b) cable depths using the PBS and APBS. The true cable depths is shown at the final time step.

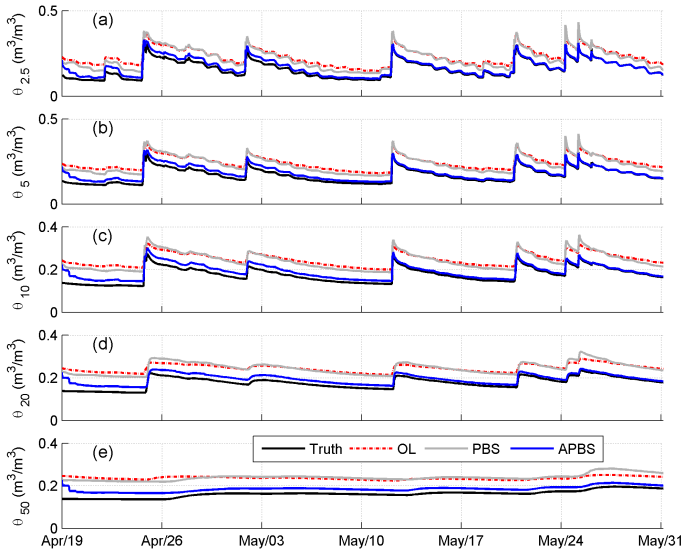

Figure 6.4: Comparison of the soil moisture estimated using PBS and APBS at 5 depths.

lation with the soil moisture in the deeper profiles. Thus, the APBS algorithm mainly benefits the root zone soil moisture estimates through model physics, i.e. increasing or decreasing soil water content at surface will eventually lead to a wetter or drier root zone soil moisture. Hence, the impact of the errors in the initial condition increases with depth (e.g. Figure 6.4 e). This may indicate longer term simulations are necessary for the PBS algorithms to yield to improvements at depth. The estimated soil moisture at depth might be improved by employing bias correction methods analogous to those employed in the previous studies [e.g. 73, 75, 99, 100].

6.3.2. A MULTIPLE TRUTH COMPARISON OF DIFFERENT PBS SCHEMES

The previous section showed results from a single synthetic case. Figures 6.5 and 6.6, and Table 6.1 summarize the results when this experiment was repeated 15 times for different synthetic “truths". Figure 6.5 shows a comparison of the true and the estimated cable depths when β is set to 1.0 (PBS), a fixed value of 0.25 (PBS$_\beta$) or determined using the adaptive PBS (APBS). The PBS$_\beta$ and APBS estimated cable depths and separation distances with similar accuracies, which were better than that of PBS approach. Similar to the illustrative case, the depth of the lower cable proved difficult to estimate.

The estimated model parameters and the correlations between the estimates and the truths are shown in Figure 6.6 and Table 6.1. In general, the PBS estimated model parameters are the least correlated with the true parameters, compared with the PBS$_\beta$ and APBS. This is consistent with a previous study [101], which demonstrated that when no tuning factor is used (i.e. PBS), the PBS algorithm will have severe weight degeneration problems, which eventually lead to unreliable estimates. The PBS$_\beta$ and APBS provide similar results. In general, the correlations between the truth and the APBS estimated

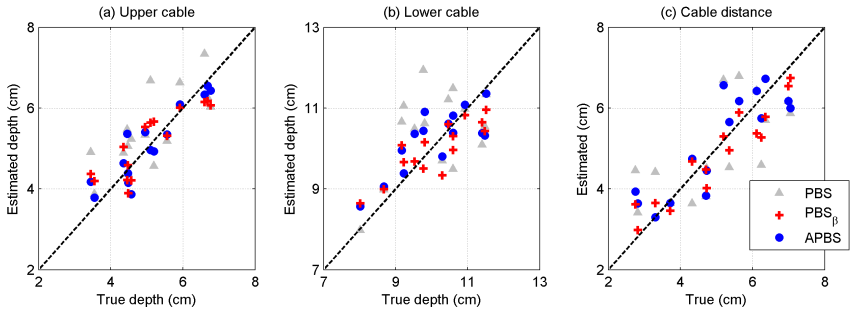

Figure 6.5: Comparison of the true and estimated cable depths (a and b), and cable distances (c) using 15 randomly sampled truths.

Table 6.1: The correlation of the estimated and the true model parameters using different PBS schemes

Par.	**PBS**	**PBS$_\beta$**	**APBS**
θ_r	0.41	0.89	0.91
θ_s	0.44	0.88	0.91
α	-0.15	0.52	0.62
n	0.46	0.93	0.95
K_s	0.54	0.69	0.88
λ_{sat}	0.13	0.74	0.55
θ_{33}	0.37	0.92	0.90
LAI	0.71	0.94	0.96

model parameters are higher than that of PBS$_\beta$.

The optimal β values from a single truth in the APBS is presented in a histogram to provide a more detailed comparison of the APBS and PBS$_\beta$ method (Figure 6.7). The optimal value of β is not a simple function of the state, forcing or parameter, but a combination of those that may lead to particle degeneration. Also, the optimal value is that which provides a particle range just wide enough to encompass the observation, so it will depend on the difference between the particle model estimates and the observation. If the observation is already within the prior range, a β value as high as 0.9 will encompass the observation. Assuming a single value of 0.25 for β reduces the likelihood that the observation is beyond the particle range, but does not ensure this to be the case. From Figure 6.7, most of the optimal β values in the APBS method are between 0 and 0.1. Using such small β values, APBS prevents the particle spread from collapsing, and ensures that the particle range is just wide enough to encompass the observations.

As a result of the similarity in the soil properties, the PBS$_\beta$ and APBS also provide very similar soil moisture estimates (Table 6.2). Using the PBS yields the poorest soil moisture estimates, the RMSE of which is approximately twice as high as that of the PBS$_\beta$ and APBS. The necessity of using the tuning factor in the PBS is also shown in the soil temperature estimates. The RMSE of PBS estimated soil temperature is above 0.5 ^{o}C for depths $\leq$ 20cm, which is even larger than the observation error. Using a tuning factor of 0.25 (PBS$_\beta$) can significantly avoid the particle weight degeneration, and leads to a sharp

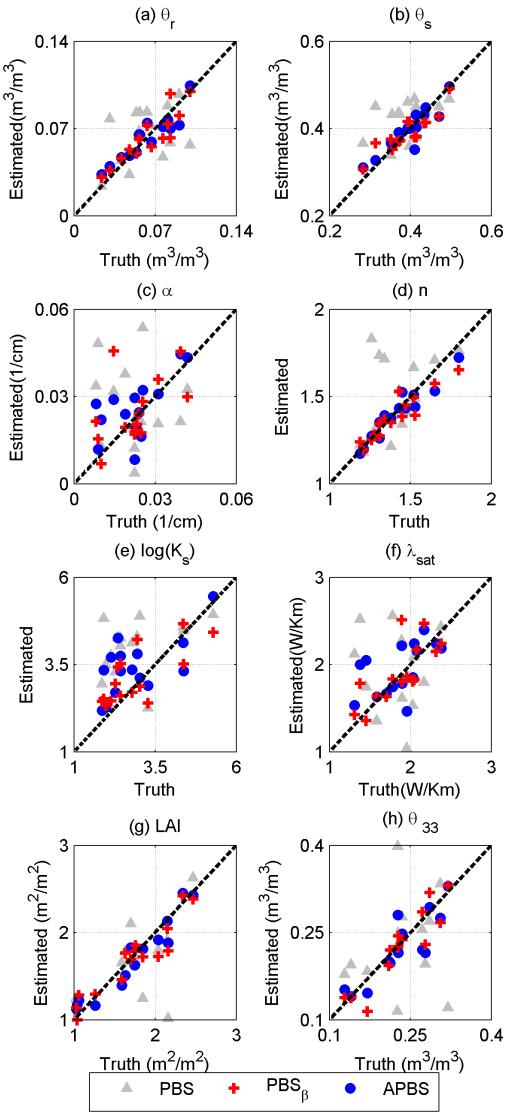

Figure 6.6: Comparison of the true and the estimated model parameters (a to g) and the θ_{33} (h) calculated using the estimated soil hydraulic properties.

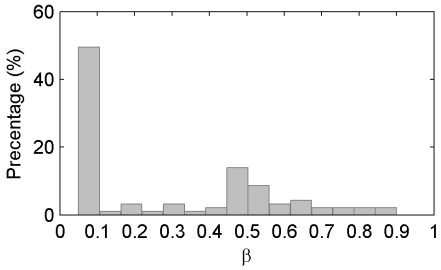

Figure 6.7: The distribution of the β values that used in the APBS method. The present values are from one single truth for illustration.

decrease in the soil temperature RMSE. When the adaptive tuning factor is used (APBS), the soil temperature RMSE were further reduced at depth above 10 cm.

