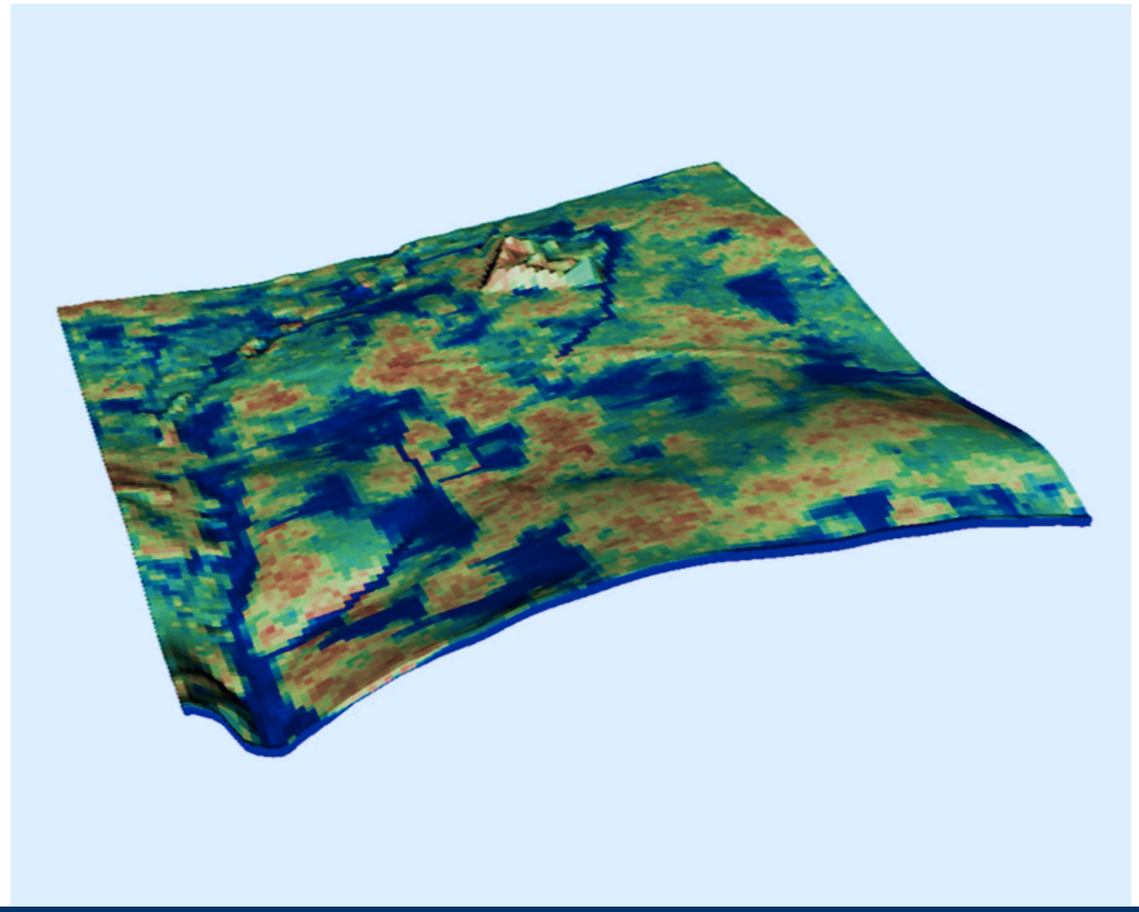

Inverse conditioning of a high resolution integrated terrestrial model at the hillslope scale: the role of input data quality and model structural errors

Sebastian Thomas Gebler

Energie & Umwelt / Energy & Environment
Band / Volume 444
ISBN 978-3-95806-372-3

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
Forschungszentrum

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment Band / Volume 444

Forschungszentrum Jülich GmbH
Institut für Bio- und Geowissenschaften
Agrosphäre (IBG-3)

Inverse conditioning of a high resolution integrated terrestrial model at the hillslope scale: the role of input data quality and model structural errors

Sebastian Thomas Gebler

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment Band / Volume 444

ISSN 1866-1793 ISBN 978-3-95806-372-3

Bibliografische Information der Deutschen Nationalbibliothek.
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte Bibliografische Daten
sind im Internet über http://dnb.d-nb.de abrufbar.

Herausgeber und Vertrieb: Forschungszentrum Jülich GmbH
Zentralbibliothek, Verlag
52425 Jülich
Tel.: +49 2461 61-5368
Fax: +49 2461 61-6103
zb-publikation@fz-juelich.de
www.fz-juelich.de/zb

Umschlaggestaltung: Grafische Medien, Forschungszentrum Jülich GmbH

Druck: Grafische Medien, Forschungszentrum Jülich GmbH

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment, Band / Volume 444

D 82 (Diss., RWTH Aachen University, 2017)

ISSN 1866-1793
ISBN 978-3-95806-372-3

Vollständig frei verfügbar über das Publikationsportal des Forschungszentrums Jülich (JuSER)
unter www.fz-juelich.de/zb/openaccess.

*A*nybody who has been seriously engaged in scientific work of any kind realizes that over the entrance to the gates of the temple of science are written *the words: 'Ye must have faith.'*

Max Planck

Contents

List of Figures

List of Tables

List of Acronyms

AWAT	Adaptive Window and Adaptive Threshold
BATS	Biosphere–Atmosphere Transport Scheme
CESM	Community Earth System Model
CLM	Community Land Model
COSMO-DE	Consortium for Small Scale Modelling – numerical weather prediction platform of the German Weather Service
DWD	Deutscher Wetterdienst (German Weather Service)
ET	Evapotranspiration
EC	Eddy Covariance
EnKF	Ensemble Kalman Filter
FAO	Food and Agriculture Organization of the United Nations
GCOSIM3D	Three Dimensional Multiple Variable MultiGaussian Conditional Sequential Simulation Program
KF	Kalman Filter
LAI	Leaf Area Index
LUA NRW	North Rhine-Westphalian State Environment Agency
LYS	Lysimeter
OASIS3-MCT	Ocean-Atmosphere-Sea-Ice-Soil coupler-Model Coupling Toolkit
PDAF	Parallel Data Assimilation Framework
PFT	Plant Functional Type
PTF	Pedo-Transfer-Function
SCE-UA	shuffle complex evolution algorithm
SWC	Soil Water Content
TB	Tipping Bucket
TERENO	Terrestrial Environmental Observatories
TFG	Terrain-Following Grid Transform
TOPMODEL	Topography based hydrological Model
TerrSysMP	Terrestrial System Modeling Platform
WMO	World Meteorological Organization

List of Symbols

a	[L^3 L^{-3}]	intercept constant [L^3 L^{-3}]
B	[-]	quotient of explained lysimeter mass variance
b	[-]	site specific wind exposition coefficient
c	[M L^2 T^{-2} θ^{-1}]	total volumetric heat capacity
E_g	[M T^{-2}]	soil/ground evaporation
E_v	[M T^{-2}]	vegetation transpiration
E_w	[M T^{-2}]	canopy evaporation from intercepted water
EBD	[M T^{-3}]	energy balance deficit
EF	[M T^{-3}]	evaporative fraction
ET_a	[M T^{-1}]	actual evapotranspiration
ET_{PM}	[L T^{-1}]	Penman-Monteith evapotranspiration
e_a	[M L^{-1} T^{-2}]	actual vapor pressure
$e°(T)$	[M L^{-1} T^{-2}]	saturation vapor pressure at a given air temperature
F	[*]	linear model operator
f_{trend}	[L^3 L^{-3}]	linear trend of soil moisture sensor
f_{ref}	[L^3 L^{-3}]	reference soil water for trend correction
G	[M T^{-3}]	soil heat flux
H	[*]	Matrix for mapping observation points to simulated states
H_g	[M T^{-3}]	sensible heat flux from soil
H_v	[M T^{-3}]	sensible heat flux from vegetated surface
HS	[M T^{-3}]	average heat storage
h	[L]	pressure head (vector)
h	[L]	pressure head (scalar)
h_{plant}	[L]	grass length
i	[-]	realization
j	[T]	time interval
K	[-]	Kalman Gain matrix

$\mathbf{K}_s$	[L T^{-1}]	saturated hydraulic conductivity (vector)
K_s	[L T^{-1}]	saturated hydraulic conductivity (scalar)
k	[-]	order of polynomial
k_k	[-]	von Karman's constant
k_r	[-]	relative permeability
L	[M T^{-1}]	amount of leachate water
$L(T)_{\mathrm{H_2O}}$	[L^2 T^{-2}]	vaporization energy at a given temperature
$\mathrm{LAI}_{\mathrm{act}}$	[-]	active leaf area of upper grass surface
LE	[M T^{-3}]	latent heat flux
LE_g	[M T^{-3}]	latent heat flux from soil
LE_v	[M T^{-3}]	latent heat flux from canopy
LE*	[M T^{-3}]	latent heat flux for a certain measurement point in time
m′	[L]	interfacial thickness
N	[-]	number of realizations
N	[-]	number of soil layers
n	[-]	pore size distribution
n_{man}	[L T$^{-1/3}$]	Manning's roughness coefficient
$\mathbf{P}$	[*]	covariance matrix of the forecast
P	[M T^{-1}]	precipitation
P^{cor}	[M T^{-1}]	corrected precipitation
ΔP	[M T^{-1}]	estimated precipitation deficit
q_o	[L T^{-1}]	surface runoff
$\mathbf{q}$	[L T^{-1}]	volume specific Darcy flux
q_d	[M T^{-2}]	surface drainage
q_e	[M T^{-1}]	source/sink for surface water and subsurface exchange
q_i	[L T^{-1}]	evaporation from intercepted water
q_o	[M T^{-2}]	surface runoff
q_r	[L T^{-2}]	rainfall rate
q_{rg}	[M T^{-2}]	runoff from glaciers, lakes or other surface types

q_{rl}	[M T^{-2}]	runoff from glaciers, lakes or other surface types
q_s	[T^{-1}]	source/sink term for pumping or injection
q_{so}	[M T^{-2}]	amount of solid precipitation
R	[*]	observation error covariance matrix
R_{spc}	[L^2 T^{-2} θ^{-1}]	specific gas constant for dry air
R^2	[-]	coefficient of determination
R_n	[M T^{-3}]	net radiation
r	[-]	number of measurements
r_a	[T L^{-1}]	aerodynamic resistance
r_i	[T L^{-1}]	stomatal resistance
r_s	[T L^{-1}]	(bulk) surface resistance
$S_{o,x}$	[-]	bed slope in x direction
$S_{o,y}$	[-]	bed slope in y direction
S_f	[-]	friction slope
S_s	[L^{-1}]	specific storage coefficient
S_w	[-]	relative saturation
S_{wst}	[M T^{-1}]	soil water storage
SWC_{tmax}	[L^3 L^{-3}]	maximum soil water content (trend correction)
SWC_t	[L^3 L^{-3}]	soil water content at time t
SWC_t^{corr}	[L^3 L^{-3}]	corrected soil water content at time t
s_{dat}	[M]	standard deviation of measured values
s_{res}	[M]	average residual
T_s	[θ]	soil/snow temperature
T	[θ]	air temperature
T_v	[θ]	virtual temperature
t	[T]	time (step)
$t_{97.5}$	[-]	Student t value for the 95 % confidence level
u_2	[L T^{-1}]	average wind speed at 2 m height

W_c	[M T^{-2}]	canopy water
W_s	[M T^{-2}]	snow water equivalent
w_i	[T]	adaptive window width
w_l	[M T^{-2}]	soil water
w_{ic}	[M T^{-2}]	amount of soil ice
$w_{\min}$	[T]	minimum window width
$w_{\max}$	[T]	window width maximum (AWAT)
$\mathbf{X}$	[*]	augmented state vector of model states and parameters
$\mathbf{x}$	[*]	model state vector
$\mathbf{x}^{\boldsymbol{a}}$	[*]	updated state vector
$\mathbf{x}_{\boldsymbol{s}}$	[*]	updated state vector
$\mathbf{x}_{\boldsymbol{\rho}}$	[*]	updated state vector
$\mathbf{Y}$	[L T^{-1}]	logarithm of soil hydraulic conductivities $\log_{10}$ $\mathbf{K}_s$ (vector)
$\mathbf{y}$	[*]	vector of measurements
$\boldsymbol{z}$	[L]	positive downward vertical coordinate
z_{m}	[L]	height of the wind measurement
z_{h}	[L]	height of the humidity measurement
z_s	[L]	depth from soil surface

*G*reek *L*etters

α	[L^{-1}]	inverse air entry suction
γ	[M]	lysimeter mass measured
$\hat{\gamma}$	[M]	fitted lysimeter mass
$\bar{\gamma}$	[M]	mean measured lysimeter mass
γ_{psy}	[M L^{-1} T^{-2} θ^{-1}]	psychrometric constant
Δ	[M L^{-1} T^{-2} θ^{-1}]	slope of the saturated vapour pressure curve
δ_i	[M]	adaptive threshold
$\delta_{\min}$	[M]	minimum provided threshold in AWAT
$\delta_{\max}$	[M]	maximum provided threshold in AWAT

ϵ	[-]	empiric precipitation type coefficient
$\boldsymbol{\varepsilon_y}$	[*]	observation error
$\boldsymbol{\varepsilon_{yi}}$	[*]	random observation error for each ensemble realization
$\varepsilon_{\mathrm{air}}$	[-]	ratio of molecular weight and water vapour (dry air)
$\varepsilon_{x,y}$	[-]	local slope angle
$\boldsymbol{\eta}$	[*]	Vector of model forcings
$\boldsymbol{\theta}$	[L^3 L^{-3}]	soil water content (vector)
θ_s	[L^3 L^{-3}]	saturated water content
θ_r	[L^3 L^{-3}]	residual saturation
$\bar{\theta}$	[L^3 L^{-3}]	mean soil water content
σ_θ	[L^3 L^{-3}]	spatial standard deviation of soil water content
λ	[M^1 L^1 T^{-3} Θ^{-1}]	thermal conductivity
$\boldsymbol{\nu}_{x/y}$	[L T^{-1}]	sectional velocity vector averaged over depth in x or y direction
$\bar{\boldsymbol{\nu}}$	[L T^{-1}]	depth averaged velocity vector
$\boldsymbol{\rho}$	[*]	model parameter vector
ρ_{H_2O}	[M L^{-3}]	density of water
φ	[-]	porosity
Ψ_s	[L]	surface ponding depth
$\boldsymbol{\nabla}$	[L^{-1}]	Nabla operator

* unit depends on application

Abstract

Understanding the soil-vegetation-atmosphere continuum is essential to improve hydrological model predictions. Particularly the characterization and prediction of the spatio-temporal variability of soil water content (SWC) and its controlling factors are of high interest for many geoscientific fields, since these patterns influence for example the rainfall-runoff response and the partitioning of the net radiation into latent and sensible heat fluxes while interacting with the vegetation cover. Within this context, this PhD thesis explores the degree of model complexity that is necessary to adequately represent heterogeneous subsurface processes, and the benefit of merging soil moisture data with an integrated terrestrial model. This includes an uncertainty analysis of model forcing (i.e. precipitation) and evaluation data (actual evapotranspiration). On this account, the fully coupled land surface-subsurface model ParFlow-CLM, which is part of the terrestrial system modeling platform (TerrSysMP), was applied to the 38 ha Rollesbroich headwater catchment located in the Eifel (Germany). Detailed long-term data for model setup, calibration, and evaluation were provided by the TERENO infrastructure initiative, the North Rhine-Westphalian State Environment Agency, and the Transregional Collaborative Research Center 32. It was expected that this combination of process orientated model and extensive observation data contributes to the understanding of the complex processes of the energy and water cycle at the hillslope, the elementary unit for the runoff generation process.

The first part compared different measurements of actual evapotranspiration (ET_a) obtained by a set of six weighable lysimeters, ET_a estimates acquired with the eddy covariance (EC), and evapotranspiration calculated with the full-form Penman-Monteith equation (ET_{PM}) for the Rollesbroich site. This comparison of ET_a included a correction of the energy balance deficit for the EC method, which is not often reported in literature and allows a deeper insight into the performance of both methods. The evaluation was conducted for the year 2012 and showed that both measurement methods are in good agreement with a total difference of 3.8 % (19 mm). The smallest relative difference (< 8 %) between the monthly ET_a of lysimeters and EC was found in summer. Both ET_a estimates were close to the ET_{PM} which was used as reference for the observations. This indicates that ET_a was energy limited but not limited by water availability during the entire evaluation period. The ET_a difference between lysimeter and EC were found to be mainly related to differences in grass height due to harvest and EC footprint. The lysimeter data further were used to estimate precipitation in combination with a filter and smoothing algorithm for high-precision lysimeters by Peters et al. (2014). The comparison with the on-site standard tipping bucket rain gauge revealed that precipitation amounts from the lysimeter were 16 % higher for 2012. Applying the correction method of Richter (1995) reduced this amount to 3 %. With the help of an on-site camera system, it was shown that dew and rime contributed 17 % to the total yearly difference between both measurement methods, while fog and drizzling rain explained another 5.5 %. During snowfall, the tipping bucket underestimated precipitation (7.9 % of the total difference). The largest proportion of the total yearly difference (36 %) was associated with snow cover without apparent snowfall. The strong overestimation of precipitation under these conditions most likely can be explained with snow drift and snow bridges. The remaining 33 % of the total yearly precipitation difference could not be explained.

In the second part, the results of TerrSysMP simulations with the 3D variably saturated groundwater flow model (ParFlow) and the Community Land Model (CLM) are compared to soil water content measurements from a wireless sensor network, actual evapotranspiration recorded by lysimeters and EC, and discharge. Therefore, the ParFlow-CLM model with a lateral resolution of 10 × 10 m and a variable vertical resolution (0.025-0.575 m) was tested with different subsurface parametrization strategies: (i) completely homogeneous, (ii) homogeneous parameters for different soil horizons, (iii) different parameters for each soil unit and soil horizon and (iv) heterogeneous stochastic realizations. The soil hydraulic parameters for these simulations were either (i) sampled from measured soil texture and the Rosetta pedotransfer functions (ROS), or (ii) estimated using a 1D subsurface model in combination with shuffle complex evolution (SCE). The results showed that spatially heterogeneous soil hydraulic properties in combination with topography dominate the

spatial variability of SWC at the hillslope scale. Whereas topography in combination with homogeneous soil hydraulic properties underestimated the observed spatial SWC variability, the full SWC variability could be explained with a heterogeneous distribution of soil hydraulic parameters. All model simulations reproduced the seasonal evapotranspiration, while the discharge showed a high bias compared to the observations. This poor performance of the uncalibrated model could be explained with a lack of information about the bedrock characteristics and the on-site drainage system. In general, a better performance was found for the simulation of the SCE setup with interpolated 1D inverse parameter estimates. Showing a higher inverse air entry parameter within the van Genuchten subsurface parameterization, the SCE-setup was in better correspondence with the observed SWC during dry periods than the simulation of the ROS-setup. This demonstrates the difficulty to transfer small scale local measurements into larger scale.

In the third part, SWC data was assimilated with Ensemble Kalman Filter (EnKF) into a Parflow-CLM Model of the Rollesbroich study site. Dense information from a soil moisture sensor network was available at the site and assimilated with the EnKF in a coupled land surface-subsurface model with 10 m lateral resolution. Different data assimilation scenarios with the EnKF were applied including the update of model states with or without updating of model parameters as well as different time update intervals for parameters and states. The ensemble size was 128 (or 256) with 3D- heterogeneous fields of saturated hydraulic conductivity and Mualem-van Genuchten parameters α and n. Furthermore, simulations were also carried out with a synthetic test case mimicking the Rollesbroich site, with the aim to get more insight in the role of model structural errors. The results of the real-world scenarios showed that the exclusive update of soil hydraulic properties was not enough to improve the SWC characterization. On average the RMSE at the sensor locations was reduced by 14 % if states and parameters were updated jointly and discharge estimation was not improved significantly. Simulations for the synthetic case showed much better results with an overall RMSE reduction by 55 % at independent verification locations in case of daily SWC-data assimilation including parameter estimation. Individual data assimilation scenarios with parameter estimation also showed NSE-increase for discharge from -0.04 for the open loop run to ca. 0.61. This shows that data assimilation in combination with high-resolution physically based models can potentially strongly improve soil moisture and discharge estimation at the hillslope scale. The large differences in the performance for the synthetic case and the real-world case can be explained with model structural errors like the representation of the on-site drainage system. Further studies should show whether the role of model structural errors was site specific, or that in general small scale processes play a more dominant role which are not captured by the modelling approach.

In summary, these studies show that integrated terrestrial high-resolution models at the hillslope scale are a valuable tool to better understand the basic hydrological processes. However, on the way to a predictive application of high resolution models many challenges remain. For more complex models a larger parameter space has to be explored (e.g. by additional soil hydraulic properties) in combination with multiple observation types. Further, an improved representation of significant processes (e.g. preferential flow) will contribute to the benefit of data-fusion techniques. On this account, high-resolution integrated terrestrial models can be used to identify significant information gaps in the existing measurement infrastructure.

Zusammenfassung

Die Weiterentwicklung von hydrologischen Modellvorhersagen erfordert ein tieferes Verständnis des Kontinuums aus Boden, Vegetation und Atmosphäre. Besonders die Charakterisierung und Vorhersage der hochgradig variablen zeitlichen und räumlichen Skalen des Bodenwassergehalts und dessen Kontrollmechanismen sind hierbei von großer Bedeutung auf vielen Gebieten der Geowissenschaften. Der Bodenwassergehalt interagiert mit der Vegetation und hat unter anderem großen Einfluss auf das Verhältnis von Niederschlag und Abfluss sowie auf die Aufteilung der Nettostrahlung in latente und sensible Wärme. Diese Dissertation beschäftigt sich mit der Frage, wieviel Modelkomplexität notwendig ist um die heterogenen Prozesse im Boden adäquat darzustellen. Des Weiteren wird untersucht welchen Nutzen die Assimilierung von Bodenfeuchtedaten in Verbindung mit einem integrierten terrestrischen Model hat. Das beinhaltet auch eine Unsicherheitsanalyse des Niederschlags sowie der aktuellen Verdunstung. Zu diesem Zweck wurde das 38 ha große Einzugsgebiet Rollesbroich, ein Graslandstandort in der Eifel (Deutschland), mit Hilfe des gekoppelten Landoberflächen- Bodenmodells ParFlow-CLM welches Bestandteil der terrestrischen System-Modellierungs-Plattform TerrSysMP ist, modelliert. Für das Einzugsgebiet standen Messdaten der TERENO Forschungsinfrastruktur Initiative, des Nordrheinwestfälischen Umweltministeriums, sowie des Transregional Collaborative Research Center 32 zur Verfügung. Von dieser Kombination aus einem physikalisch basierten Model und umfangreichen Beobachtungsdaten wird ein verbessertes Verständnis der Wasser- und Energiekreislaufs auf der Hangskala erwartet.

Im ersten Teil der Arbeit wurde die tatsächliche Evapotranspiration (ET_a), gemessen von sechs wägbaren Lysimetern, mit der ET_a einer Eddy Kovarianz Station und der nach der Penman-Monteith Gleichung berechneten Verdunstung (ET_{PM}) verglichen. Dieser Vergleich, der auch eine Korrektur des Energiebilanzdefizites für die EC-Methode enthält, wurde bisher nur in wenigen Studien untersucht und erlaubt genauere Erkenntnisse über das Potenzial beider Methoden. Die Auswertung der ET_a für das Jahr 2012 zeigt, dass beide Methoden mit einer absoluten Differenz von 3.8 % (19 mm) eine hohe Übereinstimmung zwischen den Messwerten aufweisen. In den Sommermonaten war die Übereinstimmung mit einer relativen Differenz (<8 %) besonders hoch. Die ET_a beider Methoden stimmte dabei gut mit der berechneten ET_{PM} überein, die als Referenzwert für die Messungen diente. Damit wird deutlich, dass ET_a während des gesamten Beobachtungszeitraumes von der Energieverfügbarkeit und nicht von der Wasserverfügbarkeit limitiert war. Unterschiede zwischen Lysimeter und EC wurden hauptsächlich von unterschiedlichen Graslängen nach der Mahd und dem EC Footprint verursacht. Weiterhin untersuchte die Studie das Potenzial der Lysimeterdaten zur Ermittlung des Niederschlags in Kombination mit einem Filter- und Glättungsalgorithmus von Peters et al. (2014). Im Vergleich mit einem herkömmlichen Kippwagen-Niederschlagsmesser des Untersuchungsgebiets für das Jahr 2012 zeigten die Lysimeter einen 16 % höheren Niederschlag. Diese Differenz konnte durch die Anwendung der Korrekturmethode von Richter (1995) auf den Kippwagenniederschlag auf 3 % reduziert werden. Mithilfe eines Kamerasystems konnte gezeigt werden, dass Tau- und Reifbildung 17 % der jährlichen Differenz beider Messmethoden ausmachen, Nebel und Sprühregen erklären weitere 5.5 % der Differenz. Bei Schneefall unterschätzte das Kippwagen-Messgerät den Niederschlag stark, was 7.9 % der gesamten Differenz ausmachte. Der größte Anteil der gesamten jährlichen Differenz (36 %) war jedoch mit einer vollständigen Bedeckung der Lysimeter durch Schnee verbunden. Dabei lag ohne offensichtlichen Schneefall am Untersuchungsstandort eine starke Überschätzung des Niederschlags vor, die wahrscheinlich durch Schneedrift oder Schneebrücken verursacht wurde. Die verbleibenden 33 % der gesamten jährlichen Niederschlagsdifferenz zwischen den beiden Methoden konnte nicht erklärt werden.

Im zweiten Teil der Arbeit wurden die Ergebnisse von verschiedenen TerrSysMP Simulationen mit den Bodenfeuchtemessungen eines drahtlosen Sensorennetzwerkes, der ET_a aus Lysimeter- und EC-Daten sowie Abflussdaten verglichen. Die Simulationen wurden mit ParFlow welches 3D variabel gesättigten Grundwasserfluss simuliert, und dem Community Land Model (CLM) durchgeführt. Das Modell mit einer horizontalen Auflösung von 10 x 10 m und einer variablen vertikalen Auflösung (0.025-0.575 m) wurde mit verschiedenen Strategien der Bodenparametri-

sierung getestet: (i) vollständig homogen, (ii) jeweils homogene Parameter für verschieden Bodenhorizonte, (iii) unterschiedliche Parameter für jeden Bodentyp und Bodenhorizont und (iv) heterogene stochastische Realisationen. Die hydraulischen Parameter für diese Simulationen wurden entweder (i) aus untersuchten Bodentexturen durch die Rosetta Pedotransferfunktionen (ROS) ermittelt oder (ii) mithilfe eines 1D Bodenmodells in Kombination mit einem Shuffle Complex Evolution (SCE) Algorithmus berechnet. Die Ergebnisse zeigen, dass räumlich heterogene hydraulische Bodeneigenschaften in Kombination mit Informationen über die Topographie die räumliche Variabilität der Bodenfeuchte auf Hangebene dominieren. Während die Kombination aus Topographie mit homogenen hydraulischen Bodeneigenschaften die gemessene räumliche Variabilität der Bodenfeuchte unterschätzte, konnte die gesamte Bodenfeuchte-variabilität mit einer heterogenen Verteilung der hydraulischen Bodenparameter durch die Topographie erklärt werden. Alle Modellsimulationen waren in der Lage die saisonale ET_a abzubilden, der modellierte Abfluss andererseits zeigte starke Abweichungen im Vergleich zu den Messungen. Diese Schwäche des unkalibrierten Modelles konnte mit dem Fehlen von Informationen über die Charakteristik des anstehenden Gesteins und des Drainagesystems im Untersuchungsgebiet erklärt werden. Besonders in Trockenperioden erzielte der SCE-Setup bedingt durch einen höheren inversen Lufteintrittswert innerhalb der van Genuchten Boden-parametrisierung bessere Modellierungsergebnisse gegenüber dem alternativen ROS-Setup. Dieses Ergebnis verdeutlicht die Schwierigkeit, kleinskalige lokale Messungen auf eine größere Skala zu übertragen.

Im dritten Teil der Arbeit wurden Bodenfeuchtedaten aus dem dichten Sensornetzwerk in ein hochauflösendes (10 m laterale Auflösung) integriertes terrestrisches Modell (ParFlow-CLM) des Rollesbroich Untersuchungsgebiets mit dem Ensemble Kalman Filter (EnKF) assimiliert. Dabei wurden verschiedene Varianten der Zustands- und Parameteraktualisierung auf ein Ensemble von 128 (bzw. 256) Realisierungen mit vollständig heterogenen bodenhydraulischen Eigenschaften nach Mualem-van Genuchten angewendet. Diese Aktualisierungsvarianten des EnKF beinhalteten die Aktualisierung von Zustandsvariablen mit oder ohne Aktualisierung von Modellparametern sowie verschiedene zeitliche Aktualisierungsintervalle für die Parameter. Darüber hinaus gab der Vergleich von Simulationen mit realen Messungen und Simulationen mit synthetischen Beobachtungen die Möglichkeit strukturelle Fehler im Modell zu identifizieren. Die Ergebnisse der realen Szenarien zeigen, dass die exklusive Aktualisierung der bodenhydraulischen Eigenschaften nicht ausreichte, um die SWC-Charakterisierung im Verifizierungszeitraum zu verbessern. Im Durchschnitt wurde der RMSE der Sensorstandorte um 14 % reduziert falls Zustandsvariable und Modellparameter aktualisiert wurden. Der Abfluss im Modell wurde dabei nicht signifikant verbessert. Für die synthetischen Simulationen wurden deutlich größere Verbesserungen gefunden. So verbesserte sich der RMSE an unabhängigen Untersuchungspunkten um bis zu 55 % für einzelne Simulationsszenarien mit Parameter Aktualisierung. Weiterhin konnte für einzelne Szenarien eine Verbesserung des NSE Kriteriums im Abfluss von -0.04 auf 0.61 festgestellt werden. Das zeigt, dass DA in Kombination mit Informationen aus einem dichten Sensornetzwerk und einem integrierten terrestrischen Modell das Potential besitzt Bodenfeuchte auf Feldskala zu verbessern. Der große Unterschied in der Verbesserung zwischen realen und synthetischen Versuchsszenarien können dabei mit strukturellen Modellfehlern erklärt werden. Weitere Studien werden zeigen, ob es sich dabei um ein gebietsspezifisches Problem handelt, oder ob im Modell nicht implementierte kleinskalige Prozesse verantwortlich für die strukturellen Probleme sind.

Zusammenfassend zeigen die Studien, dass hochaufgelöst integrierte terrestrische Modelle auf der Hangskala ein wertvolles Instrument sind, um die grundlegenden hydrologischen Prozesse besser zu verstehen. Jedoch bleiben viele Herausforderungen im Hinblick auf eine prädiktive Anwendung von hochauflösenden Modellen bestehen. Für komplexere Modelle muss ein größerer Parameterraum (z. B. durch zusätzliche bodenhydraulische Eigenschaften) in Kombination mit mehreren Beobachtungsarten erforscht werden. Ferner wird eine verbesserte Darstellung von signifikanten Prozessen (z. B. präferentielle Strömung) zum Nutzen von Techniken der Datenassimilierung beitragen. Vor diesem Hintergrund können hochauflösende integrierte terrestrische Modelle verwendet werden, um signifikante Informationslücken in der bestehenden Messinfrastruktur zu identifizieren.

Chapter 1 Introduction

1.1 Motivation and Background

Reliable assessment of the impact of climate change and demographic pressure on the terrestrial hydrological cycle at the local, regional and global scale is a major challenge in environmental research in the 21st century. This includes many aspects of people's everyday life, for example agricultural management, the assessment of freshwater resources and contaminant transport, as well as hazard prediction. These aspects can be assessed with a variety of stochastic and process based models. However, modeling the hydrological cycle is challenging since it shows complex interactions with the energy and biogeochemical cycles on a wide range of spatio-temporal scales in the subsurface, land-surface, and atmosphere (Gentine et al., 2012). These processes typically act on temporal scales varying between seconds and years and spatial scales varying between centimeters and thousands of kilometers (Blöschl and Sivapalan, 1995; Brutsaert, 2010). High spatio-temporal process heterogeneity is already found on the local and field scale. Atmospheric turbulence occurs at scales of meters and seconds, whereas overland flow and infiltration processes typically develop at spatial scales of magnitudes of 10^1 - 10^2 m and temporal scales of minutes to hours. Typical temporal scales for saturated and unsaturated flow processes even range from months to years (Blöschl and Sivapalan, 1995; Brutsaert, 2010; Gentine et al., 2012). The interdependence between water and energy processes across various time scales has been recognized in many studies (e.g. York et al., 2002; Yeh and Eltahir, 2005; Bierkens and Van den Hurk, 2007). Particularly, the water in the soil vadose and saturated zone was detected to be a key player within the system (Western et al., 1998; Vereecken et al., 2008; Vereecken et al., 2016). Water in soils controls the rainfall-runoff-response (Grayson et al., 1997; Robinson et al., 2008), the partitioning of net radiation in sensible, latent and ground heat flux, and various processes through the interaction between soil moisture and vegetation (e.g. plant growth, agricultural production) (Rodriguez-Iturbe et al., 2001; Western et al., 2002). Figure 1.1 illustrates the non-linear interactions between the different terrestrial compartments relevant for the hydrological cycle across different temporal scales. The incoming solar radiation is the main driver for the surface radiation budget (SRB) and interacts with soil and vegetation cover which controls the distribution of latent and sensible heat flux in the atmospheric boundary layer. This impacts the atmospheric convection and local cloud formation, which in turn affects incoming radiation. This example shows that more accurate weather and climate models require an advanced knowledge of hydrological and land surface processes taking into account also the small-scale interactions between the different compartments of soil, vegetation and atmosphere at the hillslope (Wood et al., 2011).

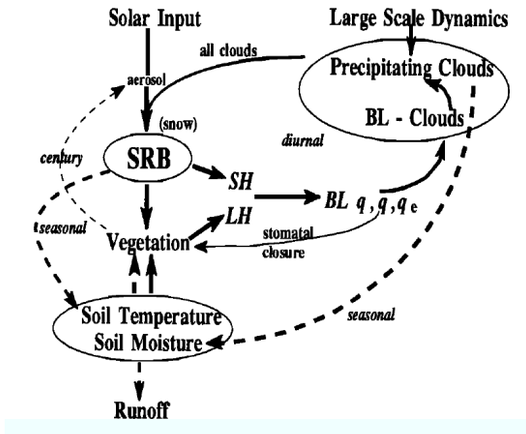

Figure 1.1: Schematic of a feedback loop within the terrestrial hydrological cycle (taken from Betts et al., 1996). Interactions of the surface radiation budget (SRB) and the atmospheric boundary layer (BL) via latent (LH) and sensible heat (SH) flux are illustrated. Water and heat fluxes within the boundary layer are represented by q, and q_e. Different temporal scales (diurnal-century) are indicated by the line style.

For the simulation of the hydrological cycle in general two contrasting model concepts exist, stochastic and process-orientated models (Brutsaert, 2010). Stochastic models rely on a black box system making use of the empirical relationships between model input (e.g. rainfall) and the output (e.g. runoff). Stochastic models are well suited for hydrological problems with limited input data and sparse information about the underlying physical processes, but are limited in terms of non-stationary effects triggering hydrologic changes which result for example from deforestation, urbanization or climate change (Brutsaert, 2010). On the contrary, process-orientated or deterministic models represent the physical processes (runoff, subsurface flow, groundwater flow, evapotranspiration) in more detail, like in the form of partial differential equations. These models rely on verified mathematical descriptions (i.e. Richard‘s equation for vadose zone flow) and relevant spatial distributed inputs (e.g. topography, vegetation and soil properties). However, the pure determinism of these models is also a disadvantage since the prediction ability strongly depends on the input data or, respectively, the available information on the model domain. Physically based integrated models gained increasing attention within the scientific community during the last two decades due to the advances in computer power and numerical methods. These models are not restricted to an individual compartment of the hydrological cycle and represent the relevant surface and subsurface processes of the terrestrial cycle in a fully coupled fashion. This includes coupled models for groundwater, land-surface and overland flow (e.g. AquiferFlow-SiB2 (Tian et al., 2012), CATHY (Bixio et al., 2002; Camporese et al., 2010), MikeShe (Abbott et al., 1986; Graham and Butts, 2005), HydroGeosphere (Therrien et al., 2010), ParFlow-CLM (Kollet and Maxwell, 2006; Maxwell and Miller, 2005), PIHM (Kumar et al., 2009; Qu and Duffy, 2007),

tRIBS+VEGGIE (Ivanov et al., 2008)) which can be additionally coupled to an atmospheric model (e.g. TerrSysMP (Shrestha et al., 2014)). Some modelling studies have already illustrated the importance of including 3d subsurface flow for feedbacks between subsurface, land-surface and atmosphere at the regional scale (e.g. Kollet and Maxwell, 2008; Ferguson and Maxwell, 2010; Tian et al., 2012; Miguez-Macho and Fan, 2012; Guay et al., 2013; Shrestha et al., 2014; Bonetti et al., 2015). By taking into account the interactions of the different terrestrial compartments, these models can be a valuable tool for process understanding, in particular for the combination of coupled high resolution subsurface and land-surface models at the hillslope scale. Specifically, the controlling mechanisms of soil moisture variability (e.g. Herbst et al., 2006; Cornelissen et al., 2014), rainfall runoff response (e.g. Herbst and Diekkrüger, 2003; Sciuto and Diekkrüger, 2010), and vegetation (e.g. Ivanov et al., 2010; Fatichi et al., 2015) can be better investigated. The fast increase in available computing power and high-performance parallel computing during the last decade allows increasing the model resolution at the hillslope scale (e.g. Cornelissen et al., 2014; Gatel et al., 2016). According to Wood et al. (2011), fully coupled process based model simulations in combination with high model resolution close to the process scale are key features in the future development of hydrological modelling. However, this effort creates also a number of challenges in terms of data need, characterization of spatial heterogeneity and parameter estimation which jointly contribute to the model uncertainty (Fatichi et al., 2016).

Process based coupled surface-subsurface models are very demanding in terms of the needed amount of input data. This concerns initial and boundary conditions as well as model forcings and model parameters. Model input is uncertain because limited information is available. Measurements which could help to constrain the model input are only available for selected points, affected by measurement errors. For example, observations of precipitation, a major component of the hydrological cycle, are still quite uncertain despite a long history of measurements (Kohnke et al., 1940; Strangeways, 2010). Correct estimates of precipitation are essential for any kind of modeling purpose (Bárdossy and Das, 2008; Stisen et al., 2012). Precipitation measurements are typically prone to a wide range of systematic errors. The accuracy of gauge based precipitation estimates is affected by wind induced catch deficiencies, evaporation, instrument calibration, splashing and malfunctions of the device (Ciach, 2003; Kampf and Burges, 2010). This may lead to a substantial precipitation under-catch (5-40 %) depending on the season and instrumental setup (Groisman et al., 1994). Precipitation undercatch related to solid precipitation (Nešpor and Sevruk, 1999) has a high impact on the water balance due to the fact that most stream flow and groundwater recharge in the mid latitudes commonly take place during winter and spring (Stisen et al., 2012). Also the inadequate representation of the spatial precipitation distribution is an issue for model calibration and application (Bárdossy and Das, 2008) and may lead to a high model bias in rainfall runoff models (Oudin et al., 2006). Paschalis et al. (2014) pointed out that even small scale rainfall patterns may have great effect on runoff production.

Another important source of uncertainty besides the model forcings are the model parameters. This is in particular the case for subsurface models. Soil hydraulic properties have been shown to be a source of large uncertainty in numerous studies (e.g. Gupta et al., 1998; Christiaens and Feyen, 2000; Christiaens and Feyen, 2001; Liu et al., 2005). In many cases soil hydraulic properties cannot be determined directly via measurements and are estimated from semi-empirical relationships. Christiaens and Feyen (2001) investigated the impact of different soil hydraulic properties calculated with different methods (lab measurements, semi-empirical texture relationships, pedo-transfer-functions, neural network approach), on the simulation results with a process-based model. They found that the uncertainties of the different parameters derived with the aforementioned methods range from 3-700 %. Particularly saturated hydraulic conductivity and residual soil water content were found to be very uncertain regarding their estimation method. Simulated discharge, groundwater depth and soil moisture content showed large differences depending on the input. Gutmann and Small (2007) compared the simulation results of two alternative model subsurface parameterizations originating either from empirical relationships with textural classes or inverse modelling. They concluded that the model parameterization by inverse optimization outperformed its alternative as latent heat flux was better reproduced. These examples illustrate that uncertainty with respect to model forcings and model parameters is important and has to be addressed by inverse modeling or data assimilation.

Many studies have shown that model resolutions closer to process scale improved the modelling of water and heat fluxes (e.g. Sulis et al., 2011; Mascaro et al., 2015; Shrestha et al., 2015; Cornelissen et al., 2016) since topography and subsurface heterogeneity were better represented. A good opportunity to investigate the effects of heterogeneity is modelling at the hillslope scale. This allows model simulations at a high spatial and temporal resolution. Given the smaller scale of the study, the computational costs are still affordable and data availability for some hill slopes might be good. For example, Maxwell and Kollet (2008) and Meyerhoff and Maxwell (2011) found a large effect of 3d heterogeneous soil hydraulic parameters on runoff, specifically the partitioning in runoff from infiltration excess (Hortonian flow) and base flow (Dunne flow) for idealized hillslope experiments. Herbst et al. (2006) pointed out that runoff results became more accurate with a higher level of subsurface complexity. Atchley and Maxwell (2011) investigated the controlling mechanisms of soil water patterns at a synthetic hillslope. They concluded that spatial patterns of soil moisture are more sensitive to the spatial distribution of soil hydraulic conductivity than topography. This is in line with the findings of Gatel et al. (2016) from real world experiments with a high resolution model of a grass buffer strip. They pointed out that the spatial patterns of soil moisture are more determined by hydraulic conductivity than micro topography. Synthetic experiments with different levels of terrain, subsurface and vegetation heterogeneity by Rihani et al. (2010) showed that the land surface energy balance is dominated by vegetation cover whereas subsurface heterogeneity and terrain are controlling water table and stream flow. Even at this high

resolution conceptual model errors can still be more detrimental than parameter errors (Liu and Gupta, 2007). For example, preferential flow and sediment transport processes at the hillslope scale can be of non-local nature (McDonnell et al., 2007; Stark et al., 2009; Foufoula-Georgiou et al., 2010; Hale et al., 2016) and therefore be underrepresented in models at field scale resolution. In this case the predictive ability of the model might be limited.

Recently, the establishment of terrestrial observatories for long-term monitoring support scientists providing long-term data of various terrestrial compartments in spatial and temporal resolution. This infrastructure is built on the basis of partnerships between study sites operated by different institutions on global level (e.g. FLUXNET (Baldocchi et al., 2001)), or continental and national level (e.g. the European Integral Carbon Observation System (ICOS); US Soil Climate Analysis Network (SCAN) (Schaefer et al., 2007); US Snowpack Telemetry network (SNOTEL) (Serreze et al., 1999), Terrestrial Environmental Observatories (TERENO) in Germany (Zacharias et al., 2011)). This infrastructure gives the possibility to combine different data types with various spatial supports for informing numerical simulation models. Furthermore, the observation accuracy can be better assessed with cross-checks on the basis of several repetitions (e.g. Hendricks Franssen et al., 2010; Kessomkiat et al., 2013) or different types of measurement devices (e.g. López-Urrea et al., 2006; Vaughan et al., 2007). For example, measurements of actual evapotranspiration are made with hydrological approaches (soil water balance, weighable lysimeters), micrometeorological methods (e.g. Bowen ratio, eddy covariance method), plant physiological approaches (e.g. sap flow, plant chambers), or remote sensing (Allen et al., 1998; Rana and Katerji, 2000). One of the most reliable methods is a weighable lysimeter which directly estimates actual evapotranspiration from mass balance considerations (Rana and Katerji, 2000). Observations of weighable lysimeters are typically prone to errors associated with wind-induced mechanical vibrations. Furthermore, differences in thermal regime and/or vegetation cover between the lysimeter and its surroundings contribute to the uncertainty of the lysimeter measurements (Zenker, 2003). The standard error of a lysimeter is estimated 10-20 % for daily actual ET estimates (Pruitt and Lourence, 1985; Rana and Katerji, 2000). As an alternative the eddy covariance technique can be applied, which estimates vertical moisture fluxes (and therefore evapotranspiration) via the covariance of water vapor density and vertical wind speed. A well-known problem of this method is the imperfect closure of the energy balance. Energy balance deficits up to 25 % were reported in many studies (e.g. Wilson et al., 2001; Hendricks-Franssen et al., 2010; Imukova et al., 2016). Although the energy balance deficit (and therefore the evapotranspiration estimate) can be corrected on the basis of the Bowen ratio or the evaporative fraction at the site (e.g. Wohlfahrt et al., 2010; Kessomkiat et al., 2013; Wohlfahrt and Widmoser, 2013; Imukova et al., 2016), it is controversially discussed how the energy balance deficit should be divided over the latent or sensible heat fluxes (Imukova et al., 2016).

The increasing availability of measurement data and the new infrastructure allow to better constrain process-based models. Different methods can be used to condition predictions with process-based models to measurement data in a stochastic framework. Examples are the generalized likelihood uncertainty estimation (GLUE) approach (Beven and Binley, 1992), the shuffled complex Evolution (SCE) (Duan et al., 1992), the SCE metropolis extension (SCE-UA) by Vrugt et al. (2003) or the dynamic identifiability analysis framework (Wagener and Gupta, 2005). These methods rely on statistical interference with Bayes' theorem estimating the posterior probability density function via Markov Chain Monte Carlo random sampling (Hastings, 1970). However, Markov Chain Monte Carlo requires a large number of model runs (10^5-10^6) for estimating the model uncertainty and for deriving an optimal parameter set. Hence, they are not suited for numerical expensive coupled models with a large parameter space. Another drawback is that uncertainty in model forcings and initial conditions is usually neglected. If a parameter is optimized under conditions where the model forcing was biased, also the estimated parameters are likely biased as they compensate for the bias in the model forcing (Stisen et al., 2012). As an alternative, sequential data assimilation (DA), specifically the Ensemble Kalman filter (EnKF) (Burgers et al., 1998; Evensen, 1994) and its variants (e.g. Ensemble Transform Kalman Filter (Bishop et al., 2001); Ensemble Square Root Filter (Tippett et al., 2003)), might be attractive methods for model calibration. These ensemble based methods allow considering parameter as well as forcing uncertainty at the same time. Each time observations are available, the model simulations are informed by the measurement data and model states (and probably also parameters) are updated by an equation which takes into account both measurement and model uncertainty. Updates are larger if the measurement uncertainty is small and the model prediction uncertainty large. The EnKF is easy to implement and the number of model runs is typically reduced to less than 1000, which is much smaller than the number of model runs required by Markov Chain Monte Carlo methods. Furthermore, EnKF is well suited for real time approaches (e.g. Hendricks Franssen et al., 2011). Numerous studies demonstrated the value of EnKF in terrestrial modelling for improving estimates of model states and parameters. Whereas groundwater models commonly assimilate pressure heads or groundwater temperature data (e.g. Hendricks Franssen and Kinzelbach, 2008; Hendricks Franssen et al., 2011; Kurtz et al., 2014), land-surface models assimilate soil moisture (e.g. Pauwels et al., 2001; Houser et al., 1998; Brocca et al., 2010, Lievens et al., 2014), land surface temperature (e.g. Ghent et al., 2010; Reichle et al., 2010; Han et al., 2013), brightness temperature (e.g. Crow and Wood, 2003; De Lannoy et al., 2007; Han et al., 2013, Lievens et al., 2016) or snow cover (e.g. Andreadis and Lettenmaier, 2006; Su et al., 2010; Xu and Shu, 2014). However, the majority of the numerous synthetic and real world studies relies either on a model representing a single compartment of the terrestrial cycle or makes strong simplifications regarding the subsurface component (e.g. 1d subsurface fluxes).

On the contrary, only a few studies with two way coupled surface-subsurface models in combination with the EnKF exist. Camporese et al. (2009) assimilated discharge and pressure head either individually or in combination for a synthetic tilted v-catchment with the aim to improve discharge and groundwater table characterization. Their model setup included components for subsurface and overland flow driven by an external upper boundary condition of rainfall and evapotranspiration. Camporese et al. (2009) found that discharge estimation can be improved already by assimilating one of the two data types, while the groundwater table could not be improved with the assimilation of discharge only. With a similar experimental setup Bailey and Baù (2012) improved the estimation of hydraulic conductivity within the test catchment. Ridler et al. (2014) presented a DA framework which is operating in combination with MikeShe including subsurface, land-surface and overland flow. They demonstrated its functionality for a catchment in Denmark. For the same catchment, Rasmussen et al. (2015) explored the relationship between ensemble size, number of observations and accuracy of the data assimilation. They showed that the ensemble size can be reduced if adaptive localization techniques are used which limit the influence of individual observations by setting the covariance to zero if the distance between observations and model grid cells is larger than a given critical distance. Shi et al. (2014) assimilated synthetic observations of discharge, water table depth, soil moisture, heat fluxes and land-surface temperature for a small headwater catchment (0.08 km^2). With a limited number of site specific observations, they showed that the estimated parameter values were close to the synthetic truth. Within this context, discharge, land surface temperature and soil moisture had more impact on the assimilation performance than the other assimilated variables. This result was also confirmed by Shi et al. (2015) in combination with real world data for the same catchment. Recently, Kurtz et al. (2016) introduced a framework which couples the Terrestrial System Modelling Platform (TerrSysMP) by Shrestha et al. (2014) with the Parallel Data Assimilation Framework (PDAF) developed by Nerger and Hiller (2013). They successfully demonstrated the capabilities of the EnKF for a virtual catchment with 20 million unknown states and parameters for coupled land surface and 3d-subsurface model simulations. These examples show that DA experiments with integrated land-surface and 3d subsurface flow models were restricted so far to mainly synthetic test cases. Furthermore, the impact of real world observation data on complex heterogeneous subsurface model structures has not been investigated within this context.

1.2 Objectives and Outline

Against this background, this work explores the uncertainties of a coupled land surface-subsurface model at the hillslope scale. Specifically, aspects of input data uncertainty, representation of heterogeneous soil hydraulic parameters and the advantage of merging soil moisture data with the coupled land surface-subsurface model are evaluated. These insights are relevant for future developments in hydrologic modeling and the representation of hillslopes in high resolution large scale physically based models. On this account, we apply ParFlow (Ashby and Falgout, 1996; Jones and Woodward, 2001; Kollet and Maxwell, 2006) and the Community Land Model (CLM) (Oleson et al., 2004; Oleson et al., 2008) at high lateral and vertical resolution to a headwater grassland catchment Rollesbroich located in the Eifel (Germany). ParFlow and CLM are fully coupled within the TerrSysMP framework (Shrestha et al., 2014) and take lateral subsurface flow, overland flow and topography into account. Long-term data for model setup, calibration and evaluation (e.g. discharge, soil moisture, evapotranspiration) at high spatial and temporal resolution are provided by the Terrestrial Environmental Observatories (TERENO) infrastructure initiative (Zacharias et al., 2011), the North Rhine-Westphalian State Environment Agency (LUA NRW) as well as Transregional Collaborative Research center (TR-32). We expect that a more realistic representation of land-surface and subsurface fluxes and feedbacks in combination with extensive observation data for model input and evaluation will contribute to a better understanding of the complex processes of the energy and water cycles at the hillslope scale.

Therefore we address the following research objectives:

(1) The estimation of uncertainty of actual evapotranspiration and precipitation measurements. Both major water balance components are essential for the reliable estimate of the water balance in a modelling approach;

(2) The exploration of the degree of model complexity that is necessary to adequately represent heterogeneous processes at the hillslope scale with a focus on soil moisture, discharge and evapotranspiration;

(3) The evaluation of the potential of the EnKF to improve the states and soil hydraulic parameters of a fully coupled land surface-subsurface model at the hillslope scale using a dense network of soil moisture sensors measuring in different depths.

Chapter 2 provides a short overview of the study site and the observations used for this PhD-thesis. Furthermore, the chapter briefly summarizes the relevant background theory of the ParFlow-CLM model, the technical features of its coupling to the PDAF module within the TerrSysMP framework as well as underlying theoretical aspects of the Ensemble Kalman filter.

Chapter 3 focuses on the measurement uncertainty of evapotranspiration and precipitation which are major drivers of the hydrological cycle. More specifically, this chapter is dedicated to the difference between actual evapotranspiration (ET_a) measured by the eddy covariance method (EC), and by a set of six weighable lysimeters. This comparison allows a deeper understanding of the uncertainty and performance of both methods, which includes also the well-known problem of energy balance closure for EC measurements. The data of the lysimeters are further used to estimate the precipitation of the Rollesbroich study site. This is achieved with a recently developed filter algorithm for high precision lysimeters by Peters et al. (2014). The obtained lysimeter precipitation is subsequently compared to the precipitation amount recorded by a standard tipping bucket device which is typically prone to errors induced by wind and evaporation loss in combination with the precipitation type (Richter, 1995; Sevruk et al., 2009). The insights of both analyses are used in chapter 4 and 5 for creating model forcing data (i.e. precipitation) as well as for evaluating model performance in terms of ET_a for the year 2012.

Chapter 4 investigates the prediction of spatial and temporal variability of soil moisture, discharge, and evapotranspiration with the coupled land surface-subsurface model ParFlow-CLM. Therefore, the ParFlow-CLM simulation results of the Rollesbroich catchment are compared with soil water contents measured by a wireless sensor network, measured ET_a, and discharge. The high resolution simulations are conducted with a horizontal lateral resolution of 10 × 10 m and a variable vertical resolution (0.025-0.575 m). To investigate the impact of heterogeneous subsurface structure four different parameterization strategies are tested: (i) completely homogeneous parameters, (ii) homogeneous parameters for different soil horizons, (iii) different parameters for each soil horizon and each soil unit, and (iv) heterogeneous stochastic realizations. Given the non-linearity of the water retention function and the variety of methods for deriving soil hydraulic properties, an assessment of the impact of alternative initial parameter datasets within the context of model complexity is essential. Thus, the soil hydraulic parameters for the simulations are either (i) sampled from measured soil texture and pedotransfer functions, or (ii) inversely estimated by shuffle complex evolution in combination with 1D subsurface models.

Chapter 5 describes the estimation of model states, fluxes and parameters of the Parflow-CLM Rollesbroich model with the help of the assimilation of soil moisture data from a soil moisture sensing network, by the EnKF. Specifically, the study explores the potential of the EnKF to improve soil water content as well as discharge and evapotranspiration. An ensemble of 128 (or 256) model realizations is generated with uncertain precipitation and uncertain, spatially heterogeneous soil and aquifer hydraulic parameters as input. Different scenarios are tested including the updating of model states with or without updating model parameters. In addition, also different time update intervals for parameters were evaluated as well as parameter damping. The

experiment was done for the real-world Rollesbroich site and a synthetic case which mimics the Rollesbroich site. This comparison gives more insight in the role of model conceptual errors.

Chapter 6 finally summarizes the main conclusions of this thesis and gives an outlook on requirements and methods for future research.

Chapter 2 Theory, Materials and Methods

2.1 Study Site and Measurement Setup

The Rollesbroich site (50° 37' 27" N, 6° 18' 17" E) is the study site in this work. It is located in the Eifel low mountains range and is part of the Terrestrial Environmental Observatories (TERENO) in Germany (Zacharias et al., 2011; Qu et al., 2016). Rollesbroich is a sub-catchment of the Rur river with an altitude ranging from 474 m to 518 m a.s.l.. The average slope at the Rollesbroich site varies between 0 and 10 %, exhibiting more flat terrain in the South and steeper topography in the North. The total catchment area is ca. 38 ha. Earlier studies (e.g. Qu et al., 2014) mentioned an area of the study site of 27 ha or ca. 31 ha (Gebler et al. 2015). However, a detailed analysis of the water balance and the flow patterns of the domain revealed a larger catchment size. The drainage system installed at the northeastern part of the study site as well as roadside ditches from the federal road, which represent the eastern bounds of the study site, additionally contribute ca. 7 ha to the catchment area. Rollesbroich has a humid temperate climate with a mean annual precipitation of 1033 mm and a temperature of 7.7 °C (period 1981-2001). These data were recorded with a meteorological station operated by the North Rhine-Westphalian State Environment Agency (LUA NRW) in ca. 4 km distance from the study site. The extensively managed grassland vegetation in Rollesbroich is dominated by smooth meadow grass (*Poa pratensis*) and perennial ryegrass (*Lolium perenne*). The prevalent soils according to the classification of the Food and Agriculture Organization (FAO) are Cambisols (gleyic) in the southern part, and Stagnosols in the northern part of the catchment (Qu et al., 2016) varying between 0.5 and 1.5 m thickness (Korres et al., 2010; Koyama et al., 2010). The underlying bedrock consists of sandstone and siltstone with a 0.1 - 0.5 m heavily weathered top layer (saprolite). A drainage system is installed close to the source of the Kieselbach at the study site which prevents flooding during periods with high groundwater tables. Depending on the position in the dentriform arrangement, the diameters of the ca. 80-year old clay pipes vary between 3 and 20 cm. Figure 2.1 shows a map of the study area and gives an overview of the installed measurement equipment relevant for this PhD thesis.

On-site precipitation measurements are conducted with a standard Hellmann type tipping bucket (TB) rain gauge (ecoTech GmbH, Bonn, Germany) with a resolution of 0.1 mm and a measurement interval of 10 minutes. The measurement elevation of 1 m above ground is in agreement with recommendations of the German weather service (DWD 1993) for areas with an altitude > 500 m a.s.l. with occasionally heavy snowfall (WMO standard is 0.5 m).

Other meteorological data are recorded at the micrometeorological tower (50° 37' 19" N, 6° 18' 15" E, 514 m a.s.l.) at the southern part of the study site. This includes incoming and outgoing longwave and shortwave radiation, wind speed, temperature, and air pressure. Radiation is recorded

with a four-component net radiometer (CNR-1, Kipp and Zonen, Delft, The Netherlands). Wind speed and direction are observed by a sonic anemometer (CSAT3, Campbell Scientific, Inc., Logan, USA) at 2.6 m height. A gas analyzer (LI7500, LI-COR Inc., Lincoln, NE, USA) measures specific humidity and also air pressure at 2.6 m above the ground surface and a second device of the same type is installed at 0.57 m. Air temperature is recorded by a HMP45C (Vaisala Inc., Helsinki, Finland) at 2.58m above the ground surface. Additionally the data of the micrometeorological tower were used to measure latent heat flux (evapotranspiration) and sensible heat flux with the eddy covariance (EC) technique at the study site.

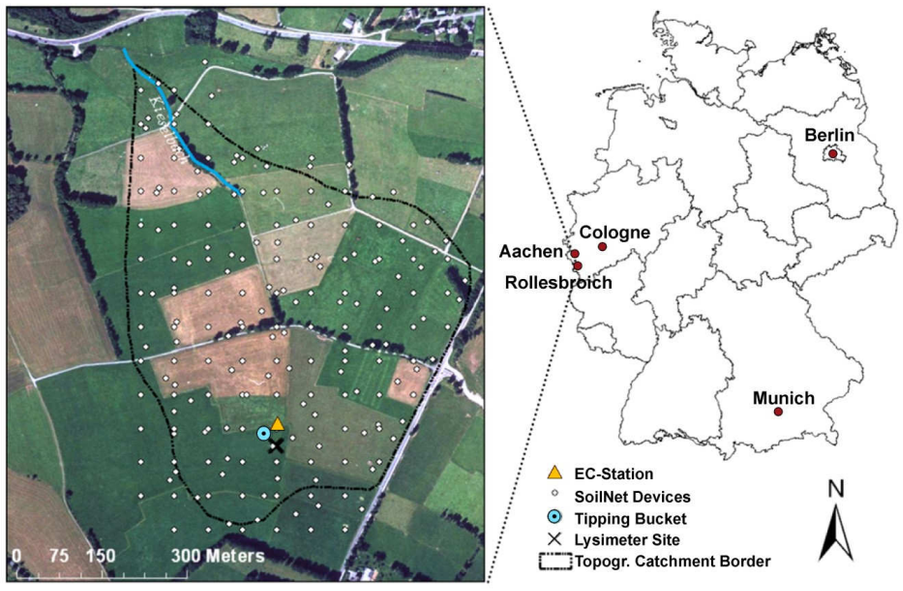

Figure 2.1: Overview of the Rollesbroich study site showing the locations of the precipitation gauge, the eddy covariance station, the lysimeters and the SoilNet devices installed at the study site. All devices are arranged within a radius of 50 meters at the southern part of the study site (modified from Gebler et al., 2015).

As an alternative to EC measurements, evapotranspiration was also recorded with a set of six weighable lysimeters (TERENO-SoilCan project, UMS GmbH, Munich, Germany) arranged in a hexagonal design ca. 30 m away from the micrometeorological tower. Each lysimeter contains an on-site silty loam soil profile and the 1 m² surface is covered with grass species. The lysimeters mimic the soil water regime of the surrounding subsurface with an automatically controlled lower boundary condition. Matric potential differences between each lysimeter and its surroundings are compensated on-demand by suction cup rakes (SIC 40, UMS GmbH, Munich, Germany), which

are able to remove or inject water from the lysimeter. Further technical details are provided by Unold and Fank (2008).

Soil water content and soil temperature measurements are performed with a wireless sensor network (SoilNet) installed at the study site (Qu et al., 2013; Qu et al., 2016). The 179 sensor locations at the Rollesbroich site contain six SPADE sensors (model 3.04, sceme.de GmbH i.G., Horn-Bad Meinberg, Germany) installed at 5, 20 and 50 cm depth. Two redundant sensors at each depth increase measurement precision and measurement volume to avoid data inconsistencies (e.g. contact issues with the soil matrix).

Discharge measurements for the on-site channel (Kieselbach) are conducted with a modified Venturi-Gauge Weir (Parshall flume) located nearby the catchment outlet and two upstream Tomson gauges (V-notches) close to the Kieselbach headwaters. Both discharge recordings are used in combined manner for model verification within a 10 min interval. Normal and high flow (18-1000 $m^3\ h^{-1}$) are optimally captured by the Parshall Flume, while V-notches better record low flow conditions (<18 $m^3\ h^{-1}$) (Qu et al., 2016).

Further details on the measurement devices installed at the study site as well as information about the data processing are provided in chapter 3-5 of this thesis.

2.2 TerrSysMP

The Terrestrial Systems Modelling Platform (TerrSysMP) (Gasper et al., 2014; Shrestha et al., 2014) consists of the regional atmospheric model (COSMO-DE) of the German Weather Service (Baldauf et al., 2011), the land-surface model CLM (Oleson et al., 2004; Oleson et al., 2008), and the subsurface model ParFlow (Maxwell and Miller, 2005; Kollet and Maxwell, 2006), which are coupled in a modular fashion using the external coupler OASIS3-MCT (Valcke, 2013; Valcke et al., 2013). In the work presented in this PhD-dissertation, we solely applied ParFlow and CLM within the TerrSysMP framework for this study, given the limited spatial extend of the Rollesbroich study site and the possibility of on-site meteorological measurements. In the following section, we present the governing equations for water and energy flow in the subsurface, and equations which govern the exchange of water and energy between the land and the atmosphere. Further, we provide a short description of the coupling between the individual subsurface and land-surface model components. For conducting data assimilation (Chapter 5), TerrSysMP was further extended by the Parallel Data Assimilation Framework (PDAF) (Nerger and Hiller, 2013) which was implemented into TerrSysMP by Kurtz et al. (2016) (Chapter 2.4).

2.2.1 ParFlow

Established by Ashby and Falgout (1996) at the Lawrence Livermore National Laboratory, ParFlow was designed to simulate large scale groundwater flow and contaminant transport in heterogeneous porous media by massively parallel computation. After the implementation of Newton–Krylov-multigrid solvers (Jones and Woodward, 2001) and a 2D overland flow boundary condition (Kollet and Maxwell, 2006), ParFlow has evolved to simulate transient subsurface flow problems with variable saturation, which are fully coupled to overland flow at a high temporal resolution (Kollet et al., 2010). Starting from predefined initial conditions, Parflow is driven by external boundary conditions and calculates subsurface water pressure as a function of time. Therefore, it solves the three-dimensional, transient Richards equation (Richards, 1931) in the mixed form (Eq. 2.1) using cell-centered finite differences in space:

$$S_s S_w(h) = \frac{\partial h}{\partial t}\,\varphi\,\frac{\partial S_w(h)}{\partial t} = \boldsymbol{\nabla}\cdot\mathbf{q} + q_s + \frac{q_e}{m'} \tag{2.1}$$

$$q = -\mathbf{K}_s k_r(h)\,\boldsymbol{\nabla}\,(h - z) \tag{2.2}$$

where q denotes the volume specific Darcy flux [L T^{-1}] (Darcy, 1856), S_s is the specific storage coefficient [L^{-1}], S_w the relative saturation [-], h the pressure head [L], t is time [T], z the positive downward vertical coordinate [L], φ the porosity [-], $\mathbf{K}_s$ the saturated hydraulic conductivity tensor [L T^{-1}], k_r the relative permeability [-], q_s a general source/sink term for pumping or injection [T^{-1}], q_e a general source/sink term representing exchange fluxes between surface water and subsurface [L T^{-1}], and m' the interfacial thickness [L].

Topography in ParFlow is implemented by an orthogonal structured grid (Maxwell and Miller, 2005) with a uniform slope for all grid cells. As an alternative, Maxwell (2013) introduced a terrain-following grid transform (TFG) together with a variable vertical grid discretization, which allows to reduce the model layers for study areas with large topographic height differences. This increases the computational efficiency as well as the accuracy of calculated fluxes at the surface water- subsurface interfaces since higher vertical resolutions can be specified for the unsaturated zone. The volume specific Darcy flux $\mathbf{q}$ then follows the local slopes and accounts for gravity contribution:

$$\mathbf{q}_{x,y} = -\mathbf{K}_s k_r(h)\,[\boldsymbol{\nabla}\,(h - z)\cos\varepsilon_{x,y}\,sin\,\varepsilon_{x,y}] \tag{2.3}$$

$$\varepsilon_x = tan^{-1}(S_{o,x}) \text{ and } \varepsilon_y = tan^{-1}(S_{o,y}) \tag{2.4}$$

where $\varepsilon_{x,y}$ [-] is the local slope angle and $S_{o,x}$ / $S_{o,y}$ are the bed slopes [-] in x or y direction.

The Mualem-van Genuchten relationships (1980) are used to describe relative saturation and permeability:

$$\theta(h) = \frac{\theta_s - \theta_r}{[1 + (\alpha\, h)^n]^{\left(1-\frac{1}{n}\right)}} + \theta_r \tag{2.5}$$

$$k_r(h) = \frac{\left\{1 - \frac{(\alpha\, h)^{n-1}}{[1 + (\alpha\, h)^n]^{\left(1-\frac{1}{n}\right)}}\right\}}{[1 + (\alpha\, h)^n]^{\frac{\left(1-\frac{1}{n}\right)}{2}}} \tag{2.6}$$

where α is the parameter for inverse air entry suction [L^{-1}], n is a factor related to pore size distribution [-], θ_s is the saturated water content [$L^3\ L^{-3}$], and θ_r is the residual saturation [$L^3\ L^{-3}$].

In ParFlow, subsurface and overland flow are coupled using a backward Euler scheme in time in combination with a free-surface overland flow boundary condition (Kollet and Maxwell, 2006). Shallow overland flow is represented by the kinematic wave approximation of the continuity equation and the momentum equation for two dimensions (Eq. 2.7). In the current formulation of the momentum equation, diffusion terms are neglected (Kollet and Maxwell, 2006). Thus, the momentum can be expressed in a condensed form (Eq. 2.8), where the bed slope (gravity forcing) term ($S_{o,i}$) [-] is equal to the friction slope term ($S_{f,i}$) [-].

$$\frac{\partial \Psi_s}{\partial t} = \nabla \bar{\mathbf{v}} \Psi_s + q_r(x) + q_e(x) \tag{2.7}$$

$$S_{f,d} = S_{o,d} \tag{2.8}$$

where Ψ_s is the surface ponding depth [L], t is time [T], $\bar{\mathbf{v}}$ the velocity vector averaged over depth [$L\ T^{-1}$], $q_r(x)$ the rainfall rate [$L\ T^{-1}$], and $q_e(x)$ represents the subsurface exchange rate. $S_{f,d}$ and $S_{o,d}$ [-] are the gravity forcing and friction (bed) slope terms with d indicating x and y directions.

Free surface flow is driven by gravity. The flow-discharge relationships for the x and y direction are expressed by Manning's equation.

$$\mathbf{v}_x = \frac{\sqrt{S_{f,i}}}{n_{man}} \Psi_s^{\frac{2}{3}} \quad \text{and} \quad \mathbf{v}_y = \frac{\sqrt{S_{f,i}}}{n_{man}} \Psi_s^{\frac{2}{3}} \tag{2.9}$$

where n_{man} is the Manning's coefficient [$L\ T^{-1/3}$] representing the surface roughness, Ψ_s is the surface ponding depth [L], $S_{f,i}$ is the friction slope term [-], and $\mathbf{v}_x$,and $\mathbf{v}_y$ are the cross sectional

velocity vector components in either x or y direction averaged over depth [L T^{-1}]. The Manning's roughness coefficient can be spatially heterogeneous (Maxwell and Miller, 2005).

ParFlow resolves streamflow without parameterized routing routines. Thus, overland flow develops gradually during model spinup through the interaction of the subsurface with a given topographic pattern (Maxwell and Miller, 2005; Kollet and Maxwell, 2008). Subsurface and overland flow are computed in globally implicit manner using the Newton-Krylov method with multigrid preconditioning as described in Jones and Woodward (2001).

2.2.2 CLM

The land surface compartment of TerrSysMP is the Community Land Model (CLM) v.3.5 (Oleson et al., 2004; Oleson et al., 2008). CLM is developed as part of the Community Earth System Model (CESM) and serves as upper boundary for ParFlow. It simulates energy and momentum transfer, as well as carbon fluxes between soil, vegetation and atmosphere and has specific modules accounting for biogeophysics, hydrology, biogeochemistry and vegetation dynamics. CLM calculates subsurface water and heat fluxes in 1D without lateral exchange between the individual grid cells, whereas ParFlow simulates subsurface water fluxes in 3D. Both models communicate via source/sink terms at the overlapping root zone layer represented by 10 soil layers of variable thickness and extent. This dynamic coupling allows a transient feedback in two ways. CLM extracts evapotranspiration from ParFlow and inserts infiltrating precipitation. ParFlow in turn updates the hydraulic pressure and soil water content for the subsurface horizons. The communication between both models is enabled via OASIS3-MCT (Ocean-Atmosphere-Sea-Ice-Soil coupler – Model Coupling Toolkit) (Valcke et al., 2013). The OASIS3-MCT library is a generic interface which exchanges predefined model variables between the individual model components. Therefore it establishes communication between the components in a parallel fashion which optionally takes into account interpolation and scaling operations during data transfer. A schematic overview of the ParFlow-CLM coupling within TerrSysMP can be found in Figure 2.2.

The total water balance in CLM according to Oleson et al. (2004) is:

$$\Delta W_c + \Delta W_s + \sum_{i=1}^{N}(\Delta w_{l,i} + \Delta w_{ic,i}) = q_{rl} + q_{so} - \mathrm{E}_v - \mathrm{E}_g - E_w - q_o - q_d - q_{rg})\Delta t \qquad (2.10)$$

where ΔW_c is the canopy water [M T^{-2}], ΔW_s is the snow water equivalent [M T^{-2}], $\Delta w_{l,i}$ is the soil water [M T^{-2}], $\Delta w_{ic,i}$ is the amount of soil ice [M T^{-2}], q_{rl} is the liquid precipitation [M T^{-2}], q_{so} is the solid precipitation [M T^{-2}], E_v is transpiration from vegetation, E_g is the soil evaporation [M T^{-2}], E_w is the evaporation from intercepted water on the canopy [M T^{-2}], q_o is the surface runoff [M T^{-2}], q_d is the surface drainage [M T^{-2}], q_{rg} is the runoff of glaciers [M T^{-2}], lakes or other surface types affected by snow [M T^{-2}], Δt is the timestep [T] and N [-] is the number of soil layers .

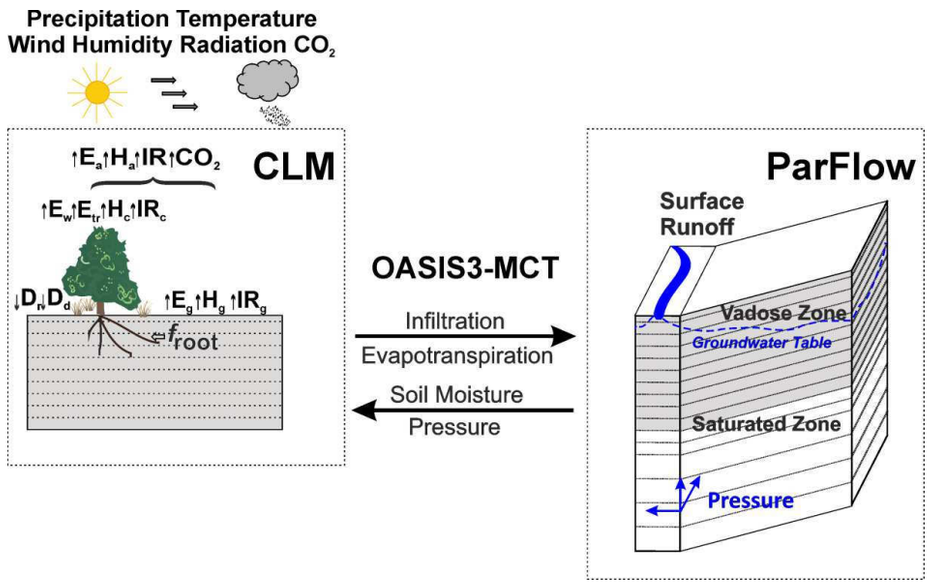

Figure 2.2: Coupling schematic of the TerrSysMP model components CLM (left) and ParFlow (right) via OASIS3-MCT. Water and heat fluxes between ground, canopy and atmosphere are calculated by CLM. This involves precipitation, radiation, and evapotranspiration partitioning as well as root water uptake among others. D_d and D_r exemplarily indicate precipitation directly falling through the canopy and the drainage from stem and foliage. IR symbolizes the outgoing long wave radiation from the ground (IR_g) and foliage (IR_c). Processes of subsurface vadose zone, saturated zone and the surface runoff are simulated by ParFlow via solving Richard's equation for water pressure in three dimensions (adapted from Gebler et al., 2017).

This shows that snow as well as the interception of water by plant foliage, through fall, stem flow are represented by the CLM hydrology module, which interact with the heat fluxes. The energy balance in CLM is implemented according to (Oleson et al., 2004):

$$R_n = \mathrm{LE}_v + \mathrm{LE}_g + H_v + H_g - G \tag{2.11}$$

where R_n is the net radiation [M T^{-3}], G the soil heat flux [M T^{-3}], LE_v is the latent heat flux from the canopy [M T^{-3})], LE_g the latent heat flux from soil [M T^{-3}], H_v the sensible heat flux for a vegetated surface [M T^{-3}], and H_g the sensible heat flux from soil [M T^{-3}].

All vertical momentum, mass and heat fluxes in CLM are described by the Monin-Obukhov similarity principle according to the solution provided by Zeng et al. (1998). The boundary layer resistance terms are taken from the Biosphere–Atmosphere Transport Scheme (BATS) developed by Dickinson et al. (1993) and Zeng et al. (2005).

CLM distinguishes between five different classes of land cover (glacier, lake, wetland, urban and vegetation) representing the land surface. Vegetation on the grid cell level is further subdivided into different plant functional types (PFT). Each PFT is associated with different plant physiological and structural parameters (e. g. aerodynamic resistance, albedo, root distribution) which are directly linked to the mass and energy transport processes between soil, land surface and atmosphere.

Soil and snow temperatures in CLM are calculated according to the heat diffusion equation for 10 soil layers and, in presence of snow, additional 5 snow layers (Oleson et al., 2004):

$$c\frac{\partial T_s}{\partial t} = \frac{\partial}{\partial z_s}\left[\lambda\frac{\partial T_s}{\partial z_s}\right] \tag{2.12}$$

where c is the total volumetric heat capacity [M L^2 T^{-2} θ^{-1}], T is the soil/snow temperature [θ], z_s is the depth [L] from the soil surface, and λ is the thermal conductivity [M^1 L^1 T^{-3} Θ^{-1}].

The overland runoff in CLM is originally represented by a TOPMODEL-based runoff scheme. However, this module is replaced by the ParFlow overland flow in case CLM runs in coupled fashion. CLM also runs in parallel with appropriate input and output file structures. Forcing data (barometric pressure, humidity, precipitation, solar radiation, and temperature) can be provided in a spatially distributed manner via netCDF file standard. More detailed technical descriptions of CLM are given by Bonan et al. (2002), Oleson et al. (2004), and Oleson et al. (2008).

2.3 Data Assimilation using the Ensemble Kalman Filter

2.3.1 Bayes' Theorem

The Ensemble Kalman Filter (EnKF) is a sequential data assimilation (DA) approach with the capability to improve model states, parameters, and fluxes using the combined information from observations and model simulations. Sequential DA methods rely on a probabilistic framework where the model predictions are repeatedly updated when observation information is available. The general theoretical basis for sequential DA is Bayes' theorem (Eq. 2.13) stating that the posterior probability density function for the states, taking into account observations, is given by the product of the prior probability density function of the model states ($\mathbf{x}^t$) and the likelihood function of the observation given the model states:

$$p(\mathbf{x}^t \mid \mathbf{y}^t) \propto p(\mathbf{y}^t \mid \mathbf{x}^t)\, p(\mathbf{x}^t) \tag{2.13}$$

where $p(\mathbf{x}^t \mid \mathbf{y}^t)$ is the posterior distribution of $\mathbf{x}^t$conditional on $\mathbf{y}^t$, $p(\mathbf{y}^t \mid \mathbf{x}^t)$ is the likelihood function of the observation given $\mathbf{x}^t$, and $p(\mathbf{x}^t)$ is the prior distribution of $\mathbf{x}^t$.

The assimilation cycle of sequential DA methods consists of (1) a forecast and (2) an analysis step, which are sequentially repeated. Eq. 2.13 can be solved analytically assuming Gaussian distributions for the prior pdf and likelihood, which renders the posterior pdf also Gaussian. The solution is the so-called Kalman Filter (Kalman, 1960).

2.3.2 The Kalman Filter

In the Kalman filter approach a linear model is assumed. The model prediction step is given by:

$$\mathbf{x}^t = \mathbf{F}^t\mathbf{x}^{t-1} + \boldsymbol{\varepsilon}_m \tag{2.14}$$

where $\mathbf{x}^t$ is the forecasted state vector, **F** the linear model operator, and $\boldsymbol{\varepsilon}_m$ is the model error at time (t).

The true model state is then related to the observations (**y)** by the observation operator matrix (**H)** mapping the true model state $\hat{\mathbf{x}}^t$ to the observation data including the observation error **(**$\boldsymbol{\varepsilon}_y$):

$$\mathbf{y}^t = \mathbf{H}^t\hat{\mathbf{x}}^t + \boldsymbol{\varepsilon}_y \tag{2.15}$$

Given that process (model) uncertainty is Gaussian distributed, the classical Kalman Filter (KF) finds the optimal solution by combining the predicted model states ($\mathbf{x}^t$**)** with the measurement data. The KF is the best linear unbiased estimator that minimizes the variance of the model prediction. In the analysis step model prediction and measurement data are combined to give the updated state vector $\mathbf{x}_i^a$ (Eq. 2.16) on the basis of a weighting with the Kalman gain **K** which weights the model predictions and the correcting influence of the data, accounting for the spatial correlation of the model states:

$$\mathbf{x}^a = \mathbf{x}^t + \mathbf{K}^t(\mathbf{y}^t - \mathbf{H}^t\mathbf{x}^t) \tag{2.16}$$

$$\mathbf{K}^t = \mathbf{P}^t\mathbf{H}^{tT}\left(\mathbf{H}^t\mathbf{P}^t\mathbf{H}^{tT} + \mathbf{R}\right)^{-1} \tag{2.17}$$

where $\mathbf{x}^a$ is the updated state vector, **y** the observation vector, **P** the model error covariance matrix, and **R** the observation error covariance matrix.

Equation 2.17 illustrates that the Kalman gain (**K)** is calculated from the covariance matrices **P** and **R**. For accurate measurements **R** is relatively small. In this case the observations have a strong impact on updating model states as **K** approaches $\mathbf{H^{-1}}$. Accordingly, more weight is put on the model forecast, if **P** is small.

The forecasted model state error covariance matrix (**P**), the analyzed model state error covariance matrix ($\mathbf{P}_a$) and the observation error covariance matrix (**R**) in the KF are given by

$$\mathbf{P}^t = \mathbf{F}^t \mathbf{P}_a^{t-1} {\mathbf{F}^t}^T + \mathbf{R} \tag{2.18}$$

$$\mathbf{P}_a^t = (\mathbf{I} - \mathbf{K}^t \mathbf{H}^t)\mathbf{P}^t \tag{2.19}$$

$$\mathbf{R} = \overline{\boldsymbol{\varepsilon}_y {\boldsymbol{\varepsilon}_y}^T} \tag{2.20}$$

where the overbar denotes the expected value.

For non-linear model dynamics the Extended Kalman Filter (EKF) can be applied (Evensen, 1994). The EKF uses Taylor expansion to approximate the forward error propagation of the covariance matrix while neglecting the contribution of higher order terms. However, the linearization of the covariance matrix potentially leads to poor evolution of the error covariance (Evensen, 2009) and requires the application of higher order closure, which is associated with high computational costs. The EKF is hence suitable for moderate non-linear problems. For high dimensional problems with large state spaces the Ensemble Kalman Filter is a stochastic alternative to overcome the limitations of the deterministic EKF (Evensen, 1994; Evensen, 2009).

2.3.3 The Ensemble Kalman Filter

The Ensemble Kalman Filter (EnKF) is an extension of the classical Kalman Filter approach. The EnKF is well suited for high-dimensional non-linear problems typically encountered for geophysical systems (Burgers et al., 1998). The EnKF utilizes an ensemble of model realizations with uncertain inputs (e.g., initial conditions, parameters, forcings) to approximate the prior and posterior pdf of the model states. It is therefore crucial that the ensemble of model runs characterizes correctly the true model prediction uncertainty and therefore the dominant sources of uncertainty (for example of input parameters and model forcings) need to be captured well.

During the forecast, each realization of the non-linear model M i.e. ParFlow-CLM is propagated forward in time (t):

$$\mathbf{x}_i^t = M(\mathbf{x}_i^{t-1}, \rho_i, \boldsymbol{\eta}_i) \tag{2.21}$$

where $\mathbf{x}_\mathbf{i}^t$ represents the model state vector at each time step when observations are available for assimilation; $\mathbf{x}_i^{t-1}$ is the state vector of the previous time step, ρ_i represents the model parameters, and $\boldsymbol{\eta}_i$ the model forcings for each ensemble realization i.

The model covariance matrix ($\mathbf{P}^t$) can be estimated from the ensemble spread prior to the update by:

$$\mathbf{P}^t = \frac{1}{\mathrm{n_{realz}} - 1} \mathbf{D}^t \mathbf{D}^{t\,T} \tag{2.22}$$

$$\mathbf{D}^t = [\mathbf{x}_1^t - \overline{\mathbf{x}}^t, \mathbf{x}_2^t - \overline{\mathbf{x}}^t, \dots, \mathbf{x}_{\mathrm{n_{realz}}}^t - \overline{\mathbf{x}}^t] \tag{2.23}$$

where $\mathbf{P}$ is the estimated model forecast covariance matrix and $\overline{\mathbf{x}}^t$ the ensemble mean vector with model states calculated over all realizations $\mathrm{n_{realz}}$.

This procedure allows including the effects of non-linear model dynamics but also implies a strong dependency between ensemble spread.

In contrast to the classical KF, the measurement data are treated as random variables (Burgers et al., 1998):

$$\tilde{\mathbf{y}}_i^t = \mathbf{y}^t + \boldsymbol{\varepsilon}_{iy} \tag{2.24}$$

where $\tilde{\mathbf{y}}_i^t$ is the vector with the perturbed observations, $\mathbf{y}^t$ is the observation vector, and $\boldsymbol{\varepsilon}_i$ is the vector with Gaussian random noise drawn from a normal distribution $\mathcal{N}(0, \varepsilon)$ with a mean of zero and a user specified observation error ($\boldsymbol{\varepsilon}_y$) for realization i at time t.

Together with the error covariance matrix $\mathbf{R}$ (Eq. 2.20), which is commonly provided by the modeler, this leads to the updated state vector $\mathbf{x}_i^a$ i.e. a weighted average of model forecast and observation for each ensemble member:

$$\mathbf{x}_i^a = \mathbf{x}_i^t + \mathbf{K}^t(\tilde{\mathbf{y}}_i^t - \mathbf{H}^t\mathbf{x}_i^t) \tag{2.25}$$

The model error covariance matrix $\mathbf{P}_a$ can be derived directly from the ensemble using $\mathbf{x}_i^a$ and the updated ensemble mean.

Real-time applications of the EnKF commonly involve only an update of the model state vector, while an additional estimation of the uncertain parameters may be beneficial for terrestrial modelling (e.g. Gutmann and Small, 2007; Liu et al., 2005) and particularly for groundwater modelling (Hendricks Franssen and Kinzelbach, 2008). On that account, the EnKF can be applied with an additional estimation of the model parameters, which is commonly conducted using an augmented state vector approach (e.g. Annan and Hargreaves, 2004; Hendricks Franssen and Kinzelbach, 2008).