Table 6.2: The RMSE of the estimated soil moisture and soil temperature using different PBS schemes at 5 depths. The values are averaged from 15 tests using randomly sampled truths

State	App.	2.5cm	5cm	10cm	20cm	50cm
Moisture	PBS	0.041	0.042	0.041	0.040	0.042
	PBS_β	0.022	0.023	0.024	0.028	0.032
	APBS	0.022	0.023	0.023	0.026	0.031
Temperature	PBS	0.656	0.541	0.509	0.505	0.494
	PBS_β	0.417	0.342	0.285	0.255	0.251
	APBS	0.339	0.275	0.258	0.253	0.249

The results from the multiple truth tests presented in this section show several key points in successfully implementing the PBS algorithm. The PBS algorithm updates the model states and parameters by placing more weight on the particles that have a better fit to the observations, and discarding the particles that have little weight. Hence, when the range of the particle estimates is small, the prior distribution mapped by the PBS is very unlikely to contain the global optimal parameter sets. As a result, the PBS estimates can be unreliable, which is consistent with a previous study [101]. The PBS_β uses a tuning factor of 0.25, which makes the particles more acceptable in the resampling process. This will allow the prior distribution for the model states and parameters to have heavier tails, which is more likely to encompass the observations. As a result, the estimates were greatly improved compared to those from the PBS (Table 6.1 and 6.2). However, until now, the tuning factors had been obtained by trial and error, and assumed to be constant in time. In order to implement the PBS to real DTS data, it is essential that the tuning factor can be determined objectively and automatically for each section of cable, and "on-the-fly" to allow for seasonal effects related to model structural error. The APBS can also be considered as a more statistically meaningful way of tuning the PBS algorithm. For each batch window, the APBS finds a β value that gives the largest reliability of the soil temperature estimates, i.e. a β value that optimally addresses uncertainties of the estimates. For example, when the particle spread is small, small tuning factors should be

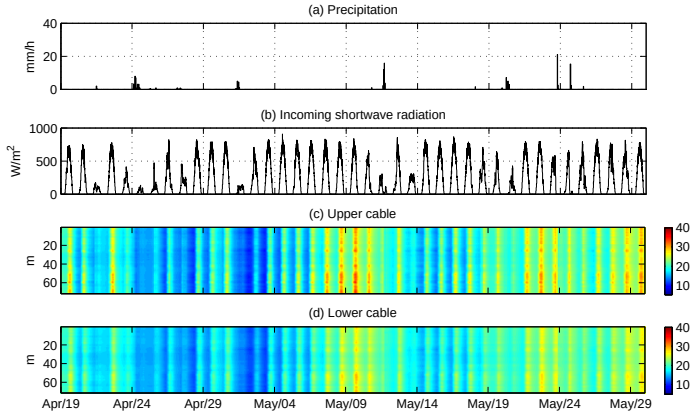

Figure 6.8: Observed precipitation (a), solar radiation (b), and soil temperature using two fiber optic cables (c and d) along a 71 m transect from April 19th and May 29th. The color represents the soil temperature (^{o}C).

used to breed the particles. Conversely, when the particle spread is large, a large tuning factor can be used to make full use of the soil temperature information. Thus, the APBS is shown to be superior than PBS$_{\beta}$ in most of the model state and parameter estimates.

6.3.3. IMPLEMENTING APBS IN A REAL DTS EXPERIMENT

Figure 6.8 shows the temperatures measured at approximately 5 cm and 10 cm using DTS from April 19 2011 to May 29 2011. The diurnal cycles of the observed soil temperatures due to solar radiation (Figure 6.8 b) are clearly visible. The daily amplitude of soil temperature is lower on days with precipitation and low radiation (e.g. April 24 to 29). Some spatial variation is observed, e.g. lower temperatures between 30 m and 42 m, and higher temperature between 50 to 60 m. These variations could be due to differences in soil texture and/or vegetation cover, and variations in cable depths.

Figure 6.9 shows the estimated cable depths along the 71 m DTS transect at the SMAP MOISST site. The average depths of the upper and lower cables are 5.28 cm and 9.35 cm respectively. The average estimate of the separation distance between the two cables is approximately 4 cm. These estimates are plausible given the plow configuration. The spatial variability can be attributed to the impact of soil roughness on the cable installation.

Histograms of the estimated saturated soil thermal conductivity (λ_{sat}) and θ_{33} along this 71 m DTS transect are shown in Figure 6.10. These are the only two parameters for which validation measurements were available at the nearby site B. The prior guessed λ_{sat} is approximately 1.9 W/mK, which is significantly higher than the measured value. The average of the APBS estimated λ_{sat} values was drawn closer to the reference value, with an average of 1.77 W/mK along this 71 m transect.

The mean of the prior estimated θ_{33} is approximately 0.2 m^3/m^3, which is significantly lower than the field measured reference value. When soil temperature observations from the DTS are assimilated, the mean of the estimated θ_{33} of this 71 m transect

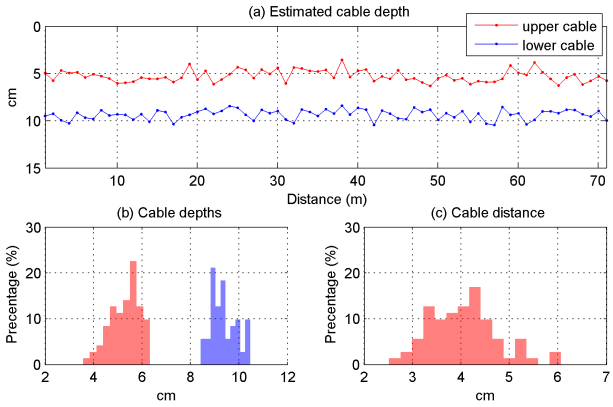

Figure 6.9: The estimated upper and lower cable depths along a 71m DTS transect (a), the distribution of the estimated 71m cable depths (b), and the distribution of the estimated cable distances (c).

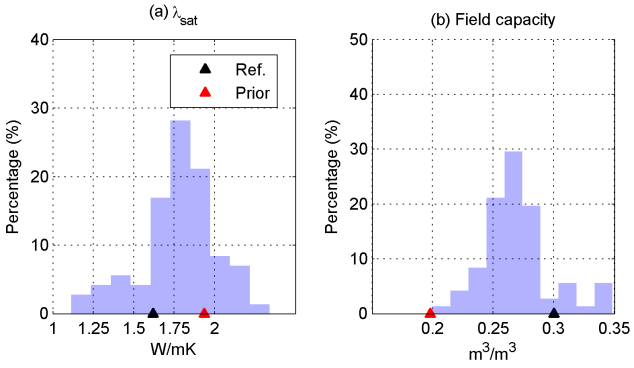

Figure 6.10: The distribution of the estimated soil thermal property (λ_{sat}, a) and soil hydraulic property (θ_{33}, b) along the 71m DTS transect.

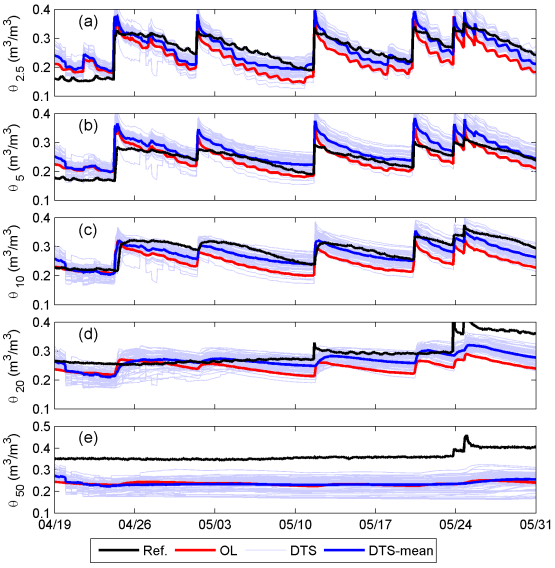

Figure 6.11: The estimated soil moisture along the 71m DTS transect at 5 depths. The black solid line is the soil moisture measured at site B. Each thin blue line represents soil moisture estimates for one meter of cable, and the mean estimated soil moisture of the 71m cable is shown as the blue solid line.

is 0.27 m^3/m^3.

The time series of the estimated soil moisture using the DTS observed soil temperature is shown in Figure 6.11, and compared to the observed soil moisture at the reference site. Due to the significantly biased estimated soil hydraulic properties (Figure 6.10 b), the OL is biased compared to the reference site. When soil temperature is assimilated, the APBS removes the biases between the OL and the reference soil moisture at depths above 20 cm. At 50 cm, the soil moisture estimates are dominated by the initial guess for soil moisture. Therefore, no significant improvement is shown in the APBS estimates, compared to the OL.

In Figure 6.12, the mean soil moisture in each meter of DTS cable is plotted against key soil properties as well as the cable depths in order to rule out any relationship between the estimated soil moisture and the estimated cable depths. From Figure 6.12 a and b it is clear that there is no correlation between the estimated cable depths and the temporal mean of the soil moisture. This means the spatial pattern in soil moisture is not merely an artefact of uncertainty in the cable depth estimation. On the other hand, the temporal mean of the soil moisture along this 71 m transect has the strongest correlation with θ_{33}. This suggests that the spatial variability of the soil moisture across this transect is primarily attributed to the spatial variability of the soil hydraulic properties.

Figure 6.13 shows the soil moisture at depths ≤ 20 cm across this transect. Soil moisture along this 71 m of DTS cable at 2.5 cm shows a fast response to the climatic forcing. The impact of the forcing dramatically decreases with depth. Hence, the rapid increases in wetness following precipitation and the slow drying down are more evident in the soil

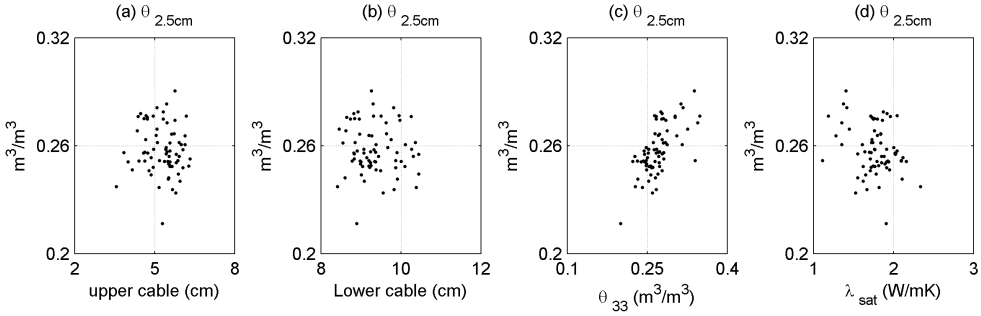

Figure 6.12: The temporal mean (TM) of soil moisture at 2.5 cm as a function of estimated cable depths (a and b), soil hydraulic (c) and thermal (d) properties. Each symbol represents the temporal mean of 2.5 cm soil moisture estimated at one meter of the DTS cable.

moisture at 2.5 cm. The horizontal stripes apparent in the soil moisture estimates represent spatial variability, which is likely due to variations in soil texture. Below 20 cm, little spatial or temporal soil moisture variability is observed. However, the variations due to spatial patterns in soil texture are still visible.