The new augmented state vector $\mathbf{x}$ is a composite of the unknown model states ($\mathbf{x}_s$) and unknown model parameters ($\mathbf{x}_\rho$):

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_s \\ \mathbf{x}_\rho \end{pmatrix} \tag{2.26}$$

where the subscript s indicates the part of the vector associated with the model states and the subscript ρ refers to the model parameters.

Consequently, the Kalman gain (**K**), the model state covariance matrix (**P**), and the observation operator (**H**) have to be adapted:

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{ss} \\ \mathbf{K}_{s\rho} \end{pmatrix} \tag{2.27}$$

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_{ss} & \mathbf{P}_{\rho s} \\ \mathbf{P}_{s\rho} & \mathbf{P}_{\rho\rho} \end{pmatrix} \tag{2.28}$$

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_s \\ \mathbf{H}_\rho \end{pmatrix} \tag{2.29}$$

where the subscript ss indicates the covariances between two states at different grid cells and the subscript $s\rho$ the evaluation of state and parameter at a grid cell.

The modified Kalman gain (**K**) then also takes into account model covariances (**P**) between model states and parameters and the augmented analysis step updates both model states and parameters. The observation operator matrix (**H**) maps the measurements (e.g. hydraulic head) on the modeled states (e.g. hydraulic head) and/or parameters (e.g. hydraulic conductivity). This may also include a non-linear transfer function between measurement (e.g. brightness temperature) and the model state (e.g. soil moisture). In case observations are not available for the model parameters, $\mathbf{H}_\rho$ (Eq. 2.29) consists of zeros and the model update is skipped.

A common problem in data assimilation is the so-called filter inbreeding problem (Hendricks Franssen and Kinzelbach, 2008). The continuous update of states and parameters from the same ensemble of realization leads to a gradual reduction of the ensemble variance. Hence, the quality of the analyzed covariance matrix deteriorates over time. Several authors suggested approaches to reduce the filter inbreeding problem. For example, Evensen (2004) and Zhang et al. (2007) applied eigenvector sampling. Anderson (2001) and Wang et al. (2007) suggested adaptive covariance inflation, which artificially increases the ensemble spread. Wen and Chen (2007) eliminated individual realizations from the ensemble which were too closely related. Hendricks Franssen and Kinzelbach (2008) showed that that filter inbreeding could be strongly reduced with a damping

factor which retricted the parameter update to a fraction of the update without damping. Other studies (e. g. Chen and Oliver, 2010; Devegowda et al., 2010; Nan and Wu, 2011; Rasmussen et al., 2015) stressed the need for localization techniques to avoid the impact of spurious correlations related to the low rank approximation of the model covariance matrix in case of small ensemble sizes. In such cases spurious correlations can be avoided with a restriction of the influence range of observation points.

Soil or aquifer hydraulic properties might in reality be non-multi-Gaussian distributed (Gómez-Hernández and Wen, 1998), but EnKF performs only optimally for Gaussian distributed states and parameters and tends to make probability density functions of states and parameters Gaussian (Zhou et al., 2011). Normal score transformation of states and/or parameters and the application of EnKF on the transformed states and parameters, with back transformation after the analysis step, can improve the characterization of non-multi-Gaussian state and parameter distributions (Zhou et al., 2011; Schöniger et al., 2012).

2.4 The Parallel Data Assimilation Framework within TerrSysMP

The Parallel Data Assimilation Framework (PDAF) by Nerger and Hiller (2013) provides routines for several well-established DA methods (e.g. the EnKF according to Evensen, 2003) and can be deployed with any sort of (parallel) geoscience model. PDAF communicates with the individual model either offline, via I/O data exchange, or online, via main memory in a parallel fashion. Online communication is more CPU efficient than offline communication due to avoiding recurrent model initializations and substantial overhead in I/O operations. The information exchange between model and PDAF includes the state vector, the measurement vector, model and measurement covariance matrices, update rules, and pre- and post-processing instructions. Within the TerrSysMP framework, PDAF is implemented as online callable routine which is controlled by the TerrSysMP-PDAF driver (i.e. main routine). A global MPI (Message Passing Interface) communicator scheme is thereby used for initialization, time-looping and finalization (Kurtz et al., 2016). This includes user defined PDAF functions (e. g. state vector definitions) as well as PDAF core functions (filter methods). The communication scheme of TerrSysMP-PDAF is specifically designed for parallel information exchange between subdivisions of models and filter as well as for the different model ensemble members. The TerrSysMP-PDAF driver also operates the individual model libraries (ParFlow and/or CLM) via the TerrSysMP wrapper interface. All terrestrial model components can be combined with PDAF in a modular fashion. Optionally a coupled model (ParFlow+CLM) using OASIS3-MCT communication or stand-alone ParFlow or CLM are available.

In the current version of TerrSysMP-PDAF, several alternatives for assimilation of observations are available. Pressure is the prognostic state variable in ParFlow and can be assimilated directly. Soil moisture can also be assimilated directly (conversion to pressure after the update) or indirectly

(updating pressure in the data assimilation step). The state vector either contains pressure or soil moisture and the correlations between pressure and soil moisture are used for pressure update with soil moisture information (augmented state vector), or, the updated soil moisture state vector is transformed to pressure using the van Genuchten relationships (Eq. 2.5, Eq. 2.6). In case CLM is applied in the stand-alone model soil moisture is updated directly.

In the current implementation, observation data are attributed to the nearest model grid cell without any interpolation or scaling. Measurement errors for individual observations can be supplied and in the current implementation correlations between measurements are not taken into account. The framework also gives the possibility of model parameter estimation; log-transformed permeability and Manning's coefficients can be updated. For parameter update the option of parameter damping is available. In case of joint update of states and parameters, the parameter update interval optionally can be detached from the interval of state variable update (e.g. daily state update with parameter update every 10 days).

Chapter 3 Actual evapotranspiration and precipitation measured by lysimeters: a comparison with eddy covariance and tipping bucket

*adapted from Gebler, S., Hendricks Franssen, H.J., Pütz, T., Post, H., Schmidt, M., Vereecken, H., 2015. Actual evapotranspiration and precipitation measured by lysimeters: A comparison with eddy covariance and tipping bucket. Hydrology and Earth System Sciences, 19: 2145-2161.

3.1 Introduction

Precise estimates of precipitation and actual evapotranspiration are important for an improved understanding of water and energy exchange processes between land and atmosphere relevant for many scientific disciplines and agricultural management. Information about measurement errors and uncertainties is essential for improving measurement methods and correction techniques as well as for dealing with uncertainty during calibration and validation of model simulations. Although first devices for modern scientific purposes were developed in Europe during the 17th century (Kohnke et al., 1940; Strangeways, 2010), the accurate estimation of precipitation (P) and actual evapotranspiration (ET_a) is still a challenge. Common precipitation measurement methods exhibit systematic and random errors depending on the device locations and climatic conditions. Legates and DeLiberty (1993) concluded from their long-term study of precipitation biases in the United States that Hellmann type gauges (US standard) undercatch precipitation amounts. Undercatch is larger in case of snowfall and larger wind speeds. Wind-induced loss is seen as the main source of error (Sevruk, 1981, 1996; Yang et al., 1998; Chvíla et al., 2005; Brutsaert, 2010). Precipitation gauges are commonly installed above ground to avoid negative impact on the measurements by splash water, hail, and snow drift. However, this common gauge setup causes wind distortion and promotes the development of eddies around the device. Wind tunnel experiments with Hellmann type gauges (Nešpor and Sevruk, 1999) have shown precipitation losses of 2–10% for rain and 20–50% for snow compared to the preset precipitation amount. In general, wind induced loss increases with installation height of the device and wind speed, and it decreases with precipitation intensity (Sevruk, 1989). Intercomparison studies between different rain gauge designs of the World Meteorological Organization (WMO) have indicated that shielded devices can considerably reduce this undercatch compared to unshielded gauges, in particular for snow and mixed precipitation (Goodison et al., 1997). Further precipitation losses that affect the rain gauge measurement are evaporation of water from the gauge surface and recording

mechanisms (Sevruk, 1981; Michelson, 2004). Moreover, measurement methods (e.g., condensation plates, optical methods) to estimate the contribution of rime, dew, and fog to the total precipitation exhibit a high uncertainty (Jacobs et al., 2006). A short-term lysimeter case study by Meissner et al. (2007) and a long-term investigation with a surface energy budget model calibrated with microlysimeters by Jacobs et al. (2006) show that rime, fog, and dew contribute up to 5% of the annual precipitation at a humid grassland site and are usually not captured by a standard precipitation gauge.

The eddy covariance (EC) method is one of the most established techniques to determine the exchange of water, energy, and trace gases between the land surface and the atmosphere. On the basis of the covariance between vertical wind speed and water vapor density, the EC method calculates the vertical moisture flux (and therefore ET) in high spatial and temporal resolution with relatively low operational costs. The size and shape of the measurement area (EC footprint) vary strongly with time (Finnigan, 2004). Under conditions of limited mechanical and thermal turbulence the EC method tends to underestimate fluxes (Wilson et al., 2001; Li et al., 2008). Energy balance deficits are on average found to be between 20 and 25% (Wilson et al., 2001; Hendricks Franssen et al., 2010), and therefore latent heat flux or actual evapotranspiration estimated from EC data shows potentially a strong underestimation. The energy balance closure problem can be corrected by closure procedures using the Bowen ratio. However, this is controversially discussed, especially because not only the underestimation of the land surface fluxes but also other factors like the underestimation of energy storage in the canopy might play a role (Twine et al., 2000; Foken et al., 2011).

As an alternative to classical rain gauges and the eddy covariance method, state-of-the-art, high-precision weighing lysimeters are able to capture the fluxes at the interface of soil, vegetation, and atmosphere (Unold and Fank, 2008). A high weighing accuracy and a controlled lower boundary condition permit high-temporal-resolution precipitation measurements at ground level, including dew, fog, rime, and snow. Additionally, ET_a can be estimated with the help of the lysimeter water balance. However, the high acquisition and operational costs are a disadvantage of lysimeters. Moreover, the accuracy of lysimeter measurements is affected by several error sources. Differences in the thermal, wind, and radiation regime between a lysimeter device and its surroundings (oasis effect) (Zenker, 2003) as well as lysimeter management (e.g., inaccuracies in biomass determination) can affect the measurements. Wind or animal-induced mechanical vibrations can influence the weighing system but can be handled by accurate data processing using filtering and smoothing algorithms (Schrader et al., 2013; Peters et al., 2014). Vaughan and Ayars (2009) examined lysimeter measurement noise for data at a temporal resolution of 1 min, caused by wind loading. They presented noise reduction techniques that rely on Savitzky–Golay (Savitzky and Golay, 1964) smoothing. Schrader et al. (2013) evaluated the different filter and smoothing

strategies for lysimeter data processing on the basis of synthetic and real measurement data. They pointed out that the adequate filter method for lysimeter measurements is still a challenge, especially at high temporal resolution, due the fact that noise of lysimeter measurements varies strongly with weather conditions and mass balance dynamics. Peters et al. (2014) recently introduced a filter algorithm for high-precision lysimeters, which combines a variable smoothing time window with a noise-dependent threshold filter that accounts for the factors mentioned above. They showed that their Adaptive Window and Adaptive Threshold (AWAT) filter improves actual evapotranspiration and precipitation estimates from noisy lysimeter measurements compared to smoothing methods for lysimeter data using the Savitzky–Golay filter or simple moving averages used in other lysimeter studies (e.g., Vaughan and Ayars, 2009; Huang et al., 2012; Nolz et al., 2013; Schrader et al., 2013).

In this work, a long-term investigation to precipitation estimation with a lysimeter is presented. One of the focal points in the study is the contribution of dew and rime to the total precipitation amount. The novelty compared to the work by Meissner et al. (2007) is the length of the study and the fact that a series of six lysimeters is used. Our work allows corroborating results from Jacobs et al. (2006), who used in their long-term study a different, more uncertain measurement method.

In the literature we find several comparisons between lysimeter measurements and standard ET calculations. López-Urrea et al. (2006) found a good agreement of FAO-56 Penman-Monteith with lysimeter data on an hourly basis. Vaughan et al. (2007) also reported a good accordance of hourly lysimeter measurements with a Penman-Monteith approach of the California Irrigation Management Information System. Wegehenkel and Gerke (2013) compared lysimeter ET with reference ET and ET estimated by a numerical plant growth model. They found that lysimeter ET overestimated actual ET, the cause being an oasis effect. On the other hand, also ET estimated by EC measurements and water budget calculations are compared in the literature. Scott (2010) found that the EC method underestimated evapotranspiration for a grassland site related to the energy balance deficit. However, only a few comparisons between ET estimated by EC and lysimeter data were found in the literature. Chávez et al. (2009) evaluated actual evapotranspiration determined by lysimeters and EC in the growing season for a cotton field site. They found a good agreement of both methods after correcting the energy balance deficit, and they suggested considering also the footprint area for EC calculations. Ding et al. (2010) found a lack of energy balance closure and underestimation of ET_a by the EC method for maize fields. An energy balance closure based on the Bowen ratio method was able to reduce the ET underestimation. Alfieri et al. (2012) provided two possible explanations for a strong underestimation of ET_a-EC compared to lysimeter ET_a: first, the energy balance deficit of the EC data, especially for those cases where EC measurements are affected by strong advection; second, deviations between the vegetation status of the lysimeter and the surrounding field. Evett et al. (2012) found an 18% underestimation of corrected ET_a-EC

compared to ET_a estimated by lysimeter and attributed the difference to differences in vegetation growth. Whereas the aforementioned studies conclude that deviations between ET_a measurements are related to vegetation differences, the EC footprint, and the ability to close the energy balance gap, the uncertainties of lysimeter measurements in this context have hardly been investigated. Lysimeter ET_a estimations often rely on relatively low temporal resolution due to challenges in noise reduction, which impedes a simultaneous estimation of both P and ET_a by lysimeters. Furthermore, studies with cost- and maintenance-intensive lysimeters are either with a few or without redundant devices, so that measurement uncertainty cannot be addressed well.

The Terrestrial Environmental Observatories (TERENO) offer the possibility of detailed long-term investigations of the water cycle components at a high spatiotemporal resolution (Zacharias et al., 2011). This study compares precipitation and evapotranspiration estimates calculated with a set of six weighing lysimeters (LYS) with nearby eddy covariance and precipitation measurements for the TERENO grassland site Rollesbroich. Additional soil moisture, soil temperature, and meteorological measurements at this TERENO test site enable a detailed analysis of differences between the different measurement techniques. The lysimeter data (ET_a-LYS) are processed with the AWAT filter (Peters et al., 2014), which allows a simultaneous estimation of P and ET_a at a high temporal resolution, and the comparison is carried out with energy-balance-corrected EC data (ET_a-EC). Actual ET estimates are additionally compared to the full-form Penman-Monteith equation (Allen et al., 1998) accounting for the effects of variable grass cover height. Precipitation measurements by a classical Hellmann type tipping bucket (TB), with and without accounting for wind- and evaporation-induced loss (Richter correction), were compared with lysimeter data for 1 year (2012).

For our study, we (1) compared precipitation measurements by lysimeters and a (unshielded) standard tipping bucket device and interpreted the differences (for example, the vegetated high-precision lysimeters potentially allow for better estimates of precipitation, accounting for dew, rime, and fog); (2) compared eddy covariance and lysimeter ET estimates and tried to explain differences in estimated values; (3) tested whether a correction of the energy balance deficit for the EC method results in an ET_a estimate which is close to the lysimeter method; and (4) analyzed the variability of the measurements by the six lysimeters under typical field conditions with identical configuration and management.

3.2 Material and Methods

3.2.1 Study Site and Measurement Setup

A detailed description of the study site characteristics with an overview of the on-site measurement devices is given in chapter 2.1. The following paragraphs provide more detailed information about the setup of lysimeter, eddy covariance and precipitation devices used in this study.

In 2010 a set of six lysimeters (TERENO-SoilCan project, UMS GmbH, Munich, Germany) was arranged in a hexagonal design around the centrally placed service unit, which hosts the measurement equipment and data recording devices. Each lysimeter contains silty-clay soil profiles from the Rollesbroich site and is covered with grass. The conditions at the lysimeters therefore closely resemble the ones in the direct surroundings (Figure 3.1). Additionally, the spatial gap between lysimeter and surrounding soil was minimized to prevent thermal regimes which differ between the lysimeter and the surrounding field (oasis effect). Every lysimeter device has a surface of 1 m^2 and a depth of 1.5 m, and is equipped with a 50 L weighted leachate tank connected via a bidirectional pump to a suction rake in the bottom of each lysimeter. To reproduce the field soil water regime, the lower boundary conditions are controlled by tensiometers (TS1, UMS GmbH, Munich, Germany) monitoring the soil matric potential inside the lysimeter bottom and the surrounding field. Matric potential differences between field and lysimeter are compensated by suction rakes (SIC 40, UMS GmbH, Munich, Germany) injecting leachate tank water into the lysimeter monolith during capillary rise or removing water during drainage conditions.

Figure 3.1: The lysimeter set-up of the Rollesbroich study site (November 2012).

The weighing precision is 100 g for the soil monolith and 10 g for the leachate tank, accounting for long-term temperature variations and load alternation hysteresis effects. For short-term signal processing the relative accuracy for accumulated mass changes of soil monolith and leachate is 10 g.

For the year 2012 measurements were made every 5 s and averaged to get minute values. In the winter season a connection between the snow lying on the lysimeter and the surrounding snow layer potentially disturbs the weighing system. A mechanical vibration plate is engaged at all lysimeter devices to prevent this situation and is activated once every 5 s between two measurements. The lysimeters are also equipped with soil moisture, matric potential, and temperature sensors at different depths (10, 30, 50, and 140 cm). Amongst others, soil temperature is determined at 10, 30, and 50 cm depth with PT-100 sensors integrated in TS1 tensiometers (UMS GmbH, Munich, Germany). A schematic overview of the lysimeter device (Figure 3.2) shows the installation locations and the different sensor types. The lysimeter site was kept under video surveillance by a camera taking a photo of the lysimeter status every hour. Further technical specifications can be found in Unold and Fank (2008).

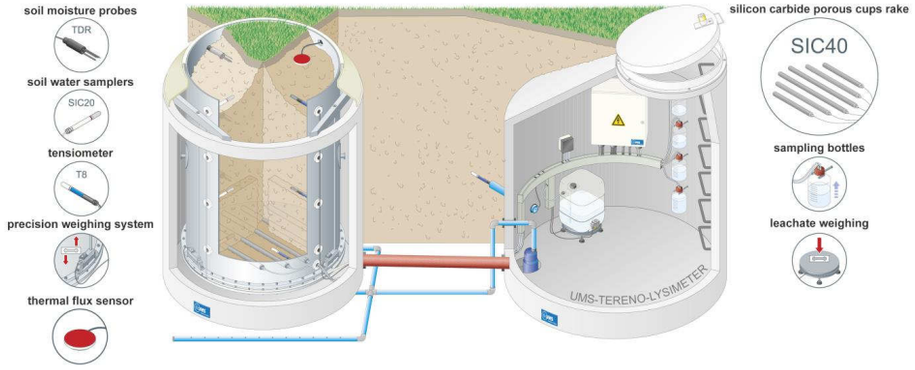

Figure 3.2: Schematic drawing of the lysimeter soil monolith (left) and service well (right) used in the TERENO-SoilCan project. The illustration of the lysimeter (left) shows the weighted soil column container with slots for soil moisture (TDR), temperature (SIS, TS1), matric potential sensors (SIS), soil water sampler (SIC20) and silicon porous suction cup rake (SIC40) installation inside and outside the monolith. The service well contains the weighted drainage tank and sampling tubes for each affiliated lysimeter (courtesy of UMS GmbH Munich, 2014, used by permission).

Latent and sensible heat fluxes were measured by an eddy covariance station at a distance of approximately 30 m from the lysimeters. The EC station (50° 37' 19" N, 6° 18' 15" E, 514 m a.s.l.) is equipped with a sonic anemometer (CSAT3, Campbell Scientific, Inc., Logan, USA) at 2.6m height to measure wind components. The open-path device of the gas analyzer (LI7500, LI-COR Inc., Lincoln, NE, USA) is mounted along with the anemometer at 2.6m above the ground surface and measures H2O content of the air. Air pressure is measured at the processing unit of the gas analyzer at a height of 0.57 m. Air humidity and temperature were measured by HMP45C (Vaisala Inc., Helsinki, Finland) at 2.58m above the ground surface. Radiation was determined by a four-component net radiometer (NR01, Hukseflux Thermal Sensors, Delft, Netherlands). Soil heat flux was determined at 0.08m depth by a pair of two HFP01 (Hukseflux Thermal Sensors, Delft, Netherlands). Precipitation measurements are made by a standard Hellmann type TB balance rain gauge (ecoTech GmbH, Bonn, Germany) with a resolution of 0.1mm and a measurement interval of 10 min. The measurement altitude of 1m above ground is in accordance with recommendations of the German Weather Service (DWD, 1993) for areas with an elevation >500 m a.s.l. and occasional heavy snowfall (WMO standard is 0.5 m). The unshielded gauge was temporarily heated during wintertime to avoid freezing of the instrument. Additional soil moisture and soil temperature measurements were carried out with a wireless sensor network (SoilNet) installed at the study site (Qu et al., 2013). The 179 sensor locations at the Rollesbroich site contain six SPADE sensors (model 3.04, sceme.de GmbH i.G., Horn-Bad Meinberg, Germany) with two redundant sensors at 5, 20, and 50 cm depth. Further technical details can be found in Qu et al. (2013). Soil water content and temperature were also measured by two sensor devices installed near the lysimeter site.

3.2.2 Data Processing

The lysimeter weighing data were processed in three steps:

1. elimination of outliers by an automated threshold filter;
2. smoothing of measurement signal with the AWAT filter routine on the basis of data at a temporal resolution of 1 min;
3. estimation of hourly precipitation and evapotranspiration on the basis of the smoothed signal.

Outliers were removed from the data by limiting the maximum weight difference between two succeeding measurements for the soil column to 5 kg and for the leachate weight to 0.1 kg. The lysimeter readings are affected by large random fluctuations caused by wind and other factors that influence the measurement. Therefore, the AWAT filter (Peters et al., 2014) in a second correction step was applied on the minute-wise summed leachate and on the weights for each individual

lysimeter. First, the AWAT routine gathers information about signal strength and data noise by fitting a polynomial to each data point within an interval of 31 min. The optimal order (k) of the polynomial is determined by testing different polynomial orders for the given interval (i.e., k: 1–6) and selecting the optimal k according to Akaike's information criterion (Akaike, 1974; Hurvich and Tsai, 1989). The maximum order of k is limited to 6 for the AWAT filter, preventing an erroneous fit caused by outliers. The average residual $s_{\mathrm{res},i}$ of measured and predicted values (Eq. 3.1) and the standard deviation of measured values $s_{\mathrm{dat},i}$ (Eq. 3.2) lead to the quotient B_i (Eq. 3.3), which gives information about the explained variance of the fit and is related to the coefficient of determination (R^2).

$$s_{\mathrm{res},i} = \sqrt{\frac{1}{r}\sum_{j=1}^{r}[\gamma_j - \hat{\gamma}_j]^2} \tag{3.1}$$

$$s_{\mathrm{dat},i} = \sqrt{\frac{1}{r}\sum_{j=1}^{r}[\gamma_j - \bar{y}]^2} \tag{3.2}$$

$$B_i = \frac{s_{\mathrm{res},i}}{s_{\mathrm{dat},i}} = \sqrt{1 - \mathrm{R}_i^2} \tag{3.3}$$

where y_j [M] is the measured data, $\hat{y}_j$ [M] the fitted value at each time interval j, $\bar{y}$ [M] the mean of the measurements and r the number of measurements within the given interval of data point i. $B_i = 0$ indicates that the polynomial totally reproduces the range of data variation in contrast to $B_i = 1$ where nothing of the variation in the data is explained by the fitted polynomial.

Second, AWAT smoothes the data using a moving average for an adaptive window width w_i [T], which is a time dependent linear function of B_i (Eq. 3.4):

$$w_i(B_i) = \max(w_{\min}, B_i w_{\max}) \tag{3.4}$$

where $w_{\max}$ [T] and $w_{\min}$ [T] are maximum and minimum provided window width. For our study $w_{\min}$ was set to 11 min, $w_{\max}$ was 61 min. A low B_i requires less smoothing and therefore small time windows, whereas a B_i close to one requires a smoothing interval close to the allowed $w_{\max}$. Third, AWAT applies an adaptive threshold δ_i (Eq. 3.5) to the data at each time step to distinguish between noise and signal related to the dynamics of mechanical disturbances:

$$\delta_i = s_{\mathrm{res},i} \cdot t_{97.5,r} \quad \text{for } \delta_{\min} < s_{\mathrm{res},i} \cdot t_{97.5,r} < \delta_{\max} \tag{3.5}$$

where δ_i [M] is a function of the interval residuals ($s_{\mathrm{res},i}$) [M] (see Eq. 3.1) and the Student t value ($t_{97.5,r}$) for the 95 % confidence level at each time step, $\delta_{\min}$ [M] is the minimum and $\delta_{\max}$ [M] is the maximum provided threshold for the mass change. The product of Student t and $s_{\mathrm{res},i}$ is a

measure for the significance level of mass changes during flux calculation. Hence, the δ_i value indicates the range ($\pm s_{\mathrm{res},i} \cdot t_{97.5,r}$), where the interval data points differ not significantly from the fitted polynomial at the 95 % confidence level. Mass changes above the adaptive threshold δ_i are significant and interpreted as signal, whereas weight differences below δ_i are interpreted as noise. The adaptive threshold is limited by $\delta_{\min}$ and $\delta_{\max}$ to guarantee that (1) mass changes smaller than the lysimeter measurement accuracy are understood as remaining noise and therefore not considered for the flux calculation and (2) noise is not interpreted as signal during weather conditions, which produce noisy lysimeter readings (i.e. thunderstorms with strong wind gusts). Lysimeter calibration tests with standard weights at the study site indicate a system scale resolution of 0.05 kg. We chose a slightly higher threshold ($\delta_{\min}$ = 0.055 kg) with an adequate tolerance for our TERENO lysimeter devices. For the upper threshold $\delta_{\max}$ = 0.24 kg was taken, similar to the example presented by Peters et al. (2014).

For the separation of precipitation and actual evapotranspiration (ET_a) AWAT assumes that increases of lysimeter and leachate weights (averaged over a period of one minute) are exclusively related to precipitation and negative differences to ET_a [M T^{-1}]. Supposing that no evapotranspiration occurs during a precipitation event and assuming a fixed water density of 1000 kg m^{-3}, precipitation (P) [M T^{-1}] can be derived from the lysimeter water balance as:

$$\mathrm{ET_a} = P - L - \frac{\mathrm{d}S_{\mathrm{wst}}}{\mathrm{d}t} \tag{3.6}$$

$$P = L + \frac{\mathrm{d}S_S}{\mathrm{d}t} \tag{3.7}$$

where L is the amount of leachate water [M T^{-1}] and $\mathrm{d}S_{\mathrm{wst}}/\mathrm{d}t$ is the change of soil water storage [M T^{-1}] with time. After smoothing the fluxes at one minute resolution were cumulated to hourly sums of P and $\mathrm{ET_a}$.

Although the six lysimeters have a similar soil profile, technical configuration and management (i.e. grass cut, maintenance), differences in measured values between lysimeters are not exclusively related to random errors. Systematic weight variations may for example be caused by soil heterogeneity, mice infestation and differences in plant dynamics. In this study precipitation measured by lysimeter and TB are compared, as well as evapotranspiration measured by lysimeter and eddy covariance. The precipitation or $\mathrm{ET_a}$ averaged over the six redundant lysimeters are used in this comparison. We assume that the lysimeter average of six redundant lysimeter devices is the most representative estimation for the lysimeter precipitation and actual evapotranspiration (unless specified otherwise).

3.2.3 Eddy Covariance Data

Eddy covariance raw measurements were taken with a frequency of 20 Hz and fluxes of sensible heat (*H*) and latent heat (LE) were subsequently calculated for intervals of 30 minutes by using the TK3.1 software package (Mauder and Foken, 2011). The complete post-processing was in line with the standardized strategy for EC data calculation and quality assurance presented by Mauder et al. (2013). It includes the application of site specific plausibility limits and a spike removal algorithm based on median absolute deviation of raw measurements, a time lag correction for vertical wind speed with temperature and water vapor concentration based on maximizing cross-correlations between the measurements of the used sensors, a planar fit coordinate rotation (Wilczak et al., 2001), corrections for high frequency spectral losses (Moore 1986), the conversion of sonic temperature to air temperature (Schotanus et al., 1983) and the correction for density fluctuations (Webb et al,. 1980). Processed half hourly fluxes and statistics were applied to a three-class quality flagging scheme, based on stationarity and integral turbulence tests (Foken and Wichura, 1996) and classified as high, moderate and low quality data. For this analysis only high and moderate quality data were used, while low quality data were treated as missing values. To assign half hourly fluxes with its source area the footprint model of Korman and Meixner (2001) was applied.

Almost every eddy covariance site shows an unclosed energy balance, which means that the available energy (net radiation minus ground heat flux) is found to be larger than the sum of the turbulent fluxes (sensible plus latent heat flux) (Foken, 2008; Foken et al., 2011). In this study the energy balance deficit (EBD) was determined using a 3-h moving window around the measurements (Kessomkiat et al., 2013):

$$\mathrm{EBD}_{3\mathrm{h}} = R_{\mathrm{n}-3\mathrm{h}} - (G_{3\mathrm{h}} + \mathrm{LE}_{3\mathrm{h}} + H_{3\mathrm{h}} + \mathrm{HS}_{3\mathrm{h}}) \tag{3.8}$$

where $R_{\mathrm{n}-3\mathrm{h}}$ is average net radiation [M T^{-3}], $G_{3\mathrm{h}}$ is average soil heat flux [M T^{-3}], $\mathrm{LE}_{3\mathrm{h}}$ is average latent heat flux [M T^{-3}], $H_{3\mathrm{h}}$ is average sensible heat flux [M T^{-3}], and $\mathrm{HS}_{3\mathrm{h}}$ is average heat storage (canopy air space, biomass and upper soil layer above ground heat flux plate) [M T^{-3}]. All these averages are obtained over a three hour period around a particular 30 min EC-measurement. The moving window of three hours is a compromise between two sources of error. First, it guarantees a relatively small impact of random sampling errors and therefore increases the reliability of the EBD calculation. Second, the relatively short interval ensures that the calculations are not too much affected by non-stationary conditions. It was assumed that the energy balance deficit is caused by an underestimation of the turbulent fluxes and therefore the turbulent fluxes are corrected according to the evaporative fraction.

The evaporative fraction (EF) was determined for a time window of seven days:

$$\mathrm{EF} = \frac{\overline{\mathrm{LE}}_{7\mathrm{d}}}{\overline{\mathrm{LE}}_{7\mathrm{d}} + \overline{H}_{7\mathrm{d}}} \tag{3.9}$$

where $\overline{\mathrm{LE}}_{7\mathrm{d}}$ and $\overline{H}_{7\mathrm{d}}$ [M T^{-3}] are the latent and sensible heat fluxes averaged over seven days. The chosen time period increases the reliability for EF calculation compared to single days. Dark days with small fluxes may not give meaningful results. Kessomkiat et al. (2013) investigated the impact of the time window on the calculation of the EF and found that a moving average over seven days gives good results, whereas a too short time window of one day gives unstable, unreliable results.

The energy balance corrected latent heat flux was determined by redistribution of the latent heat on the basis of the calculated evaporative fraction:

$$\mathrm{LE}^{*}_{0.5\mathrm{h}} = \mathrm{LE}_{0.5\mathrm{h}} + \mathrm{EBD}_{3\mathrm{h}}(\mathrm{EF}) \tag{3.10}$$

where $\mathrm{LE}^{*}_{0.5\mathrm{h}}$ is the latent heat flux (for a certain measurement point in time; i.e. a 30 minutes period for our EC data). The EBD is added to the uncorrected LE according to the partitioning of heat fluxes in the EF. Further details on the EBD correction method can be found in Kessomkiat et al. (2013).

In this study, also the evapotranspiration (ET_a-EC) calculated with the original latent heat flux (not corrected for energy balance closure) will be presented for comparison. Furthermore, the most extreme case would be that the complete EBD is linked to an underestimation of the latent heat flux. Some authors argue (Ingwersen et al., 2011) that the EBD could be more related to underestimation of one of the two turbulent fluxes than the other turbulent flux. Therefore, as an extreme scenario the complete EBD is assigned to the correction of the latent heat flux.

ET_a-EC is calculated from the latent heat flux according to:

$$\mathrm{ET}_{\mathrm{a}} = \frac{\mathrm{LE}^{*}_{\mathrm{h}}}{L(T_{\mathrm{h}})_{\mathrm{H_2O}} * \rho_{\mathrm{H_2O}}} \tag{3.11}$$

where ET_a is ET_a-EC [L T^{-1}], $\mathrm{LE}^{*}_{\mathrm{h}}$ is latent heat flux [M T^{-3}], $\rho_{\mathrm{H_2O}}$ is the density of water [M L^{-3}] and $L(T_{\mathrm{h}})_{H_2O}$ is the vaporization energy [L^2 T^{-2}] at a given temperature.

The lysimeters are thought to be representative for the EC footprint, although size and shape of the EC footprint are strongly temporally variable. However, the EC footprint is almost exclusively constrained to the grassland and the lysimeters are also covered by grass.

3.2.4 Grass Reference Evapotranspiration

The measurements of ET_a by the EC-method and lysimeters were in this study compared with evapotranspiration calculated with full-form Penman-Monteith equation as presented by Allen et al. (1998). This approach accounts for vegetation and ground cover conditions during crop stage considering bulk surface and aerodynamic resistances for water vapor flow. The calculations were adapted for hourly intervals according to Eq. 3.12:

$$\mathrm{ET_{PM}} = \frac{0.408\Delta(R_\mathrm{n} - G) + \gamma \frac{3600\varepsilon_{air}}{T_\mathrm{vh} R_\mathrm{spc}(r_\mathrm{a} u_2)} u_2 (e°(T_\mathrm{h}) - e_\mathrm{a})}{\Delta + \gamma_{psy}(1 + \frac{r_\mathrm{s}}{r_\mathrm{a}})} \tag{3.12}$$

where $\mathrm{ET_{PM}}$is the hourly Penman-Monteith evapotranspiration [L T^{-1}], R_n is net radiation at the grass surface [M T^{-3}], G is soil heat flux density [M T^{-3}], T_vh is mean hourly virtual temperature [θ], R_spc is the specific gas constant for dry air [L^2 T^{-2} θ^{-1}], r_a is the aerodynamic resistance [T L^{-1}], r_s is the (bulk) surface resistance [T L^{-1}], ε_{air} is the ratio molecular weigth of water vapour (dry air) [-], T_h is mean hourly air temperature [θ], Δ slope of the saturated vapour pressure curve at T_h [M L^{-1} T^{-2} θ^{-1}], γ_{psy} is psychrometric constant [M L^{-1} T^{-2} θ^{-1}], $e°$ (T_h) is saturation vapour pressure for the given air temperature [M L^{-1} T^{-2}], e_a is average hourly actual vapour pressure [M L^{-1} T^{-2}], and u_2 is average hourly wind speed [L T^{-1}] at 2 m height. All required meteorological input parameters for calculating $\mathrm{ET_{PM}}$ were taken from the EC station. The wind speed data were corrected to 2 m using the FAO-standard wind profile relationship of Allen et al. (1998).

We approximated aerodynamic resistance (r_a), (bulk) surface resistance (r_s) and leaf area index (LAI) with help of grass height according to Allen et al. (2006):

$$\mathrm{r_a} = \frac{ln\left[\frac{z_\mathrm{m} - \frac{2}{3} h_{plant}}{0.123\ h_{plant}}\right] ln\left[\frac{z_\mathrm{h} - \frac{2}{3} h_{plant}}{0.1\ (0.123\ h_{plant})}\right]}{{k_k}^2 u_2} \tag{3.13}$$

$$\mathrm{r_s} = \frac{\mathrm{r_i}}{\mathrm{LAI_{act}}} \tag{3.14}$$

$$\mathrm{LAI_{act}} = (0.3\ \mathrm{LAI}) + 1.2 = 0.5\ (24\ h_{plant}) \tag{3.15}$$

where z_m is the height of the wind measurement [L], z_h is the height of the humidity measurement [L], h_{plant} is the grass length [L] at the lysimeter, k_k is the von Karman's constant [-], $\mathrm{r_i}$ the stomatal resistance [T L^{-1}], and $\mathrm{LAI_{act}}$ the active leaf area index taking into account that only the upper grass surface contributes to heat and vapor transfer [-]. For our calculations we assume a fixed stomatal resistance for a well-watered grass cover of 100 s m^{-1} in accordance to Allen et al.

(1998). The grass length at the lysimeters was estimated with the help of maintenance protocols and the surveillance system. Grass lengths between two measurement intervals were linearly interpolated on a daily basis.

3.2.5 Precipitation Correction

A precipitation correction according the method of Richter (1995) was applied (Eq. 3.16/3.17) on a daily basis to account for wind, evaporation and wetting losses of the tipping bucket precipitation:

$$P^{\mathrm{cor}} = P + \Delta P \tag{3.16}$$

$$\Delta P = bP^{\epsilon} \tag{3.17}$$

where P^{cor} is the corrected daily precipitation [M T^{-1}], P is the measured tipping bucket precipitation [M T^{-1}], ΔP the estimated precipitation deficit [M T^{-1}], b the site specific wind exposition coefficient [-], and ϵ the empiric precipitation type coefficient [-].

This correction method is widely used for German weather service stations and relies on empirical relationships of precipitation type and wind exposition, without using direct wind measurements. In order to determine both empirical coefficients, we categorized the precipitation type with the help of air temperatures on a daily basis. It was assumed that temperatures below 0 °C result in solid precipitation, temperatures between 0 °C and 4 °C give mixed precipitation and air temperatures above 4 °C only liquid precipitation. Furthermore, the rain gauge is located in an open area and the summer period was defined from May to September and the winter period from October to April. The corresponding correction coefficients were calculated according to Richter (1995) and are provided in Table 3.1.

Table 3.1: Site specific wind exposition coefficient b [-] and empiric precipitation type coefficient ϵ [-] for different precipitation types at an open space gauge location.

Precipitation Type	***b***	***ϵ***
liquid (summer)	0.345	0.38
liquid (winter)	0.34	0.46
mixed	0.535	0.55
snow	0.72	0.82

3.3 Results and Discussion

3.3.1 Precipitation Measurements

Table 3.2 shows the monthly precipitation sums measured by the tipping bucket (TB) and calculated from the lysimeter balance data for the year 2012. The precipitation difference between both devices for the year 2012 is 145.0 mm implying a 16.4 % larger average lysimeter precipitation than TB. For the individual lysimeters the yearly precipitation ranges from 996.2 mm to 1037.7 mm (-3.0 to +1.0 % compared to the lysimeter average). This implies that the minimum and maximum precipitation differences between individual lysimeters and TB were 114.1 mm (12.9 %) resp. 155.6 mm (17.6 %), where precipitation for lysimeters was always higher than for TB. The monthly precipitation sums for the period April-October measured by the tipping bucket are smaller than the ones from the lysimeter average and differences range between 1 % in July and 42 % in September. The winter months show higher relative differences. The highest difference was found in March 2012, when the lysimeters registered an amount of precipitation double as large as the TB. The precipitation sums measured by lysimeter and tipping bucket correlate well on an hourly basis, especially from April to October with R^2 varying between 0.74 (Apr) and 0.99 (May), but with the exception of September (0.58). For winter months the explained variance is smaller with a minimum of 13% for February 2012.

Table 3.2: Monthly precipitation sums for lysimeter, tipping bucket, corrected tipping bucket (TBcorr) data and a comparison between the hourly precipitation values of lysimeter and uncorrected TB in terms of coefficient of determination (R^2), root mean square error and other statistics at the Rollesbroich study site for 2012. Missing data % refers to the percentage of hourly precipitation data not available for comparison.

Month	Lysimeter Average [mm]	Min. / Max. Lysimeter [mm]	Tipping Bucket [mm]	Tipping Bucket corrected [mm]	R^2	RMSE	LYS/TB %	LYS/ TB_{corr} %	Missing Data %
Jan	70.9	57.6 / 79.3	94.0	110.7	0.48	0.30	75.6	64.0	11.2
Feb	36.2	31.4 / 48.9	21.1	26.0	0.13	0.32	171.6	139.2	46.1
Mar	17.3	16.2 / 18.8	5.1	7.3	0.18	0.16	339.2	237.0	16.4
Apr	72.5	71.1 / 74.6	65.3	78.2	0.90	0.09	111.0	92.7	0.0
May	90.7	89.4 / 94.1	79.3	88.8	0.99	0.09	114.4	114.4	0.0
Jun	139.9	137.5 / 143.1	134.7	147.2	0.96	0.21	103.9	95.0	0.0
Jul	148.5	146.3 / 152.2	147.0	159.2	0.95	0.28	101.0	93.3	0.0
Aug	105.7	100.4 / 109.4	84.5	91.9	0.94	0.15	125.1	115.0	0.0
Sep	36.5	23.5 / 39.2	25.6	30.5	0.58	0.13	142.6	119.7	0.0
Oct	67.5	65.7 / 69.5	66.2	75.2	0.74	0.23	102.0	89.8	13.4
Nov	55.3	52.7 / 56.9	38.3	45.8	0.84	0.08	144.4	120.7	0.0
Dec	186.0	178.5 / 194.4	121.0	136.1	0.30	0.35	153.7	136.7	0.0
SUM /MEAN	**1027.1**	**996.2 / 1037.7**	**882.1**	**996.9**	**0.88**	**0.47**	**116.4**	**103.0**	**7.1**

The period April-August shows the smallest precipitation differences among the six lysimeters with monthly values of ±5 % in relation to the lysimeter average. In contrast, February, September, and December exhibit the highest absolute and relative precipitation differences among lysimeters with Variations between -13 and 13 mm (±35 %) with respect to the mean. Figure 3.3 shows the absolute daily differences in precipitation between lysimeter and TB measurements. It shows that the cases where lysimeters register slightly higher monthly precipitation sums than TB are related to single heavy rainfall events (June, July). In contrast, especially for February, the beginning of March, and the first half of December, larger fluctuations in differences between daily precipitation measured by TB and lysimeter are found, with less precipitation for TB than for lysimeters most of the days.

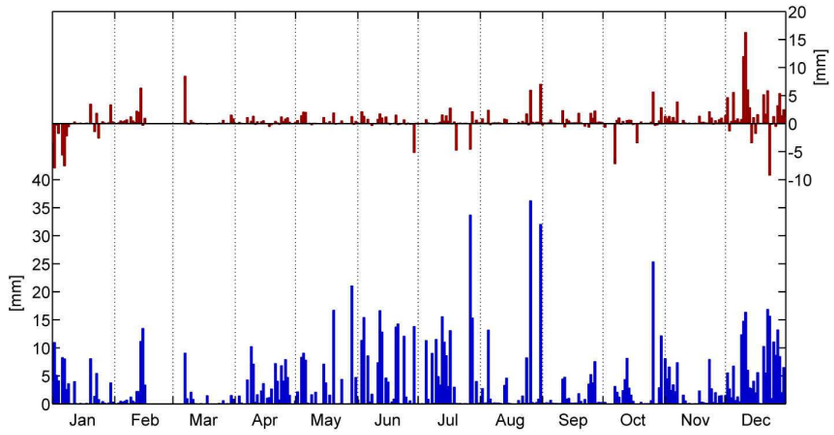

Figure 3.3: Daily precipitation sums of tipping bucket (blue) and difference in precipitation measurements between lysimeter and TB (red) at the Rollesbroich study site for 2012.

These periods coincide with freezing conditions and frequent episodes with sleet or snowfall. According to Nešpor and Sevruk (1999) these weather conditions are typically associated with a large tipping bucket undercatch because snowflakes are easier transported with the deformed wind field around a rain gauge. The surveillance system, which is installed at the lysimeter site, gives support for these findings. For example, a sleet precipitation event on March 7th explains 70 % (8.5 mm) of the monthly precipitation difference between lysimeter and TB. At this day the wind speed during the precipitation event was relatively high (4.4 m s^{-1}) and precipitation intensity varied between 0.6 and 2.9 mm h^{-1}. In general, winter measurement inaccuracies can be caused by frozen sensors and snow or ice deposit on the lysimeter surface. This situation may cause ponding effects close to the soil surface in the lysimeter and superficial runoff. In order to further address the lysimeter uncertainty, we calculated the average cumulative drainage and soil water storage with

minimum and maximum ranges for the individual lysimeters (Figure 3.4). The soil water storage was determined by the remaining term of the water balance on a daily basis. The total drainage, averaged over the six lysimeters was 411.2 mm for 2012 with a variation between 385.5 and 440.4 mm. The soil moisture storage change over the year varies between -5.1 mm to 28.3 mm with an average of +11.2 mm. The assessment of drainage volumes and changes in soil water storage was somewhat hampered by erroneous data related to drainage leakage (January) or system wide shut down due to freezing. However, the uncertainty in the water balance during those periods should have a minor effect on the short term calculations of lysimeter *P* and ET_a.

In order to explain differences in precipitation amounts between lysimeter and tipping bucket, the contribution of dew and rime to the total yearly precipitation amount was determined. The hourly data of lysimeter and TB were filtered according meteorological criteria. First, meteorological conditions were selected which favor the formation of dew, rime, fog and mist. Selected were small precipitation events between sunset and sunrise associated with high relative humidity (> 90 %), negative net radiation and low wind speed (< 3.5 m s^{-1}). Under these meteorological conditions it is probable that dew or rime is formed after sunset and before sunrise on cloud free days. For these days the difference in precipitation between TB and lysimeter is calculated if TB shows no precipitation signal or if the lysimeter has no precipitation signal. For the first case (*P*-TB=0) the total amount of the lysimeter precipitation is 24.5 mm, which contributes 16.9 % to the total yearly precipitation difference with the TB (and 2.4 % of the yearly lysimeter precipitation). The period

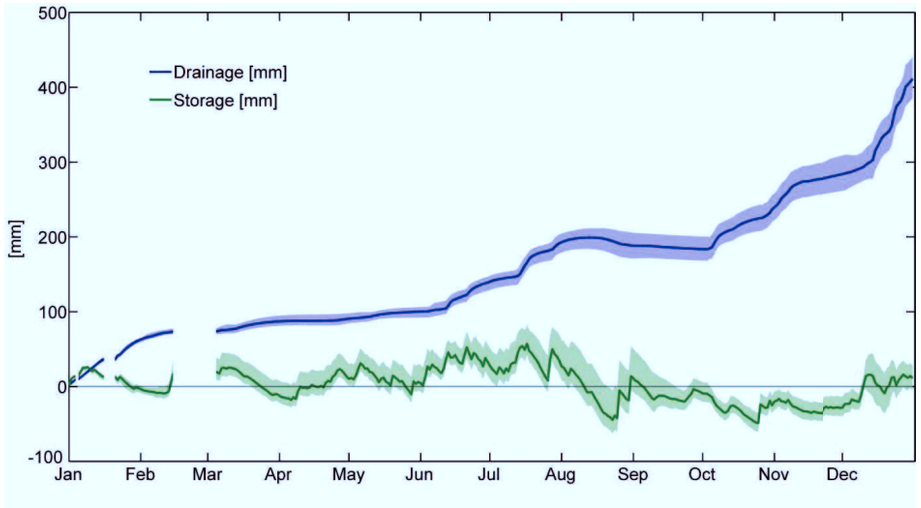

Figure 3.4: Cumulated average of lysimeter drainage and soil moisture storage on a daily basis. The colored areas indicate the range of minimum and maximum cumulated drainage and soil water storage for the individual lysimeters.

from April to August shows in general smaller precipitation amounts related to such situations. In contrast, likely dew and rime conditions where lysimeter precipitation is zero have a registered amount of TB-precipitation of 1.7 mm, which is only 0.2 % of the total measured TB amount for the considered period. A closer inspection of the precipitation data shows that both devices are able to capture dew and rime. However, a delay of some hours between TB and lysimeters was found. It is supposed that dew or fog precipitation was cumulating in the TB device until the resolution threshold of 0.1 mm was exceeded. This indicates that the TB resolution of 0.1 mm is too coarse to detect small dew and rime amounts in a proper temporal assignment. This confirms the expected ability of the lysimeter to measure rime and dew better than Hellman type pluviometers or tipping bucket devices. The surveillance system was used to check whether indeed dew/rime was formed on the before-mentioned days. On days which fulfilled the criteria and air temperatures close to or below 0 °C rime was seen on the photos. For days that fulfilled the conditions and temperatures above 0 °C camera lenses were often covered with small droplets.

Weather conditons with drizzle or fog occur frequently at the study site. This is related to humid air masses from the Atlantic which are transported with the dominating Southwestern winds and lifted against the hills in this region. The surveillance system was used to detect fog and drizzle situations during the year 2012. For those situations, a difference in precipitation between TB and lysimeters of 8 mm was found, which contributes 5.5 % to the yearly difference of both devices. Fig. 6 illustrates the example of May 5 – May 6 2012. The hourly photos of the site show drizzle, light rain and fog for this period. For both days the air temperature is close to the dew point temperature. The precipitation difference between tipping bucket and lysimeter over this period was 4.0 mm (Σ TB: 12.8 mm, Σ LYS: 16.8). The maximum difference was 0.5 mm and found at 6 h on the 5th of May in combination with fog. On May 5 during these conditions hourly TB precipitation is often zero and LYS mean precipitation rates are small (0.02 - 0.2 mm hr^{-1}). The comparison of individual lysimeter devices shows that not every lysimeter exceeds the predefined lower threshold of 0.055 mm for the AWAT filter (i.e. 5th of May 15:00, 6th of May 01:00-03:00 LT). However, in these cases at least three lysimeters show a weight increase, which supports the assumption that a real signal was measured instead of noise.

With the purpose of explaining the remaining difference in precipitation amount between TB and lysimeter, the relationship between wind speed and the precipitation differences was examined. The determined precipitation differences could in theory be explained by undercatch related to wind (Sevruk, 1981, 1996). It was checked whether correcting the tipping bucket data (TB_{corr}) according to the method of Richter (1995) could reduce the precipitation difference between lysimeter and TB. The total precipitation sum after correction is 996.9 mm for 2012, only 3% smaller than the yearly lysimeter average and within the range of the individual lysimeters.

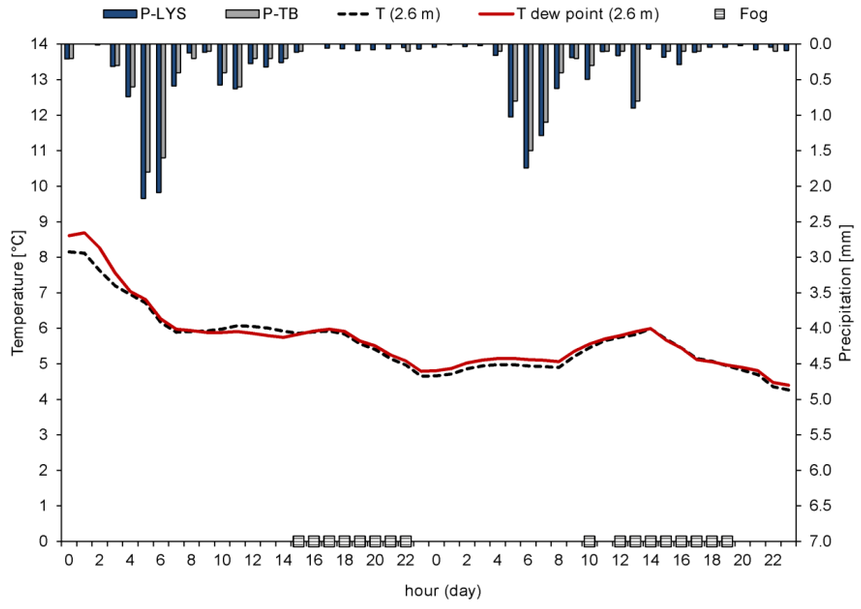

Figure 3.5: Precipitation, temperature and dew point temperature from May 5 – May 6 2012 at the Rollesbroich site. The fog symbol indicates the hours with fog occurrence (detected with installed surveillance system) for the investigated period.

The correction of TB data in general decreased the differences in the winter period (January – March, November - December). However, for the summer period the monthly precipitation sum of TB_{corr} mainly overestimated precipitation and tended to slightly increase the precipitation differences. In order to explore this relation further we examined the correlation between wind speed and precipitation residuals and found almost no correlation (Figure 3.6). A possible explanation is that other potential dew or rime situations are not properly filtered by the used criteria (e.g, dew occurs in case the net radiation is slightly positive or close to zero). Additionally, the correlation between undercatch and wind speed is dependent on precipitation type, intensity and drop size, for which information was limited during the investigation period. To investigate these relations we used the classification of precipitation types as outlined before. The contribution of liquid precipitation to total yearly precipitation is 80.9 % for the TB and 74.7 % for the lysimeters. The relative amount of solid precipitation was also different between the two measurement methods. Whereas for the lysimeters 7.8 % (79.7 mm) was classified as solid precipitation, the TB had only 0.6 % (5.6 mm) during periods with temperature < 0 °C. In relation to the total precipitation difference of 145 mm this means that 51 % of the difference was associated with solid precipitation events and 37 % with liquid precipitation events, which indicates the relatively large contribution of solid precipitation events to the total difference. The transition range (0-4 °C)

makes up 12 % of the total difference. Moreover, it was found that 78.7 % of the solid precipitation came along with small precipitation intensities (< 1.0 mm h^{-1}) and low wind speeds (< 2.0 m s^{-1}).

The surveillance system allowed to further investigate these large precipitation differences for air temperatures below zero. The snow depth at the lysimeters and surrounding areas is also an indication of precipitation amounts, assuming that 1 cm snow height corresponds to 1 mm precipitation. This method revealed that for conditions of light to moderate snowfall (< 4 mm h^{-1} precipitation intensity) the TB had a precipitation undercatch in January, February and December of 11.4 mm (7.9 % of total precipitation difference). The registered precipitation amount of the lysimeter under those conditions was realistic. However, during periods where the lysimeters were completely covered by snow (e.g. 1 – 15 February) precipitation estimates by lysimeter (up to 16 mm d^{-1} difference with tipping bucket) could not be confirmed by the camera system and were most probably influenced by snow drift or snow bridges. These situations explain 35.8 % (51.9 mm) of the total precipitation difference for 2012. For solid precipitation events a relationship (R^2=0.5) between precipitation differences and wind speed was found, but the number of datapoints was very limited (n=7). For conditions of liquid precipitation no correlation was found between residuals and wind speed (R^2<0.02).

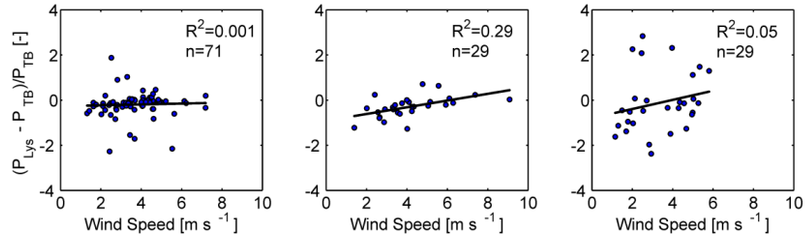

Figure 3.6: Relationship between wind speed and precipitation residuals relative to TB precipitation on a daily basis. The relationsships are classified according precipitation intensities of 1-5 mm (a), 5-10 mm (b), and > 10 mm (c). Potential rime and dew situation are excluded from the calculation.

3.3.2 Comparison of Evapotranspiration

In general, the yearly sums of ET_{PM} and ET_a-LYS were slightly higher than ET_a-EC; 6.1 % for ET_{PM} and 2.4 % for ET_a-LYS. The minimum ET_a of the individual lysimeter measurements (ET_a-LYSmin) is 467.1 mm, which is 7.9 % smaller than the lysimeter average (507.4 mm); the maximum (ET_a-LYSmax) is 523.1 mm (+ 3.1 %). This indicates that in general over the year 2012 evapotranspiration was limited by energy and not by water, as actual evapotranspiration was close to a theoretical maximum value for well watered conditions as estimated by ET_{PM}. This also

implies that our assumption of a stomatal resistance corresponding to well-watered conditions was justified. Water stress conditions would lead to decreased plant transpiration rates and increased stomatal resistance. Table 3.3 lists the evapotranspiration results of January – December 2012. In 2012 ET_{PM} was always close to ET_a-LYS and ET_a-EC and there are no months that ET_{PM} is clearly larger than measured actual evapotranspiration by lysimeter and eddy covariance. Root mean square errors of hourly ET_a sums vary between 0.01 mm h^{-1} in winter and 0.11 mm h^{-1} in summer months and are in phase with the seasonal ET dynamics.

We focus now on the comparison of monthly ET_a-LYS and ET_a-EC sums within the investigated period. During winter periods with low air temperatures and snowfall ET_a-LYS and ET_a-EC showed larger relative differences. For the period March to May ET_a-LYS and ET_a-EC differ approx. 6 % and ET_a-LYS exceeds ET_a-EC from June to August by 12 %. The larger difference in August (23 %) explains the yearly difference between ET_a-EC and ET_a-LYS. Hourly actual evapotranspiration from lysimeter and hourly actual evapotranspiration from EC are strongly correlated, but correlation is lower in the winter months. The registered monthly ET by the yearly ET amount is ca. 18 % reduced compared to gap free ET estimations.

Table 3.3: Monthly ET_a (by lysimeter and EC), ET_{PM} sums and R^2 between different ET data products on an hourly basis for 2012. Missing data % refers to the percentage of hourly ET data (ET_a-EC, ET_a-LYS) between sunrise and sunset not available for comparison. Hence, the total

	2012													
	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**	**Sum**	**Mean**
ET_a-EC [mm]	5.2	1.3	27.8	38.4	84.3	62.7	80.3	94.2	56.0	25.2	9.3	3.6	**488.**	
ET_{PM} [mm]	3.9	1.5	30.5	37.5	84.2	69.7	84.0	113.5	58.9	24.6	9.0	2.5	**519.8**	
ET_a-LYS [mm]	2.5	2.2	26.4	35.6	80.2	65.7	82.7	121.7	52.7	23.9	7.6	5.9	**507.4**	
Min. / Max. ET_a-LYS [mm]	2.1 / 2.7	1.3 / 3.1	25.9 / 26.8	34.4 / 37.6	75.2 / 85.2	62.1 / 68.2	67.8 / 91.0	116.8 / 125.2	49.6 / 58.8	21.9 / 27.1	6.8 / 8.9	3.0 / 8.7	**467.1 / 523.1**	
R^2 ET_a-EC - ET_a-LYS	0.02	0.02	0.82	0.76	0.79	0.84	0.86	0.86	0.66	0.66	0.39	0.06		**0.81**
R^2 ET_a-LYS – ET_{PM}	0.13	0.00	0.87	0.82	0.86	0.91	0.89	0.92	0.78	0.70	0.41	0.08		**0.89**
R^2 ET_a-EC – ET_{PM}	0.12	0.00	0.94	0.93	0.95	0.90	0.89	0.88	0.88	0.82	0.73	0.44		**0.91**
Missing Data %	33.2	36.9	8.1	23.5	21.5	26.5	21.9	12.9	14.0	25.8	25.0	45.3	**24.5**	

different lysimeters shows the largest variations in July with amounts that are up to 14.0 mm lower and 8.0 mm higher than the ET averaged over all six lysimeters.

Figure 3.7 shows the cumulative curve of the daily ET_a-LYS and ET_a-EC compared to ET_{PM} for 2012. From end of March 2012 the sums of ET_a-LYS and ET_a-EC tend to converge, but at the end of May ET_a-EC exceeds ET_a-LYS. In June and July ET_a-LYS and ET_a-EC are very similar, but in August ET_a-LYS is larger than ET_a-EC. After August the difference between ET_a-LYS and ET_a-EC does not increase further. The area in grey represents the range of minimum and maximum cumulative ET_a-LYS, measured by individual lysimeters. Until August ET_a-EC and ET_{PM} are slightly higher or close to the maximum measured ET_a-LYS. In August ET_{PM} increases further, wheras ET_a-EC falls below the minimum lysimeter value. Additionally, Figure 3.7 shows the course of the ET_a-EC without correction for EBD and for ET_a-EC max.. ET_a-uncorr is ca. 411 mm over this period, whereas ET_a-EC max is 567 mm, which shows the large potential uncertainty of the EC-data. The comparison illustrates that the application of the Bowen ratio correction to the EC data results in an actual evapotranspiration estimate close to the actual evapotranspiration from the lysimeter, whereas ET_a-EC uncorr is much smaller than the lysimeter evapotranspiration.

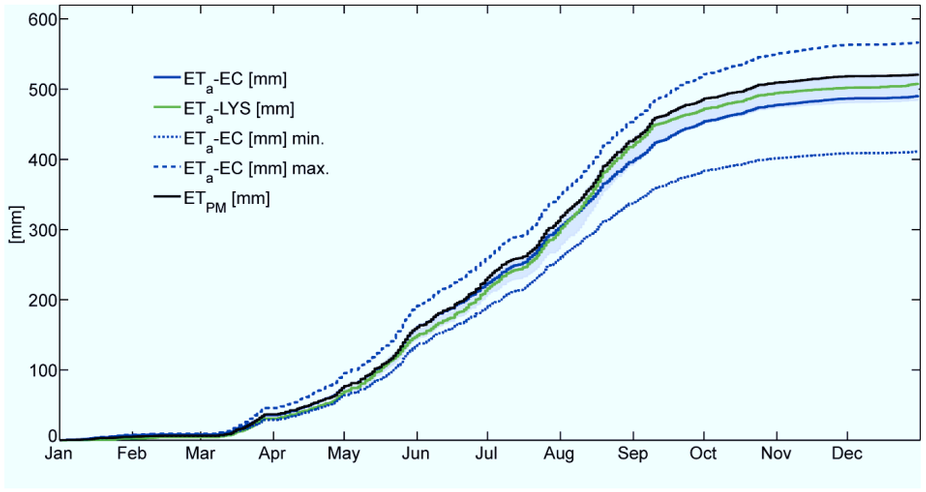

Figure 3.7: Cumulative ETa-LYS, ETa-EC (corrected according to Bowen ratio), ET_{PM} on hourly asis for 2012. Displayed are also ETa-EC max. and ETa-EC min. The area in grey shows the range of minimum and maximum cumulated ETa for the individual lysimeters.

Table 3.4 lists the monthly latent heat fluxes, the corrected LE fluxes (on the basis of the Bowen ratio) and the mean differences between both. It was found that the absolute difference is between 29.8 W m⁻² (August 2012) and 3.2 W m⁻² (February 2012). The EBD ranges from 12.6 % - 24.2 % for the period April to September. The yearly maximum was found in February with 36.9 %. EB deficits are site-specific, but these findings confirm the importance of EC data correction as suggested by Chavez et al. (2009).

In order to explain the differences between ET_{PM}, ET_a-EC and ET_a-LYS, we investigated the variations in radiation, vegetation and temperature regime and their impact on ET in more detail. The albedo could be estimated according to the measured outgoing shortwave radiation at the EC-station divided by the incoming shortwave radiation, also measured at the EC-station. The yearly mean albedo is 0.228, which is close to the assumed albedo of 0.23 for grassland. However, some periods (i.e. periods with snow cover) have a much higher albedo. Although albedo variations between different vegetation growth stages at different fields at the study site were considered as explanation for differences in ET_a, we assume similar albedo for ET_a-EC and ET_a-LYS measurement due to the central location of of the radiation measurements between the relevant fields.

The grass length is related to the LAI, which impacts water vapor flow at the leaf surface. Under well-watered conditions more surface for plant transpiration leads in general to higher transpiration rates by decreasing the bulk surface resistance. Figure 3.8 shows that the grass length measured at the Rollesbroich site is up to 80 cm before cutting. Unfortunately, grass height measurements are

Table 3.4: Measured mean monthly latent heat fluxes and corrections for EBD for 2012.

Month	Mean LE [W m^{-2}]	Mean LE corr. [W m^{-2}]	Differences LE corr. - LE	Difference mean LE corr. - LE %
Jan	21.9	29.8	7.9	36.2
Feb	8.7	11.9	3.2	36.9
Mar	78.1	94.0	15.9	20.4
Apr	86.4	101.8	15.3	17.7
May	138.7	164.6	25.9	18.7
Jun	111.8	125.8	14.0	12.6
Jul	136.3	157.2	20.9	15.3
Aug	151.6	181.4	29.8	19.6
Sep	104.0	129.2	25.2	24.2
Oct	61.3	79.6	18.3	29.9
Nov	24.4	32.1	7.7	31.4
Dec	22.0	28.3	6.3	28.5
SUM/MEAN	**78.8**	**94.6**	**15.9**	**24.3**

not available for the lysimeters but only for the surrounding field. It is assumed, on the basis of information from the video surveilance system, that grass heights generally are in good agreement between lysimeters (lysimeter site) and the surrounding field (lysimeter field), which allows a reconstruction of the grass length illustrated in Figure 3.8. However, the grass harvesting dates of lysimeters and surrounding field deviate in August and September and are given for the lysimeters in Figure 3.8.

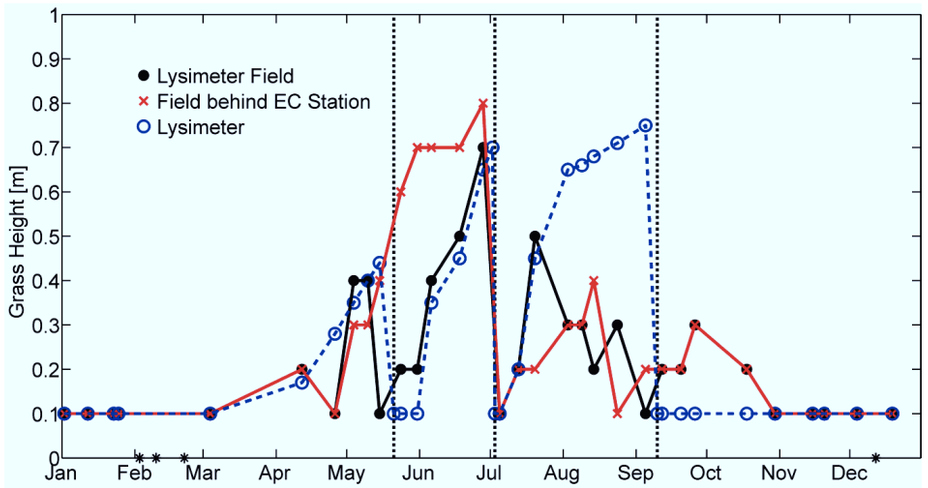

Figure 3.8: Grass heights at the lysimeter field, the lysimeter devices, and the field behind the EC station for 2012. The grass length at the lysimeter devices was reconstructed by comparing grass length measurements of the lysimeter field with the observations of the surveillance system. The star (*) indicates the presence of a snow cover. Grass cutting dates on lysimeter devices are marked by dashed lines.

Figure 3.9 illustrates the differences of the measured daily ET_a sums between lysimeter and EC. High positive and negative differences up to 2.1 mm/day were found from March 2012 – September 2012. In general, the differences of ET_a-LYS and ET_{PM} show smaller fluctuations than the differences of ET_a-EC and ET_{PM}. It was found that lysimeter harvesting affects the differences between ET_a-LYS and ET_{PM}/ET_a-EC. The differences were positive before harvesting and negative after harvesting indicating ET_a reduction due to the grass cutting effects. For the period from the 21st of May to the 3rd of July, a period with high grass length differences (Figure 3.8) between the lysimeter site and the field behind the EC-station, ET_a differences (ET_a-EC - ET_a-LYS) and grass length differences show a good correlation (R^2=0.58), which is illustrated in Figure 3.10. During the period with maximum grass length difference (24 May – 1 June) ET_a-EC is 26 % higher than ET_a-LYS.

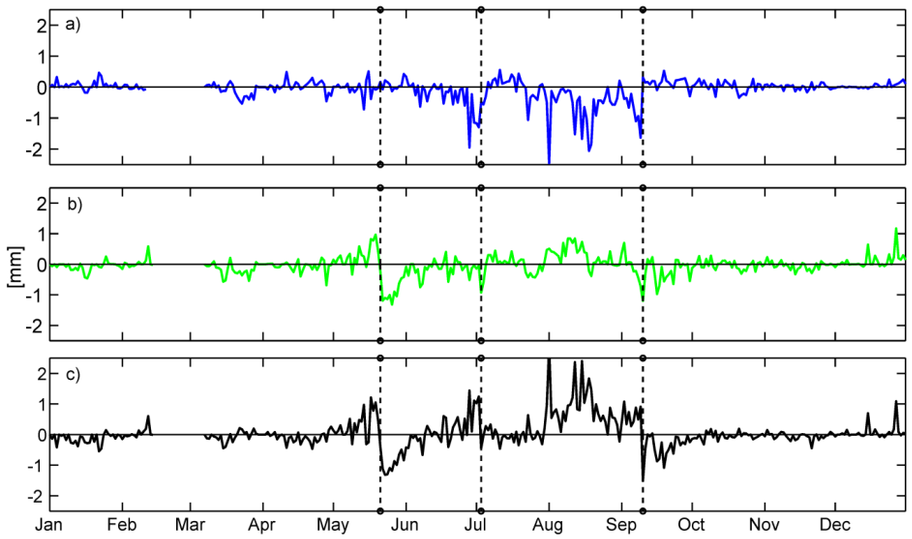

Figure 3.9: Differences between daily ET for 2012. Displayed are ET_a-EC – ET_{PM} (a), ET_a-LYS – ET_{PM} (b) and ET_a-LYS – ET_a-EC (c). The dashed lines indicate harvest at lysimeters.

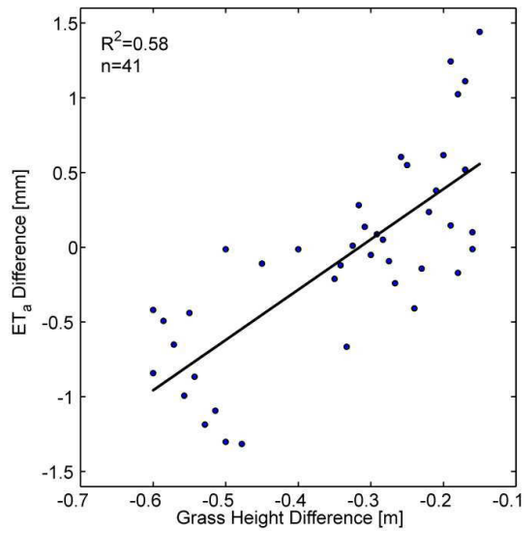

Figure 3.10: Relationship between grass length difference (between the lysimeters and the field behind the EC-device) and ET_a difference measured by lysimeters and EC station from May 21- July 3.

The differences between ET_a-EC and ET_{PM} do not show such a significant correlation with grass heights, although the relationship in August is in correspondence with the differences of ET_a-EC and ET_a-LYS. This could be related to the EC-footprint, because the EC station is centrally located in between the two investigated fields with different grass lengths. The EC-footprint might also include other surrounding fields with different grass heights. 80 % of the EC footprint is located within a radius of 100 m of the EC tower, and 70 % in a radius of 40 m, which is the approximate lysimeter distance. Therefore, the ET_a-EC estimations represent a spatial mean of a wider area, where cutting effects are averaged compared to the lysimeter point measurements. Figure 3.11 shows the mean hourly ET_a rates of lysimeter and EC as well as the ET_{PM} for 2012. In general, the daily courses and the daily maxima of ET_a-LYS, ET_{PM} and ET_a-EC correspond well. ET_a-EC shows higher peaks at noon in May and September compared to ET_a-LYS, but corresponds well to ET_{PM}. In contrast, ET_a-LYS exhibits the highest rates from June to August. The absence of a harvest of the lysimeter in August and the first September decade (in contrast to the surrounding fields) leads to potentially increased lysimeter ET_a measurements as compared to the surroundings due to an island position.

In order to examine whether lysimeter measurements could have been affected by a soil temperature regime different from the field, the temperature regimes of the lysimeters were compared to the field temperature. Figure 3.12 shows the daily mean soil temperature differences between the lysimeters, a nearby SoilNet device (SN 30) and the mean of all available SoilNet

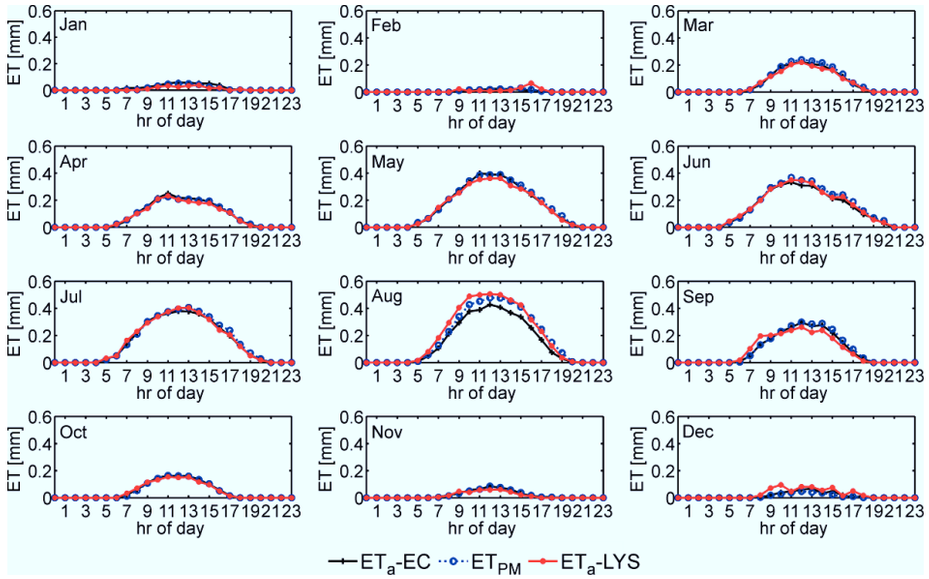

Figure 3.11: Mean daily cycle of ET_a-LYS, ET_a-EC and ET_{PM} for 2012.

devices installed at the southern study site. SoilNet temperatures were measured 5 cm below surface; lysimeter temperature measurements were conducted with SIS sensors in 10 cm depth. The temperature differences between the lysimeter and the nearby SoilNet device and the SoilNet mean are less than 1 K, which is as well the range of variation of the SoilNet device with respect to the SoilNet mean. In general the temperature differences increase until noon and then decrease again. Positive differences from May to July indicate higher lysimeter soil temperatures than the surroundings. However, a clear indicator for a bias caused by an oasis effect in the lysimeter measurements was not found. Feldhake and Boyer (1986) describe the effect of soil temperature on evapotranspiration for different grass types, which allow an estimation of ET_a increase caused by a differing lysimeter temperature regime. They showed that daily ET_a rates can increase with an increase of soil temperature (i.e. daily Bermuda grass ET_a rate increases from 4.3 mm/day to 6.4 mm/day (49 %) for a soil temperature increase from 13 to 29 °C). We used this linear relationship to roughly estimate the effect on ET_a for the period May – August on a daily basis. For this period the measured soil temperature with SN(30) for daylight hours ranged between 9.5 and 15.1 °C and between 9.3 and 15.5 °C for the lysimeter mean (SIS sensors). The mean difference is 0.67 K. This results in a total ET_a increase of 8.8 mm or 2.5 % in relation to the total ET_a-LYS of 349 mm on the basis of hourly ET. Therefore, the effect of increased soil temperature in the lysimeter is most probably limited, but not negligible.

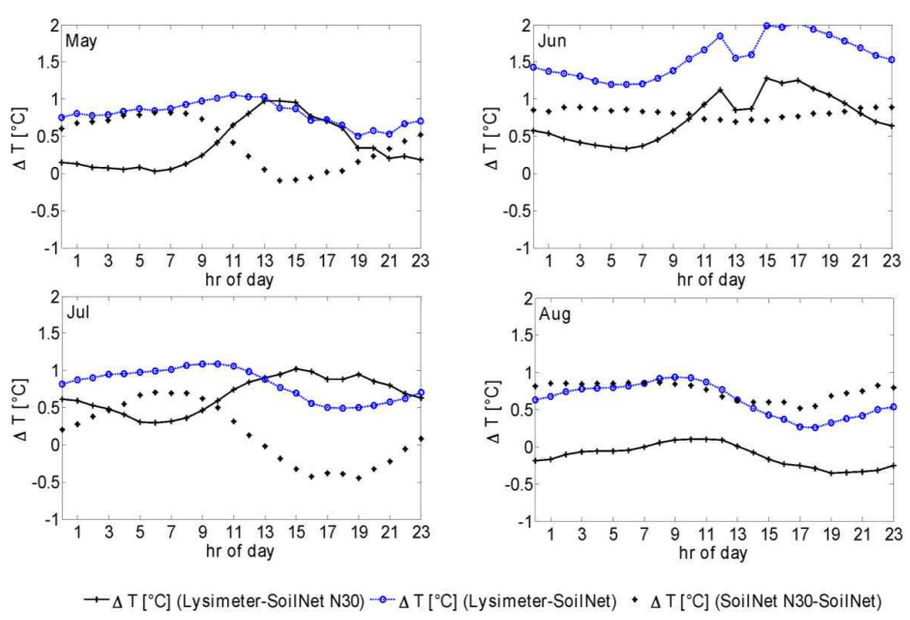

Figure 3.12: Differences in daily mean soil temperature (averaged over the six lysimeters), a nearby SoilNet device (SN 30) and the mean of all available SoilNet devices located at the study site.

3.4 Conclusions

This study compares evapotranspiration and precipitation estimates calculated using a set of six redundant weighable lysimeters with nearby eddy covariance and precipitation measurements at a TERENO grass land site in the Eifel (Germany) for one year (2012). The lysimeter data at a temporal resolution of one minute are processed with the AWAT filter (Peters et al., 2014), which takes account of the lysimeter noise due to random fluctuations caused by changing weather conditions. Additional precipitation measurements were conducted with a classical unshielded Hellmann type tipping bucket and compared with lysimeter data. For the ET_a comparison eddy covariance (EC) data is corrected for the energy balance deficit using the Bowen ratio method. Additionally, evapotranspiration and the evapotranspiration according the full-form Penman-Monteith equation were calculated.

The estimated hourly precipitation amounts derived by lysimeter and tipping bucket data show significant differences and the total precipitation measured by the lysimeter is 16.4 % larger than the tipping bucket amount. The relative differences in the monthly precipitation sums are small in the summer period, whereas high differences are found during the winter season. The winter months with solid precipitation exhibit the lowest correlations between lysimeter and tipping bucket amounts. Precipitation was measured by six different lysimeters and yearly amounts for individual lysimeters showed variations of -3.0 to 1.0 % compared to the yearly precipitation mean over all lysimeters. An additional comparison with corrected tipping bucket precipitation measurements according to the method of Richter (1995) shows in general a decrease of the monthly and yearly difference, which was 3 % after correction. In order to explain the differences in precipitation between the devices the contribution of dew, rime and fog to the yearly precipitation was analyzed. This was done by filtering the data for typical weather conditions like high relative humidity, low wind speed and negative net radiation which promote the development of dew and rime. For the identified cases a check was made with a visual surveillance system whether dew/rime was visible. During these conditions the lysimeter shows clearly larger precipitation amounts than the TB, which explains 16.9 % of the yearly precipitation difference. Fog and drizzling rain conditions, additionally identified with the help of the on-site camera system, explain another 5.5 % of the yearly precipitation differences. These findings indicate an improved ability of the lysimeters to measure dew and rime as well as fog and drizzling rain. The remaining 78 % of the precipitation difference between lysimeters and tipping bucket is strongly related to snowfall events, as under those conditions large differences were found. Lysimeter precipitation measurements are affected by a relatively high measurement uncertainty during winter weather conditions similar to TB and other common measurement methods. Thus, the limitations for the lysimeter precipitation measurements during those periods need further investigation. We found that during conditions where the lysimeters were completely covered by

snow, lysimeter records were unreliable, and contributed to 36 % of the total precipitation difference.

Actual evapotranspiration measured by the eddy covariance method (ET_a-EC) and lysimeter (ET_a-LYS) showed a good correspondence for 2012, with larger relative differences and low correlations in winter in contrast to high correlations and smaller relative differences in summer. The variability of ET_a of the individual lysimeters in relation to the lysimeter average was -7.9 to 3.1 % in 2012 with larger absolute differences in summer. Both ET_a-EC and ET_a-LYS were close to the calculated Penman-Monteith evapotranspiration (ET_{PM}), which indicates that evapotranspiration at the site was energy limited. The differences between ET_a-LYS, ET_a-EC and ET_{PM} were mainly related to harvesting management at the study site. A relationship between grass length at the lysimeter and differences between ET_{PM} and ET_a-LYS was found. Variable grass cutting dates for different fields around the EC-station and the lysimeter harvest lead to differences in actual evapotranspiration up to 2.1 mm day^{-1} for periods with larger grass length discrepancies.

The correction of the energy balance deficit with the Bowen ratio method resulted in ET_a-EC which was close to ET_a-LYS. If the correction was not applied, ET_a-EC was 16 % smaller than for the case where it was applied. In contrast, if the EB-deficit was completely attributed to the latent heat flux ET_a was 15.7 % larger than for the default case. These results point to the importance of adequate EC data correction.

Chapter 4 High resolution modelling of soil moisture patterns with TerrSysMP: A comparison with sensor network data

*adapted from Gebler, S., Franssen, H. J. H., Kollet, S. J., Qu, W., Vereecken, H., 2017. High resolution modelling of soil moisture patterns with TerrSysMP: A comparison with sensor network data. Journal of Hydrology, 547, 309-331.

4.1 Introduction

In an interacting system including lower atmosphere, vegetation, vadose zone and shallow and deep groundwater, the spatio-temporal variability of soil hydrological processes is of increasing research interest (Legates et al., 2011; Simmer et al., 2015). The soil water content (SWC) strongly affects the water and energy cycles by controlling the rainfall-runoff-response (Grayson et al., 1997; Robinson et al., 2008), partitioning net radiation in latent and sensible heat fluxes (Western et al., 2002; Robinson et al., 2008), and through interactions with the vegetation cover (Rodriguez-Iturbe et al., 2001). Thus, the characterization and prediction of the spatial and temporal patterns of soil water content is essential for understanding and quantifying the water, energy, and biochemical cycles of a given system (Schume et al., 2003; Ivanov et al., 2010). This insight is of great importance for many scientific and applied purposes (e. g. hazard prediction, soil degradation, and agricultural management) on local, regional and global scale (Vereecken et al., 2016). More specifically, advanced knowledge about the SWC spatio-temporal dynamics can support researchers in the optimization and uncertainty estimation of hydrological models (Heuvelink and Webster, 2001; Heathman et al., 2003), the construction and improvement of sensor networks (Heathman et al., 2009), and the calibration and validation of remote sensing products (Famiglietti et al., 1999; Choi and Jacobs, 2007; Rötzer et al., 2014; Greifeneder et al., 2016).

Soil moisture patterns can be related to topography (Grayson et al., 1997; Hu et al., 2010), vegetation cover (Gómez-Plaza et al., 2001; Schume et al., 2004), climatic conditions (Western et al., 2004; Martínez et al., 2014), antecedent SWC (Pan and Peters-Lidard, 2008), soil properties (Vereecken et al., 2007; Wang and Franz, 2015) and/or hysteresis effects (Vivoni et al., 2010; Rosenbaum et al., 2012). A common measure characterizing these patterns is the relationship between mean SWC ($\bar{\theta}$) and its corresponding spatial variability expressed as spatial standard deviation (σ_θ) or coefficient of variation. Recently, also fuzzy theory and empirical orthogonal function analysis were used in SWC data analysis and modelling studies (e. g. Graf et al., 2014;

Koch et al., 2015; Schröter et al., 2015; Hohenbrink et al., 2016). Previous modelling and field studies showed that the relationship between $\bar{\theta}$ and σ_θ usually can be described as an upward convex parabolic shape curve, where the spatial SWC variability increases during wetting from dry to intermediate soil moisture conditions and decreases for further wetting (Western et al., 2004; Choi and Jacobs, 2007). However, contributions of the different individual physical SWC controls are under discussion in the literature. For example, Hu and Islam (1998) and Vereecken et al. (2007) reported a causal relationship between SWC variability and soil properties like soil texture and structure, which are directly linked to soil hydraulic properties (e.g., saturated hydraulic conductivity, porosity, air entry pressure). Further, Rosenbaum et al. (2012) reported soil hydraulic properties in combination with antecedent SWC state mainly affecting short-term SWC during a wetting period. This is in line with the findings of Martini et al. (2015), Poltoradnev et al. (2016), and Wiekenkamp et al. (2016). During intensive precipitation events the SWC variability either increased due to the occurrence of preferential flow or decreased in case of intermediate or dry SWC state. Martínez García et al. (2014) investigated the role of vegetation and soil texture on the relationship of mean soil moisture and its variability using a one dimensional subsurface model with daily atmospheric forcing. They found soil texture, more than climate conditions, determining this relationship due to an observed relationship of the hydraulic conductivity and spatial variability. Manfreda et al. (2007) found a decrease of spatial SWC variability during wetting while soil texture and evapotranspiration increased SWC varibiulity during drying. They pointed out that the impact of different controling factors might also vary with depth. In contrast, Mascaro and Vivoni (2016) indicated vegetation dominating the SWC variability during the growing season for two semiarid study sites, whereas terrain and soil have a larger impact on SWC variability during periods with low vegetation cover. Despite these efforts made in identifying and characterizing the controls of SWC variability, the controversial results hamper a generalization of the relationship between $\bar{\theta}$ and its corresponding spatial variability. The highly non-linear interactions of the various controlling mechanisms make it challenging to quantify the contribution range of individual factors controlling SWC variability. However, studies often rely on simplified models with idealized test cases away from real natural surface and subsurface conditions. The SWC variability then is potentially affected by an underrepresentation and simplification of subsurface structure (Ivanov et al., 2010).