Figure 6.14 a and b show the evolution of the statistics of the soil moisture spatial variability along this transect, taking 2.5 cm soil moisture for illustration. No clear temporal pattern is shown in the standard deviation of soil moisture, except that the spikes after rainfall. In data assimilation, no perturbation for precipitation is used for days with zero precipitation, while a large multiplicative error is used for large precipitation events. Hence, the larger standard deviation of the soil moisture is likely to occur during or shortly after rainfall events. The standard deviation of the soil moisture is not shown to be a strong function of the areal mean soil moisture (Figure 6.14 c). It also seems the soil moisture standard deviation slightly increases with increased mean soil moisture, while several previous studies show soil moisture standard deviation reaches the peak value when areal mean soil moisture is in the range of 0.17 to 0.23 m^3/m^3 [e.g. 104, 105]. However, various relationships between soil moisture and standard deviation at intermediate scales were also found in previous studies, as summarized in [106]. This might indicate the standard deviation - mean soil moisture relationship is sensitive to observation errors, soil properties, and climatic forcing of the study area. The validity of the standard deviation - mean areal soil moisture relationship detected along this transect may need to be validated in a further study.

The coefficient of variation (CV) of the soil moisture along this transect has a clear temporal signature (Figure 6.14 b). In general, the CV of soil moisture at 2.5 cm is low during rainfall events (e.g. May 1). The value of CV increases during the dry down period (e.g. May 1 to 12). This makes sense as soil moisture is essentially reinitialized by precipitation events, with sharply increased soil water content and decreased CV. The CV of the soil moisture will start to increase again in the subsequent dry down process. CV is also shown to be a strong function of the areal mean soil moisture (Figure 6.14 d), which is consistent with previous studies [e.g. 105, 107, 108]. This highlights the potential of using DTS to observe and monitor the temporal and spatial soil moisture variability.

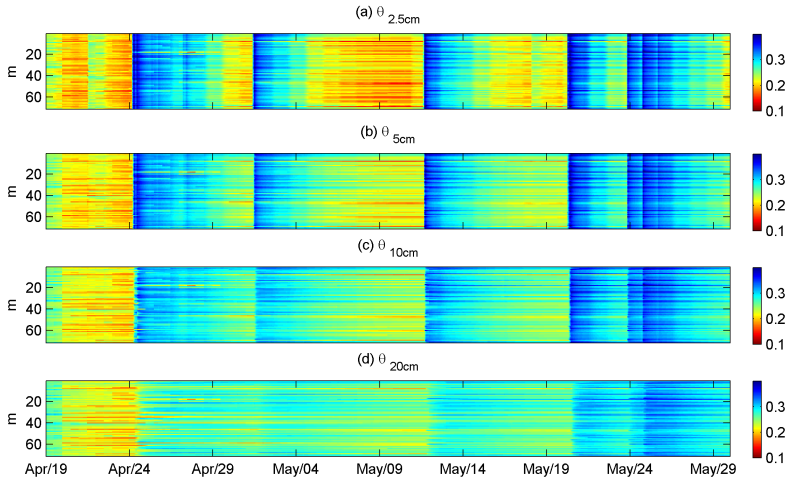

Figure 6.13: Soil moisture estimates along a 71m DTS transect at four depths (b to e). The color denotes the value of soil moisture (m^3/m^3).

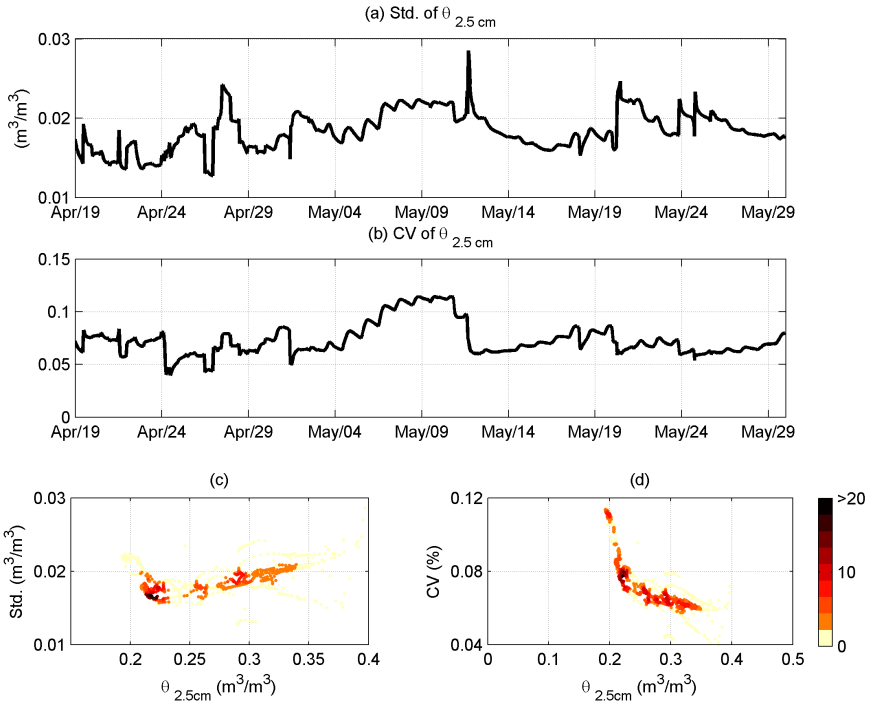

Figure 6.14: The time series of the standard deviation (a) and coefficient of variation (b) of 2.5 cm soil moisture. The standard deviation (c) and coefficient of variation (d) as a function of areal mean soil moisture.

6.4. CONCLUSION

In this study, we demonstrated the potential of mapping meter resolution soil moisture and soil properties along a 71 m DTS fiber optic cable using an adaptive Particle Batch Smoother (APBS). The estimated soil moisture and properties were evaluated using field measurements at a nearby site. Results show that the coefficient of variation and areal mean soil moisture relationship are consistent with previous studies. This indicates that the high resolution (spatial: 1 m, and temporal: 1 hour) soil moisture pattern mapped using the DTS has the potential of providing insight into the spatial and temporal evolution of soil moisture.

Consistent with previous studies, and also with the synthetic tests presented in this study, soil thermal properties are more difficult to be accurately estimated than the soil hydraulic properties. Including additional observations (e.g. Active DTS) may further constrain the estimated model parameters.

Due to the difficulties in installing the cable at this site with relatively uneven terrain and dense vegetation, the present study used data collected by only part of the DTS cables. The maturity of fiber optic cable installation techniques in recent years [109], may be helpful in providing high quality DTS observed soil temperature for cables up to kilometers in length. Using hourly measured DTS soil temperature data at two shallow depths already yields quite promising results. This suggests that hourly DTS data is sufficient, which will significantly reduce the data management and processing efforts in the large scale DTS experiments.

In this study, an adaptive PBS approach (APBS) was presented that automatically tunes the PBS to avoid severe weight degeneracy and improve parameter estimation. This obviates the need to assume that the tuning factor is a constant, and allows the tuning factor to be determined “on-the-fly" in a statistically meaningful way. The proposed APBS approach should also be valuable in other smoother applications in which parameter estimation might be desirable e.g. streamflow prediction, snow water equivalent estimation, and soil moisture reanalysis [e.g. 53, 57, 59, 110].

7

CONCLUSION

7.1. KNOWLEDGE GENERATED AND ORIGINAL CONTRIBUTIONS

7.1.1. ON SOIL MOISTURE MAPPING AND SCALING

The necessity of monitoring high spatial resolution soil moisture over large areas is presented in the first chapter, in the context of understanding soil moisture spatial variability, soil moisture scaling and validating large scale soil moisture measurement techniques. The key motivation of this thesis is to develop Passive DTS into a viable tool for high resolution, continuous, long-term soil moisture monitoring up to the kilometer scale.

Extracting soil moisture and soil property information using data assimilation methods was tested using both synthetic and real world data. The soil moisture spatial variability mapped using DTS also reflects a pattern consistent with previous studies derived using traditional point scale sensors [105]. High resolution soil moisture mapping with point sensors is logistically difficult and economically infeasible. Hence, it is difficult to explore the temporal evolution of soil moisture spatial patterns using traditional point scale measurement tools. DTS can provide continuous fine resolution soil moisture measurements at large scales. This means DTS will be a powerful tool for investigating the factors that affect the soil moisture spatial variability, e.g. heterogeneity of precipitation and seasonal vegetation changes.

7.1.2. ON SOIL MOISTURE ESTIMATION USING SOIL TEMPERATURES

Though the first attempt of estimating soil moisture using soil temperature observations dates back to 1970s, great challenges were presented in these methods, as was discussed in Chapter 1. Data assimilation methods outperform the typical traditional approaches, e.g. approaches related to soil thermal inertia and optimizing analytical soil heat transfer equations, in several aspects.

First, a forward model, which is the state of the art vadose zone model, is used for providing prior guessed soil moisture in the data assimilation methods. This allows us to incorporate meteorological forcing information into the estimation system, which yields

more informative prior soil moisture estimates. Conversely, moisture/thermal properties are estimated as a free parameter in the traditional methods, which means the prior guesses in the traditional methods are less informative.

Second, soil vapor transfer is taken as a key link between soil temperature and soil moisture in the data assimilation method, which is ignored in the traditional methods. Since the vapor transfer processes were neglected in the traditional approaches, the estimated soil thermal properties were usually biased. Hence, the soil moisture can not be effectively estimated using soil temperatures in the traditional methods.

Third, data assimilation can better handle different types of uncertainties. Soil moisture was solved in a deterministic way in the traditional methods. The relationship between soil moisture and soil thermal properties, the soil temperature observations, and the analytical solutions of the heat conduction equation were all considered to be perfect, and hence all the errors (uncertainties) will be accumulated in the final soil moisture estimates. This can often lead to physically unreasonable soil moisture estimates. In data assimilation, all types of uncertainties were considered, and the prior distribution was updated according to the observations. It is essentially equivalent to adjusting the weights of the prior guesses. Therefore, data assimilation methods are more capable of handling the uncertainties, and the final soil moisture estimates are always physically reasonable.

7.1.3. ON DATA ASSIMILATION ALGORITHMS

A Particle Batch Smoother (PBS) is presented to extract information of the variable of interest from a sequence of observations. The PBS algorithm outperforms the sequential particle filtering algorithm, as tested in this thesis. Development of the PBS for soil moisture estimation based on DTS is the main scientific contribution of this thesis. The PBS overcomes most problems associated with earlier methods. This is because a sequence of observations contains more information than measurements at instantaneous points. This thesis also demonstrated that joint/dual state-parameter estimation usually yields significantly improved estimates compared to the cases in which only model states are updated. This is because improved model parameters can significantly improve the performance of the forward model, which allows the prior estimates to better encompass the “truth". When model parameters are updated together with model states, a potential concern is the weight degeneracy problem in the PBS. When the PBS overconfidently fits the model parameters to the model errors, the model parameters may converge to incorrect values. As a consequence, the prior estimates will be less capable to encompass the observations and lead to severe weight degeneracy problems, and the uncertainty of the posterior estimates is underestimated. Reliability addresses whether the uncertainties of the estimates are reasonable. Consequently, the weight degeneracy problem can be avoided by maximizing the reliability of the estimates within each batch window. Based on this new concept, an adaptive Particle Batch Smoother (APBS) is proposed, which can automatically tune the PBS and yield robust estimates without weight degeneracy.