Recent integrated process based models could potentially give better estimates of soil moisture patterns (Chen and Hu, 2004) and the water balance (Maxwell and Miller, 2005; Kollet and Maxwell, 2008) to improve the feedbacks of subsurface regarding latent and sensible heat fluxes (Kollet, 2009), although they are computationally intensive. The parameterization of these models strongly influences the soil water states and fluxes calculated by the coupled simulations (Teuling and Troch, 2005). This is still a challenge and requires a broad data basis for model parameterization and validation (Cornelissen et al., 2014). In case the spatial variability and

structure of these data is not properly implemented in the model it is a potential bottleneck (Beven, 1996).

A number of simulation studies using integrated models were dedicated to simulating high resolution soil moisture patterns or dynamics of discharge and evapotranspiration. For example, Herbst and Diekkrüger (2003) modeled the spatio-temporal soil water content variability in a subcatchment of the river Rhine at Berrensiefen (North Rhine-Westphalia, Germany) with SWMS_3D (Šimůnek, et al., 1995) driven by atmospheric forcings for a two month period. In their geostatistical analysis, they found an inverse relationship of SWC semi variance and mean SWC during dry periods. At the same study site and using a similar model configuration, Herbst et al. (2006) compared five different scenarios of subsurface complexity ranging from homogeneous subsurface to a subsurface simulated with conditional stochastic geostatistical simulation with layered vertical and horizontal variations. They found a moderate but significant effect of the representation of subsurface complexity on runoff generation. The simulated runoff amount became more accurate with increasing complexity while root mean square errors decreased. Fang et al. (2015) compared soil moisture dynamics modeled with ParFlow-CLM (Kollet and Maxwell, 2006; Maxwell and Miller, 2005) and measurements of a sensor network for a small forested headwater catchment in the Eifel (Germany) with focus on fast lateral flow above a bedrock layer. They concluded that spatial SWC patterns could be better reproduced with a layered subsurface setup with zones of heterogeneous soil units and porosity variations. Although the total SWC variability could not be reproduced, they found an improved reproduction during wet seasons compared to dry seasons and suggested a better representation of heterogeneous soil hydraulic parameters for further studies. Limitations in reproducing SWC variability were potentially related to model parameterization and not adequately representing vertical bypass flux and lateral drainage. Cornelissen et al. (2014) investigated the impact of model resolution and bedrock hydraulic conductivity on water balance and SWC variability at the same study site. The introduction of a bedrock layer was almost without any effect on the soil moisture variability and dynamics, but Cornelissen et al. (2014) recommended a minimal model resolution of 25 m for a detailed reproduction of soil moisture patterns. In contrast, Ivanov et al. (2010) were more focused on vegetation parameterization and the hysteretic cycle on the hillslope scale. They suggested a homogenizing effect of vegetation on SWC variability affecting the topographic redistribution of precipitation. Fatichi et al. (2015) differentiated between biotic and abiotic factors controlling soil moisture variability in a synthetic hillslope model study. They studied different subsurface heterogeneities using as forcing measured meteorological data from different climate zones. While abiotic controls like soil hydraulic properties dominated in wet climates, biotic controls had a stronger impact in dry climates. They also pointed out that heterogeneity of soil properties or other soil characteristics (e. g. micro-topography, litter layer) increased the variability of the soil water content on the level of field observations, but also potentially remove the various signatures of

other (biotic) controlling factors. Another example for modelling high resolution soil moisture patterns on a regional scale was presented by Mascaro et al. (2015). They concluded that SWC patterns are related to soil hydraulic parameters, while vegetation and terrain elevation were more affecting runoff and evapotranspiration. These examples show that the holistic view on SWC variability and water balance components within an integrated hydrologic system is associated with uncertainty. For instance, subsurface soil hydraulic properties are usually derived using pedo-transfer functions (PTF) and observed soil data, or alternatively, by inverse methods. Several studies (e. g. Gutmann and Small 2007; Hohenbrink and Lischeid 2014) have shown that subsurface parameterization highly affects even less complex hydrologic model systems. Given the variety of PTFs (e. g. Rawls and Brakensiek, 1989; Schaap et al., 2001; Wösten et al., 2001), uncertainty of soil hydraulic parameters in combination with model complexity potentially have a high impact on model SWC variability and other water balance components.

Our study compares simulated spatio-temporal distributions of soil water content, evapotranspiration, and discharge with measurements for a headwater grassland sub-catchment in the Eifel (Germany) for a two-year period (2011-2012). We perform model calculations with the Terrestrial Systems Modeling Platform (TerrSysMP) developed by Shrestha et al. (2014) with a focus on soil properties as controlling factors. TerrSysMP contains the integrated physically based models ParFlow (Ashby and Falgout, 1996; Jones and Woodward, 2001; Kollet and Maxwell, 2006) and the Community Land Model (Oleson et al., 2004; Oleson et al., 2008) in a fully coupled manner taking lateral subsurface flow, overland flow and topography into account, which might result in redistribution of soil moisture. Detailed long term data for model parameterization and evaluation (e.g., high resolution soil moisture data in space and time, evapotranspiration measured by the eddy covariance method and lysimeters, discharge measurements) are provided by the infrastructure initiative Terrestrial Environmental Observatories (TERENO) (Zacharias et al., 2011). The simulations are performed with different parameterizations of soil hydraulic properties ranging from homogeneous to fully heterogeneous, using geostatistical simulations at a resolution of 10×10 m. We sampled hydraulic conductivity and related Mualem-van Genuchten parameters from probability density functions originating from two alternative initial datasets: (i) measurement data and pedotransfer functions; and (ii) estimated soil hydraulic parameters by 1D inverse modelling. Compared to previous studies, we extend analyses by taking into account the two alternative methods to sample hydraulic conductivity and Mualem-van Genuchten parameters. In addition, the uncertainty in the estimation of soil hydraulic properties is also taking into account, by conducting an ensemble of model simulations for each complexity scenario.

With this experimental setup we want to address the following research questions:

(1) How do the different subsurface heterogeneity scenarios affect the reproduction of spatio-temporal patterns of soil water content, discharge and evapotranspiration?

(2) To what degree can soil water content variability be explained by subsurface heterogeneity?

(3) What differences in model output arise due to the alternative methods to estimate prior distributions of soil hydraulic properties?

4.2 Study Site and Data

4.2.1 The Rollesbroich Catchment

A description of the TERENO Rollesbroich catchment is given in chapter 2.1. Figure 4.1 provides an overview of measurement equipment used in this study. It also shows a map of the study site and the location of the drainage system installed to avoid persistent saturation of the upper soil horizons and flooding. The diameters of the clay pipes of the ca. 80-year old system range from 3 to 20 cm and increase in the direction of the system outlet close to the Kieselbach source.

All atmospheric data were recorded at the EC-tower location for the period 2011- 2012 and were used as hourly meteorological forcing for the CLM model simulations (Figure 4.2). For 2010 and for gap filling, we used off-site data from the nearby LUA NRW station. Discharge, evapotranspiration, and soil water content data are the basis for verification of model predictions.

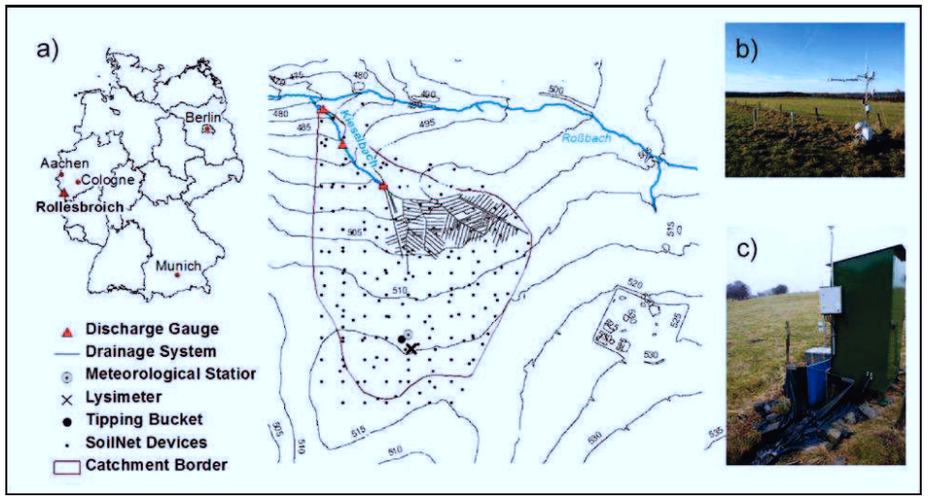

Figure 4.1: Map of the Rollesbroich study site (a), Germany, showing the locations of the lysimeter, SoilNet devices, discharge gauges, the drainage system and the meteorological station. The EC-tower and the tipping bucket device are located within flat terrain of the southern test site whereas the Venturi-gauge (c) is positioned at the catchment outlet in the north.

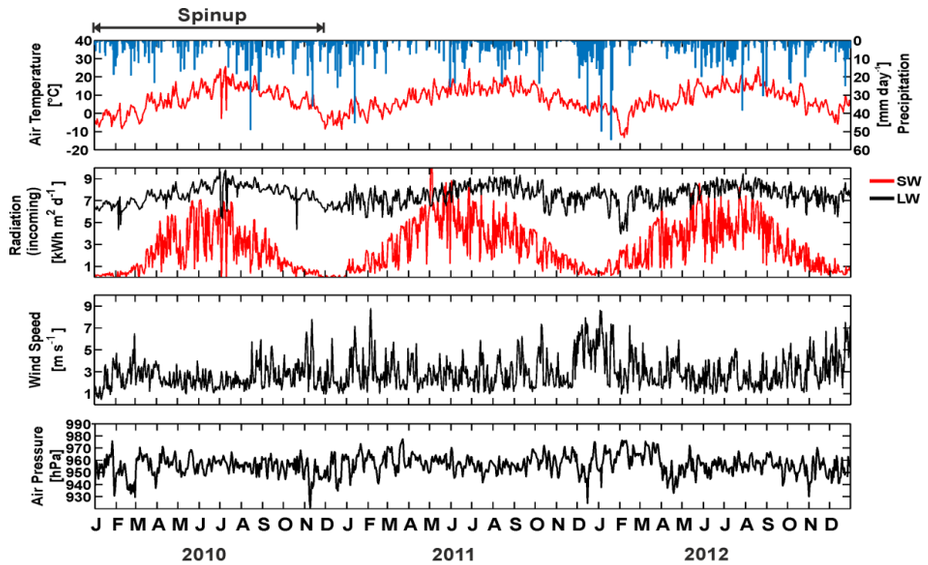

Figure 4.2: Daily precipitation sums and daily average air temperature, incoming shortwave (SW) and longwave (LW) radiation and wind speed used as forcing for the CLM model for the period 2010 – 2012. Data from the year 2010 were used for model spin up.

4.2.2 Data Processing

The comparison of the on-site tipping bucket gauge (TB) with lysimeter data at the Rollesbroich site for 2012 (Gebler et al., 2015) indicated an improved precipitation estimate using the correction method of Richter (1995), which accounts for evaporation, wind and wetting losses on a daily basis (Eq. 4.1 and Eq. 4.2). As the TerrSysMP model input requires hourly data, we redistributed the corrected daily precipitation amount according the hourly precipitation fraction (Eq. 4.3).

$$P^{\mathrm{cor}} = P + \Delta P \tag{4.1}$$

$$\Delta P = bP^{\epsilon} \tag{4.2}$$

$$P_h^{\mathrm{cor}} = \frac{P_d^{\mathrm{cor}}}{P_d} P_h \tag{4.3}$$

where P_d^{cor} is the corrected daily precipitation [M T^{-1}], P_d is the measured daily tipping bucket precipitation [M T^{-1}], ΔP_d is the estimated precipitation deficit [M T^{-1}], b the site specific wind exposition coefficient [-], ϵ the empirical precipitation type coefficient [-], P_h is the measured hourly tipping bucket precipitation, and P_h^{cor} is the corrected hourly precipitation [M T^{-1}]. This is based on the assumption that precipitation errors at the study site are mainly affected by wind and

precipitation type. Diurnal variation in evaporation loss of precipitated water from the rain gauge is neglected. If evaporation loss would be important it should have been considered that this loss is larger during daytime and in summer than at nighttime and in winter. In order to estimate wind exposition coefficient (b), we classified the rain gauge location as open area for the complete measurement period similar to Gebler et al. (2015). The precipitation type (ϵ) was categorized by daily on-site measurements of air temperature. All specific location characteristics are provided in Table 3.1.

The long term analysis of the SPADE sensor SoilNet data revealed for several locations a positive upward trend in measured soil water content. For the measurement period 2011 – 2012 this trend is indicated by rising maximum soil water contents at soil saturation conditions during wet periods. The drift also continued in 2013, a relatively dry year, where the majority of the sensors gradually went out of service. The soil moisture trend was potentially caused by enlarged power consumption due to lingering moisture infiltration into measurement boxes and therefore not accounted for in the sensor calibration of Qu et al. (2016).

After taking out unreliable data, we calculated a linear trend (f_{trend}) for each sensor between the highest soil water content in December 2011 and December 2012:

$$f_{\mathrm{trend}}(t) = a + \frac{\mathrm{SWC}_{\mathrm{tmax12}} - \mathrm{SWC}_{\mathrm{tmax11}}}{t_{\mathrm{max12}} - t_{\mathrm{max11}}}\, t \tag{4.4}$$

Where SWC_{tmax11} [L^3 L^{-3}] is maximum SWC in December 2011 at time t_{max11} [T], SWC_{tmax12} [L^3 L^{-3}] is maximum SWC in December 2012 at time t_{max12} [T], t [T] is the respective time step, and a the intercept constant [L^3 L^{-3}] .

The correction is based on the assumption that we find similar SWC during the wet periods in both years. The reference (f_{ref}) is the average maximum soil water content over the years 2011 and 2012:

$$f_{\mathrm{ref}} = \frac{\mathrm{SWC}_{\mathrm{tmax12}} + \mathrm{SWC}_{\mathrm{tmax11}}}{2} \tag{4.5}$$

The onset of the before mentioned moisture issue affecting the sensors cannot be determined precisely and independent reference SWC datasets were not available in an adequate spatial resolution, and therefore the definition of the reference was somewhat arbitrary.

Thus, SWC differences between these observation points indicate a linear trend, which was used to correct measured soil water contents (Eq. 4.6).

$$SWC_t^{\mathrm{corr}} = SWC_t + f_{\mathrm{ref}} - f_{\mathrm{trend}}(t) \tag{4.6}$$

where SWC_t^{corr} is the corrected hourly soil water content [L^3 L^{-3}] at time t [T], SWC_t is the measured soil water content [L^3 L^{-3}] at time t, $f_{\mathrm{ref}}(t)$ the estimated reference SWC, and $f_{\mathrm{trend}}(t)$ the trend function

The trend correction procedure was conducted at 82 SoilNet locations at all three depths individually and for both sensors installed at a given location and depth. For the evaluation of the simulations, we mainly use these SoilNet locations from the southern part in order to have continuous time series. SWC data was not available for February and beginning of March 2012 because measurements are unreliable during frozen soil conditions. Data from the northern part were not available before August 2012 and went partly out of operation in mid-2013.

Measurements of actual evapotranspiration (ET) measurements were conducted with a set of six weighable lysimeters (TERENO-SoilCan project, UMS GmbH, Munich, Germany) arranged in a hexagonal design at the southern part of the site ca. 30 m away from the EC-tower. Each lysimeter contains an on-site silty loam soil profile and the 1 m² surface is covered with grass species. The lysimeters resemble the field soil water regime, the lower boundary condition is monitored by tensiometers (TS1, UMS GmbH, Munich, Germany). The matric potential differences between lysimeter and surroundings are automatically compensated by suction rakes (SIC 40, UMS GmbH, Munich, Germany) injecting or removing water from the lysimeter. Further technical details can be found in Unold and Fank (2008). For comparison we used an hourly average actual ET derived from lysimeter weight data (temporal resolution: 1 min) of all six lysimeters. The data processing includes outlier elimination and signal smoothing for the individual lysimeter devices with an adaptive time window according to the method of Peters et al. (2014).

In addition to the lysimeter measurements, sensible and latent heat fluxes were measured by the eddy covariance (EC) station. The EC data were corrected for the energy balance deficit according to the method of Kessomkiat et al. (2013). In a first step, the energy balance deficit was determined over a 3-hour moving window surrounding an EC-measurement. The evaporative fraction was calculated for an interval of 7 days surrounding the particular EC-observation. Finally, the half-hourly latent heat flux as measured by EC was corrected for the energy balance deficit, determined over the 3-hour moving window, by adding part of the energy deficit to the latent heat flux. The added part is the evaporative fraction times the energy balance deficit. Lysimeter and eddy covariance measurements were only available for 2012. Further details on the data processing of EC and lysimeter actual evapotranspiration can be found in Gebler et al. (2015).

To establish a water balance for the observed data in 2012, we filled data gaps (ca. 14 % of all data) with ET data calculated with the full-form Penman-Monteith equation (ET_{PM}) as presented by Allen et al. (1998) on an hourly basis. These gaps mainly occur outside the growing season with low grass height. More detailed information on the ET_{PM} calculation for the Rollesbroich study site including variable bulk and surface resistance of the plants can also be found in Gebler et al. (2015). All required meteorological input for ET_{PM} calculation was taken from the on-site measurements. This includes also measured net radiation and soil heat flux. The wind speed data were corrected to the 2 m FAO-standard using the wind profile relationship according to Allen (1998).

Hourly discharge measurements for the Kieselbach are conducted with a Venturi-Gauge Weir close to the catchment outlet and two upstream Tomson gauges close to the headwaters of the Kieselbach.

4.3 Methods

4.3.1 Model setup

For this study, ParFlow and CLM were applied within the TerrSysMP framework (Shrestha et al., 2014) using the external coupler OASIS3-MCT (Valcke, 2013). A model description can be found in Chapter 2.2. The following paragraphs hence focus on the model setup including model domain, vegetation parameterization and subsurface parameter sampling.

The model domain for the Rollesbroich study site is 1280 × 1120 m with a total model depth of 3.2 m. It was built with a lateral spatial resolution of 10 × 10 m and a variable vertical resolution associated with the terrain following grid (TFG). The upper layer has a depth of 0.025 m, the ten layers below a vertical resolution of 0.05 m and the 12th layer a depth of 0.1 m. Further below, the vertical model resolution increases from 0.2 m (layer 13 – 17), to 0.5 m (layer 18 – 19), and 0.575 m (layer 20) at the bedrock horizon. The lower boundary at 3.2 m soil depth was made impermeable. The thickness of the 10 layer CLM subsurface was adapted to match the ParFlow counterparts. This results in 128 x 112 grid cells with 20 layers (286720 rectilinear elements). The catchment area of 38 ha was extended further for practical reasons in eastern, western and southern direction including parts of the Rollesbroich urban area, streets and other artificial structures. In order to determine the flow paths and location of streams, the topographic slopes in *x* and *y* direction were calculated from a digital elevation model (DEM). To avoid sinks in the flow path, pit filling was carried out with a four neighbor pit filling algorithm. To avoid disconnection of the channels caused by scale discrepancies of the channel bed (0.3 – 1 m width) and the 1 × 1 m resolution of the elevation map, the slopes within Kieselbach and Roßbach locations are connected manually, with the help of a mask of the stream network in the domain.

The hourly spatially uniform forcing data (Figure 4.2) were from preprocessed and corrected meteorological data measured on-site (2011-2012) and from the nearby LUA NRW (2010). Land use was C3-grass throughout the model domain with a leaf area index ranging from 0.3 – 3.0 [-] according the annual cycle. The grass rooting depth was set to 0.5 m in CLM with rooting function parameters (roota, rootb) having values of 10.6 (roota) and 6.0 (rootb). With this rooting parameterization we find 90% of all roots within the upper 30 cm of the soil profile, which is in accordance with literature values (e. g. Brown et al., 2010). Along the eastern, western, northern and southern ParFlow domain boundaries a no flow condition was set. Overland flow was parameterized with a Manning's roughness value of 0.001 h m$^{-1/3}$ accounting for the small streambed of 30 cm with high bank vegetation density. Model spin-up was conducted for a period of one year (2010) beginning with an initial hydrostatic equilibrium condition with a groundwater table at 1.5 m depth. The validation period was May 2011 – Dec 2012 for SWC and Jan 2011 – Dec 2012 for discharge and ET.

In order to investigate the sensitivity of the model output to the parameterization of the van Genuchten subsurface hydraulic properties we tested two alternative methods for subsurface sampling:

1. Van Genuchten hydraulic parameters (van Genuchten, 1980) estimated from texture and bulk density measurements using a ROSETTA (Schaap et al., 2001) pedo-transfer function (PTF). We henceforth denote this method as ROS-setup.
2. Optimized van Genuchten hydraulic parameters estimated by 1D-inverse modelling with HYDRUS-1D (Šimůnek et al., 2008) and the shuffle complex evolution algorithm (SCE-UA) by Duan et al. (1992). Input for these simulations were the joint probability hydraulic parameter distributions according to Carsel and Parrish (1988) as provided in Table 4.1. In the following, we refer to this method as SCE-setup.

Table 4.1: Mean values, standard deviations, and correlations between transformed soil hydraulic parameters for a silty loam soil according to Carsel and Parrish (1988).

Parameter	**Mean**	**Standard deviation**	K_s	θ_r	α	n
K_s [cm day^{-1}]	2.69	1.23	1.00	-	-	-
θ_r [cm^3 cm^{-3}]	0.06	0.02	-0.36	1.00	-	-
α [cm^{-1}]	0.02	0.01	0.98	-0.30	1.00	-
n [-]	1.85	0.12	0.73	-0.59	0.78	1.00

Figure 4.3 gives an overview of the subsurface sampling workflow for both methods. In the first case (ROS), the Mualem-van-Genuchten subsurface parameters were obtained using texture and bulk density information of 202 soil samples at three depths (Schiedung, 2015) and the ROSETTA pedo-transfer software (Schaap et al., 2001). We made use of the ROSETTA H3 PTF, which relies on bootstrap and neural network methods and additionally provides information of the PTF parameter uncertainty (Schaap et al., 2001). Table 4.2 shows the range of the derived van Genuchten parameters at the study site for each horizon. As an alternative to the on-site measurements, inverse modelling can be used to estimate soil hydraulic properties (Duan et al., 1992; Vrugt et al., 2003). Qu et al. (2014) showed that spatial variability at the Rollesbroich site can also be captured with a HYDRUS-1D approach. For the second case (SCE), the range of optimized van Genuchten parameter sets for the 82 Rollesbroich locations can be found in Table 4.2. Further details on model and optimization setup are provided by Qu et al. (2014). As the SCE-UA parameter estimation typically comes along without estimation of uncertainty for soil hydraulic properties, we considered the discrepancy between the spatial support at the optimization location and the model discretization. We hence performed ordinary block kriging (Burgess and Webster, 1980) using the VESPER (Whelan et al., 2002) software in an unconditional manner at 10×10 m model resolution for all soil hydraulic parameters. Ordinary block kriging is a standard

method for the interpolation of soil properties which considers the difference in the spatial support of the measurement and the grid cell. The estimated block kriging uncertainty at the grid support scale shows a similar uncertainty as the uncertainties of the ROSETTA estimation (e. g. α: ± 1.3 cm^{-1} (SCE); ± 1.5 cm^{-1} (ROS)).

The subsequent random sampling of van Genuchten soil hydraulic properties consists of four steps (Figure 4.3). First, for the upper three layers, parameter values were randomly sampled from a multivariate normal distribution taking into account mean and (co-)variance from the ROS and SCE datasets. From these datasets the covariances between the individual soil hydraulic properties were derived. This was done separately for each location and at each depth. For the underlying siltstone and sandstone bedrock (horizon 4), soil hydraulic properties were randomly sampled for each sensor location following the parameter ranges (Table 4.2) of Bogena (2003). Second, the logarithmic hydraulic conductivity ($\log_{10}K_s$) of each location for the layers at 5, 20, and 50 cm was additionally perturbed with a value taken from a normal distribution with mean equal to zero and a standard deviation of ± 0.25 $\log_{10}K_s$. With this procedure, we account for extra model uncertainty which may originate from an underestimation of uncertainty by the pedotransfer functions and 1D inversion.

After these first two steps we have defined mean values at each location for $\log_{10}K_s$, which vary between soil horizons and soil units. Third, a spatial heterogeneous field of $\log_{10}K_s$ was generated for each soil layer with sequential Gaussian simulation using GCOSIM3D (Gómez-Hernández and Journel, 1993), conditioned to the hydraulic conductivities estimated from the local probability density function at the sensor locations. The variogram parameters (range, sill, nugget) were obtained from fitting the experimental semi-variograms with an exponential model for each soil layer (horizon 1-3). This was done separately for the ROS and SCE-setups. For the bedrock layer (horizon 4) only very limited information regarding the spatial dependence of the hydraulic properties was available. Therefore, we used the variogram parameters of horizon 3 (ROS- or SCE-setup) for the bedrock stochastic simulations. Horizon 4 was subdivided into two parts. The upper bedrock parameterization (0.2 - 0.5 m thickness) mimics the high porosities and conductivities of the weathered sandstone and siltstone which makes lateral preferential flow though subsurface cracks and macropores very likely (Guo et al., 2014; Liu and Lin, 2015). Therefore for this layer the lateral conductivity was enhanced by a factor of 10, which also takes the contribution of lateral flow by the drainage system into account. In contrast, the lower bedrock parameterization emulates the lower hydraulic conductivity of solid rock. In a last step, $\log_{10}\alpha$, $\log_{10}n$, θ_r and θ_s were estimated on the basis of their relations with $\log_{10}K_s$. These were determined from the multivariate normal distributions determined in the first step, for each soil horizon. The spatially heterogeneous $\log_{10}K_s$, renders $\log_{10}\alpha$, $\log_{10}n$, θ_r, and θ_s also spatially variable. Table 4.3 provides an overview of the van Genuchten soil hydraulic parameters after sampling.

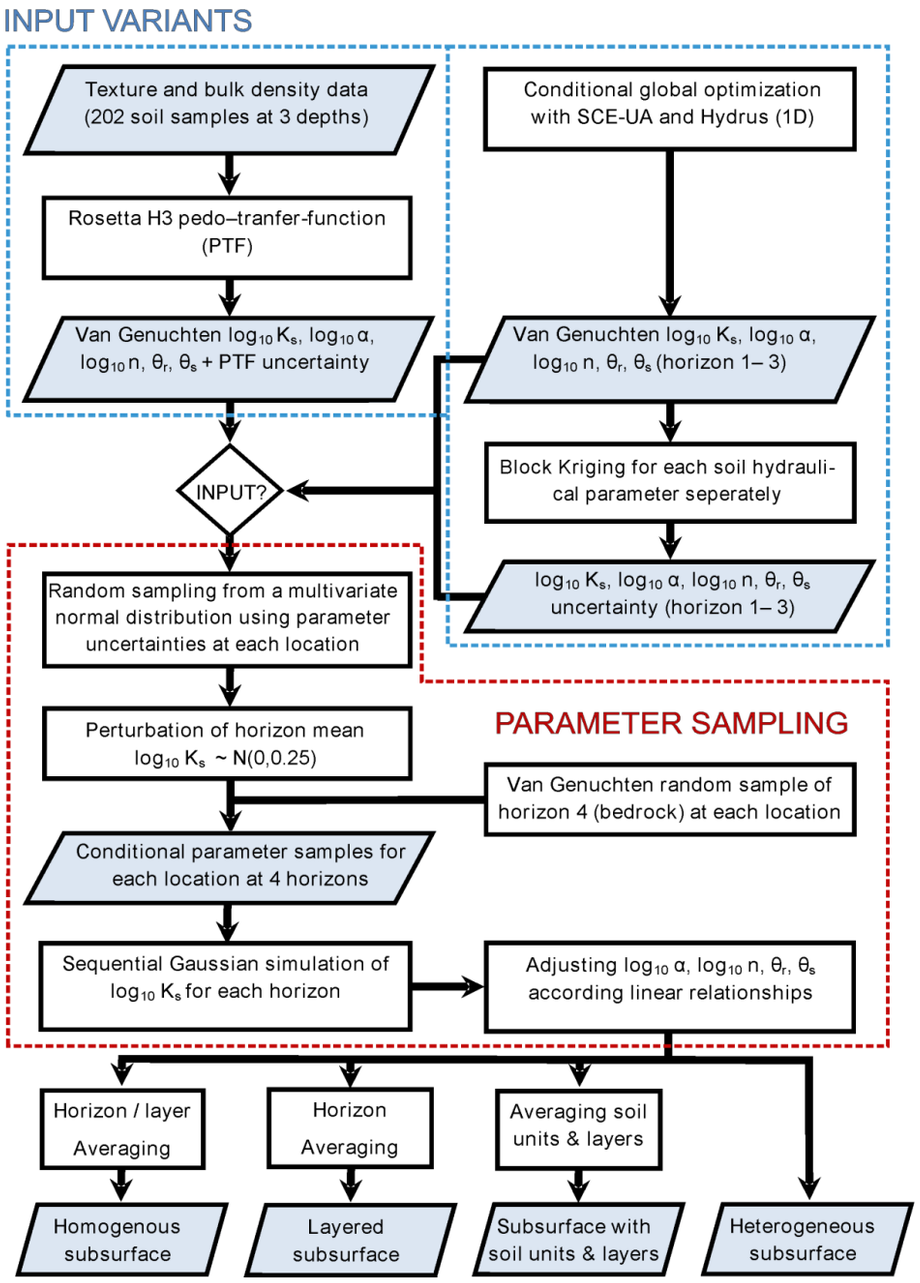

Figure 4.3: Flowchart of four different subsurface parameter setups. The dashed line in blue indicates the two alternative ROSETTA (left) and SCE-UA (right) inputs. The red dashed line marks random sampling steps similar for both input variants.

Table 4.2: Range of saturated hydraulic conductivity (Ks), inverse air entry suction (α), residual water content (θ_r), water content at saturation (θ_s) and a factor related to pore size distribution (n) for three soil horizons and the bedrock layer at the Rollesbroich site for the ROS- and SCE-scenarios. The upper table shows the parameter range of ROSETTA pedo-transfer function, including the estimation of mean bedrock according to Bogena (2003), while the lower table indicates the range of SCE-UA optimization results at each soil horizon.

ROSETTA	**Horizon 1**	**Horizon 2**	**Horizon 3**	**Bedrock**
Depth [m]	0.15	0.35	1.0 – 1.5	3.2
K_s [cm/day]	42.7 – 495.0	6.6 – 261.6	3.06 – 117.5	0.9 – 0.002
α [1/cm]	0.004 – 0.008	0.004 - 0.007	0.004 – 0.02	0.005
θ_r [cm³/cm³]	0.06 – 0.1	0.05 – 0.09	0.04 – 0.09	0.07
θ_s [cm³/cm³]	0.45 – 0.66	0.36 – 0.57	0.29 – 0.55	0.3
n	1.6 – 1.7	1.5 – 1.7	1.3 – 1.7	1.5 -2.0
SCE-UA Optimization				
K_s [cm/day]	0.15 – 86.5	0.93 – 19.44	0.04 – 20.21	0.9 – 0.002
α [1/cm]	0.007 – 0.5	0.009 – 0.084	0.004 – 0.030	0.005
θ_r [cm³/cm³]	0.02 – 0.08	0.03 – 0.09	0.02 – 0.08	0.07
θ_s [cm³/cm³]	0.33 – 0.61	0.28 – 0.56	0.17 – 0.54	0.3
n	1.6 – 2.1	1.4 – 1.9	1.5 – 2.2	1.5 -2.0

Table 4.3: Range of mean and standard deviation of saturated hydraulic conductivity (Ks), mean inverse air entry suction (α), mean residual water content (θ_r), mean water content at saturation (θ_s), a factor related to pore size distribution (n) and range of spatial correlation for three soil horizons and the bedrock layer after random sampling for ROSETTA (upper) and SCE-UA (lower).

ROSETTA	**Horizon 1**	**Horizon 2**	**Horizon 3**	**Bedrock**
Depth [m]	0.15	0.35	1.0 – 1.5	3.2
Corr. Range ($\log_{10} K_s$) [m]	23.6 – 67.5	16.5 – 47.1	12.4 – 28.05	15.6 – 45.9
K_s [cm/day]	32.36 – 1441.44	3.48 – 206.21	3.34 – 84.81	0.01 – 0.04
α [1/cm]	0.005 – 0.008	0.005 – 0.006	0.005 – 0.011	0.005 – 0.052
θ_r [cm³/cm³]	0.11 – 0.18	0.12 – 0.19	0.11 – 0.19	0.17 – 0.22
θ_s [cm³/cm³]	0.52 – 0.58	0.40 – 0.45	0.33 – 0.42	0.22 – 0.28
n	1.5 – 1.8	1.5 – 1.8	1.5	1.6 – 2.5
SCE-UA Optimization				
Range [m]	16.5 – 29.4	15.7 – 33.3	14.1 – 47.8	14.7 – 31.8
K_s [cm/day]	0.33 – 20.4	0.55 – 16.85	0.12 – 8.35	0.02 – 0.04
α [1/cm]	0.011 – 0.017	0.016 – 0.017	0.019 – 0.02	0.01 – 0.07
θ_r [cm³/cm³]	0.003 – 0.18	0.03 – 0.26	0.03 – 0.28	0.17 – 0.22
θ_s [cm³/cm³]	0.39 – 0.53	0.35 – 0.49	0.26 – 0.44	0.23 – 0.29
n	1.7 – 2.0	1.5 – 1.9	1.6 – 1.9	1.5 – 2.2

Different averaging procedures, applied on the generated random fields of the Mualem-van Genuchten parameters thus result in four different conceptualizations of the subsurface heterogeneity:

1. A setup with spatially homogenous K_S for the complete subsurface (Figure 4.4, Setup A) except for a different, also spatially homogeneous K_S for the bedrock layer. The saturated hydraulic conductivity was calculated by the harmonic mean of K_S for the individual soil layers weighted by the layer thickness. All other soil hydraulic parameters were calculated by the arithmetic mean of parameter values for individual layers, and again weighted by layer thickness. This is in accordance with the guidelines by Zhu and Mohanty (2002), who suggested arithmetic means for n and α highly correlated with K_S.

2. A layered setup (Figure 4.4, Setup B) using spatially homogeneous subsurface parameters $\log_{10} K_S$, $\log_{10} \alpha$, $\log_{10} n$, θ_r , θ_s) for each of the four horizons. In contrast to setup A, K_S is different for each of the four horizons. The horizontal spatial averages were calculated from the arithmetic means (θ_r,θ_s, $\log_{10} \alpha$, $\log_{10} n$) or geometric means ($\log_{10} K_S$). The lateral K_S is increased for the bedrock horizon following setup A.

3. A layered setup with spatially heterogeneous, $\log_{10} K_S$, $\log_{10} \alpha$, $\log_{10} n$, θ_r , θ_s (Figure 4.4, Setup C). These parameters are different for each of the 25 soil units of the Rollesbroich soil map. For each of these 25 texture zones four soil horizons were averaged similarly to setup B assuming homogeneous soil hydraulic properties for each layer.

4. A layered setup with spatially heterogeneous $\log_{10} K_S$, $\log_{10} \alpha$, $\log_{10} n$, θ_r , θ_s (Figure 4.4, Setup D).

To address the uncertain spatial distribution of soil hydraulic parameters, we created an ensemble of 32 stochastic realizations for each of the four setups. Figure 4.4 provides an overview of the different simulation setups.

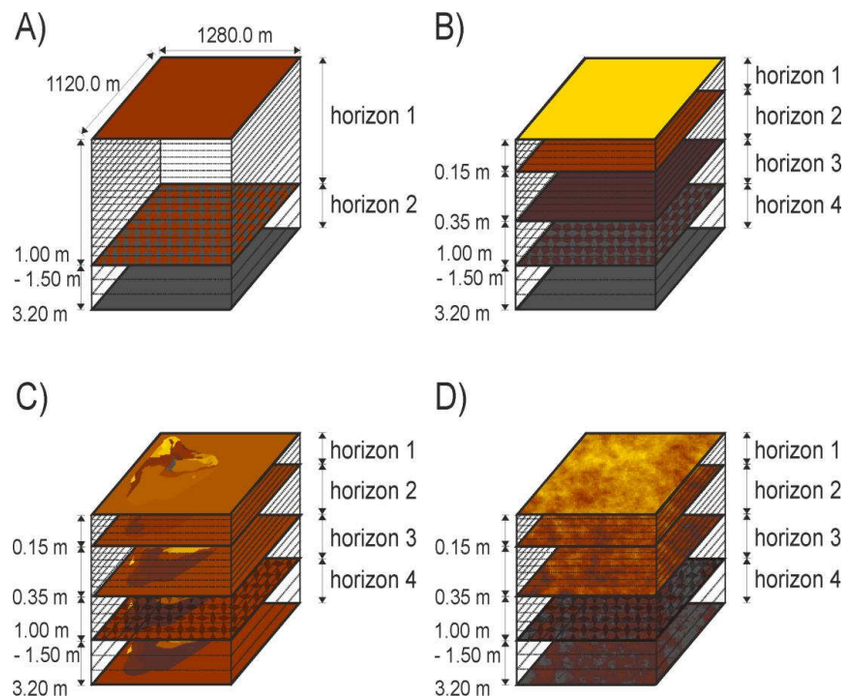

Figure 4.4: Schematic overview of the four investigated parameterizations for subsurface hydraulic parameters: homogeneous (A), layered subsurface (B), soil units and layers (C), and heterogeneous layers (D). The vertical horizon extents are also indicated. The hatchings between the third and fourth horizon indicate the variable bedrock depth at 1.0 – 1.5 m.

Table 4.4: Resolution, soil hydraulic parameterization and land cover of conducted ParFlow-CLM simulations. The appreviations "ROS" and "SCE" indicate soil hydraulic parameters obtained from pedo-transfer functions (ROS) or optimized (SCE) soil hydraulic properties (e.g., scenario B10ros stands for a layered medium (B) and soil hydraulic parameters estimated from texture measurements and pedotransfer function (ROS))

Symbol	Lateral Resolution (dx, dy) [m]	Variable vertical Resolution (dz) [m]	Subsurface Complexity Level	Landuse	Number of simulations
A10ros	10 × 10	0.025 – 0.575	A	grassland	32
B10ros	10 × 10	0.025 – 0.575	B	grassland	32
C10ros	10 × 10	0.025 – 0.575	C	grassland	32
D10ros	10 × 10	0.025 – 0.575	D	grassland	32
A10sce	10 × 10	0.025 – 0.575	A	grassland	32
B10sce	10 × 10	0.025 – 0.575	B	grassland	32
C10sce	10 × 10	0.025 – 0.575	C	grassland	32
D10sce	10 × 10	0.025 – 0.575	D	grassland	32

4.3.2 Model Validation Strategy

For the two-year validation period (2011-2012), the performance of the different setups was compared with measured daily evapotranspiration, daily mean discharge at the catchment outlet as well as SWC at all functioning SoilNet sensor locations at different depths. The model performance was investigated with the Nash-Sutcliffe efficiency index (NSE), root mean squared error (RMSE), percent bias (PBIAS) or absolute bias (BIAS). The (NSE) is given by

$$\mathrm{NSE} = 1 - \left(\sum_{t=1}^{n_{\mathrm{tstep}}} \left(y_t^{\mathrm{sim}} - y_t^{\mathrm{obs}}\right)^2 \Big/ \sum_{t=1}^{n_{\mathrm{tstep}}} \left(y_t^{\mathrm{obs}} - \bar{y}^{\mathrm{obs}}\right)^2 \right) \tag{4.7}$$

where $\bar{y}_t^{\mathrm{sim}}$ simulated ensemble mean values, y_t^{obs} are the observed data, $\bar{y}^{\mathrm{obs}}$ the mean of the observed data, and *t* the time. Possible values for NSE range between -∞ and 1.0. Values ≤ 0.0 indicate unacceptable simulation performance whereas values > 0.5 indicates good performance (Moriasi et al., 2007; Nash and Sutcliffe, 1970). Absolute bias (BIAS) and percentage bias (PBIAS) are given by:

$$\mathrm{BIAS} = \frac{1}{n_{\mathrm{tstep}}} \sum_{t=1}^{n_{\mathrm{tstep}}} \left(y_t^{\mathrm{sim}} - y_t^{\mathrm{obs}}\right) \tag{4.8}$$

$$\mathrm{PBIAS} = 100 \times \left(\sum_{t=1}^{n_{\mathrm{tstep}}} \left(y_t^{\mathrm{sim}} - y_t^{\mathrm{obs}}\right) \Big/ \sum_{t=1}^{n_{\mathrm{tstep}}} y_t^{\mathrm{obs}} \right) \tag{4.9}$$

Positive values indicate an overestimation, whereas negative values imply an underestimation compared to the observed data.

In addition to common goodness of fit indices, wavelet and cross-wavelet methods were deployed to analyze signals and their strength in both the frequency and time domain, which can unravel lagged signal response and correlations (Graf et al., 2014; Rahman et al., 2014). For detailed evaluation of SWC and its spatial variance, we use a cross-wavelet spectrum to indicate temporal variance as function of frequency. This approach considers scale and direction of signal correlation at the same time. For the wavelet transform, we use a Morlet wavelet (Torrence and Compo, 1998) assuming an equally spaced time series for a Gaussian modulated signal. For the detailed mathematical description of wavelet and cross-wavelet analysis we refer at this point to Torrence and Compo (1998), and Si (2008). Cross-wavelet analyses that deal with soil moisture can be found in Si (2008), Graf et al. (2014), and Fang et al. (2015), amongst others.

4.4 Results

4.4.1 Soil Moisture

4.4.1.1 Average Soil Water Content

The temporal evolution of measured and simulated soil water content for the scenarios A, B, C and D is illustrated for SCE-setup ( ) and the ROS-setup (Figure 4.6). The measurements at the site show high SWC during winter months and periods with precipitation (e. g. summer 2012). The SWC seasonality is captured by all simulations. The highest deviation compared to the observed SWC can be found for the A10 scenarios (Table 4.4). The SCE and ROS-setup both underestimate measured SWC at 5 cm, but the mean SWC of A10ros is close to the measurements (Table 4.5). On the contrary, the soil water contents at 20 and 50 cm depth are overestimated by the model simulations of both setup alternatives. A bias of up to 0.15 cm^3 cm^{-3} (A10ros) can be found for individual realizations at 50 cm depth and SWC reaches saturation during wet periods. This shows that homogenous A10 setups without distinct soil horizons are strongly affected by parameter averaging, although the simulation can match serendipitously individual soil layer averages. For example, the stochastic realizations of A10ros for 5 cm depth are an exception with high NSE (0.79; maximum 0.82), low bias (-0.01 cm^3 cm^{-3}), and low RMSE (0.032 cm^3 cm^{-3}), while the soil layers at 20 cm (NSE maximum: -4.0) and 50 cm (NSE maximum: -72.7) are poorly represented. With the introduction of soil horizons into subsurface parameterization (B10), the overall model bias was reduced (0.01 – 0.05 cm^3 cm^{-3}) and led to a performance improvement for the ROS and the SCE-setup and all soil horizons. Simulated SWCs by B10sce fit the observed data already well at 5 and 20 cm (NSE: 0.67 - 0.76). B10ros simulations also were improved but overestimated the average SWC for all three soil layers. In particular B10ros simulations of the uppermost layer at 5 cm were not able to resemble well the dry periods May-June 2011 and August-September 2012. Simulations of the C10 complexity level exhibit in general slight improvements for the ROS and SCE-setup compared to the B10 scenario. For individual layers of the SCE-setup (e.g. 5 cm, 50 cm) even a reduction of simulation performance was found, indicating that the division in soil units added little additional skill to the simulations. The D10-simulations show the best representation of the SWC-dynamics for all three depths. The comparison of the ROS- and SCE-setups reveals that the D10sce simulations better reproduce the dry periods in summer 2011 and Aug. – Sep. 2012, whereas D10ros overestimates SWC during these periods, especially at 50 cm depth. On the other hand, the mean SWC of the wet period in winter is underestimated by D10sce, and D10ros is in better agreement with the measurements. These findings illustrate that the different stochastic simulation runs for both setups show an increasing ability to reproduce seasonal trends, if the complexity of the subsurface parameterization is increased. This is also supported by performance indices listed in Table 4.5 The NSE increases with increasing model complexity, whereas BIAS and RMSE decrease. For the SCE-setup the mean NSE increases from -0.04 (A10) to 0.78 (D10) at

5 cm, 0.31 (A10) to 0.70 (D10) at 20 cm, and -33.0 (A10) to 0.29 (D10) at 50cm depth. For the ROS-setup the mean NSE rises from 0.29 (B10) to 0.62 (D10) at 5 cm, -4.0 (A10) to 0.68 (D10) at 20 cm, and -73.8 (A10) to -4.0 (D10) at 50cm depth.

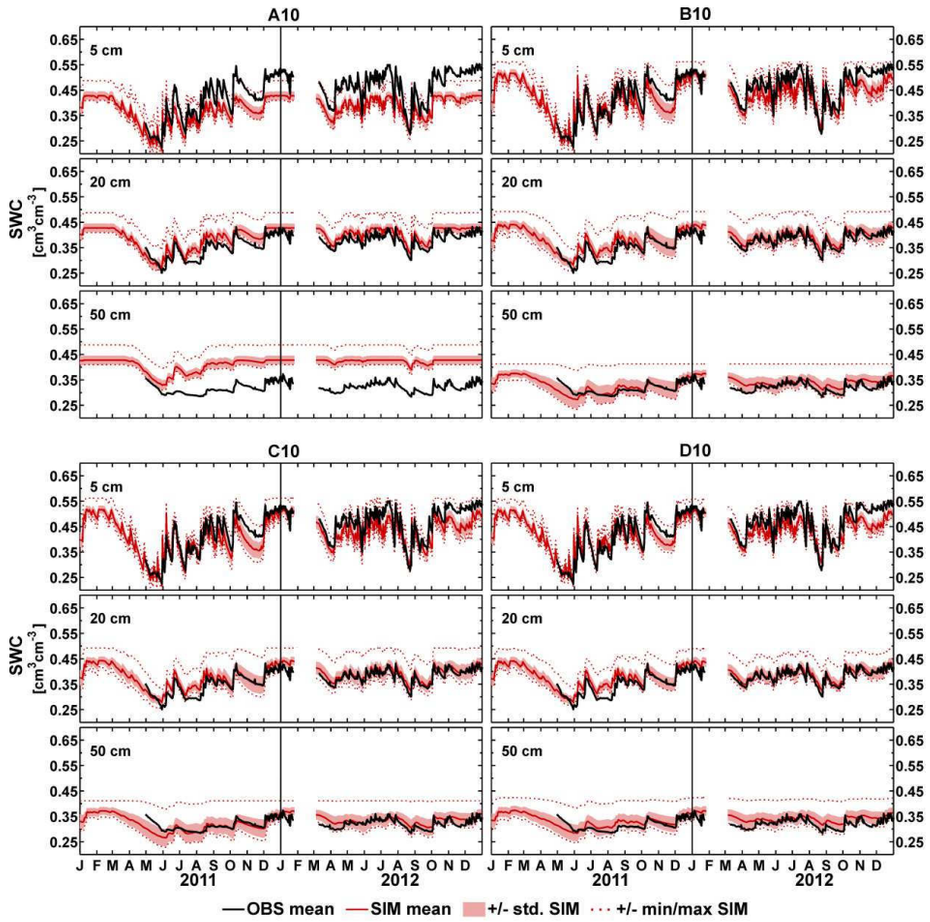

Figure 4.5: Mean soil water content as measured by SoilNet (black) and simulated by Parflow-CLM according to the complexity levels A, B, C, and D of the SCE-setup. From top to bottom, average soil water contents at 5 cm, 20 cm and 50 cm depth for the years 2011 and 2012. Standard deviation, minimum and maximum SWC of the 32 setup realizations are also indicated.

The ensemble spread is higher for the SCE-setup than for the ROS-simulations. Within each setup the spread slightly increases from scenario D10 to B10, whereas A10 shows a smaller spread. The standard deviation of the NSE indicating the variation of performance over the stochastic realizations is in correspondence with these findings, indicating a smaller variation among realizations for the ROS-simulations. To determine the realization with the best overall soil

moisture characterization, we calculated the sum of RMSE at 5, 20, and 50 cm depth. Realization #8 of the SCE-setup and setup D10 showed the best performance with an RMSE of 0.035 $cm^3\ cm^{-3}$ at 5 cm, 0.025 $cm^3\ cm^{-3}$ at 20 cm, and 0.034 $cm^3\ cm^{-3}$ at 50 cm depth. For the D10 ROS-setup realization #13 performed best with an RMSE of 0.04 $cm^3\ cm^{-3}$ at 5 cm, 0.028 $cm^3\ cm^{-3}$ at 20 cm, and 0.04 $cm^3\ cm^{-3}$ at 50 cm depth. This shows that individual realizations of heterogeneous fields for the soil hydraulic parameters are able to characterize average soil moisture well, even though they are not calibrated to measurement data.

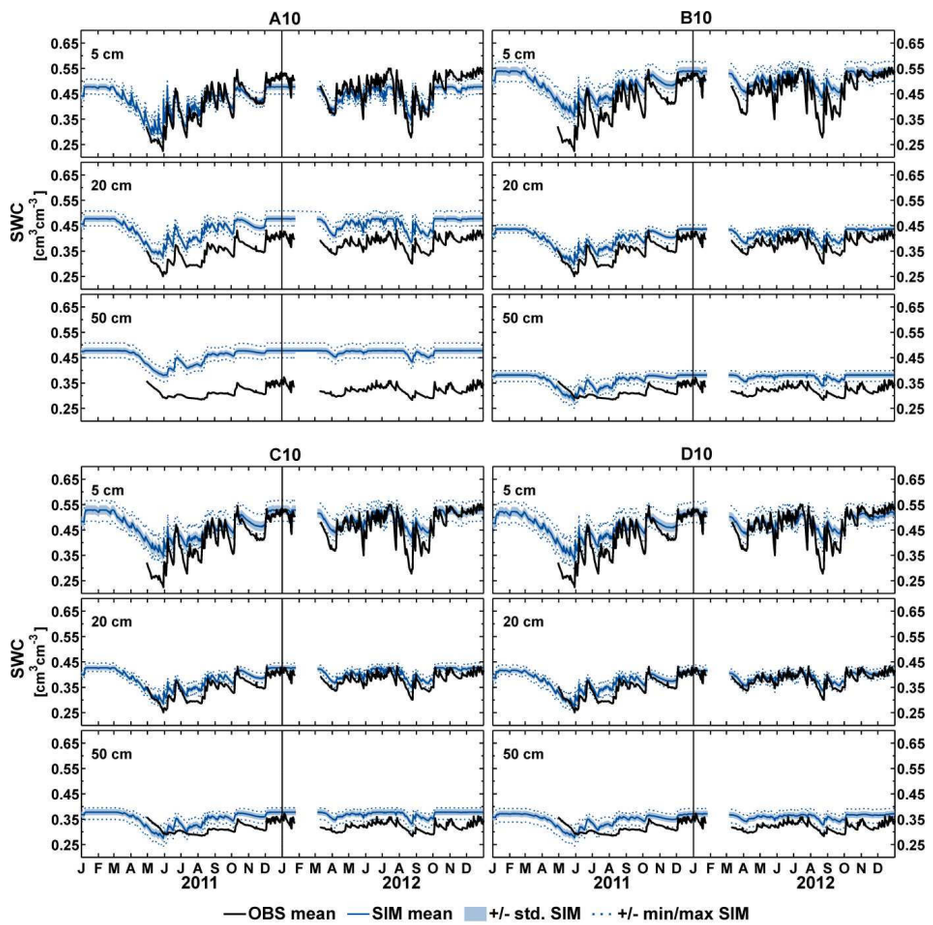

Figure 4.6: Mean soil water content as measured by SoilNet (black) and simulated by Parflow-CLM according to the complexity levels A, B, C, and D of the ROS-setup. From top to bottom, average soil water contents at 5 cm, 20 cm and 50 cm depth for the years 2011 and 2012. Standard deviation, minimum and maximum SWC of the 32 setup realizations are also indicated.

Table 4.5: Model performance regarding SWC indicated by mean, standard deviation, maximum Nash-Sutcliffe Efficiency (NSE), average bias and root mean squared error. The accuracy indices were calculated with modeled and observed spatial mean soil-water-content in 5, 20 and 50 cm depth for different model subsurface parameterizations (A, B, C, and D).

		Scenario	SCE				ROS			
			A10	**B10**	**C10**	**D10**	**A10**	**B10**	**C10**	**D10**
SWC	**5 cm**	**NSE mean [-]**	-0.04	0.76	0.72	0.78	0.79	0.29	0.57	0.62
		NSE std. [-]	0.08	0.06	0.07	0.04	0.01	0.04	0.03	0.03
		NSE max. [-]	0.13	0.80	0.77	0.82	0.82	0.41	0.64	0.68
		BIAS mean [$cm^3\ cm^{-3}$]	-0.07	-0.02	-0.03	-0.02	-0.01	0.05	0.03	0.03
		RMSE mean [$cm^3\ cm^{-3}$]	0.070	0.035	0.037	0.033	0.032	0.058	0.045	0.042
	20 cm	**NSE mean [-]**	0.31	0.67	0.72	0.70	-4.00	-0.38	0.36	0.68
		NSE std. [-]	0.11	0.09	0.07	0.06	0.30	0.09	0.05	0.04
		NSE max. [-]	0.47	0.83	0.84	0.81	-3.85	-0.20	0.45	0.72
		BIAS mean [$cm^3\ cm^{-3}$]	0.03	0.01	0.01	0.01	0.08	0.04	0.03	0.02
		RMSE mean [$cm^3\ cm^{-3}$]	0.030	0.020	0.019	0.019	0.082	0.043	0.029	0.021
	50 cm	**NSE mean [-]**	-33.0	-0.33	-0.08	-0.29	-73.8	-8.5	-6.5	-4.0
		NSE std. [-]	1.52	0.18	0.08	0.35	2.55	0.65	0.69	0.78
		NSE max. [-]	-30.3	-0.04	0.03	0.31	-72.7	-7.68	-5.74	-3.52
		BIAS mean [$cm^3\ cm^{-3}$]	0.1	0.01	0.01	0.01	0.14	0.05	0.04	0.03
		RMSE mean [$cm^3\ cm^{-3}$]	0.098	0.019	0.018	0.018	0.146	0.052	0.046	0.038

Figure 4.7 displays the spatial RMSE-pattern for the D10-setups. The smallest errors can be found at 20 cm depth, in particular for the SCE-setup. The ROS-setup exhibits in general higher RMSEs and some locations with very high RMSE. For 5cm and 20cm depth outliers are found in the northern and western part of the catchment section, with a poor SWC reproduction for both parameterization setups. The local RMSEs do not show a pronounced vertical correlation as sensors at a particular depth are not affected by outliers in under- or overlying layers. Major differences between simulation results from different model setups instead are found mainly at locations with larger differences in soil hydraulic properties between the model setups. For example, sensor #7 in the south-eastern part of the study site exhibits major differences in α (ROS: 0.00 cm^{-1}; SCE: 3.0 cm^{-1}), n (ROS: 1.7; SCE: 2.1), and θ_s (ROS: 0.42 cm cm^{-1}; SCE: 0.26 cm cm^{-1}) whereas K_s (ROS: 24 cm d^{-1}; SCE: 20.2 cm d^{-1}) and θ_r (ROS: 0.06 cm cm^{-1}; SCE: 0.02 cm cm^{-1}) are within the same order of magnitude. This illustrates the high dependence of the local RMSE at individual locations on the parameterization of soil hydraulic properties.

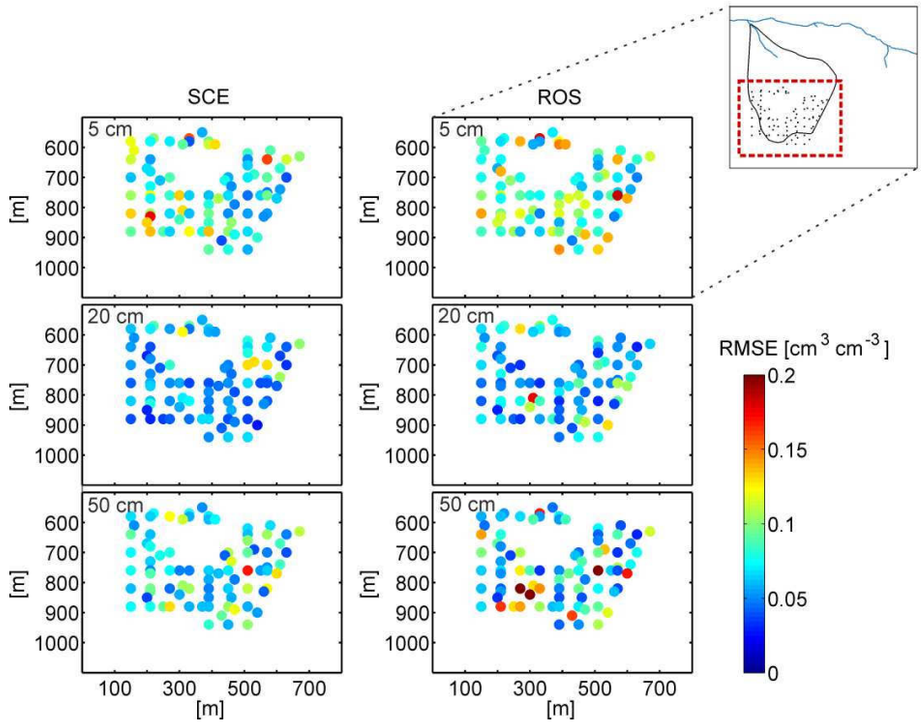

Figure 4.7: RMSE for SWC and the setup D10 (both SCE and ROS) at the individual SoilNet locations at 5, 20, and 50 cm depth. Only results for the southern study site are shown.

4.4.1.2 Soil Moisture Variability

Figure 4.8 illustrates the mean and standard deviation of SWC for the evaluation period (2011-2012) for both the SCE- and ROS-setups. It gives an additional overview of the spatial SWC patterns for the entire model domain. In general, we find higher SWC mean in the ROS model at all three depths. The SCE mean reflects a higher degree of spatial organization related to the subsurface parameterization than the ROS-setup. For both alternatives, SWC is high near the channel, the Kieselbach source including the drainage system area and the center of the test site, whereas low SWC arise in the west and the zone between the center and northeast of the test site. At these locations, also the highest temporal variability within the study site was found. Low standard deviations mainly indicate wet areas of the domain. For these we find relative similar structures in the SCE- and the ROS- setup, representing also the Roßbach and Kieselbach.

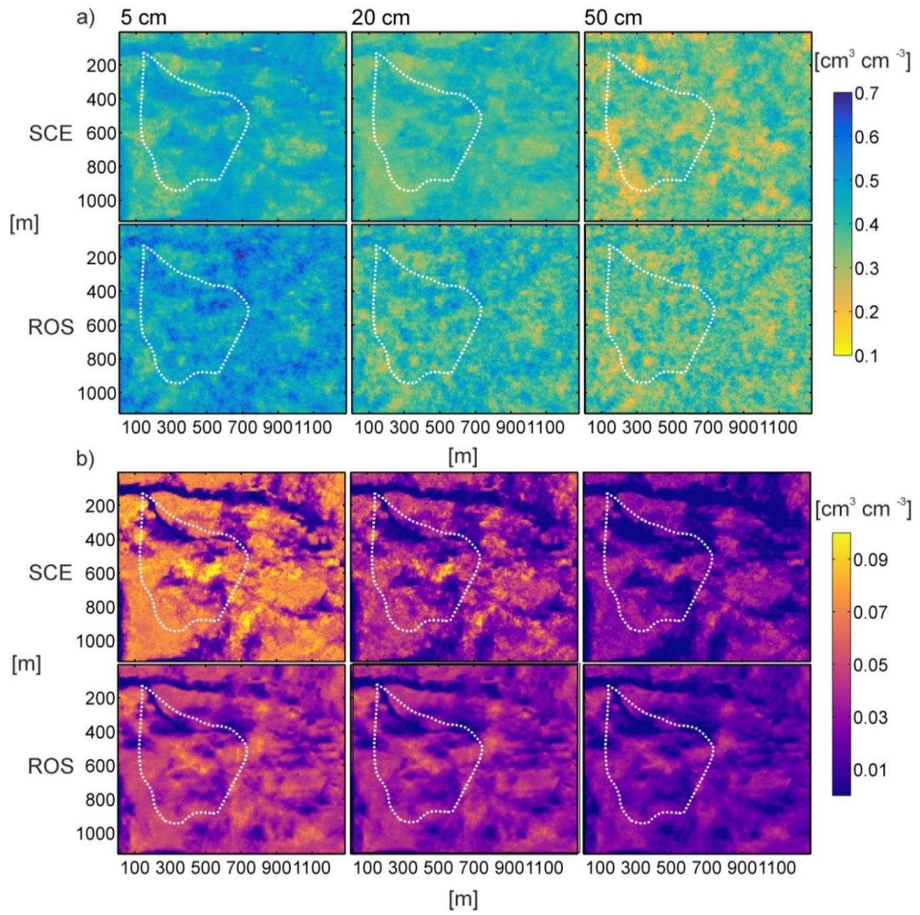

Figure 4.8: Modelled average SWC (a) and standard deviation (b) at 5, 20 and 50 cm depth for the period 2011-2012 and the averages for the scenarios D10sce and D10ros. The test site is indicated by a white dashed line.

Figure 4.9 shows the relationship between spatial mean soil water content ($\bar{\theta}$) and spatial standard deviation (σ_θ) of SWC for measured and simulated data. Simulation results for the different model realizations and the three different depths are shown. For the period 2011 – 2012, we find σ_θ values for measured SWC of 0.06 – 0.08 $cm^3\ cm^{-3}$ for 5 cm depth, 0.04 – 0.06 $cm^3\ cm^{-3}$ at 20 cm, and 0.06 – 0.08 $cm^3\ cm^{-3}$ at 50 cm depth. The comparison of uncorrected and trend corrected SWC shows that the overall range of $\bar{\theta}$ is slightly reduced (~0.05 $cm^3\ cm^{-3}$) similar to σ_θ (~0.005 $cm^3\ cm^{-3}$). Whereas the σ_θ values show no clear trend with depth, the range between maximum and minimum $\bar{\theta}$ decreases from 0.4 $cm^3\ cm^{-3}$ to 0.1 $cm^3\ cm^{-3}$ for the layers from 5 to 50 cm. It is found that the σ_θ of SWC is a function of $\bar{\theta}$, showing a parabolic shape with a general trend of decreasing spatial

variability with increasing SWC, despite a local minimum between 0.3 and 0.4 cm^3 cm^{-3}. The simulations of the A10ros and the A10sce setup clearly underestimate measured spatial SWC variability with σ_θ values of 0.02 – 0.04 cm^3 cm^{-3}. In scenario B10, the ROS- and the SCE-setup exhibit a SWC variability different from each other. Whereas the σ_θ of the SCE-simulations is 0.04 – 0.06 cm^3 cm^{-3}, for the ROS-setup, σ_θ remains at the low level of A10 scenario with only slight improvements. The SWC variability in the C10 scenario further shows only marginal increase when compared to their B10 simulation counterparts of the ROS and the SCE-setup. On the contrary, the heterogeneous D10sce and D10ros are in good correspondence with the parabolic shape of the measured pattern and without an offset, in particular for low and intermediate $\bar{\theta}$. For most scenarios (A10sce –D10sce; A10-C10ros) at 5 cm depth a strong decline of σ_θ values for high $\bar{\theta}$ values can be observed, which is not found in the measurements.

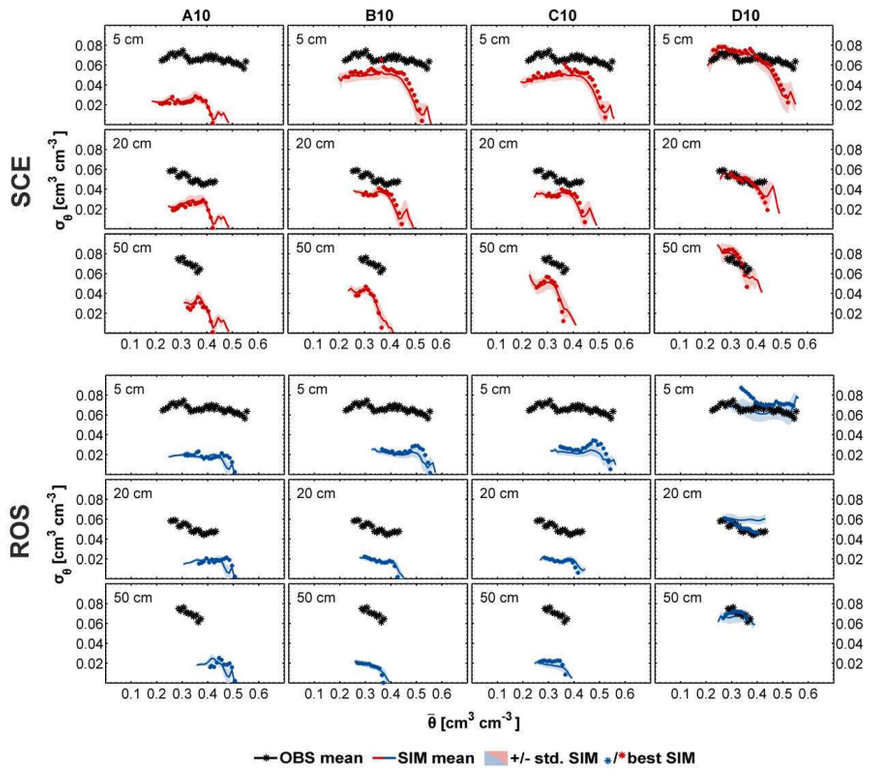

Figure 4.9: Relationship between spatial mean soil water content ($\bar{\theta}$) and standard deviation of soil water content (σ_θ), both for measured data from SoilNet (black), modeled data according the SCE-setup (red), and modeled data following the ROS-setup (blue). Results for three different depths (5 cm, 20 cm, and 50 cm) are shown for the four setups with 10 × 10 m lateral resolution (A10, B10, C10 and D10). The soil water content at the x-axis was classified into intervals of 10^{-2} [cm³ cm-3]. The standard deviations on the y-axis are the values for each depth and SWC interval.

Figure 4.10 provides additional information on the best D10 simulations of the ROS and SCE-setups with focus on differentiation between wetting and drying events. The SCE-simulation #8 for scenario D10 better resembles the range for σ_θ and $\bar{\theta}$ in 5 and 20 cm depth of the observations than the ROS-simulation (D10ros s13) which shows a smaller σ_θ and $\bar{\theta}$ range for wetting and drying. Figure 4.10 further illustrates the differences in the relationship between $\bar{\theta}$ and σ_θ for the ROS- and the SCE-simulation. The decline of σ_θ values for high $\bar{\theta}$ values is either not present (D10ros 5 cm) or less pronounced in the ROS simulation (D10ros 20, 50 cm) than in the SCE-simulation. The relationship of $\bar{\theta}$ and σ_θ is quite similar for wetting and drying events and both for the ROS- and the SCE-setup. An exception is D10ros s13, where we find a definite linear shape for drying events, while wetting events show larger scatter, but do not cover the range of observations.

We further examined the temporal course of the σ_θ simulation average (Figure 4.11) using cross-wavelet analysis. In combination with daily precipitation, we expect more information about the temporal evolution of modelled σ_θ compared to the observations during wetting and drying events. Furthermore, the time-window and direction of the correlation between modelled and observed data can unravel time lagged differences between the modelled σ_θ and observed σ_θ. Although both the SCE and ROS-setup capture the general σ_θ variability level over the entire evaluation period well, we find high seasonal differences in σ_θ between measured and modeled data during both years of the study period. The modelled σ_θ by the ROS-setup shows smaller variations over the year than the observations. On the other hand, the modelled σ_θ by the SCE-setup exhibits high fluctuations in particular through the summer period 2011, associated with strong precipitation events. For these events as well as for the winter period 2011/12 σ_θ of the SCE-setup decreases to 50 % of the observed level. This reduction can also be found in the ROS-setup but is less distinct there. The variability over the individual ROS-simulations is high over the entire evaluation period. On the contrary, the SCE-setup has maximum variability of only ± 0.01 $cm^3\ cm^{-3}$ between the scenario realizations during dry periods (May-June 2011, October-December 2011). For these periods the ROS- and the SCE-setup cover or partly overestimate the σ_θ observations. The SCE-setup thereby shows high but time-lagged or inverse correlation with the observed σ_θ ($R^2 > 0.8$). These time-lagged or inverse correlations with observed σ_θ occur over longer time intervals (up to 32 days) than between the ROS-setup and the observations. During these dry periods, the simulated σ_θ signal turns upward and recovers from the σ_θ decrease which was simulated in relation to intensive precipitation events. However, in 2012 the correlation for the SCE-based σ_θ is almost synchronous with the observations in wet periods. This indicates a high temporal coherence between measurement data and the SCE-setup simulations. From July to August 2012, σ_θ of observation and SCE-setup simulations exhibit a similar decrease of σ_θ. This is in contradiction to the ROS-setup exhibiting short periods with inverse or shifted correlations indicating less signal coherence of σ_θ.

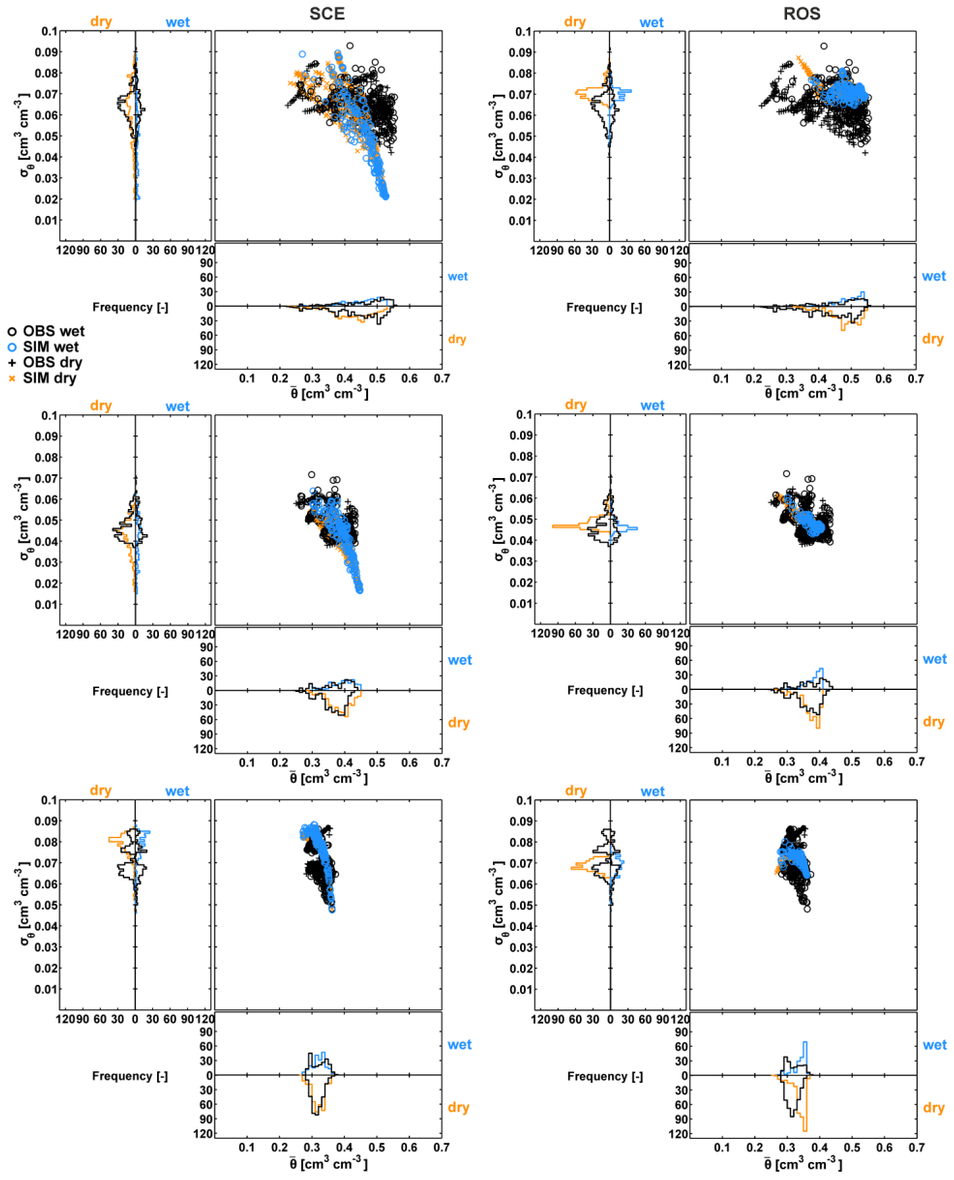

Figure 4.10: Standard deviation of soil water content (σ_θ) as function of mean soil water content ($\bar{\theta}$) for measured SoilNet data, modeled data from the best D10sce simulation s8(left) and best D10ros simulation s13(right). The histograms of SWC and SWC standard deviation also indicate the frequencies of observed and modelled data split into wetting and drying events.

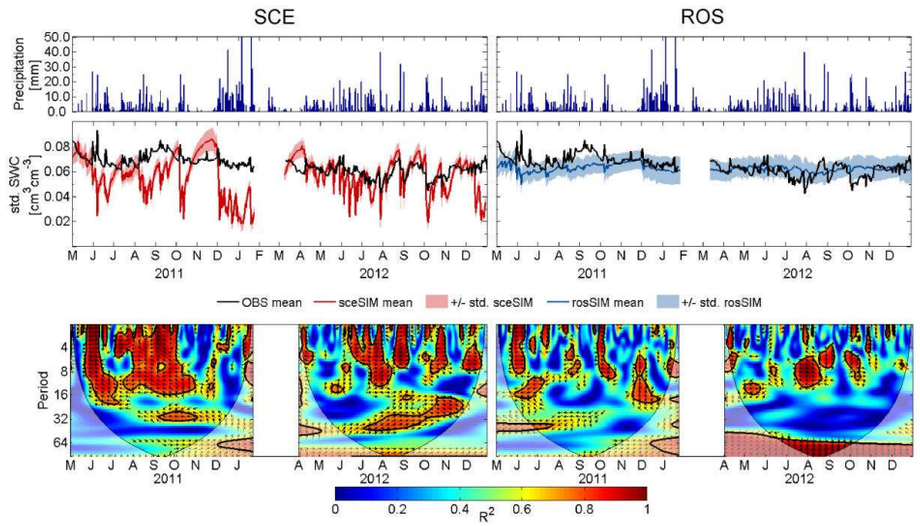

Figure 4.11: Observed and simulated σ_θ with uncertainty for the SCE and ROS setups (middle row of figures) and scenario D10 at 5 cm depth using wavelet coherence analysis (lowest row of figures). The local R2 in domain and frequency is indicated by color codes and significant correlations at the 95 % level between simulation and observations. The arrows indicate the phase angle for R²> 0.5. Arrow on the right connotes synchronous correlation, whereas an arrow on the left point out anti correlation. Downward or upward arrows indicate a delayed signal correlation. Precipitation (P) is also indicated.

4.4.2 Evapotranspiration Dynamics

The simulated evapotranspiration is in good correspondence with the measured actual ET (ET_a). Both the SCE and ROS setups have high NSE, small bias and RMSE (Table 4.6) in comparison with the observed lysimeter (NSE: 0.76 - 0.77; bias: -0.12 - -0.07 mm d^{-1}; RMSE: 0.64 - 0.66 mm d^{-1}) and EC data (NSE: 0.64 - 0.66; bias: -0.16 - -0.11 mm d^{-1}; RMSE: 0.75 - 0.76 mm d^{-1}). The different simulation scenarios for each of the setups only show very small NSE differences (~10^{-3}). The performance indices do not show a clear trend as function of the complexity of the subsurface parameterization. The simulations of the ROS-setup exhibit a slightly better performance than their SCE counterparts. The simulations of all scenarios slightly underestimate the monthly ET of lysimeter and EC data, however, in June 2012 simulations overestimate ET compared to lysimeter (0.15 mm d^{-1}) and EC data (0.19 mm d^{-1}). Figure 4.12 reveals that daily ET_a EC (and also ET_a LYS; not shown) show more variability in ET between different days in the summer than simulated ET. While the simulations underestimate measured peak ET_a on summer days with (intensive) precipitation events, they overestimate ET observations on dry days. In contrast, the ET-dynamics for the different ROS- and SCE-scenarios hardly differ from each other. Daily ET of ROS-simulations is slightly higher than for SCE comparing scenarios with the same complexity. The

highest ET_a difference (~ 1mm d^{-1}) between ROS and SCE is found for the dry period April – June 2011 for the scenarios B, C, and D. The difference in actual ET between the A10 SCE- and ROS-scenarios and the layered SCE-scenarios (B10 - D10) can be explained by lower SWC in the latter, which result in somewhat reduced transpiration and soil evaporation.

Table 4.6: Accuracy indices calculated with modeled and observed mean catchment evapotranspiration for subsurface parameterizations A10, B10, C10, and D10. Mean, standard deviation NSE, average bias and root mean squared error are shown for lysimeter and EC evapotranspiration for 2012.

Scenario			SCE				ROS			
			A10	**B10**	**C10**	**D10**	**A10**	**B10**	**C10**	**D10**
ET	**NSE mean**	**LYS**	0.76	0.76	0.76	0.76	0.77	0.77	0.77	0.77
	[-]	**EC**	0.64	0.64	0.64	0.64	0.65	0.65	0.65	0.65
	NSE std	**LYS**	0.003	0.003	0.003	0.004	0.001	0.002	0.002	0.002
	[-]	**EC**	0.003	0.003	0.003	0.003	0.001	0.001	0.001	0.001
	BIAS mean	**LYS**	-0.09	-0.11	-0.11	-0.12	-0.078	-0.069	-0.072	-0.076
	[mm d^{-1}]	**EC**	-0.13	-0.15	-0.15	-0.16	-0.12	-0.11	-0.11	-0.12
	RMSE mean	**LYS**	0.65	0.66	0.66	0.66	0.64	0.64	0.64	0.64
	[mm d^{-1}]	**EC**	0.75	0.76	0.76	0.76	0.75	0.75	0.75	0.75

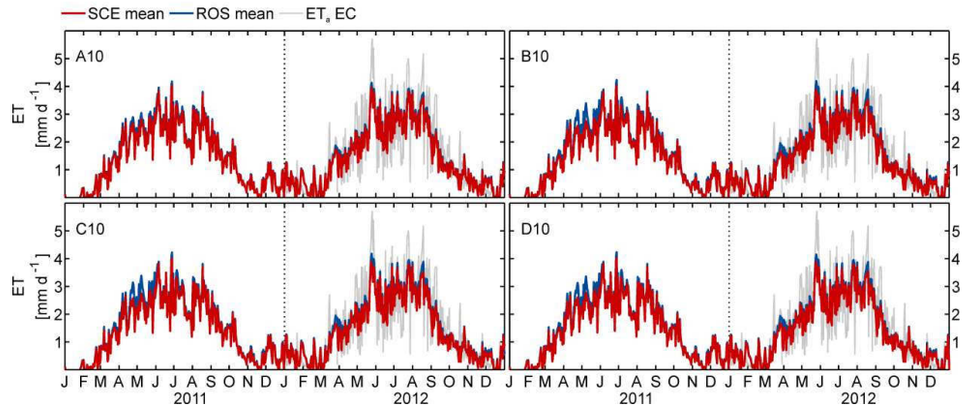

Figure 4.12: Simulated mean daily ET for different complexity levels (A10-D10) of the SCE- and ROS-setup. Observed ETa EC is also indicated for 2012.

4.4.3 Discharge Dynamics

The discharge dynamics of 2011 and 2012 are illustrated in Figure 4.13. The discharge measurements at the catchment outlet show a period with high flow during winter rain and snow melt periods (Dec – Apr) and low flow conditions from May – November. Exceptions were June/July 2012 and end of August 2012 due to high precipitation and thunderstorms leading to flooding.

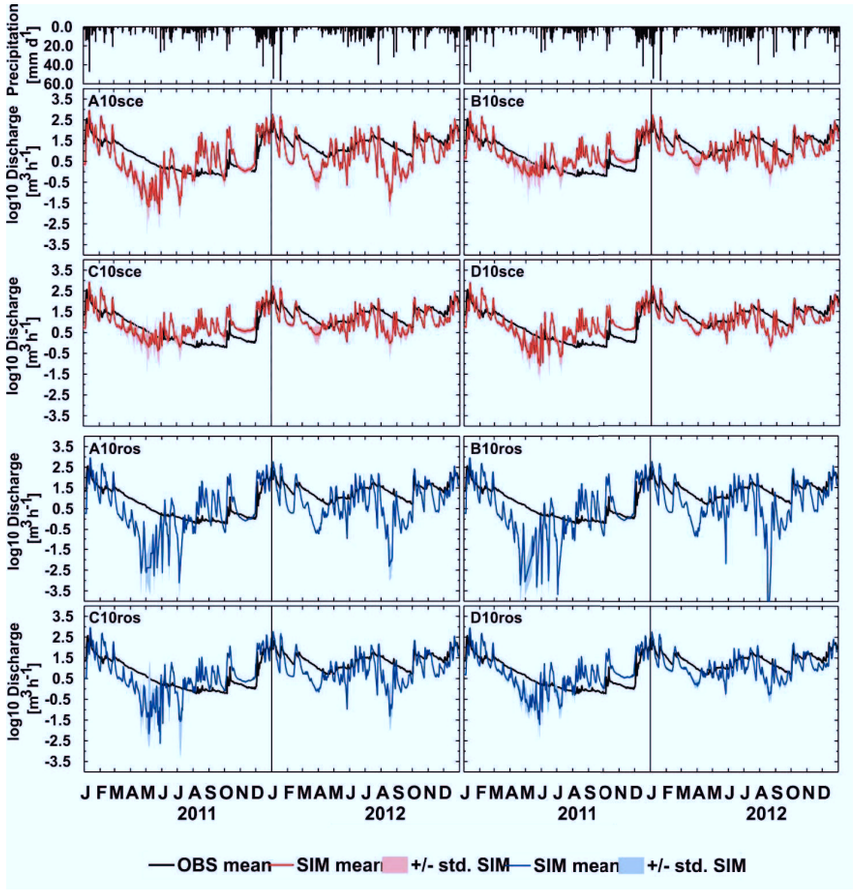

Figure 4.13: Observed and simulated daily mean logarithmic discharge at the catchment outlet for 2011-2012. Shown are alternative subsurface setups (SCE/ROS) with scenarios of different complexity (A10-D10). Precipitation for this period is indicated as well.

Mean Nash-Sutcliffe-Efficiencies (Table 4.7) ranging from -1.71 - -0.93, and high model bias (30 – 14 %), indicate a poor model performance for discharge simulation for both model setups (ROS/SCE) and for all scenarios (A10 – D10). With increasing complexity we found a slight improvement of these performance indices for both the ROS- and SCE- setup (e. g. RMSE A10sce: 8.01 $m^3 h^{-1}$; RMSE D10sce: 3.69 $m^3 h^{-1}$). The model simulations tend to overestimate discharge peaks particularly during winter conditions and thunderstorms. By removing these events (ca. 34 days), an increasing NSE (0.34 – 0.41) indicates a better performance for the remaining data. However, during low flow conditions in dry periods (e. g. Aug. –Sep. 2011 and 2012) the simulations overestimate the measured discharge. The tendency of the Kieselbach running dry during summer seasons (which is not observed in the data) is found for scenarios with less complexity. The highest variability among the individual realizations for the different scenarios is found in summer 2011, for both the SCE and ROS setups.

Table 4.7: Discharge model performance at the catchment outlet for subsurface parameterizations A10, B10, C10, and D10 indicated by mean, standard deviation, maximum Nash-Sutcliffe Efficiency (NSE), average bias and root mean squared error.