7.2. FUTURE RESEARCH

7.2.1. VALIDATION DATA

As demonstrated in this thesis, assimilating DTS observed soil temperatures into a vadose zone model can yield high spatial resolution soil moisture, soil thermal and hydraulic properties, and soil surface energy fluxes. The data assimilation method developed in this thesis was primarily tested and validated using synthetic tests and data collected from point sensors. To further utilize the DTS estimates, validation data along the DTS cables should be independently measured.

Soil moisture point sensors should be used for monitoring DTS estimated soil moisture at a few locations along the DTS cables. This is useful for validating the DTS soil moisture estimates. Field campaigns are also needed to map the surface soil moisture along the DTS cables. These soil moisture measurements can be used for evaluating whether the soil moisture spatial variability detected by the DTS is appropriate. To use the DTS as a general soil hydrology research tool, the soil hydraulic properties estimated by the DTS also require further validation. Hence, soil hydraulic property measurements at locations with distinctive soil textures are clearly needed. Soil energy fluxes can also be estimated every meter along the DTS cables. Independent soil energy fluxes should be measured at point scales (e.g. using heat pulse sensors [111]) to validate the DTS estimated fluxes.

7.2.2. COMBINING ACTIVE AND PASSIVE DTS

Active DTS, in which heat pulses are applied to the DTS cables, is more directly linked to soil moisture. As shown in Chapter 1, there are two categories of Active DTS methods. The first category of Active DTS methods relates soil heat pulses to soil moisture measurements using empirical equations. The second category of Active DTS is more physically based, which solves the soil thermal conductivity using the soil temperature change after heating. The soil thermal conductivity is then used for solving soil moisture, provided the soil thermal conductivity to soil moisture relationship is known.

When the Active DTS uses more physically based methods, soil thermal conductivity derived from the Active DTS can be directly assimilated. To optimally combine the Active and Passive DTS methods, the error of the Active DTS estimate soil thermal conductivity has to be accurately quantified. When soil moisture is inferred using empirical equations in the Active DTS, the uncertainties of this empirical equation have to be considered. This empirical equation is essentially a measurement operator in data assimilation, which links the model estimates (soil moisture) to the observations (heat pulses). Quantifying the uncertainty of this empirical equation, particular the uncertainties associated to different soil textures and soil moisture contents, is necessary.

7.2.3. DATA ASSIMILATION ALGORITHM

The weight degeneracy problem in the PBS method was handled using a tuning factor in Chapter 4 to 6. Using the concept of Gaussian mixture may further improve the PBS method and avoid the weight degeneracy problem. The PBS approximates the prior distribution using step functions, and describes the prior in a discrete form. Consequently, the PBS is likely to have weight degeneracy problems, when observations are at

the tails of the prior distribution. In Gaussian mixture approaches, the prior distribution is mapped using a set of Gaussian distributions with small standard deviations (bandwidth) [86, 92]. A Gaussian mixture batch smoother can be considered as a generalized batch smoother, which is equivalent to the ES when bandwidth is 1, and equivalent to the PBS when the bandwidth is 0. To successfully use the Gaussian mixture approach, an optimal bandwidth that can approximate the prior with good accuracies and heavy tails has to be investigated.

In sequential data assimilation, biases in the model forecasts should be removed. In this thesis, particularly the synthetic tests, the biases in the model forecasts are mainly attributed to the model parameters, which were implicitly handled when the parameters are jointly updated with the model states (Chapter 4 to 6). Compared with using bias correction techniques [e.g. 73, 74], a key benefit of using joint model state and parameter estimation is that model parameters can also be inferred from DTS observations. Further, the biases can be removed in a more physically meaningful way [e.g. 73, 74]. However, when the biases are attributed to the model structures or the observations, a bias correction algorithm should be implemented [75, 76, 112, 113]. We recognize that inclusion of a bias correction scheme would alter the results in Chapter 2 and 3.

The fiber optic cables have a protection layer of a few millimeters. When the soil temperature changes faster than the time required for reaching the thermal equilibrium between the fiber optic cables and the soil, the soil temperature observations will have temporally correlated errors. These and other autocorrelated observation errors are not considered in this thesis for simplicity. Further, an observation error with a standard deviation of 0.5 ^{o}C, which is significantly larger than the DTS instrument error (approximately 0.1 ^{o}C), was used to represent the observation error related to representativeness. However, when observations are more frequently assimilated, e.g. every minute, the autocorrelation of the observation errors may have significant impacts on the PBS. A data assimilation scheme similar to that presented in [74], might be useful in handling the autocorrelated errors.

7.2.4. INTERMEDIATE SCALE SOIL MOISTURE MEASUREMENTS VALIDATION

DTS will be useful in calibrating intermediate scale soil moisture measurement techniques, e.g. Cosmic ray probe and GPS-R. As discussed in the first chapter, both techniques are affected by the vegetation cover, soil surface roughness and soil moisture conditions. Further, the penetration depths of the two techniques also vary with different soil moisture conditions.

As shown in this thesis, the Passive DTS estimated soil moisture is usually most accurate at depths above 10 cm, with a vertical resolution of 1 cm. Since the GPS-R only has a penetration depth of approximately 5 cm, DTS will be particularly suitable for calibrating the GPS-R techniques.

To accurately calibrate the Cosmic ray probe, the accuracy of the DTS estimated soil moisture at depths has to be guaranteed. Root zone soil moisture, particularly soil moisture below 50 cm usually shows less responses to the meteorological forcing, the error of which comes mainly from the initial conditions. Hence, a longer spin-up time for the forward model is required to correct the errors in the initial conditions.

7.2.5. RELEVANCE TO GENERAL DATA ASSIMILATION PROBLEMS

It is expected that the PBS/APBS algorithm will also be suitable for a wide range of hydrological data assimilation studies. The PBS algorithm has already been applied to snow water equivalent (SWE) estimation [110], which shows that the PBS outperforms the ES algorithm. Lu et al. [114] also demonstrated that the PBS algorithm can provide accurately estimated ground surface energy fluxes by assimilating land surface temperature. In addition to SWE and land surface energy fluxes estimation, the PBS algorithm could be applied to land surface soil moisture estimation by assimilating microwave remote sensing observations. As shown in a previous study [59], assimilating a sequence of observations is more suitable for soil moisture estimation. However, when the ES was used, the soil moisture estimates prior to precipitation events were drawn to wetter values, which is physically unreasonable [59]. This is partly because the ES projects the joint distribution of the estimated soil moisture within a batch window into a multivariate Gaussian distribution, which is violated in reality. The PBS algorithm maps the entire distribution using random sampling points, and adjusts weights of different particle simulated soil moisture values. This guarantees all the posterior estimates are physically reasonable. Hence, the PBS algorithm may provide superior estimates.

In streamflow data assimilation studies, the Particle Filter (PF) was frequently used to assimilate streamflow observations to update the states and the parameters in the hydrological model. However, a sequence of streamflow observations contains more information of hydrological model parameters. Hence, the PBS should be more suitable than the PF in hydrological model parameter estimation. Further, the filtering algorithm cannot consider the time lag between the model states and discharge [115]. Since all the observations within a batch window are used to update the model states within this window, the time lag between the states and the discharge is implicitly considered in the PBS.

When applying the PBS/APBS algorithm to other hydrological data assimilation problems the optimal window length should be investigated. When a short window length is used, the correlation of different model states, and the correlation of model states and parameters cannot be sufficiently considered. An extensively long window length will significantly increase the dimension of the estimation problems, and lead to weight degeneracy problems. The optimal window length should be determined by testing the algorithm using multiple-truth synthetic tests, and checking the accuracy and the reliability of the estimates in real world data.

7.2.6. RELEVANCE TO HYDROLOGY AND HYDROMETEOROLOGY

From the hydrology perspective, this thesis provided a tool for collecting model parameter information and calibration/validation data for distributed hydrological modeling. Distributed hydrological modeling may contain substantial uncertainties when the model parameters are poorly known. The DTS method can provide spatially distributed soil property and vegetation information over large areas, which may benefit the parameterization of distributed hydrological models. Further, the hydrological models are commonly calibrated and validated using only streamflow data. The soil moisture data derived from the DTS can also be used for calibrating and validating the hydrological models. This may allow us to improve the performance of different hydrological model

structures.

This thesis may also improve our understanding of groundwater recharge. Groundwater recharge can be interpreted using tracers or simulated using Darcy's law, provided the soil hydraulic properties are known [116]. The groundwater recharge estimated at a few points were usually used to represent the areal mean, which is partly caused by the difficulty in intensively sampling the tracer data or the soil hydraulic properties over a large area. This will ignore the spatial variability of the recharge rates and may result in inaccurate ground water recharge estimates [116]. The DTS method, as shown in this thesis, can provide information of soil hydraulic properties and the change of soil moisture profiles at intermediate scales with sub-meter resolution. This information can be helpful in investigating the spatial variability and the areal mean of the ground recharge rates.

Applying the proposed methodology to the traditional point scale soil temperature sensors globally could yield new ground based soil moisture and energy flux products at large scales. For example, applying this methodology to the International Soil Moisture Network (ISMN), and/or FLUXNET measured soil temperature data, we may derive ground based soil moisture, soil property, and soil surface energy flux datasets over the globe. The derived datasets are independent from the point sensor soil moisture values, and also independent from flux tower measured fluxes. Hence, these datasets can be used for quality control, and filling the temporal and spatial gaps of the ground based soil moisture and/or energy fluxes measurements.

REFERENCES

[1] D. Entekhabi, I. Rodriguez-Iturbe, and F. Castelli, "Mutual interaction of soil moisture state and atmospheric processes," Journal of Hydrology, vol. 184, no. 1, pp. 3–17, 1996.