Scenario		SCE				ROS			
		A10	**B10**	**C10**	**D10**	**A10**	**B10**	**C10**	**D10**
Q	**NSE mean [-]**	-1.70	-1.05	-1.08	-0.93	-1.74	-1.71	-1.59	-1.30
	NSE std. [-]	0.04	0.29	0.26	0.22	0.02	0.03	0.05	0.13
	NSE max. [-]	-1.64	-0.52	-0.55	-0.37	-1.71	-1.68	-1.51	-1.03
	BIAS mean %	30.5	15.7	14.6	14.0	29.6	28.0	24.1	22.3
	RMSE mean [$m^3 h^{-1}$]	8.01	4.12	3.89	3.69	7.77	7.35	6.32	5.83

4.4.4 Water Balance

Table 4.8 provides an overview of different water balance components including discharge, evapotranspiration and precipitation for the two evaluated years. Both, the ROS- and the SCE-setup simulations, show a small inter-annual variability as well as small differences between the individual complexity levels. Compared to the observations of ET-EC and ET-LYS in 2012 the simulations underestimate measured ET by up to 55 mm (D10sce). The annual measured sum of discharge is 429.3 mm in 2011 and 797.6 mm in 2012. Precipitation was also larger in 2012 (1251 mm) than in 2011 (1084 mm). Whereas the model simulations on average strongly overestimate the discharge for 2011 (up to ca. 300 mm difference), the discharge is underestimated by most simulations (ca. 160 mm difference) in 2012. The residuals of the observed water balance in 2012 are -7.7 % for the EC data and -9.3 % for the lysimeter data. For the simulations, we found water balance residuals up to -16 % (A10ros 2011) and +8 % (C10sce 2012). Another important difference between reality and simulations is probably related to snow cover, which leads to

delayed infiltration and discharge affecting the subsequent annual water balance. We estimated a potential storage error up to 8 % related to the unknown exact partitioning of mixed precipitation in liquid or solid. The model also does not account for frozen soils and impaired runoff related to frozen soils which might lead to further balance deviations. Thus, the water balance residuals are within an acceptable range compared to the observed residuals. The water balance for the ROS-setup has smallest residuals for all complexity stages in 2012 (0.5 – 3.2 %) and higher residuals in 2011 (-15.9 - -5.0 %), which is related to the higher discharge and evapotranspiration. In contrast, the SCE-setup exhibits intermediate negative and positive residuals in both years (-15.5 – 0.7). With decreasing subsurface heterogeneity, we find a decreasing positive residual for 2012, while the negative residual for the SCE and ROS-setup increased for 2011. These residuals are mostly related to the quality of the simulated discharge.

Table 4.8: Observed and simulated yearly water balance components precipitation (P), evapotranspiration (ET) and surface runoff (Q) of the Rollesbroich catchment for different subsurface scenarios.

OBSERVATION

Period	P [mm]	ET [mm]	Q [mm]	P-ET-Q [mm]	P-ET-Q %
2011	1,083.6	n/a	429.3	n/a	n/a
2012	1,251.7	566 (LYS) / 552 (EC)	797.6	-111.9 (LYS) / -97.9 (EC)	-8.9 (LYS) / -7.8 (EC)

TerrSysMP-SIMULATION

Scenario	Period	P [mm]	ET [mm]	Q [mm]	P-ET-Q [mm]	P-ET-Q %
A10ros	2011	1,083.6	523.3	732.2	-171.9	-15.9
	2012	1,251.7	530.2	714.8	6.7	0.5
A10sce	2011	1,083.6	512.5	738.6	-167.5	-15.5
	2012	1,251.7	524.5	718.4	8.8	0.7
B10ros	2011	1,083.6	531.4	720.4	-168.2	-15.5
	2012	1,251.7	533.4	708.6	9.7	0.8
B10sce	2011	1,083.6	505.5	648.5	-70.4	-6.5
	2012	1,251.7	517.5	643.2	91.0	7.3
C10ros	2011	1,083.6	529.8	695.9	-142.1	-13.1
	2012	1,251.7	532.5	687.6	31.6	2.5
C10sce	2011	1,083.6	504.1	643.8	-64.2	-5.9
	2012	1,251.7	516.9	635.6	99.1	7.9
D10ros	2011	1,083.6	529.1	679.5	-124.9	-11.5
	2012	1,251.7	531.0	680.3	40.4	3.2
D10sce	2011	1,083.6	505.3	632.8	-54.5	-5.0
	2012	1,251.7	516.2	634.7	100.8	8.1

4.5 Discussion

Our findings suggest that spatial heterogeneous soil hydraulic parameters and subsurface complexity dominate the observed spatial variability of SWC at the small headwater catchment scale. The spatial SWC variability for the individual soil horizons increases as function of heterogeneity of soil hydraulic parameters for both setups (ROS/SCE). For the complex scenario D10, with fully heterogeneous fields of the soil hydraulic parameters, the SWC variability for the ROS and the SCE-setups are similar and comparable to the measurements. On the contrary, for the A10 scenario both setup alternatives (SCE/ROS) exhibit on average a similar low σ_θ (0.02 cm^3 cm^{-3}) which is much lower than σ_θ for the measurements. Given the model setup for this catchment with spatially uniform vegetation and atmospheric forcings, the SWC variability for homogeneous simulations can be seen as base variability induced by topography. This is 25 – 40 % of the σ_θ detected by the observations (i.e., 0.06–0.08 cm^3 cm^{-3}). The introduction of a layered structure (B10) leads to a significant increase of SWC variability, whereas the introduced additional zonal heterogeneity of C10 shows only a minor impact. In fact fully spatially heterogeneous fields are needed to explain the observed SWC spatial variability.

Selected simulations with a satisfying reproduction of SWC at 5, 20 and 50 cm depth give further insight in the limitations and uncertainty of the model simulations. Whereas the D10 SCE-simulation is able to reproduce the relationship of mean SWC and σ_θ for drying events, this is less the case for wetting events. During some precipitation events (e.g. thunderstorms during the dry summer period 2011) the decreasing σ_θ for both modelling setups contrasts with an observed increase in σ_θ. These results indicate that the ability of the simulations to reproduce the observed σ_θ is related to the antecedent soil moisture state. For the nearby Wüstebach site, Rosenbaum et al. (2012) and Wiekenkamp et al. (2016) reported that infiltration excess decreases observed σ_θ during wet periods. In contrast, an increase of σ_θ during intensive rain events and dry conditions was attributed to preferential flow. This was also observed in other field studies by Martini et al. (2015) and Poltoradnev et al. (2016). Our simulation results of both setups show that during these situations, infiltration excess leads to a decreasing σ_θ for the simulations, which do not take fast bypass water flow within macrospores into account. However, the SCE-setup partly reproduces temporal dynamics of σ_θ in particular for periods with average soil moisture conditions in 2012. The ROS-simulations, on the contrary, poorly represent the temporal evolution of σ_θ, although having an overall σ_θ level similar to the observations. These findings indicate that the SWC-variability is dominated by spatially variable soil hydraulic parameters, although impacts due to spatial variable atmospheric forcing (i.e. precipitation) or vegetation variations in the uniform grass cover cannot be excluded completely.

Furthermore, these conclusions are based on the assumption that after trend correction the remaining observation error is related to small random noise, which does not affect the overall level

of spatial variation. Systematic bias of individual sensors or strong white noise lead to a potential overestimation of spatial SWC variability and therefore to overestimating the subsurface heterogeneity in the simulations. Although an error in absolute SWC measurements cannot be completely excluded, it seems unlikely given the number of redundant sensors and trend correction for individual sensors.

The results further indicate that 1D inversely estimated soil hydraulic properties of the SCE-setup result in a better characterization of SWC at catchment scale than their counterparts estimated from soil texture data via pedo-transfer function (ROS). The range of RMSE for SCE-setup simulations at the individual locations (5 cm: 0.033 $cm^3\ cm^{-3}$; 20 cm: 0.019 $cm^3\ cm^{-3}$; 50 cm: 0.018 $cm^3\ cm^{-3}$) is similar or smaller than for the ROS-setup (5 cm: 0.042 $cm^3\ cm^{-3}$; 20 cm: 0.021 $cm^3\ cm^{-3}$; 50 cm: 0.038 $cm^3\ cm^{-3}$) suggesting more accurate results for the SCE-setup. The good agreement between observed and simulated spatio-temporal SWC (scenario D10, in particular for the SCE-setup) is in correspondence with the findings of Qu et al. (2014). This implies that lateral water movement plays a minor role for soil moisture characterization of the upper soil horizons (5, 20, 50 cm) as soil hydraulic parameters obtained from 1D inverse calibrations also performed relatively well as hydraulic parameters for the 3D distributed model. The layered SCE-setup model scenarios were able to simulate the relatively high spatial SWC variability at 50 cm compared to the 5 and 20 cm observations. According to Qu et al. (2014), this pattern can be attributed to the pedological situation of the study site.

While lateral flow is not important for the upper layers, it is for the shallow groundwater present at the site. The sapprolite layer drains the water fast from the overlaying soil layers and routes it laterally to the catchment outlet. The strong lateral flow component is also important in areas where the drainage system is installed. For the modelling of the Rollesbroich site, the lateral flow therefore becomes more important for the bedrock horizon. This is different from the nearby Wüstebach site, a forested headwater catchment of similar size and pedological characteristics (Rosenbaum et al., 2012). For the Wüstebach site, Cornelissen et al. (2014) and Fang et al. (2015) found only a minor importance of the bedrock horizon and its model parameterization. Cornelissen et al. (2014) reported a slight improvement for the modelling at the Wüstebach site (lower peak discharges more in correspondence with the data) if available bedrock information was considered.

The ROS and the SCE-setup result in total yearly ET close to observed ET of lysimeter and EC-station, which is also close to potential ET. ROS gives a consistently higher ET than the SCE scenarios (15 - 25 mm/year) due to higher SWCs of the ROS-setup compared to the SCE SWCs. This leads in general to higher ET closer to the measured ET. This can be attributed to the smaller values for the soil hydraulic parameter α (air entry pressure) and in spite of higher values for the saturated hydraulic conductivity. A number of factors can potentially explain the difference to observed ET: errors in the meteorological forcing data (especially incoming shortwave radiation

and wind speed), vegetation parameters or model structural errors. For example, an underestimation of the LAI or rooting depth potentially leads to underestimation of transpiration and interception loss. Although we adjusted rooting depth for the site specific grass species, the C3 grass parameterization in CLM relies on global average characteristics. It is also found that land surface models differ greatly in the semi-empirical relationships of ET-reduction as function of soil saturation, which points to poor constraints on the ET-process and the possibility of a significant model structural error. Finally, processes not accounted for in the simulations, such as grass harvest and maintenance, may result in deviations between measurement data and model. Grass harvest was already found to explain differences between ET_a of the lysimeters and the eddy covariance EC method (Gebler et al., 2015). It should be noted that maintenance most likely affected SWC variability as the investigated area is characterized by uniform grass cover.

For both subsurface model setups (SCE/ROS) and all complexity levels, runoff could not be reproduced well, although the SCE-setup showed a marginal but insignificant better performance compared to the ROS-setup. With increasing complexity of the subsurface parameterization the performance of the ROS and the SCE-setups slightly improved in our simulations. This could be explained with spatial averaging in particular for the homogeneous scenario, which led for this scenario to reduced infiltration rates for the uppermost layers. This results in increased runoff excess and higher runoff peaks for moderate and intensive precipitation events. On the other hand, for the complex heterogeneous scenario D10, a downhill re-infiltration after infiltration excess becomes more likely. Uncertainties with respect to the infiltration properties can therefore be better compensated with increasing complexity or variability of soil hydraulic parameters (Herbst et al., 2006). The comparison of the simulated ET time series with observations also indicates that runoff is affected by short-term ET fluctuations. During summer days with precipitation events, ET_a LYS and ET_a EC are up to 20 % higher than the simulated ET for both setups which also accounts for intercepted water. This partly explains simulated peak runoff in summer as an extra amount of water potentially contributes to direct runoff. Generally, the poor model performance of this physically based hydrological model for reproducing discharge can be attributed to model parameter uncertainty, structural model issues or insufficient or incorrect other input information (Binley and Beven, 1992; Vázquez et al., 2008). For example, preferential flow due to macropores also strongly affects the infiltration process (Weiler and Naef, 2003) and is potentially underrepresented for this grassland site with intensively rooted topsoil layers. Furthermore, the contradiction between the relatively good reproduction of SWC for the upper soil layers (5, 20, 50 cm) and poor discharge simulations suggests a potential issue with the lower subsurface layers. Interflow and groundwater flow at the study site are affected by bedrock characteristics and the installed drainage system, whose parameterization in the model is affected by substantial uncertainty due to a lack of information. This also affects base flow as fractured bedrock might act as additional storage (Hale et al., 2016). Another aspect that can contribute to the poor discharge

simulation is the discretization of the channel bed. The 2d sheet flow in ParFlow is directly linked to the pressure head of the individual grid cell. Given the size of the small Kieselbach channel (0.3-1.0 m), discharge can be affected by uncertainty in topography although calculated from 1×1 m DEM. This includes also the interaction between surface water and groundwater. An underestimation of the water level in a too broad channel leads to a reduction of the vertical pressure gradient. Subsequently, less water re-infiltrates into the subsurface from the channel. It is also important to remind that discharge calibration (e. g. adjustment of Manning's roughness coefficient) was not an objective in this study. Accordingly, the full parameter space relevant for discharge parameter calibration was not explored, and much better results could be achieved by parameter calibration.

4.6 Conclusions

This study investigated the effect of different soil hydraulic parameterization schemes of the ParFlow-CLM component of the TerrSysMP framework for modelling a grassland headwater catchment in the Eifel (Germany). Therefore model runs at 10 × 10 m lateral resolution, a variable vertical resolution (0.025-0.575 m), and different complexity levels regarding the parameterization of subsurface hydraulic parameters were conducted. For each model complexity level (homogeneous, homogeneous layers, homogeneous layers differing between different soil units, and fully heterogeneous subsurface generated by stochastic simulations), we performed calculations with 32 stochastic realizations. These stochastic realisations had different values for saturated hydraulic conductivity and the Mualem-vanGenuchten parameters. The simulations with the four different complexity levels of subsurface soil hydraulic parameterization were conducted with two different setups. One setup was on the basis of measured soil hydraulic properties (ROS) and the other setup on the basis of inversely estimated soil hydraulic properties from 1D simulations (SCE). We subsequently investigated the model performance to reproduce soil water content, evapotranspiration, discharge and the overall water balance with common performance indices (Nash-Sutcliffe efficiency, RMSE, and model bias) as well as cross-wavelet analysis. In particular, we focused on the spatial and temporal soil moisture patterns and spatial soil water content variability in the southern part of the study site.

The results indicate that heterogeneous soil hydraulic parameters together with topography can explain the overall spatial variability of SWC. Given the fact that homogeneous soil hydraulic parameters alone only explained a small portion of the spatial variability of SWC, it can be argued that a heterogeneous distribution of soil hydraulic parameters is more important to explain the spatial variability of SWC than topography. All model simulations are able to reproduce the temporal seasonal variability of ET. However, short-term ET dynamics were underestimated compared to the observations affecting the poor discharge reproduction in summer. In contrast, discharge could not be reproduced well, showing particularly high model bias during winter high flow events. This can probably be related to the lack of information about bedrock characteristics and its interaction with the on-site drainage system. In addition, discharge is highly sensitive to soil hydraulic parameters and preferential flow and therefore other parameter configurations could result in a better performance than the calculated uncalibrated model stochastic realizations. The analysis nevertheless shows the enormous challenge to obtain good discharge predictions with a physically based hydrological model.

The SCE-setup outperforms in general the ROS-setup. The higher inverse air entry parameter used for the SCE-simulations induced better SWC dynamics during dry periods illustrating the difficulty to transfer small scale measurements of soil hydraulic parameters to the larger scale and that 1D inversely estimated parameters result in a better performance. It also illustrates that interpolated 1D

inverse parameter estimates result in an acceptable performance for the catchment, in spite of the fact that lateral flow processes were not considered in the 1D calibrations.

Our study shows that process based, fully integrated models applied at the small catchment scale, in combination with various types of measurement data available at a high spatio-temporal resolution can potentially be a valuable tool not only for unraveling structural model errors, but also for improving the monitoring network. However, for physically based integrated modelling a better fusion with measurement data is needed, which can for example be achieved with data assimilation. This would give further insights in structural model deficits and increase data value.

Chapter 5 Assimilation of high resolution soil moisture data into an integrated terrestrial model for a small-scale head-water catchment

*submitted to Water Resources Research

5.1 Introduction

Modelling the soil water content is of high interest for various geoscientific research fields. The soil water content (SWC) influences the water and energy cycles at the local, regional, and global scale. It controls for example the partitioning of net radiation into latent, sensible and ground heat flux as well as the partitioning of precipitation into infiltration and runoff (Grayson et al., 1997; Robinson et al., 2008). Thus, a precise characterization and prediction of the soil water content patterns is essential for understanding and quantifying the water and energy cycles for applications like weather prediction, flood prediction and real time irrigation scheduling.

On this account, recent modelling studies apply integrated terrestrial model platforms (e.g. AquiferFlow-SiB2 (Tian et al., 2012), CATHY (Bixio et al., 2002; Camporese et al., 2010), MikeShe (Abbott et al., 1986; Graham and Butts, 2005), ParFlow-CLM (Maxwell and Miller, 2005; Kollet and Maxwell, 2006), ParFlow-WRF (Maxwell et al., 2011), PIHM (Qu and Duffy, 2007; Kumar et al., 2009), TerrSysMP (Shrestha et al., 2014)) for modelling two way feedbacks between different terrestrial compartments. Coupled land surface-subsurface models are well suited for in-depth investigation of SWC spatial variability in combination with in-situ SWC data at the hillslope scale. Examples are studies on the controls of SWC variability (e.g. Herbst et al., 2006; Ivanov et al., 2010; Cornelissen et al., 2014; Fatichi et al., 2015) as well as the relation between SWC distribution and the rainfall runoff response (e.g. Herbst and Diekkrüger, 2003; Sciuto and Diekkrüger, 2010). These studies demonstrated that fully coupled land surface-subsurface models can be a valuable tool for a better understanding of the small scale processes at the hillslope scale. However, due to the high compute needs of these models, studies often were conducted with uncalibrated parameters and without model uncertainty estimation. The uncertainty of modelled SWC is related to erroneous model forcings (e.g. precipitation), model parameters (e.g. soil hydraulic properties), model structural deficits, and initial conditions.

With the help of data assimilation (DA) techniques, model predictions can be improved by merging (uncertain) observation data with uncertain model predictions (Burgers et al., 1998; Vrugt et al., 2005). Ensemble based sequential DA methods rely on a probabilistic framework where an ensemble of model realizations is propagated forward and updated each time observation data are available. One of the most commonly applied algorithms is the ensemble Kalman Filter (Evensen, 1994; Burgers et al., 1998). EnKF is also used for updating model parameters, for example by an augmented state vector approach (Chen and Zhang, 2006; Hendricks Franssen and Kinzelbach, 2008). This is especially important for subsurface applications, since parameter uncertainty is an important source of uncertainty for subsurface terrestrial system models. The computational costs for EnKF are affordable as the needed ensemble size is in general not too large (less than 1000 ensemble members) and parallelization to speed up the simulations is trivial (Pauwels and De Lannoy, 2009; Kurtz et al., 2016). This is one of the reasons why the EnKF has been used in combination with different types of terrestrial system models (e.g. atmospheric models, hydrological models, land surface models) and various kinds of observation data.

Only in relatively few studies at th hillslope scale DA was combined with integrated terrestrial system models. Nevertheless, some authors were able to show the high potential of EnKF in combination with integrated hydrological models. For example, Camporese et al. (2009) found that streamflow predictions improved by assimilating different combinations of pressure head and streamflow data. They did the DA-experiments for a synthetic tilted v-catchment with a fully coupled surface water-groundwater flow (CATHY) model including 3D subsurface flow and found that pressure head data always improved the characterization of streamflow, whereas standalone streamflow assimilation showed less improvement. This was also confirmed by Bailey and Baù (2012) who estimated the spatial distribution of hydraulic conductivity by the assimilation of streamflow and water table data in a similar test case. Shi et al. (2014) assimilated six data types (discharge, groundwater table depth, SWC, land-surface temperature, sensible heat, latent heat, and transpiration) with the EnKF and the land surface-subsurface model Flux-PIHM mimicking a small forested catchment. They estimated several, but spatially homogeneous, van Genuchten soil hydraulic properties and found an improvement of the model simulations in association with a high sensitivity of the subsurface soil hydraulic properties (i.e. van Genuchten air entry and shape parameters). Shi et al. (2015) confirmed these findings also for experiments with real-world observations of the Shale Hills catchment in Pennsylvania, United States. Both studies by Shi et al. (2014, 2015) were conducted with a simplified representation of subsurface flow, namely 2D groundwater flow and 1D flow in the unsaturated zone.

Some studies with integrated hydrologic models focused more on experimental design. For instance, in a synthetic experiment which mimicked the Biosphere 2 Landscape Evolution Observatory (LEO) hillslopes, Pasetto et al. (2015) estimated saturated hydraulic conductivity and

investigated the impact of the number and spatial distribution of assimilated SWC data for a hillslope on the characterization of the hydrologic response of the CATHY model with a 3D synthetic subsurface. If the number of sensors fell below a certain threshold (100 instead of 496) the model was not able to reproduce the synthetic truth by SWC assimilation with EnKF. Rasmussen et al. (2015) investigated the relationship between the number of ensemble members and the number of pressure head measurement data on the estimation of the discharge rate with help of the transform Ensemble Kalman Filter. They performed this study for the agriculture dominated Karup catchment in Denmark and concluded that less head observations required more ensemble members to reproduce the synthetic discharge and pressure head observations. Assimilating discharge observations in combination with parameter estimation required a larger ensemble size than the assimilation of groundwater table measurements. Zhang et al. (2015) found that the EnKF performance is very sensitive to the estimation of the initial model parameter uncertainty. The majority of these examples are limited to synthetic test cases or, in case of real-world case studies, strong simplifications of the subsurface compartment. In particular, complex subsurface structures (i.e. fully 3D heterogeneous fields of subsurface hydraulic properties) and their impacts on the EnKF performance within a fully coupled land surface-subsurface model at high resolution have not yet been investigated.

DA-studies in combination with highly non-linear vadose zone flow models and heterogeneity of soil hydraulic parameters face additional challenges. EnKF shows optimal performance for Gaussian distributions, but strongly skewed pressure head distributions in the vadose soil zone can be expected under very dry conditions (Erdal et al., 2015; Zhang et al., 2018). Furthermore, it is important that the spatial structure of the heterogeneous soil hydraulic parameters (i.e., horizontal layers) is well captured. Otherwise a calibration with EnKF leads to poor results (Erdal et al., 2014). Data assimilation experiments for different soil types (e.g. Li and Ren, 2011; Montzka et al., 2011; Tran et al., 2014) also indicated that besides stratification and heterogeneity the capacity of DA to adequately characterize model states and parameters depends on the soil type and its related hydraulic properties. Within this context, Li and Ren (2011) indicated that the joint update of multiple hydraulic properties (i.e. van Genuchten inverse air entry (α) and shape parameter (n) in combination with hydraulic conductivity) by the EnKF using a 1D subsurface hydraulic model is more efficient than updating hydraulic conductivity alone. Other examples for the joint update of hydraulic conductivity, α, and n in combination with the EnKF and a 1D subsurface model are given by Wu and Margulis (2011, 2013). Montzka et al. (2011, 2013) used the particle filter for updating these soil hydraulic parameters in a 1D-model. The vast majority of 2D and 3D-studies where soil hydraulic properties are estimated only update saturated hydraulic conductivity, since the estimation of other soil hydraulic parameters is associated with numerical instability (Rasmussen et al. 2015; Vereecken et al. 2016). Exceptions are studies by Shi et al. (2014, 2015) and Passetto et al. (2015). However, these studies made various simplifications. Shi et al. (2014,

2015) assumed spatially homogeneous zones within their subsurface setup. Passetto et al. (2015) estimated 3D spatially distributed hydraulic conductivity and porosity under assumption of perfect knowledge of the remaining hydraulic parameters. Chaudhuri et al. (2018) estimated 3D spatially distributed hydraulic conductivity and Mualem-van Genuchten parameters with an iterative filter for a synthetic case. The impact of only updating saturated hydraulic conductivity on the performance of an integrated 3D terrestrial model for a real-world case, which includes also other spatial heterogeneous and uncertain soil hydraulic parameters (e.g. van Genuchten α, n) has not yet been explored.

This gap is covered in this study where SWC observations were assimilated in the Terrestrial Systems Modeling Platform (TerrSysMP) developed by Shrestha et al. (2014) which was coupled to the PDAF data assimilation framework (Nerger and Hiller, 2013) by Kurtz et al. (2016). More specifically, the physically based model ParFlow (Ashby and Falgout, 1996; Jones and Woodward, 2001; Kollet and Maxwell, 2006) and the Community Land Model (Oleson et al., 2004; Oleson et al., 2008) of TerrSysMP were used. ParFlow and CLM are fully coupled and take into account 3D subsurface flow, overland flow and topography. Datasets collected in the context of the infrastructure initiative Terrestrial Environmental Observatories (TERENO) (Zacharias et al., 2011), and the Transregional Collaborative Research center (TR-32) (Simmer et al., 2015) were used for assimilation and also verification. In particular, SWC, evapotranspiration and discharge measurements were available for this study in a headwater grassland catchment in the Eifel (Germany). The simulations were made with fully 3D heterogeneous fields of soil hydraulic properties which results in a model system with 0.3 million unknowns. Data assimilation experiments were both made for the real-world case and a synthetic case which mimics the real-world case, with the aim to unravel the role of model structural errors. In the DA-experiments different combinations of state and parameter updating were explored. The following research questions are addressed with the conducted experiments:

(1) Can the characterization of the hydrology at the hill slope scale (SWC, evapotranspiration, discharge) be improved with a combination of physically based integrated hydrological models and a dense network of SWC data, using data assimilation?
(2) Is the joint update of model states and saturated soil hydraulic conductivity sufficient to adjust the model simulations closer to the observations or is it necessary to include additional soil hydraulic properties (e.g. inverse air entry) within the parameter update to constrain this complex non-linear coupled model system with spatially heterogeneous soil hydraulic properties?
(3) Is there a systematic difference in the performance of data assimilation between synthetic experiments and real-world experiments which points to processes which are not captured by the high resolution integrated model?

5.2 Study Site and Data

5.2.1 The Rollesbroich Catchment

The Rollesbroich study site (50° 37' 27" N, 6° 18' 17" E) is located in the Eifel low mountain range (Germany). This sub-catchment of the river Rur has an area of ca. 38 ha, an altitude varying between 474 m and 518 m a.s.l., and slopes between 0 and 10 %. As a part of the TERENO infrastructure, the Rollesbroich test site is an example for an extensively managed grassland site dominated by smooth meadow grass (Poa pratensis) and perennial ryegrass (Lolium perenne). A drainage system in the northwestern part of the study site prevents flooding through high groundwater tables. The outlet of the drainage system is located close to the Kieselbach source. The diameters of the ca. 80 years old clay pipes vary between 3 and 20 cm. Figure 5.1 provides an overview map of the study site and the measurement equipment used in this study. Further details on the catchment characteristics can be found in Chapter 2.1.

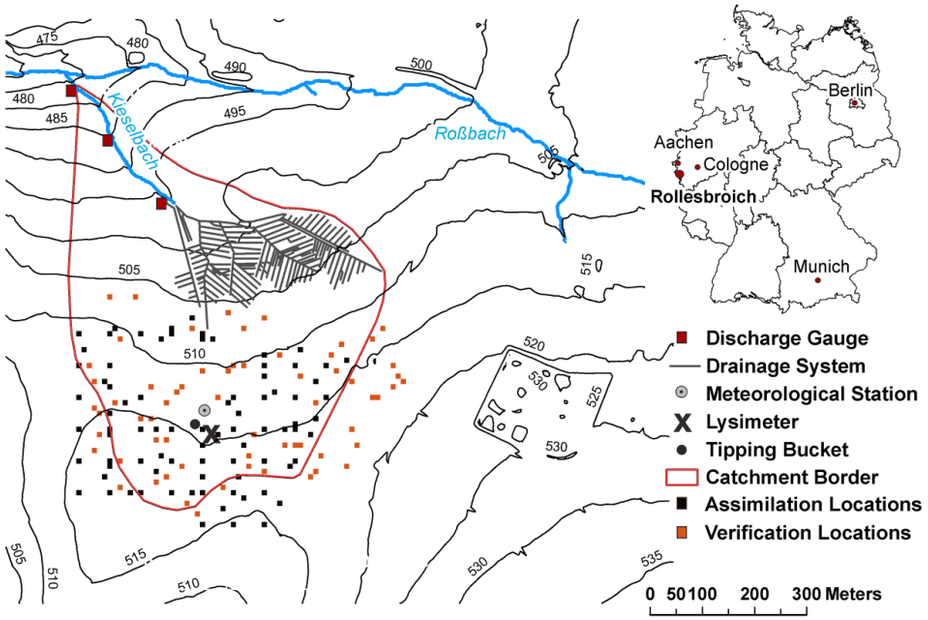

Figure 5.1: Rollesbroich overview map showing the locations of discharge gauges, lysimeter, meteorological station and drainage system. The locations for SWC assimilation (SoilNet) and verification are also indicated.

5.2.2 Observation Data

Precipitation measurements at the study site are performed with an interval of 10 minutes and a resolution of 0.1 mm using a standard Hellmann type tipping bucket balance (TB) rain gauge (ecoTech GmbH, Bonn, Germany). The device setup 1 m above ground is in agreement with the policy of the German weather service for areas with heavy snowfall and altitude > 500 m a.s.l. (WMO guideline recommends 0.5 m). To avoid measurement errors due to instrument freezing, the device is temporally heated during the winter period.

Latent and sensible heat fluxes are measured by an eddy covariance (EC) station at the southern part of the study site (50° 37' 19" N, 6° 18' 15" E, 514 m a.s.l.) and used for verification in this project. At this location also temperature and air humidity are recorded by a HMP45C, Vaisala Inc., Helsinki, Finland at 2.58 m above ground. Furthermore, a four-component net radiometer (CNR-1, Kipp and Zonen, Delft, The Netherlands) measuring incoming and outgoing longwave and shortwave radiation, a sonic anemometer (CSAT3, Campbell Scientific, Inc., Logan, USA) recording wind speed and direction, and a gas analyzer (LI7500, LI-COR Inc., Lincoln, NE, USA) for specific humidity and air pressure measurements are installed there at 2.6 m above ground surface. These on-site data were used as hourly forcing for the CLM model simulations for 2011. For 2010 and for gap filling off-site meteorological data from the nearby LUA NRW station were applied instead.

The SWC at the study site is measured with a wireless sensor network (SoilNet) (Qu et al., 2013). SoilNet data were used in the data assimilation and for model verification. At each of the 179 sensor locations two redundant SPADE ring oscillator sensors (model 3.04, sceme.de GmbH i.G., Horn-Bad Meinberg, Germany) are vertically installed at 5, 20 and 50 cm depth within a distance of ca. 10 cm to each other. This increases measurement volume and precision and helps to avoid data inconsistencies (e.g. contact issues with the soil matrix). Additional technical details can be found in Qu et al. (2013) and Qu et al. (2016). A long term evaluation of the SWCs measured with the SPADE sensors at the study site indicated a drift in SWC-values for individual SoilNet locations. The SWC data exhibited a gradual increase of maximum SWCs during saturated soil conditions for the period 2011-2013, which was not in line with the annual precipitation trend (Gebler et al. 2017). A linear SWC trend correction function was applied which takes the maximum SWC of the wet periods 2011 and 2012 into account. Trend correction was initially conducted for 82 SoilNet locations located in the southern part of the study site and separately for all three sensor depths (Gebler et al. 2017). This dataset was further reduced to 61 sensor locations, since only gap free SWC time series were used for DA. The sensors of the northern part of the study site were not included in our experiments as SWC-values for the northern part of the study site were not available before 2012.

Discharge measurements, used for verification in this study, for the Kieselbach are conducted with a Venturi-Gauge weir close to the catchment outlet and two upstream Tomson gauges close to the headwaters of the Kieselbach (Qu et al. 2016).

5.3 Materials and Methods

5.3.1 TerrSysMP

For this study, the ParFlow and CLM components of the Terrestrial Systems Modelling Platform (TerrSysMP) (Shrestha et al., 2014) were applied in combination with the Parallel Data Assimilation Framework (PDAF) (Nerger and Hiller, 2013), which was implemented by Kurtz et al. (2016) using the external coupler OASIS3-MCT (Valcke, 2013). A detailed description of the individual model components is given in Chapter 2 of this thesis.

5.3.2 State and Parameter estimation with the EnKF

In this study, the EnKF was applied to the ParFlow-CLM model of the Rollesbroich study site. The EnKF is an ensemble based sequential DA method to improve the characterization of model states (and possibly parameters) on the basis of optimally combining information from model simulations and observations.

The assimilation cycle of the EnKF is a two-step process consisting of a forecast and an analysis, which are repeated sequentially. In the forecast step, each stochastic realization of the ParFlow-CLM model (M) is propagated forward in time:

$$\mathbf{x}_i^t = M(\mathbf{x}_i^{t-1}, \rho_i, \boldsymbol{\eta}_i) \tag{5.1}$$

where $\mathbf{x}_i^t$ represents the predicted model state vector at time step t, $\mathbf{x}_i^{t-1}$ is the model state vector of the previous time step t-1, ρ_i represents all the model parameters, , and $\boldsymbol{\eta}_i$ are the model forcings for each ensemble realization i.

The prognostic variable in Parflow is pressure. However, SWC is assimilated in this study which is used to update pressure indirectly via the correlation of both variables. The model state vector hence consists of the pressure head ($\mathbf{h}$) and the soil water content ($\boldsymbol{\theta}$).

The model state vector at time t is then defined as:

$$\mathbf{x}^t = \begin{pmatrix} \mathbf{h}^t \\ \boldsymbol{\theta}^t \end{pmatrix} \tag{5.2}$$

For the EnKF the ensemble mean represents the best estimate of the true model state. The model covariance matrix is determined from the ensemble of forward models runs at each time step by

$$\mathbf{P}^t = \frac{1}{n_{\text{realz}} - 1} \sum_{i=1}^{\text{n}_{\text{realz}}} \left(\mathbf{x}_i^t - \overline{\mathbf{x}}^t\right)\left(\mathbf{x}_i^t - \overline{\mathbf{x}}^t\right)^T \tag{5.3}$$

where $\mathbf{P}$ is the model covariance matrix calculated from the ensemble of model forecasts and $\overline{\mathbf{x}}^t$ contains the average model states calculated over all realizations n_{realz}.

The SWC observations are perturbed according (Burgers et al., 1998):

$$\tilde{\mathbf{y}}_i^t = \mathbf{y}_i^t + \boldsymbol{\varepsilon}_{iy} \tag{5.4}$$

where $\tilde{\mathbf{y}}_i^t$ is a vector with perturbed observations, $\mathbf{y}_i^t$ is the observation vector, and $\boldsymbol{\varepsilon}_{iy}$ is a vector whose elements contain a value drawn from a normal distribution $\mathcal{N}(o, \sigma)$ with mean zero and standard deviation σ.

With the EnKF, the model states are updated with help of SWC observations. This can be done by only updating the states:

$$\mathbf{x}_i^a = \mathbf{x}_i^t + \mathbf{K}^t(\tilde{\mathbf{y}}_i^t - \mathbf{H}^t\mathbf{x}_i^t) \tag{5.5}$$

where $\mathbf{x}_i^a$ is the updated state vector [L] and $\mathbf{K}$ the Kalman Gain. The Kalman Gain $\mathbf{K}$ is calculated according:

$$\mathbf{K}^t = \mathbf{P}^t\mathbf{H}^{t^T}\left(\mathbf{H}^t\mathbf{P}^t\mathbf{H}^{t^T} + \mathbf{R}\right)^{-1} \tag{5.6}$$

where $\mathbf{H}$ is the measurement operator and $\mathbf{R}$ is the observation error covariance matrix.

The covariance matrix $\mathbf{P}$ consists then of the cross-covariances between SWC and pressure head at the measurement locations:

$$\mathbf{P}^t = \begin{pmatrix} \mathbf{P}_{\mathbf{hh}}^t & \mathbf{P}_{\mathbf{h\theta}}^t \\ \mathbf{P}_{\mathbf{\theta h}}^t & \mathbf{P}_{\mathbf{\theta\theta}}^t \end{pmatrix} \tag{5.7}$$

As highlighted in the introduction, for data assimilation problems involving groundwater and soil it is important to update uncertain parameters (e.g. Chen and Zhang, 2006; Hendricks Franssen and Kinzelbach, 2008; Chen et al., 2011).

The augmented state vector if both states and parameters are updated is:

$$\mathbf{x}^t = \begin{pmatrix} \mathbf{h}^t \\ \boldsymbol{\theta}^t \\ \mathbf{Y}^t \end{pmatrix} \tag{5.8}$$

where **x** is now the augmented state vector including pressure heads (**h)** [L], the soil water content (**θ**) [L^3 L^{-3}] and the logarithm of soil hydraulic conductivities (**Y=** log_{10} $\mathbf{K}_s$ [L T^{-1}]).

In case parameters are updated, the covariance matrix **P** consists of the states and parameters (**h, θ, Y**) at the observation locations ($\tilde{\boldsymbol{y}}$)

$$\mathbf{P}^t = \begin{pmatrix} \mathbf{P}_{\mathbf{hh}}^t & \mathbf{P}_{\mathbf{h\theta}}^t & \mathbf{P}_{\mathbf{hY}}^t \\ \mathbf{P}_{\boldsymbol{\theta}\mathbf{h}}^t & \mathbf{P}_{\boldsymbol{\theta\theta}}^t & \mathbf{P}_{\boldsymbol{\theta}\mathbf{Y}}^t \\ \mathbf{P}_{\mathbf{Yh}}^t & \mathbf{P}_{\mathbf{Y\theta}}^t & \mathbf{P}_{\mathbf{YY}}^t \end{pmatrix} \tag{5.9}$$

In some of the scenarios where both states and parameters are updated, a damping factor (α) was used to minimize the impact of filter inbreeding (Hendricks Franssen and Kinzelbach, 2008). This results in the following updating equation for the joint state-parameter estimation problem:

$$\mathbf{x}_i^a = \mathbf{x}_i^t + \alpha \mathbf{K}(\tilde{\mathbf{y}}_i^t - \mathbf{H}\mathbf{x}_i^t) \tag{5.10}$$

where $\mathbf{x}_i^a$ is the updated augmented state vector after data assimilation, and α [-] the damping factor.

Although the damping factor and other simple inflation methods have been applied successfully (e.g. Hendricks Franssen and Kinzelbach, 2008; Erdal et al., 2014), these methods are subject of controversial discussion (Houtekamer and Zhang, 2016). The reduced impact of the error statistics in the EnKF potentially obscures other sources of error and therefore can lead to inconsistencies in the derivation of the Kalman Gain (Houtekamer and Zhang, 2016).

For conducting DA with TerrSysMP, TerrSysMP-PDAF has been used, which is a compute efficient implementation showing a very favorable scalability. For the different variants of the PDAF module within TerrSysMP as well as additional technical details see Kurtz et al. (2016).

5.3.3 Setup of Data Assimilation Experiments

The Rollesbroich model domain covers an area of 1280 × 1120 m with a lateral resolution of 10 × 10 m and a total depth of 3.2 m. The ParFlow subsurface domain was built with a variable vertical resolution making use of the terrain following grid (TFG) approach (Maxwell, 2013). From surface to bedrock the model resolution was gradually reduced with depth. The first layer with a vertical extend of 0.025 m is followed by 10 layers of 0.05 m vertical resolution and a 12^{th} layer

with 0.1 m depth. The vertical resolution then further decreases from 0.2 m (layer 13 – 17) to 0.5 m (layer 18 – 19), and 0.575 m (layer 20) for the deepest layer (bedrock).

The subsurface parameterization scheme mainly follows the heterogeneous subsurface setup by Gebler et al. (2017). For the assimilation experiments, the subsurface was structured into three soil horizons and two bedrock layers with spatially heterogeneous stochastic fields of van Genuchten hydraulic properties. These were derived using information from the 1D SCE-UA optimization approach for the different sensor locations (Qu et al., 2014) in the upper soil horizons or, respectively, random samples for the underlying bedrock layers. For the weathered siltstone and sandstone bedrock layers preferential flow though macropores and subsurface flow is very likely. To mimic these characteristics, the hydraulic conductivity of the bedrock was increased to the magnitude of a sandy soil to enable an adequate drainage of the upper soil horizons and a high lateral flow component in the bedrock layer in order to increase the efficiency of the drainage system. For the bedrock layers the setup differs from Gebler et al. (2017), since technical reason prevented an increase of the lateral flow component in TerrSysMP-PDAF. Instead the bedrock was subdivided into two separate layers and the characteristics of the bedrock were changed. The hydraulic conductivity of the lower bedrock horizon (1.2 - 1.5 m thickness) was more increased than the upper bedrock horizon (0.2 - 0.5 m thickness). This guarantees a realistic drainage of the lower soil horizon. Figure 5.2 illustrates the simulation setup with the different horizons.

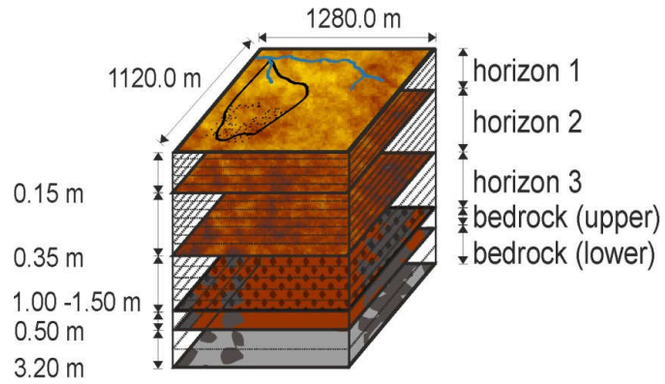

Figure 5.2: Schematic overview of the five heterogeneous horizons of soil hydraulic parameters within the Parflow model. Study site borders (black line drawn inside horizon 1), channels (blue) and locations of the SoilNet devices (black dots) are also indicated. The hatchings between the third and the upper bedrock horizon represent the variable bedrock depth at 1.0 – 1.5 m.

The default CLM subsurface setup with 10 exponentially increasing subsurface layers was adapted to match the before mentioned extent of corresponding ParFlow model layers. This results in a model domain of 128 × 112 × 20 (286720) grid cells and includes the Rollesbroich catchment area of 38 ha as well as parts of the Rollesbroich residential area, streets and other anthropogenic artificial structures. The slopes in x and y direction, which represent the topographic driven overland flow in the domain, were calculated with the help of a digital elevation model (DEM). Further details of the DEM preparation can be found in Gebler et al. (2017). A Manning's roughness value of 0.001 h $m^{-1/3}$ accounts for high (bank) vegetation density and the small stream channel bed of approx. 30 cm. The lower boundary was set impermeable at 3.2 m soil depth.

The CLM vegetation parameterization mainly relies on standard spatial uniform CLM C3-grass parameters with an annual leaf area index (LAI) varying between 0.3 and 3.0. The CLM default root distribution parameters (roota, rootb) were adapted in accordance with the modifications made for the CLM subsurface extent (roota: 10.6; rootb: 6.0). This modification, which assigns 90 % of grass roots within the upper 30 cm soil, is in line with literature values (e.g. Brown et al., 2010). No flow boundary conditions were imposed for the northern, eastern, southern and western domain boundaries.

Precipitation and potential evapotranspiration from on-site measurements (2011) and the nearby LANUV station (2010) were used as input to CLM in hourly time steps (Gebler et al. 2017).

In total, a period of two years (2010-2011) was simulated with ParFlow-CLM.