[2] W. Crow and D. Ryu, "A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals," Hydrology and Earth System Sciences, vol. 13, no. 1, pp. 1–16, 2009.

[3] A. W. Western, R. B. Grayson, G. Blöschl, G. R. Willgoose, and T. A. McMahon, "Observed spatial organization of soil moisture and its relation to terrain indices," Water resources research, vol. 35, no. 3, pp. 797–810, 1999.

[4] S. I. Seneviratne, T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, "Investigating soil moisture–climate interactions in a changing climate: A review," Earth-Science Reviews, vol. 99, no. 3, pp. 125–161, 2010.

[5] M. Hirschi, S. I. Seneviratne, V. Alexandrov, F. Boberg, C. Boroneant, O. B. Christensen, H. Formayer, B. Orlowsky, and P. Stepanek, "Observational evidence for soil-moisture impact on hot extremes in southeastern Europe," Nature Geoscience, vol. 4, no. 1, pp. 17–21, 2010.

[6] S. I. Seneviratne, D. Lüthi, M. Litschi, and C. Schär, "Land–atmosphere coupling and climate change in europe," Nature, vol. 443, no. 7108, pp. 205–209, 2006.

[7] R. D. Koster, S. P. P. Mahanama, B. Livneh, D. P. Lettenmaier, and R. H. Reichle, "Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow," Nature Geoscience, vol. 3, no. 9, pp. 613–616, 2010.

[8] T. E. Ochsner, M. H. Cosh, R. H. Cuenca, W. A. Dorigo, C. S. Draper, Y. Hagimoto, Y. H. Kerr, E. G. Njoku, E. E. Small, and M. Zreda, "State of the art in large-scale soil moisture monitoring," Soil Science Society of America Journal, vol. 77, no. 6, pp. 1888–1919, 2013.

[9] S. B. Idso, R. D. Jackson, and R. J. Reginato, "Compensating for environmental variability in the thermal inertia approach to remote sensing of soil moisture," Journal of Applied Meteorology, vol. 15, no. 8, pp. 811–817, 1976.

[10] S. Idso, T. Schmugge, R. Jackson, and R. Reginato, "The utility of surface temperature measurements for the remote sensing of surface soil water status," Journal of Geophysical Research, vol. 80, no. 21, pp. 3044–3049, 1975.

[11] J. C. Price, "Thermal inertia mapping: a new view of the earth," Journal of Geophysical Research, vol. 82, no. 18, pp. 2582–2590, 1977.

[12] A. Verhoef, "Remote estimation of thermal inertia and soil heat flux for bare soil," Agricultural and forest meteorology, vol. 123, no. 3, pp. 221–236, 2004.

[13] T. Schmugge, "Remote sensing of surface soil moisture," Journal of Applied Meteorology, vol. 17, no. 10, pp. 1549–1557, 1978.

[14] B. Saltzman and J. A. Pollack, "Sensitivity of the diurnal surface temperature range to changes in physical parameters," Journal of Applied Meteorology, vol. 16, no. 6, pp. 614–619, 1977.

[15] E. G. Njoku and D. Entekhabi, "Passive microwave remote sensing of soil moisture," Journal of hydrology, vol. 184, no. 1, pp. 101–129, 1996.

[16] F. T. Ulaby, P. C. Dubois, and J. van Zyl, "Radar mapping of surface soil moisture," Journal of Hydrology, vol. 184, no. 1, pp. 57–84, 1996.

[17] J.-P. Wigneron, J.-C. Calvet, T. Pellarin, A. Van de Griend, M. Berger, and P. Ferrazzoli, "Retrieving near-surface soil moisture from microwave radiometric observations: current status and future plans," Remote Sensing of Environment, vol. 85, no. 4, pp. 489–506, 2003.

[18] N. Ahmed, "Estimating soil moisture from 6· 6 ghz dual polarization, and/or satellite derived vegetation index," International Journal of Remote Sensing, vol. 16, no. 4, pp. 687–708, 1995.

[19] Y.-A. Liou, S.-F. Liu, and W.-J. Wang, "Retrieving soil moisture from simulated brightness temperatures by a neural network," Geoscience and Remote Sensing, IEEE Transactions on, vol. 39, no. 8, pp. 1662–1672, 2001.

[20] J.-P. Wigneron, A. Chanzy, J.-C. Calvet, and N. Bruguier, "A simple algorithm to retrieve soil moisture and vegetation biomass using passive microwave measurements over crop fields," Remote Sensing of Environment, vol. 51, no. 3, pp. 331–341, 1995.

[21] F. Del Frate, P. Ferrazzoli, and G. Schiavon, "Retrieving soil moisture and agricultural variables by microwave radiometry using neural networks," Remote sensing of environment, vol. 84, no. 2, pp. 174–183, 2003.

[22] E. G. Njoku, T. J. Jackson, V. Lakshmi, T. K. Chan, and S. V. Nghiem, "Soil moisture retrieval from amsr-e," Geoscience and Remote Sensing, IEEE Transactions on, vol. 41, no. 2, pp. 215–229, 2003.

[23] M. Owe, R. De Jeu, and J. Walker, "A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index," Geoscience and Remote Sensing, IEEE Transactions on, vol. 39, no. 8, pp. 1643–1654, 2001.

[24] S. Reynolds, "The gravimetric method of soil moisture determination part ia study of equipment, and methodological problems," Journal of Hydrology, vol. 11, no. 3, pp. 258–273, 1970.

[25] D. Robinson, C. Campbell, J. Hopmans, B. Hornbuckle, S. B. Jones, R. Knight, F. Ogden, J. Selker, and O. Wendroth, "Soil moisture measurement for ecological and hydrological watershed-scale observatories: A review," Vadose Zone Journal, vol. 7, no. 1, pp. 358–389, 2008.

[26] K. M. Larson, E. E. Small, E. D. Gutmann, A. L. Bilich, J. J. Braun, and V. U. Zavorotny, "Use of gps receivers as a soil moisture network for water cycle studies," Geophysical Research Letters, vol. 35, no. 24, 2008.

[27] K. M. Larson, E. E. Small, E. Gutmann, A. Bilich, P. Axelrad, and J. Braun, "Using gps multipath to measure soil moisture fluctuations: Initial results," GPS solutions, vol. 12, no. 3, pp. 173–177, 2008.

[28] K. M. Larson, J. J. Braun, E. E. Small, V. U. Zavorotny, E. D. Gutmann, and A. L. Bilich, "Gps multipath and its relation to near-surface soil moisture content," Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, vol. 3, no. 1, pp. 91–99, 2010.

[29] M. Zreda, D. Desilets, T. Ferré, and R. L. Scott, "Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons," Geophysical research letters, vol. 35, no. 21, 2008.

[30] M. Zreda, W. Shuttleworth, X. Zeng, C. Zweck, D. Desilets, T. Franz, and R. Rosolem, "Cosmos: the cosmic-ray soil moisture observing system," Hydrology and Earth System Sciences, vol. 16, no. 11, pp. 4079–4099, 2012.

[31] M. H. Cosh, T. J. Jackson, R. Bindlish, and J. H. Prueger, "Watershed scale temporal and spatial stability of soil moisture and its role in validating satellite estimates," Remote Sensing of Environment, vol. 92, no. 4, pp. 427–435, 2004.

[32] M. H. Cosh, T. J. Jackson, S. Moran, and R. Bindlish, "Temporal persistence and stability of surface soil moisture in a semi-arid watershed," Remote Sensing of Environment, vol. 112, no. 2, pp. 304–313, 2008.

[33] J. Martínez-Fernández and A. Ceballos, "Temporal stability of soil moisture in a large-field experiment in spain," Soil Science Society of America Journal, vol. 67, no. 6, pp. 1647–1656, 2003.

[34] A. Bárdossy and W. Lehmann, "Spatial distribution of soil moisture in a small catchment. part 1: geostatistical analysis," Journal of Hydrology, vol. 206, no. 1, pp. 1–15, 1998.

[35] Q. ZHU and H. Lin, "Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes," Pedosphere, vol. 20, no. 5, pp. 594–606, 2010.

[36] W. T. Crow, A. A. Berg, M. H. Cosh, A. Loew, B. P. Mohanty, R. Panciera, P. de Rosnay, D. Ryu, and J. P. Walker, "Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products," Reviews of Geophysics, vol. 50, no. 2, p. RG2002, 2012.

[37] W. T. Crow, D. Ryu, and J. S. Famiglietti, "Upscaling of field-scale soil moisture measurements using distributed land surface modeling," Advances in Water Resources, vol. 28, no. 1, pp. 1–14, 2005.

[38] S. C. Steele-Dunne, M. M. Rutten, D. M. Krzeminska, M. Hausner, S. W. Tyler, J. Selker, T. a. Bogaard, and N. C. van de Giesen, "Feasibility of soil moisture estimation using passive distributed temperature sensing," Water Resources Research, vol. 46, p. W03534, Mar. 2010.

[39] C. Sayde, C. Gregory, M. Gil-Rodriguez, N. Tufillaro, S. Tyler, N. van de Giesen, M. English, R. Cuenca, and J. S. Selker, "Feasibility of soil moisture monitoring with heated fiber optics," Water Resources Research, vol. 46, p. W06201, June 2010.

[40] J. S. Selker, L. Thévenaz, H. Huwald, A. Mallet, W. Luxemburg, N. Van De Giesen, M. Stejskal, J. Zeman, M. Westhoff, and M. B. Parlange, "Distributed fiber-optic temperature sensing for hydrologic systems," Water Resources Research, vol. 42, no. 12, p. W12202, 2006.

[41] F. Ciocca, I. Lunati, N. Van de Giesen, and M. B. Parlange, "Heated optical fiber for distributed soil-moisture measurements: A lysimeter experiment," Vadose Zone Journal, vol. 11, no. 4, 2012.

[42] R. H. Reichle, "Data assimilation methods in the Earth sciences," Advances in Water Resources, vol. 31, no. 11, pp. 1411–1418, 2008.

[43] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics," Journal of Geophysical Research: Oceans (1978–2012), vol. 99, no. C5, pp. 10143–10162, 1994.

[44] S. A. Margulis, D. McLaughlin, D. Entekhabi, and S. Dunne, "Land data assimilation and estimation of soil moisture using measurements from the southern great plains 1997 field experiment," Water resources research, vol. 38, no. 12, pp. 35–1, 2002.

[45] R. H. Reichle, R. D. Koster, J. Dong, and A. A. Berg, "Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation," Journal of Hydrometeorology, vol. 5, no. 3, pp. 430–442, 2004.

[46] R. H. Reichle and R. D. Koster, "Global assimilation of satellite surface soil moisture retrievals into the nasa catchment land surface model," Geophysical Research Letters, vol. 32, no. 2, 2005.

[47] W. T. Crow, W. P. Kustas, and J. H. Prueger, "Monitoring root-zone soil moisture through the assimilation of a thermal remote sensing-based soil moisture proxy into a water balance model," Remote Sensing of Environment, vol. 112, no. 4, pp. 1268–1281, 2008.

[48] C. Draper, R. Reichle, G. De Lannoy, and Q. Liu, "Assimilation of passive and active microwave soil moisture retrievals," Geophysical Research Letters, vol. 39, no. 4, 2012.

[49] B. F. Zaitchik, M. Rodell, and R. H. Reichle, "Assimilation of grace terrestrial water storage data into a land surface model: Results for the mississippi river basin," Journal of Hydrometeorology, vol. 9, no. 3, pp. 535–548, 2008.

[50] B. A. Forman, R. Reichle, and M. Rodell, "Assimilation of terrestrial water storage from grace in a snow-dominated basin," Water Resources Research, vol. 48, no. 1, 2012.

[51] B. A. Forman and R. Reichle, "The spatial scale of model errors and assimilated retrievals in a terrestrial water storage assimilation system," Water Resources Research, vol. 49, no. 11, pp. 7457–7468, 2013.

[52] M. P. Clark, D. E. Rupp, R. A. Woods, X. Zheng, R. P. Ibbitt, A. G. Slater, J. Schmidt, and M. J. Uddstrom, "Hydrological data assimilation with the ensemble kalman filter: Use of streamflow observations to update states in a distributed hydrological model," Advances in water resources, vol. 31, no. 10, pp. 1309–1324, 2008.

[53] H. Moradkhani, K.-L. Hsu, H. Gupta, and S. Sorooshian, "Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter," Water Resources Research, vol. 41, no. 5, 2005.

[54] C. Montzka, H. Moradkhani, L. Weihermüller, H.-J. H. Franssen, M. Canty, and H. Vereecken, "Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter," Journal of Hydrology, vol. 399, pp. 410–421, Mar. 2011.

[55] C. Montzka, J. P. Grant, H. Moradkhani, H.-J. H. Franssen, L. Weihermüller, M. Drusch, and H. Vereecken, "Estimation of radiative transfer parameters from l-band passive microwave brightness temperatures using advanced data assimilation," Vadose Zone Journal, vol. 12, no. 3, 2013.

[56] A. H. Weerts and G. Y. El Serafy, "Particle filtering and ensemble kalman filtering for state updating with hydrological conceptual rainfall-runoff models," Water Resources Research, vol. 42, no. 9, 2006.

[57] C. M. DeChant and H. Moradkhani, "Examining the effectiveness and robustness of sequential data assimilation methods for quantification of uncertainty in hydrologic forecasting," Water Resources Research, vol. 48, no. 4, 2012.

[58] H. Moradkhani, C. M. DeChant, and S. Sorooshian, "Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-markov chain monte carlo method," Water Resources Research, vol. 48, no. 12, 2012.

[59] S. Dunne and D. Entekhabi, "An ensemble-based reanalysis approach to land data assimilation," Water resources research, vol. 41, no. 2, 2005.

[60] S. Bateni and D. Entekhabi, "Surface heat flux estimation with the ensemble kalman smoother: Joint estimation of state and parameters," Water Resources Research, vol. 48, no. 8, 2012.

[61] P. J. Van Leeuwen and G. Evensen, "Data assimilation and inverse methods in terms of a probabilistic formulation," Monthly Weather Review, vol. 124, no. 12, pp. 2898–2913, 1996.

[62] G. Evensen, "The ensemble Kalman filter: Theoretical formulation and practical implementation," Ocean dynamics, vol. 53, no. 4, pp. 343–367, 2003.

[63] J. Simunek, M. Sejna, H. Saito, M. Sakai, and M. T. Van Genuchten, "The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media," University of California, Riverside, 2009.

[64] M. T. Van Genuchten, "A closed-form equation for predicting the hydraulic conductivity of unsaturated soils," Soil Science Society of America Journal, vol. 44, no. 5, pp. 892–898, 1980.

[65] M. G. Schaap, F. J. Leij, and M. T. van Genuchten, "Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions," Journal of hydrology, vol. 251, no. 3, pp. 163–176, 2001.

[66] G. S. Campbell, Soil physics with BASIC: transport models for soil-plant systems, vol. 14. Elsevier, 1985.

[67] S.-O. Chung and R. Horton, "Soil heat and water flow with a partial surface mulch," Water Resources Research, vol. 23, no. 12, pp. 2175–2186, 1987.

[68] H. Saito, J. Šimůnek, and B. P. Mohanty, "Numerical Analysis of Coupled Water, Vapor, and Heat Transport in the Vadose Zone," Vadose Zone Journal, vol. 5, no. 2, p. 784, 2006.

[69] H. Moradkhani, S. Sorooshian, H. V. Gupta, and P. R. Houser, "Dual state–parameter estimation of hydrological models using ensemble Kalman filter," Advances in Water Resources, vol. 28, no. 2, pp. 135–147, 2005.

[70] S. Tyler, D. Holland, V. Zagorodnov, A. Stern, C. Sladek, S. Kobs, S. White, F. Suárez, and J. Bryenton, "Using distributed temperature sensors to monitor an antarctic ice shelf and sub-ice-shelf cavity," Journal of Glaciology, vol. 59, no. 215, pp. 583–591, 2013.

[71] N. van de Giesen, S. C. Steele-Dunne, J. Jansen, O. Hoes, M. B. Hausner, S. Tyler, and J. Selker, "Double-ended calibration of fiber-optic raman spectra distributed temperature sensing data," Sensors, vol. 12, no. 5, pp. 5471–5485, 2012.

[72] S. Lu, T. Ren, Y. Gong, and R. Horton, "An improved model for predicting soil thermal conductivity from water content at room temperature," Soil Science Society of America Journal, vol. 71, no. 1, pp. 8–14, 2007.

[73] D. P. Dee and A. M. Da Silva, "Data assimilation in the presence of forecast bias," Quarterly Journal of the Royal Meteorological Society, no. 545, pp. 269–295, 1998.

[74] G. Evensen, "The ensemble kalman filter: Theoretical formulation and practical implementation," Ocean dynamics, vol. 53, no. 4, pp. 343–367, 2003.

[75] G. J. De Lannoy, R. H. Reichle, P. R. Houser, V. Pauwels, and N. E. Verhoest, "Correcting for forecast bias in soil moisture assimilation with the ensemble kalman filter," Water Resources Research, vol. 43, no. 9, 2007.

[76] G. J. De Lannoy, P. R. Houser, V. Pauwels, and N. E. Verhoest, "State and bias estimation for soil moisture profiles by an ensemble kalman filter: Effect of assimilation depth and frequency," Water resources research, vol. 43, no. 6, 2007.

[77] P. S. Arya, Introduction to micrometeorology, vol. 79. Academic press, 2001.

[78] S. Dunne and D. Entekhabi, "Land surface state and flux estimation using the ensemble kalman smoother during the southern great plains 1997 field experiment," Water resources research, vol. 42, no. 1, 2006.

[79] G. Boni, F. Castelli, and D. Entekhabi, "Sampling strategies and assimilation of ground temperature for the estimation of surface energy balance components," Geoscience and Remote Sensing, IEEE Transactions on, vol. 39, no. 1, pp. 165–172, 2001.

[80] J. A. Vrugt, C. J. ter Braak, C. G. Diks, and G. Schoups, "Hydrologic data assimilation using particle markov chain monte carlo simulation: Theory, concepts and applications," Advances in Water Resources, vol. 51, pp. 457–478, 2013.

[81] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," Handbook of Nonlinear Filtering, vol. 12, pp. 656–704, 2009.

[82] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," Signal Processing, IEEE Transactions on, vol. 50, no. 2, pp. 174–188, 2002.

[83] S. Eslamian, Handbook of Engineering Hydrology: Modeling, Climate Change, and Variability. Handbook of Engineering Hydrology, CRC Press, 2015.

[84] S. Madadgar and H. Moradkhani, "Improved bayesian multimodeling: Integration of copulas and bayesian model averaging," Water Resources Research, vol. 50, no. 12, pp. 9586–9603, 2014.

[85] M. Thyer, B. Renard, D. Kavetski, G. Kuczera, S. W. Franks, and S. Srikanthan, "Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using bayesian total error analysis," Water Resources Research, vol. 45, no. 12, 2009.

[86] A. S. Stordal, H. A. Karlsen, G. Nævdal, H. J. Skaug, and B. Vallès, "Bridging the ensemble kalman filter and particle filters: the adaptive Gaussian mixture filter," Computational Geosciences, vol. 15, no. 2, pp. 293–305, 2011.

[87] X. Han, H.-J. H. Franssen, C. Montzka, and H. Vereecken, "Soil moisture and soil properties estimation in the community land model with synthetic brightness temperature observations," Water Resources Research, vol. 50, no. 7, 2014.

[88] H. Yan, C. M. DeChant, and H. Moradkhani, "Improving soil moisture profile prediction with the particle filter-markov chain monte carlo method," Geoscience and Remote Sensing, IEEE Transactions on, vol. 53, no. 11, pp. 6134–6147, 2015.

[89] A. Verhoef, B. J. van den Hurk, A. F. Jacobs, and B. G. Heusinkveld, "Thermal soil properties for vineyard (efeda-i) and savanna (hapex-sahel) sites," Agricultural and Forest Meteorology, vol. 78, no. 1, pp. 1–18, 1996.

[90] H. Su, Z.-L. Yang, G.-Y. Niu, and C. R. Wilson, "Parameter estimation in ensemble based snow data assimilation: A synthetic study," Advances in Water Resources, vol. 34, no. 3, pp. 407–416, 2011.

[91] A. Aksoy, F. Zhang, and J. W. Nielsen-Gammon, "Ensemble-based simultaneous state and parameter estimation in a two-dimensional sea-breeze model," Monthly weather review, vol. 134, no. 10, pp. 2951–2970, 2006.

[92] A. S. Stordal and R. J. Lorentzen, "An iterative version of the adaptive Gaussian mixture filter," Computational Geosciences, vol. 18, no. 3-4, pp. 579–595, 2014.

[93] M. Leisenring and H. Moradkhani, "Analyzing the uncertainty of suspended sediment load prediction using sequential data assimilation," Journal of hydrology, vol. 468, pp. 268–282, 2012.

[94] X. Yang and T. Delsole, "Using the ensemble kalman filter to estimate multiplicative model parameters," Tellus A, vol. 61, no. 5, pp. 601–609, 2009.

[95] K. W. Oleson, D. M. Lawrence, B. Gordon, M. G. Flanner, E. Kluzek, J. Peter, S. Levis, S. C. Swenson, E. Thornton, J. Feddema, et al., "Technical description of version 4.0 of the community land model (clm)," 2010.

[96] P. J. Sellers, "Canopy reflectance, photosynthesis and transpiration," International Journal of Remote Sensing, vol. 6, no. 8, pp. 1335–1372, 1985.

[97] G. B. Bonan, "Land surface model (lsm version 1.0) for ecological, hydrological, and atmospheric studies: Technical description and users guide. technical note," tech. rep., National Center for Atmospheric Research, Boulder, CO (United States). Climate and Global Dynamics Div., 1996.

[98] J. Dong, S. C. Steele-Dunne, T. E. Ochsner, and N. van de Giesen, "Determining soil moisture by assimilating soil temperature measurements using the ensemble kalman filter," Advances in Water Resources, vol. 86, pp. 340–353, 2015.

[99] D. Ryu, W. T. Crow, X. Zhan, and T. J. Jackson, "Correcting unintended perturbation biases in hydrologic data assimilation," Journal of Hydrometeorology, vol. 10, no. 3, pp. 734–750, 2009.

[100] A. Monsivais-Huertero, J. Judge, S. Steele-Dunne, and P.-W. Liu, "Impact of bias correction methods on estimation of soil moisture when assimilating active and passive microwave observations," Geoscience and Remote Sensing, IEEE Transactions on, vol. 54, no. 1, pp. 262–278, 2016.

[101] J. Dong, S. C. Steele-Dunne, T. E. Ochsner, and N. van de Giesen, "Estimating soil moisture and soil thermal and hydraulic properties by assimilating soil temperatures using a particle batch smoother," Advances in Water Resources, vol. 81, pp. 104–116, 2016.

[102] W. Rawls, D. Brakensiek, and K. Saxton, "Estimation of soil water properties," Trans. ASAE, vol. 25, no. 5, pp. 1316–1320, 1982.

[103] J. Dong, S. C. Steele-Dunne, J. Judge, and N. van de Giesen, "A particle batch smoother for soil moisture estimation using soil temperature observations," Advances in Water Resources, vol. 83, pp. 111–122, 2015.

[104] H. Vereecken, T. Kamai, T. Harter, R. Kasteel, J. Hopmans, and J. Vanderborght, "Explaining soil moisture variability as a function of mean soil moisture: A stochastic unsaturated flow perspective," Geophysical Research Letters, vol. 34, no. 22, 2007.

[105] J. S. Famiglietti, D. Ryu, A. A. Berg, M. Rodell, and T. J. Jackson, "Field observations of soil moisture variability across scales," Water Resources Research, vol. 44, no. 1, p. W01423, 2008.

[106] L. Brocca, R. Morbidelli, F. Melone, and T. Moramarco, "Soil moisture spatial variability in experimental areas of central italy," Journal of Hydrology, vol. 333, no. 2, pp. 356–373, 2007.

[107] L. Brocca, F. Melone, T. Moramarco, and R. Morbidelli, "Spatial-temporal variability of soil moisture and its estimation across scales," Water Resources Research, vol. 46, no. 2, 2010.

[108] K. R. Bell, B. Blanchard, T. Schmugge, and M. Witczak, "Analysis of surface moisture variations within large-field sites," Water Resources Research, vol. 16, no. 4, pp. 796–810, 1980.

[109] C. Sayde, J. B. Buelga, L. Rodriguez-Sinobas, L. El Khoury, M. English, N. van de Giesen, and J. S. Selker, "Mapping variability of soil water content and flux across 1–1000 m scales using the actively heated fiber optic method," Water Resources Research, vol. 50, no. 9, pp. 7302–7317, 2014.

[110] S. A. Margulis, M. Girotto, G. Cortés, and M. Durand, "A particle batch smoother approach to snow water equivalent estimation," Journal of Hydrometeorology, vol. 16, pp. 1752–1772, 2015.

[111] X. Xiao, R. Horton, T. Sauer, J. L. Heitman, and T. Ren, "Cumulative soil water evaporation as a function of depth and time," Vadose Zone Journal, vol. 10, no. 3, pp. 1016–1022, 2011.

[112] V. R. Pauwels, G. J. DeLannoy, H.-J. Hendricks Franssen, and H. Vereecken, "Simultaneous estimation of model state variables and observation and forecast biases using a two-stage hybrid kalman filter," Hydrol. Earth Syst. Sci., vol. 17, pp. 3499–3521, 2013.

[113] C. Draper, R. Reichle, G. De Lannoy, and B. Scarino, "A dynamic approach to addressing observation-minus-forecast bias in a land surface skin temperature data assimilation system," Journal of Hydrometeorology, vol. 16, no. 1, pp. 449–464, 2015.

[114] Y. Lu, J. Dong, S. C. Steele-Dunne, and N. van de Giesen, "Estimating surface turbulent heat fluxes from land surface temperature and soil moisture observations using the particle batch smoother," Submitted to Water Resources Research, 2016.

[115] Y. Li, D. Ryu, A. W. Western, and Q. Wang, "Assimilation of stream discharge for flood forecasting: The benefits of accounting for routing time lags," Water Resources Research, vol. 49, no. 4, pp. 1887–1900, 2013.

[116] J. J. De Vries and I. Simmers, "Groundwater recharge: an overview of processes and challenges," Hydrogeology Journal, vol. 10, no. 1, pp. 5–17, 2002.

[117] M. H. Cosh, T. Ochsner, J. Basara, and T. J. Jackson, "The smap in situ soil moisture sensor testbed: Comparing in situ sensors for satellite validation," in Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International, pp. 699–701, IEEE, 2010.

[118] S. Fuhlendorf and D. Engle, "Application of the fire–grazing interaction to restore a shifting mosaic on tallgrass prairie," Journal of Applied Ecology, vol. 41, no. 4, pp. 604–614, 2004.

[119] R. A. McPherson, C. A. Fiebrich, K. C. Crawford, R. L. Elliott, J. R. Kilby, D. L. Grimsley, J. E. Martinez, J. B. Basara, B. G. Illston, D. A. Morris, et al., "Statewide monitoring of the mesoscale environment: A technical update on the oklahoma mesonet.," Journal of Atmospheric & Oceanic Technology, vol. 24, no. 3, 2007.

[120] J. Judge, J. Casanova, T.-y. Lin, K. Tien, M.-y. Jang, O. Lanni, and L. Miller, "Field observations during the second microwave, water, and energy balance experiment (microwex-2): from march 17 through june 3, 2004," tech. rep., Circular, 2005.

ACKNOWLEDGEMENTS

I enjoyed my Ph.D. life during the past four years. My Ph.D. research cannot be finished without the support from my promoter Prof. Nick van de Giesen, co-promoter Dr. Susan Steele-Dunne, and all my colleagues in the group.

Here, I would first thank Prof. Nick van de Giesen for offering me the opportunity of doing research in this group. During the past four years, Nick was always generous in supporting me to attend international conferences, which provided me the unique opportunities of international exposure. Nick always gave me enough academic freedom and supports to my Ph.D. research.

Dr. Susan Steele-Dunne gave me nearly daily supervisions at the start of my Ph.D. I always got fast and valuable feedback from Susan, and it was a real pleasure to work with her. In addition to my Ph.D. research, Susan also offered great help in my other academic related activities, and I am grateful to all the efforts she made.

I would like to thank Prof. Tyson Ochsner for offering me opportunities of working in his group. I start to feeling a little bit more confident to say “I have the soil physics background" after joining his soil physics class. I also thank Dr. Andres Patrignani, Dr. Yohannes Yimam, Dr. Xiuli Xin, Sonisa Sharma, Jordan Kay and all the other group members for helping me with the field works when I was in Oklahoma.

I had a lot of valuable discussions about DTS with Prof. John Selker and Dr. Chadi Sayde. I really appreciate the efforts they made for the collaborations. Thanks for hosting me in Oregon, and I really enjoyed myself there.

I appreciate the colleagues from both the US and the Netherlands for making experimental data available to me. Soil moisture and DTS soil temperature data from SMAP MOISST were collected with the help of Dr. Mike Cosh, Dr. Lynn McKee, Prof. Scott Tyler, Dr. Christine Hatch, Jop Jansen, Lucas Williamson, Geano Dong. I would also thank Prof. Jasmeet Judge and her group from University of Florida for sharing the MicroWEX-2 data.

Dr. Olivier Hoes, Dr. Wim Luxemburg, Dr. Thom Bogaard, Koen Hilgersom and Tim van Emmerik offered great help in my field work, and I appreciate their supports in the field. Many thanks to Betty Rothfusz, Lydia de Hoog, Petra Jorritsma, and Luz Ton-Estrada for managing the administrative issues, which made my life infinitely easier. I appreciate the conversations and the discussions with my officemate Frank Annor, Nay Myo Lin and Lan Wang. I would give a particular thank to Remko Nijzink for helping me in translating part of my thesis into Dutch. Without his help, I won't be able to submit a complete thesis.

Thank all my Chinese colleagues that I met in Delft, Prof. Huayang Cai, Dr. HongKai Gao, Wei Shao, Prof. Zongji Yang, Prof. LingNa Wei, Yingrong Wen, Yang Lu, Yongwei Liu, Zhilin Zhang. They made my life a lot more enjoyable. In particular, I would thank my friend Yerong Wu for constantly helping me with Linux system and high performance computers.

I am grateful to my Ms.C. promoter, Prof. Keli Zhang at Beijing Normal University. Without his support and encouragement, I won't be able to pursue a doctoral degree in the Netherlands. I would also like to particularly thank Prof. Guanghui Zhang for his academic supports when I was pursuing my Ms.C. degree.

Finally, I would like to thank my family. Their love and support helped me to get through the tough periods.

Jianzhi Dong
Delft, June 2016

A

DATA COLLECTION

A.1. MARENA OKLAHOMA MOISST SITE

All data used in this study were collected at the Soil Moisture Active Passive (SMAP) Marena Oklahoma In-Situ Sensor Testbed (MOISST). This is a field site located at the Oklahoma State University Range Research Station approximately 13 km southwest of Stillwater, Oklahoma. The goal of the testbed is to facilitate the validation of remote sensing measurements from SMAP with in-situ sensors distributed around the globe [117]. Figure A.1 shows the layout of the site. The predominant soil series is Grainola silty clay loam (Fine, mixed, active, thermic Udertic Haplustalfs), which are moderately deep, well drained soils formed in material weathered from shale. However, the soil texture varies significantly with depth and landscape position, as shown at the four observation sites (Figure A.2). Vegetation across much of the site is typical of tallgrass prairie with some localized areas representative of cross timbers vegetation [118].

Five-minute meteorological data including precipitation, solar radiation, humidity, air temperature, and wind speed were obtained from the Oklahoma Mesonet site at Marena during the study period [119]. Hourly soil temperature and moisture data were measured at 2.5, 5, 10, 20 and 50 cm at Site A, B and D using the Hydra probes (Hydra Probe II, Stevens Water Inc., Portland, OR, USA). Site C has similar soil moisture and temperature measurement depths, but the deepest sensor is located at 40 cm.

Soil hydraulic and thermal property data were also collected at the four sites to validate the estimated soil properties. Field capacity (θ_{33}), approximated by the soil water content at a matric potential of -33 kPa, was considered as an indicator of the accuracy of the estimated soil hydraulic properties. Soil thermal properties from 5 to 10 cm were measured at different soil water contents using a dual-probe heat-pulse sensor (KD2Pro, Decagon Devices, Pullman, WA) at site A, B and C. The measured soil thermal properties will be used to evaluate the estimated soil thermal conductivity curve.

The DTS equipment, solar panels and two calibration baths were located at Site B. Fiber-optic cables (50/125 μm multimode) were installed at approximately 5, 10 and 15 cm along the path (658m in total). Dong et al. [98] showed that soil temperatures at two

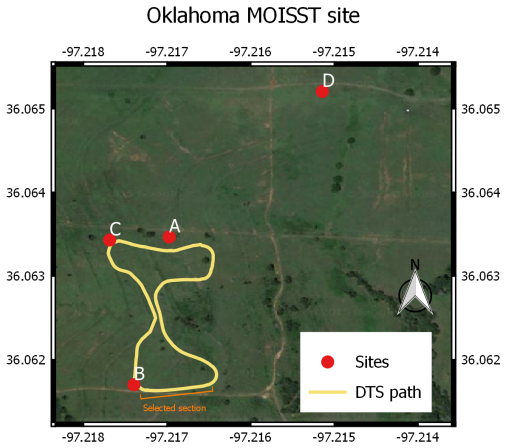

Figure A.1: Field layout of DTS cable at Oklahoma MOISST site. Soil moisture and temperature point sensors are installed at Sites A, B, C and D. The DTS equipment was housed at Site B.

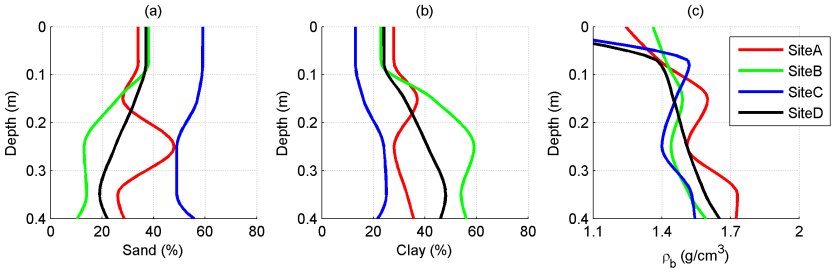

Figure A.2: The measured soil texture (a and b) and soil bulk density (c) at different depths.

depths provide sufficient information to estimate soil moisture and properties. So, this thesis only uses soil temperature observed by the upper two cables.

Cable temperatures were measured using an Oryx DTS (Sensornet, UK), with an integration time of five minutes, and calibrated using the approach outlined by [71]. The relatively high bulk density and clay content of the subsoil, the terracing on the western side of the loop, some rocks near the surface on the eastern side, and the thick vegetation cover made cable installation difficult. Several sections of the cable had to be hand-dug after the plow pass (crossing terrace berms and/or at tight turns). In these sections, the cables depths are too irregular to be used. This thesis is limited to the DTS data collected along a 71 m transect extending eastwards from site B (See Figure A.1). This is one of the smoother sections because the surface was flat and the tractor and plow could follow a straight line. Furthermore, the soil texture along this transect is comparable to that at Site B, which also facilitates validation of the results. The DTS temperature data used here were collected from 19 April to 31 May 2011.

A.2. MICROWEX-2 EXPERIMENT

Data from the Microwex-2 experiment was conducted from 17 March to 3 June 2004, by the Center for Remote Sensing, Agricultural and Biological Engineering Department, at the Plant Science Research and Education Unit of the University of Florida, Gainesville. A full description of this experiment is given by Judge et al. [120]. Soil temperature and moisture data were collected at 5 depths (2, 4, 8, 32 and 64 cm), from 26 March. Meteorological data (e.g. precipitation, air temperature, relative humidity and wind speed) were collected every 15 minutes. The sand, silt and clay content of 89.5%, 3.4% and 7.1% mean the soil at the site is classed as sand in the USDA soil textural classification. The corn was planted on 18 March. A period with negligible biomass (26 March to 10 April) is used in Chapter 2.

CURRICULUM VITÆ

Jianzhi DONG

06-Feb-1987 Born in Shenyang, China.

EDUCATION

2012–2016 Ph.D. Water resources management
Delft University of Technology, Delft, the Netherlands
Promotor: Prof. dr. Nick van de Giesen
co-Promotor: Dr. Susan Steele-Dunne

2009–2012 Ms.C. in Soil erosion
Beijing Normal University, Beijing, China
Promotor: Prof. dr. Keli Zhang

2005–2009 Bs.C. in Physical Geography
Sun Yat-Sen University, Guangzhou, China

AWARDS

2012 China Scholarship Council (CSC) Ph.D. fellowship

2012 Zhou Tingru academic excellence award at Beijing Normal University

2012 Graduate student excellence award at Beijing Normal University

2011 Graduate student excellence award at Beijing Normal University

2010 Graduate student excellence award at Beijing Normal University

2008 Professional scholarship at Sun Yat-Sen University

2006 Professional scholarship at Sun Yat-Sen University

TALKS AND ORAL PRESENTATIONS

Dong JZ, Susan C. Steele-Dunne, and Nick van de Giesen. High resolution soil moisture mapping using Distributed Temperature Sensing. EGU, Vienna, Austria, April 2016.

Dong JZ, Susan C. Steele-Dunne, and Nick van de Giesen. Soil moisture and properties

estimation by assimilating soil temperatures using particle batch smoother: A new perspective for DTS. AGU, San Francisco, U.S., December 2015.

Dong JZ, Susan C. Steele-Dunne, and Nick van de Giesen. A Particle Batch Smoother for soil moisture determination by assimilating soil temperatures. EGU, Vienna, Austria, April 2015.

Dong JZ, Susan C. Steele-Dunne, and Nick van de Giesen. Soil moisture determination through assimilation of temperature data. Boussinesq Lecture 2014, Amsterdam, the Netherlands, 23 October 2014.

Dong JZ, Susan C. Steele-Dunne, and Nick van de Giesen. Data assimilation of DTS for large scale distributed soil moisture observations. Computational Methods in Water Resources, Stuttgart, Germany, 10-13 June 2014.

Dong JZ, Susan C. Steele-Dunne, Tyson E. Ochsner, Christine Hatch, John Selker, Scott Tyler, Michael H. Cosh, and Nick van de Giesen, Soil moisture estimation using Passive DTS: Theory and field application. The 2014 Workshop at SMAP MOISST: Advancing Soil Moisture Science and Applications, Oklahoma, USA, 4-5 June 2014

LIST OF PUBLICATIONS

7. **Dong Jianzhi**, Susan C. Steele-Dunne Tyson E. Ochsner, Christine E. Hatch, John Selker, Scott Tyler, Michael H. Cosh, and Nick van de Giesen. Mapping high resolution soil moisture and properties using distributed temperature sensing data and an adaptive particle batch smoother, *Water Resources Research*, accepted.

6. Lu Yang, **Dong Jianzhi**, Susan C. Steele-Dunne, Nick van de Giesen. Estimating surface turbulent heat fluxes from land surface temperature and soil moisture observations using the particle batch smoother, *Water Resources Research*, under revision.

5. **Dong Jianzhi**, Susan C. Steele-Dunne, Tyson E. Ochsner, Nick van de Giesen. Determining soil moisture and soil properties in vegetated areas by assimilating soil temperatures, *Water Resources Research*, 52, 4280-4300.

4. **Dong Jianzhi**, Susan C. Steele-Dunne, Tyson E. Ochsner, Nick van de Giesen. Estimating soil moisture and soil thermal and hydraulic properties by assimilating soil temperatures using a particle batch smoother, 2016, *Advances in Water Resources*, 91, 104-116.

3. **Dong Jianzhi**, Susan C. Steele-Dunne, Tyson E. Ochsner, Nick van de Giesen. Determining soil moisture by assimilating soil temperature measurements using the Ensemble Kalman Filter, 2015, *Advances in Water Resources*, 85, 340-353.

2. **Dong Jianzhi**, Susan C. Steele-Dunne, Jasmeet Judge, Nick van de Giesen. A particle batch smoother for soil moisture estimation using soil temperature observations, 2015, *Advances in Water Resources*, 83, 111-222.

1. **Dong Jianzhi**, Zhang Keli, Guo Zongling. Runoff and soil erosion from highway construction spoil deposits: A rainfall simulation study, 2012, *Transportation research Part: D* 17, 8–14.