5.3.4 PDAF Setup and Assimilation Scenarios

The PDAF (v1.0) module coupled to TerrSysMP was used for the assimilation of SWC observations from SoilNet infrastructure installed at the Rollesbroich study site. Figure 5.3 schematically illustrates the operation mode of PDAF with the EnKF when applied with ParFlow-CLM. States and parameters of the different ParFlow-CLM model runs are collected by PDAF after a predefined assimilation interval. By assimilating SWC observations either model states or states and parameters are updated and passed back to the ParFlow-CLM realizations which are then propagated further in time.

First, ParFlow-CLM open loop simulations (Table 5.1) for the entire ensemble and the entire simulation period (January 2010 – December 2011) were performed with this setup. This period includes the model spin-up which was conducted for a period of one year (2010) beginning with an initial hydrostatic equilibrium condition and a groundwater table at 1.5 m depth.

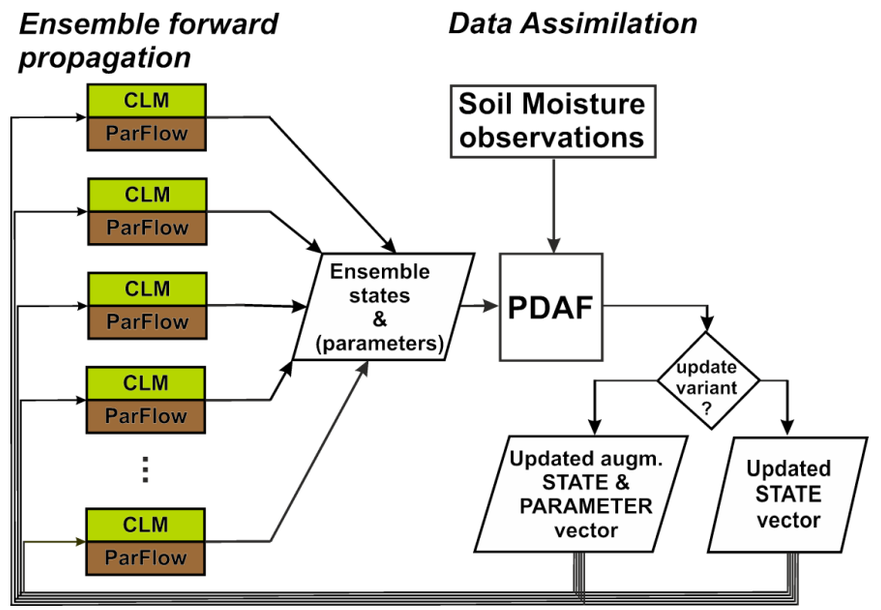

Figure 5.3: Flowchart of the subsurface parameter sampling creating a heterogeneous layered subsurface for each Parflow-CLM realization (modified from Gebler et al., 2017).

Second, several assimilation variants (Table 5.1) were tested with the EnKF assimilating SWC data taken from 61 locations in the southern part of the study site at 5, 20, and 50 cm depth. More specifically these variants are:

(1) state update on daily basis
(2) Joint update of states and hydraulic conductivity on daily basis
(3) Joint update of states and hydraulic conductivity where states are updated daily and hydraulic conductivity only each five days. In addition a damping factor (α=0.1) was applied.

In order to investigate potential model structural errors or methodical issues with the EnKF, the aforementioned variants were tested with synthetic data mimicking the real-world data at identical measurement locations. This includes also a scenario using deterministic van Genuchten α and n for the entire ensemble. Furthermore, the impact of ensemble size was explored. The selected synthetic reference truth was an ensemble member of the open loop ensemble run and synthetic measurement data were extracted from this reference run at 61 observation locations at 5, 20, and 50 cm depth. The assimilation period for real-world and synthetic experiments was from May 1, 2011 - December 31, 2011. The synthetic case was further evaluated against 63 random locations

(Figure 5.1) different from the assimilated locations in 5, 20, and 50 cm depth. In addition two layers of the upper and lower bedrock in 130 and 190 cm depth were examined.

The measurement error was set to 0.04 cm^3 cm^{-3} and assumed to be spatially uncorrelated for synthetic and real-world scenarios. Surface cells with overland flow (i.e. with saturated conditions) were excluded from the update. This avoids numerical instability by an on/off switching of the surface runoff generation which is caused by the perturbation of the pressure field after state update.

Table 5.1: Simulation scenarios for different synthetic and real-world data assimilation experiments. The open loop simulations which refer to synthetic reference (SYN_OL) or real-world observation data (REAL_OL) are also indicated. The Assimilation period for the DA scenarios is May¬ December 2011.

Scenario	Observation	Realizations	Parameter Update Frequency [d]	Parameter Damping
REAL_OL	SoilNet	128	n/a	n/a
REAL_ST	SoilNet	128	n/a	No
REAL_PAR_1	SoilNet	128	1	No
REAL_PAR_5d	SoilNet	128	5	Yes
SYN_OL	synthetic	128	n/a	n/a
SYN_ST	synthetic	128	n/a	No
SYN_PAR_1	synthetic	128	1	No
SYN_PAR*αn*_1 (determ. α, n)	synthetic	128	1	No
SYN_PAR_5d	synthetic	128	5	Yes
SYN_OL_256	synthetic	256	n/a	n/a
SYN_PAR*αn*_5d_256 (determ. α, n)	synthetic	256	5	Yes

5.3.5 Ensemble generation

In order to account for input uncertainties, model forcings as well as soil hydraulic parameters of the ParFlow-CLM model were perturbed. This resulted in 128 (or 256) different stochastic model realizations for the synthetic and real-world experiments. Hourly precipitation, assumed homogeneous for the catchment, was perturbed according a normal distribution by multiplying with a Gaussian random value ($\mathcal{N}$ (1, 0.15)). As precipitation measurements were available in the catchment, it was assumed that the uncertainty with respect to this forcing was mainly related to measurement uncertainty and therefore relatively small which justifies the Normal assumption.

The creation of an ensemble of uncertain subsurface hydraulic properties departed from van Genuchten soil hydraulic properties estimated by HYDRUS-1D (Šimůnek et al., 2008) and the

shuffle complex evolution algorithm (SCE-UA) (Duan et al. (1992) for 82 locations in the southern part of the study area. This approach was already demonstrated to capture the spatial variability at the Rollesbroich test site (Qu et al., 2014) and tested in combination with a 3D ParFlow-CLM model (Gebler et al., 2017). On this account, Gebler et al. (2017) extended the initial approach by an uncertainty assessment of the estimated soil hydraulic properties (Figure 5.4) which was not provided by the original SCE-UA parameter estimation. Therefore, the discrepancy between the measurement support (at the optimization location) and the model discretization was considered on the basis of ordinary block kriging (Burgess and Webster, 1980) using the VESPER (Whelan et al., 2002) software in an unconditional manner at 10 × 10 m model resolution for all soil hydraulic parameters. The estimated uncertainty was similar to the uncertainty that alternatively was derived with the help of measured texture data and the ROSETTA (ROS) pedo-transfer-function (Schaap et al., 2001) (e.g. van Genuchten inverse air entry α: ± 1.3 cm^{-1} (SCE-UA); ± 1.5 cm^{-1} (ROS)) as clarified in Gebler et al. (2017). These data (optimized SCE-UA parameter with uncertainty estimation by block kriging) were the basis for the sampling of 128 stochastic realizations with 3D heterogeneous fields of soil hydraulic properties which were used in the data assimilation approach.

The subsequent sampling of van Genuchten soil hydraulic properties (Figure 5.4) consists of four major steps. First, for the upper three layers, samples were taken from the multivariate distribution of the soil hydraulic parameters for each model realization. This distribution is defined by the individual parameter mean values and (co-)variances, which were given by the optimized 1D parameter set and the available uncertainty at each spatial location and depth. For the underlying siltstone and sandstone bedrock (horizon 4, 5), the hydraulic properties were randomly sampled at each location on the basis of parameter ranges provided by Bogena (2003). The bedrock hydraulic properties were further artificially increased to the magnitude of a sandy soil with the aim to mimic the hydraulic characteristics of the sapprolite and the on-site drainage system.

Second, the logarithmic hydraulic conductivity ($\log_{10}K_s$) at each location of the upper layers (5, 20, and 50 cm) was perturbed with a spatially homogeneous normal distributed Gaussian random value $\mathcal{N}$ (0, 0.25) to create more variable logarithmic hydraulic conductivity averages between different realizations for the individual soil horizons. With this procedure, additional model uncertainty was taken into account which may originate from an underestimation of the 1D inversion uncertainty for hydraulic conductivity.

Third, a spatial heterogeneous field of $\log_{10}K_s$ was generated with sequential Gaussian simulation using GCOSIM3D (Gómez-Hernández and Journel, 1993) independently for each soil layer using the generated $\log_{10}K_s$ samples for the individual sensor locations. The required variogram parameters (range, sill, nugget) were estimated by fitting the experimental semi-variograms for the different soil layers to an exponential model for each soil layer (horizon 1-3). For the upper and

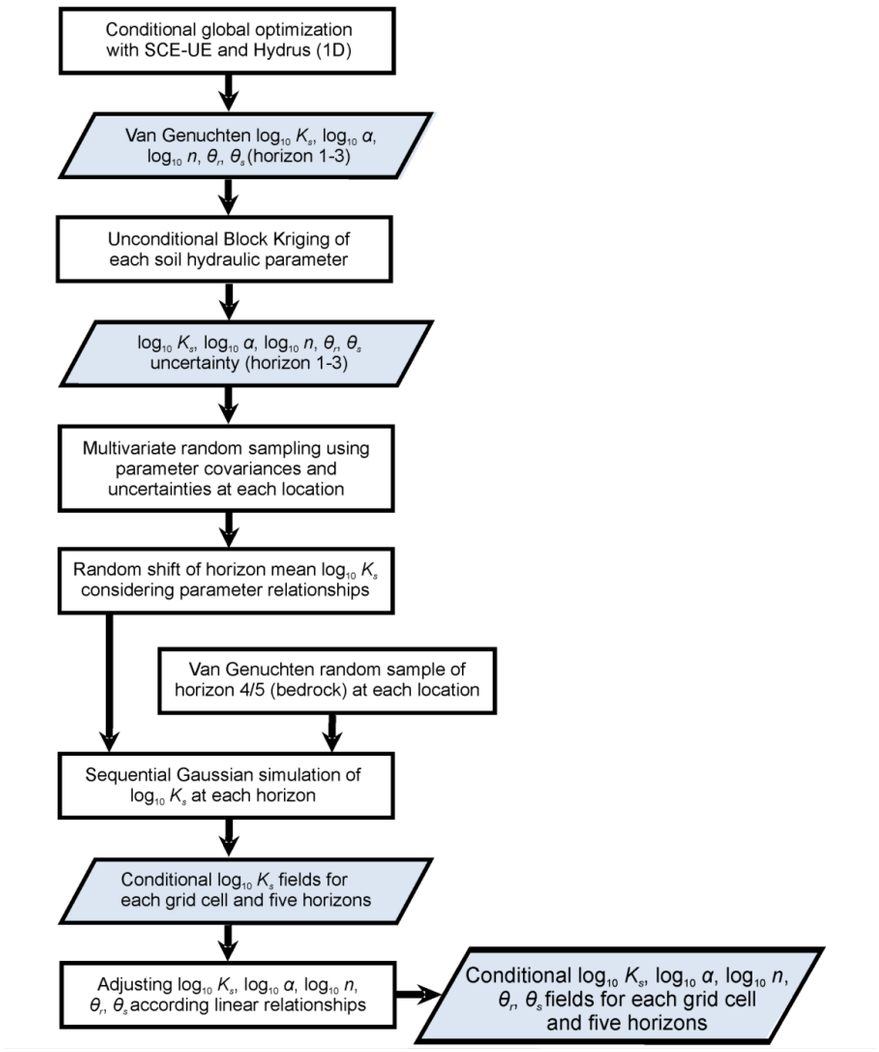

Figure 5.4: Flowchart of the subsurface parameter sampling creating a heterogeneous layered subsurface for each Parflow-CLM realization (modified from Gebler et al., 2017).

lower bedrock layer only very limited information regarding the spatial dependence of the hydraulic properties was available. Therefore, the variogram parameters of horizon 3 for the bedrock stochastic simulations were used.

In the fourth and last step, $\log_{10} \alpha$, $\log_{10} n$, θ_r and θ_s were adapted according their relations with $\log_{10} K_S$. The spatially heterogeneous $\log_{10} K_S$ fields, obtained in the third step, render $\log_{10} \alpha$,

$\log_{10} n$ also spatially variable using the multivariate relationships of $\log_{10} K_s$ with the other variables. This results in stochastic realizations for the different soil and bedrock layers of $\log_{10} K_s$, $\log_{10} \alpha$, $\log_{10} n$, θ_r and θ_s. For the different scenarios, 128 (or 256) realizations were used in the data assimilation procedure.

5.3.6 Performance Validation

The performance of the open loop and different DA scenarios was evaluated with measured or synthetic observations of daily SWC, daily ET and daily discharge. The Nash-Sutcliffe efficiency index (NSE), the root mean squared error (RMSE) and bias (BIAS) were calculated for each simulation scenario. The NSE was calculated according:

$$\mathrm{NSE} = 1 - \left(\sum_{t=1}^{n_{\mathrm{tstep}}} \sum_{i=1}^{n_{\mathrm{sample}}} \left(\bar{y}_{i,t}^{\mathrm{sim}} - y_{i,t}^{\mathrm{obs}} \right)^2 \Big/ \sum_{t=1}^{n_{\mathrm{tstep}}} \sum_{i=1}^{n_{\mathrm{sample}}} \left(y_{i,t}^{\mathrm{obs}} - \bar{y}_i^{\mathrm{obs}} \right)^2 \right) \tag{5.11}$$

where $\bar{y}_{i,t}^{\mathrm{sim}}$ are the simulated ensemble mean values at the observation (or verification) locations for the i^{th} sample at time t, $y_{i,t}^{\mathrm{obs}}$ are the observed data (or synthetic observations from the synthetic reference), $\bar{y}_i^{\mathrm{obs}}$ the average of the observed data over time, n_{sample} the number of samples, and n_{tstep} the total number of time steps. The NSE range is between -∞ and 1.0. Negative values indicate unacceptable simulation performance whereas positive values suggest good performance (Nash and Sutcliffe, 1970; Moriasi et al., 2007).

The model bias (BIAS) and the relative model bias in percent (PBIAS) are given by:

$$\mathrm{BIAS} = \frac{1}{(n_{\mathrm{sample}} \times n_{\mathrm{tstep}})} \sum_{t=1}^{n_{\mathrm{tstep}}} \sum_{i=1}^{n_{\mathrm{sample}}} \left(\bar{y}_{i,t}^{\mathrm{sim}} - y_{i,t}^{\mathrm{obs}} \right) \tag{5.12}$$

$$\mathrm{PBIAS} = 100 \times \left(\sum_{t=1}^{n_{\mathrm{tstep}}} \sum_{i=1}^{n_{\mathrm{sample}}} \left(\bar{y}_{i,t}^{\mathrm{sim}} - y_{i,t}^{\mathrm{obs}} \right) \Big/ \sum_{t=1}^{n_{\mathrm{tstep}}} \sum_{i=1}^{n_{\mathrm{sample}}} y_{i,t}^{\mathrm{obs}} \right) \tag{5.13}$$

The RMSE (5.14) was calculated to compare observed values with the ensemble mean from the simulations:

$$\mathrm{RMSE} = \sqrt{\frac{1}{(n_{\mathrm{sample}} \times n_{\mathrm{tstep}})} \sum_{t=1}^{n_{\mathrm{tstep}}} \sum_{i=1}^{n_{\mathrm{sample}}} \left(\bar{y}_{i,t}^{\mathrm{sim}} - y_{i,t}^{\mathrm{obs}} \right)^2} \tag{5.14}$$

The RMSE was also calculated for individual locations. Additionally, the mean RMSE over all locations at individual soil layers was used for the evaluation of the model simulations.

5.4 Results

5.4.1 Real-world Experiments

5.4.1.1 Soil Water Content

In this section results are discussed for SWC characterization with the help of assimilation of data for the real-world case. In section 5.4.2 results for the synthetic experiments follow. Figure 5.5 shows the temporal evolution of the mean SWC for the open loop simulations and three different data assimilation experiments for the assimilated locations in 2011. The measured mean SoilNet SWC at the study site for this period is high during winter and periods with intensive rainfall or thunderstorm events. The smallest SWC for 2011 is reached in May-June. This basic seasonality is captured well by the open loop and DA simulations. The comparison with the measured average SWC (Table 5.2) shows that the open loop simulation (REAL_OL) is close to the measurements at 5 cm depth (NSE: 0.894; RMSE: 0.026 cm^3 cm^{-3}), but not at 20cm and 50cm depth with negative NSE (20 cm: -0.269; 50 cm: -7.711), high RMSE (20 cm: 0.048 cm^3 cm^{-3}; 50 cm: 0.059 cm^3 cm^{-3}) and significant bias (20cm: 0.046 cm^3 cm^{-3}; 50 cm: 0.057 cm^3 cm^{-3}). The data assimilation scenarios REAL_ST (state updating only) and REAL_PAR_1 (daily joint state-parameter updating) and REAL_PAR_5d (daily joint updates of states) and parameters (each 5 days with damping factor) showed an improved SWC average compared to the open loop for May and June 2011 for the 5 cm layer, but these scenarios diverge from the observations afterwards until the end of

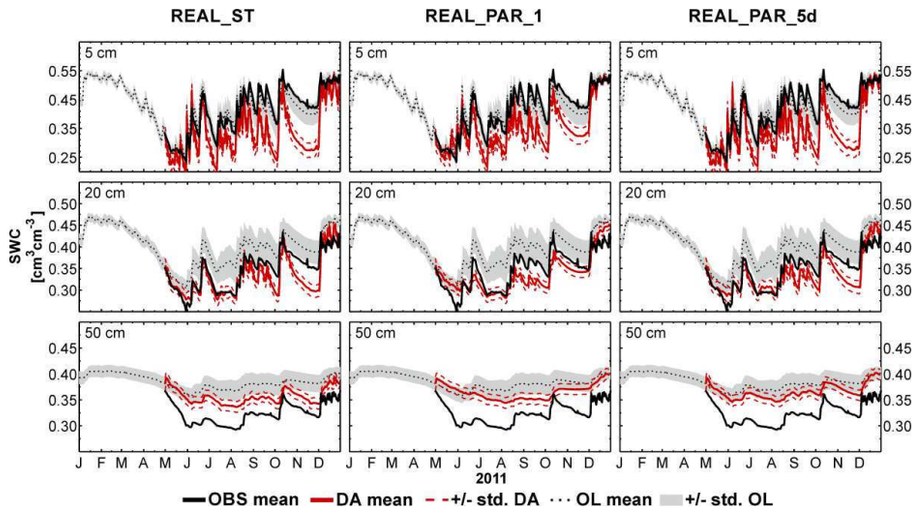

Figure 5.5: Mean SWC and standard deviation of SWC for the open loop and the real-world data assimilation experiments for the assimilation locations. Mean SoilNet observations are indicated as well.

Table 5.2: Performance measures for the reproduction of SWC for the real-world experiments. Shown are performance statistics for the open loop run (REAL_OL) and the DA experiments for 2011.

		REAL_OL	REAL_ST	REAL_PAR_1	REAL_PAR_5d
5 cm	**NSE [-]**	0.894	-0.429	0.007	0.436
	BIAS [cm³ cm⁻³]	0.002	-0.079	-0.064	-0.041
	RMSE [cm³ cm⁻³]	0.026	0.095	0.079	0.053
20 cm	**NSE [-]**	-0.269	0.599	0.596	0.576
	BIAS [cm³ cm⁻³]	0.046	-0.013	-0.009	-0.025
	RMSE [cm³ cm⁻³]	0.048	0.027	0.027	0.030
50 cm	**NSE [-]**	-7.711	-2.000	-3.935	-1.372
	BIAS [cm³ cm⁻³]	0.057	0.033	0.042	-0.017
	RMSE [cm³ cm⁻³]	0.059	0.034	0.044	0.017

November 2011. The average SWC at 20 and 50 cm depth is better characterized than in the open loop run for the three data assimilation scenarios. For the 50 cm layer the high initial model bias gradually decreased over the entire assimilation period. Whereas the three data assimilation scenarios for 5 and 20 cm depth give very similar results (October, November), at 50 cm depth the data assimilation scenarios including parameter updates (REAL_PAR_1, REAL_PAR_5d) outperform data assimilation with state update only. REAL_PAR_5d shows the best reproduction of the measurement values. The bias for SWC-characterization at 50 cm depth decreases from 0.057 cm^3 cm^{-3} (REAL_OL) to 0.017 cm^3 cm^{-3} (REAL_PAR_5d).

Figure 5.6 displays the RMSE for individual SoilNet observations at 5, 20, and 50 cm depth. The real-world experiments REAL_ST and REAL_PAR_1 show decreasing simulation performance (larger RMSE than in open loop) for 62-67 % of the sensor locations at 5 cm depth. Only less frequent assimilation in combination with joint state-parameter estimation result in a small decrease in layer-averaged RMSE of the individual sensors (REAL_OL: 0.0936 cm^3 cm^{-3}; REAL_PAR_5d: 0.0858 cm^3 cm^{-3}) at this depth. However, an improvement was found for SWC characterization at 20 cm and 50 cm depth for all real-world DA-experiments (16-28 % RMSE-reduction at 20 cm and 8-18 % RMSE- reduction at 50 cm). REAL_PAR_5d reproduces SWC-profiles the best over all three layers showing a reduction by 14 % for the RMSE averaged over all sensor locations. For REAL_ST the RMSE decreases by 6 % whereas REAL_PAR_1 shows only 2 % RMSE decrease. RMSE decreases or increases thereby do not show vertical correlations. Spatial RMSE-structures are not very similar for different simulation scenarios. In particular, the spatial RMSE structure for REAL_PAR_5d differs from REAL_ST and REAL_PAR_1.

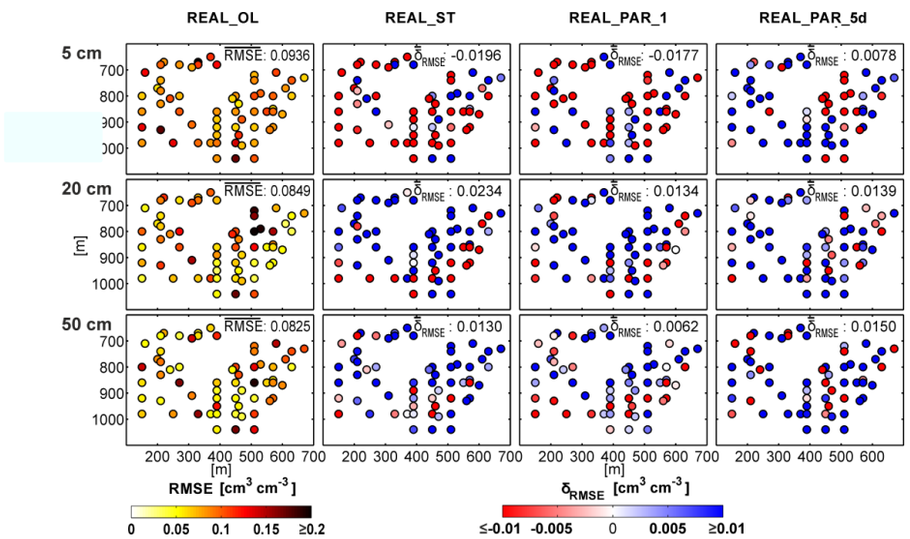

Figure 5.6: RMSE of SWC at individual locations for the open loop runs (left column) and changes in RMSE (increase implies improvement and decrease implies impoverishment) for three data assimilation scenarios (three columns on the right) of the real-world case for 2011.

5.4.1.2 Parameter Estimation

Figure 5.7 illustrates the temporal evolution of the hydraulic conductivity ($\log_{10}K_s$) estimates for the real-world experiments. For REAL_PAR_1 the ensemble spread rapidly decreases within the first weeks of assimilation. In contrast, the ensemble spread of REAL_PAR_5d with less frequent parameter updating (each five days only) still exhibits ca. 25 % of the initial ensemble spread at the end of the assimilation period. The different data assimilation scenarios show large variations in $\log_{10}K_s$ estimates for the real-world case. In particular, $\log_{10}K_s$ of REAL_PAR_1 varies between the individual horizons and the final parameter estimates show major differences in comparison with REAL_PAR_5d. The temporal evolution of $\log_{10}K_s$ for REAL_PAR_1 shows also some instability. The $\log_{10}K_s$ of REAL_PAR_1 for the second and third layer first rapidly decreases resp. increases until the end of May. Afterwards it increases resp. decreases again until it converges to a final parameter estimate which is reached at the end of July. The final estimates differ up to 0.6 $\log_{10}K_s$ (m h^{-1}) between REAL_PAR_1 and REAL_PAR_5d.

Figure 5.8 shows the differences of the ensemble average hydraulic conductivity fields, as updated in the data assimilation scenarios, compared to the ensemble average hydraulic conductivity of the open loop run, for the real-world case. The figure shows that for REAL_PAR_1 parameter updates are stronger than for REAL_PAR_5d. REAL_PAR_1 exhibits high $\log_{10}K_s$ contrasts between the sensor locations but also for the bedrock layer between different areas of the northern study site. Also REAL_PAR_5d shows positive and negative changes in $\log_{10}K_s$ which are more limited spatially.

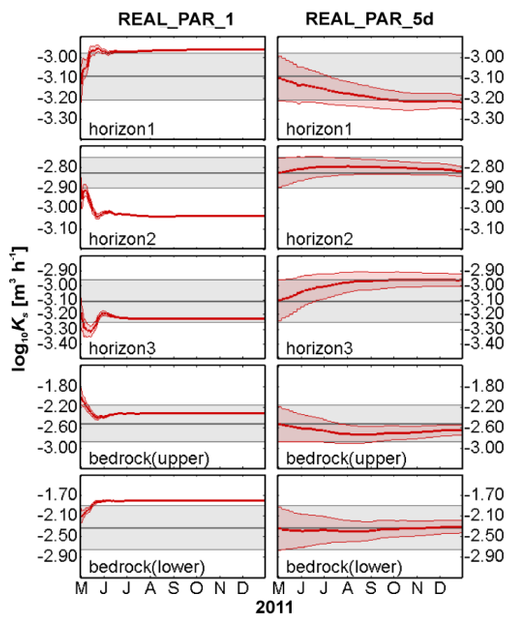

Figure 5.7: Mean and standard deviation of soil hydraulic conductivity for real-world experiments (red) against the open loop parameter set (grey) for different subsurface horizons of the Rollesbroich catchment. Please note that the used scales on the vertical axis differ for the different horizons.

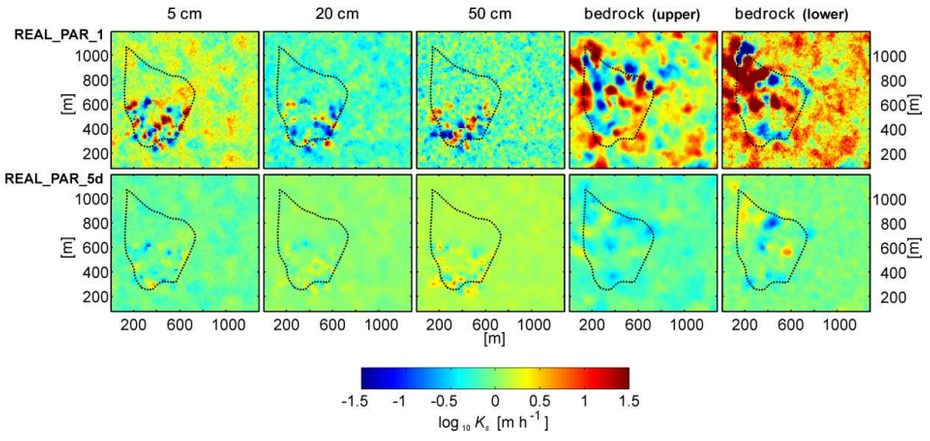

Figure 5.8: The ensemble averaged perturbation applied to the hydraulic conductivity field ($\log_{10}K_S$) during data assimilation for the real-world simulation scenarios. Indicated are results for the soil layers at 5, 20, and 50 cm depth as well as the upper and lower bedrock layer. The dashed line marks the test site borders within the model domain.

5.4.1.3 Discharge

Figure 5.9 shows the discharge dynamics for the real-world experiments for 2011. In general, the observations at the study site show high flow during winter rain and snowmelt, and low flow from May-November. This seasonality is captured by the open loop simulations, although discharge is overestimated during low flow conditions and for individual winter peaks. These findings are also reflected by the performance measures for the open loop simulations in Table 5.3. In comparison with the on-site measurements, the open loop simulations show a high relative bias (64.98 %) and high RMSE (15.77 $m^3\ h^{-1}$). The DA simulations exhibit a very different behavior. With the more intensive adjustment of SWC in the upper horizons, also discharge is affected by the assimilation cycles. This leads to absence of streamflow in the REAL_ST and the REAL_PAR_5d in November 2011. As a consequence, both scenarios exhibit a high relative bias (REAL_ST: -88.84 %; REAL_PAR_5d: -75.00 %) indicating a discharge under-estimation. The performance of REAL_ST (NSE: 0.05; RMSE: 20.43 $m^3\ h^{-1}$) is even lower than for the open loop run (NSE: 0.49; PBIAS: 64.98 %; RMSE: 15.77 $m^3\ h^{-1}$). Parameter estimation (REAL_PAR_1) results in an improved modelling of discharge (NSE: 0.67; PBIAS: -23.55 %; RMSE: 12.08 $m^3\ h^{-1}$). In particular the bias during low flow periods is strongly reduced for these simulations. However, the peak discharge of REAL_PAR_1 is still overestimated during the entire assimilation period.

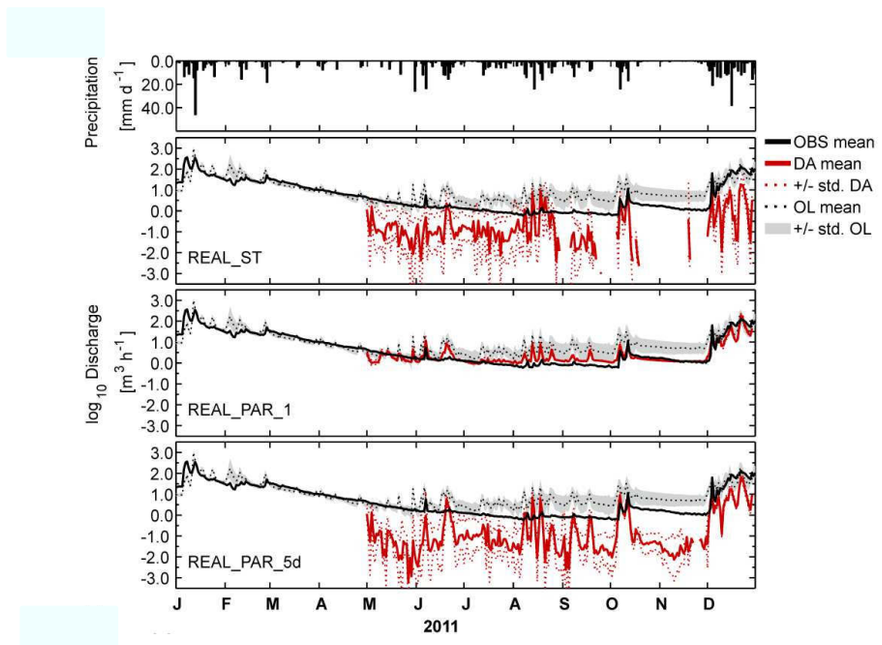

Figure 5.9: Simulated and observed daily mean logarithmic discharge for the assimilation period at the Rollesbroich catchment outlet. Open loop simulations are shown together with different updating scenarios for the real-world experiments. Precipitation is also indicated on a daily basis.

Table 5.3: Discharge model performance for the assimilation period at the catchment outlet for different update scenarios with real-world observations in 2011. The model performance is indicated by mean NSE, standard deviation of NSE, average bias and root mean square error (RMSE).

	REAL_OL	REAL_ST	REAL_PAR_1	REAL_PAR_5d
NSE [-]	0.49	0.05	0.67	0.41
PBIAS %	64.98	-88.84	-23.55	-75.00
RMSE [$m^3 h^{-1}$]	15.77	20.43	12.08	9.17

5.4.1.4 Water Balance

The water balance gaps for the real-world experiments are mainly related to discharge differences. Table 5.4 gives an overview of the different measured and modeled water balance components at the study site for 2011, where all model simulations overestimate discharge. The EnKF does not preserve the water balance and can add/extract water in correspondence with the measurements. Hence large differences up to 224 mm can be found for the DA experiments in the assimilation period. Higher balance gaps are particularly found for the real-world experiments with the strong adaptation of the pressure heads within the simulations (REAL_ST, REAL_PAR_5d).

Measured evapotranspiration was not available for 2011. However, given the range of lysimeter observations for 2012 (467-523 mm with some gap filling according Gebler et al., 2015) and a possible reduction of actual ET in 2011 due to water limitation in May-June, the simulated ET (467-488 mm) is within an appropriate range.

Table 5.4: Observed and simulated yearly water balance components: tipping bucket precipitation (P), evapotranspiration (ET) and surface runoff (Q) of the Rollesbroich catchment for different real-world simulation scenarios.

OBSERVATION

Period	P [mm]	ET [mm]	Q [mm]	P-ET-Q [mm]	
2011	953	n/a	392	n/a	n/a

REAL-WORLD EXPERIMENTS

Scenario	Period	P [mm]	ET [mm]			Q [mm]			P-ET-Q [mm]	
		Sum	Sum	Max	Min	Sum	Max	Min	Mean	%
REAL_OL	2011	953	488	513	437	546	1118	165	-81	-8.5
REAL_ST	2011	953	467	483	444	272	447	139	214	22.5
REAL_PAR_1	2011	953	471	486	444	342	558	180	140	14.7
REAL_PAR_5d	2011	953	467	483	445	262	423	137	224	23.5

5.4.2 Synthetic Experiments

5.4.2.1 Soil Water Content

In this section results are discussed for SWC characterization for the synthetic case. The open loop simulations (SYN_OL) show a good performance for all soil layers (NSE: 0.480-0.759) (Table 5.5). After data assimilation with state updating only (SYN_ST), and data assimilation scenarios with joint state-parameter updating (SYN_PAR_1, SYN_PARαn_1, and SYN_PAR_5d) the average SWC in 5 and 20 cm depth is almost perfectly matched (NSE≥0.9) and the RMSE decreased by 0.017-0.030 $cm^3\ cm^{-3}$ (47-88 %) compared to the open loop run depending on scenario and layer. In contrast, at 50 cm depth all synthetic DA simulations show larger deviations from the synthetic truth than their open loop counterparts (NSE -0.772 - -0.341; RMSE 0.014-0.016 $cm^3\ cm^{-3}$). The synthetic simulations with 256 (SYN_PARαn_5d_256) showed similar results to SYN_PARαn_1, which indicates that the performance is not much influenced by ensemble size.

Table 5.5: Performance measures for the reproduction of location averaged SWC for the synthetic experiments. Shown are performance statistics for the open loop run (SYN_OL) and the DA experiments for the assimilation locations.

		SYN_OL	SYN_ST	SYN_PAR_1	SYN_PARαn_1	SYN_PAR_5d	SYN_PARαn_5d (256)
5 cm	**NSE [-]**	0.751	0.978	0.971	0.926	0.973	0.934
	BIAS [$cm^3\ cm^{-3}$]	0.035	0.006	0.008	0.019	0.007	0.017
	RMSE [$cm^3\ cm^{-3}$]	0.036	0.011	0.012	0.019	0.012	0.018
20 cm	**NSE [-]**	0.480	0.936	0.931	0.993	0.924	0.990
	BIAS [$cm^3\ cm^{-3}$]	0.031	0.011	0.010	0.003	0.012	-0.004
	RMSE [$cm^3\ cm^{-3}$]	0.034	0.012	0.012	0.004	0.013	0.005
50 cm	**NSE [-]**	0.759	-0.341	-0.527	-0.772	-0.609	-1.112
	BIAS [$cm^3\ cm^{-3}$]	-0.003	-0.013	-0.014	-0.015	-0.015	-0.017
	RMSE [$cm^3\ cm^{-3}$]	0.006	0.014	0.014	0.016	0.015	0.017

Figure 5.10 shows the temporal evolution of the mean SWC for the open loop simulations and four different data assimilation experiments for the verification locations in 2011. The differences between the individual data assimilation scenarios are rather small. The different DA-runs show larger differences in terms of SWC differences in May as well as in October and November 2011. Another notable difference is the reduced ensemble spread of SYN_PARαn_1 which is indicated by a ca. 75 % smaller standard deviation of mean SWC compared to other synthetic scenarios. SYN_OL shows a good performance (Table 5.6) for the upper soil layers (NSE: 0.571-0.793). In contrast, negative NSE is found at 130 and 190 cm depth (NSE: -6.761- -2.333). After data

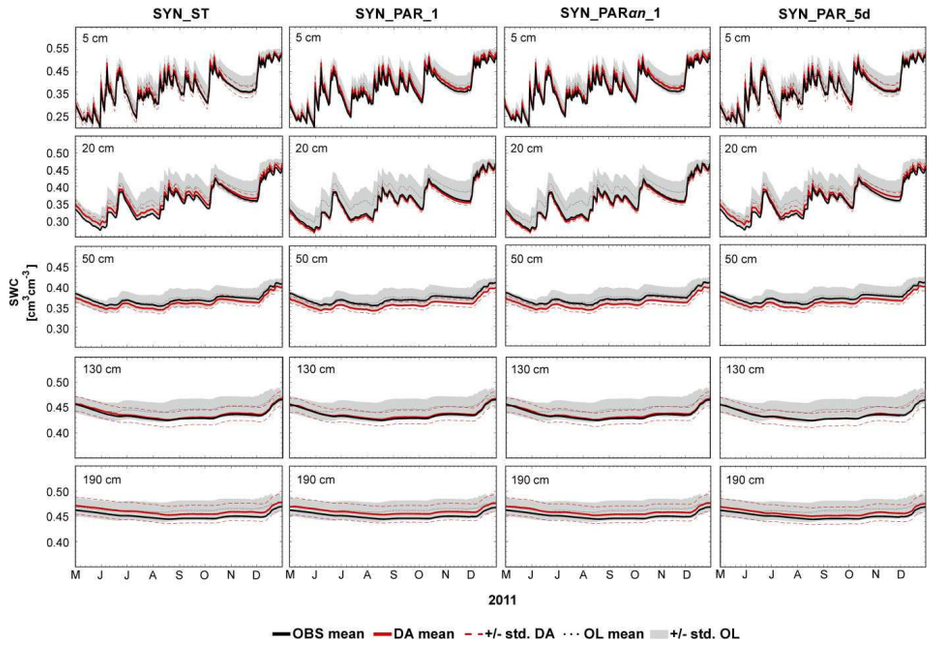

Figure 5.10: Mean SWC and standard deviation of SWC for the open loop run and the synthetic experiments at the verification locations. Mean synthetic reference observations are indicated as well.

assimilation, performance improvement is found for the mean SWC at 5, 20, 130 and 190 cm depth indicated by RMSE decrease by 0.002-0.020 cm^3 cm^{-3} (21-63 %). Particularly the mean SWC in 5, 20, and 130 cm shows a very good fit (NSE 0.885-0.944) after DA. For the bedrock layers the RMSE decrease ranges between 31-84 %. In contrast, at 50 cm depth all assimilation scenarios show larger deviations from the open loop. For this layer there is a decrease of performance for mean SWC at the verification locations (NSE: -1.363 - -0.159; BIAS: -0.009-0.013 cm^3 cm^{-3}).

Hereinafter, the RMSE at the individual assimilation locations is discussed. The open loop simulation for the synthetic case has smaller layer-averaged RMSE over the observation locations (0.044 – 0.056 cm^3 cm^{-3}) than the open loop simulation for the real-world case (RMSE: 0.0825 – 0.094 cm^3 cm^3). These numbers suggest that 40% of layer-averaged RMSE over the observation locations for the real-world case could be related to structural model errors. Structural model errors can for example be systematic errors in model forcings or values for soil hydraulic parameters, or processes which are not captured by the simulation model or not considered in the model set-up.

Table 5.6: Performance measures for the reproduction of SWC for the synthetic experiments. Shown are performance statistics for the open loop run (SYN_OL) and the DA experiments for the verification locations.

		SYN_OL	SYN_ST	SYN_PAR_1	SYN_PARαn_1	SYN_PAR_5d
5 cm	**NSE [-]**	0.793	0.914	0.885	0.887	0.913
	BIAS [cm³ cm⁻³]	0.028	0.012	0.008	0.008	< 0.001
	RMSE [cm³ cm⁻³]	0.030	0.020	0.023	0.022	0.020
20 cm	**NSE [-]**	0.571	0.892	0.944	0.944	0.900
	BIAS [cm³ cm⁻³]	0.024	0.007	-0.006	0.007	0.006
	RMSE [cm³ cm⁻³]	0.027	0.014	0.010	0.010	0.013
50 cm	**NSE [-]**	0.602	-0.159	-1.363	-1.098	-0.450
	BIAS [cm³ cm⁻³]	0.003	-0.009	-0.013	-0.013	-0.010
	RMSE [cm³ cm⁻³]	0.006	0.009	0.013	0.013	0.010
130 cm	**NSE [-]**	-2.813	0.897	0.927	0.911	0.984
	BIAS [cm³ cm⁻³]	0.013	0.002	0.002	0.002	< 0.001
	RMSE [cm³ cm⁻³]	0.015	0.002	0.002	0.002	< 0.001
190 cm	**NSE [-]**	-6.761	-2.276	-2.403	-1.873	-0.854
	BIAS [cm³ cm⁻³]	0.013	0.009	0.009	0.008	0.006
	RMSE [cm³ cm⁻³]	0.013	0.009	0.009	0.008	0.006

Figure 5.11 shows the RMSE at the different subsurface layers at the individual verification locations. For the open loop simulation the layer-averaged RMSE (0.0311-0.0492 cm^3 cm^{-3}) at the verification locations is 12-40 % smaller than the RMSE at the assimilation locations. The mean RMSE over all verification locations decreases between 40 % (SYN_ST) and 55 % (SYN_PARαn_1). For the individual layers, the mean RMSE averaged over the locations decreases between 20 and 66 % compared to the open loop run depending on scenario and layer. The scenarios with parameter estimation shows in almost all layers a better performance than SYN_ST with state update only. SYN_PARαn_1 performs best in terms of layer-averaged RMSE for the individual locations over all five soil layers showing the most significant performance increase. The RMSE reduction at 130 and 190 cm depth is smaller than in the upper soil layers. Particularly the RMSE reduction at 190 cm is relatively small (20-34 %). These layers also show more locations (58-63 %) with decreasing performance. In contrast, fewer locations with decreasing performance are found at the upper soil layers particularly for SYN_PARαn_1 (0.0-1.5 %). In summary, all DA-scenarios give very similar results in terms of average SWC in the assimilation period.

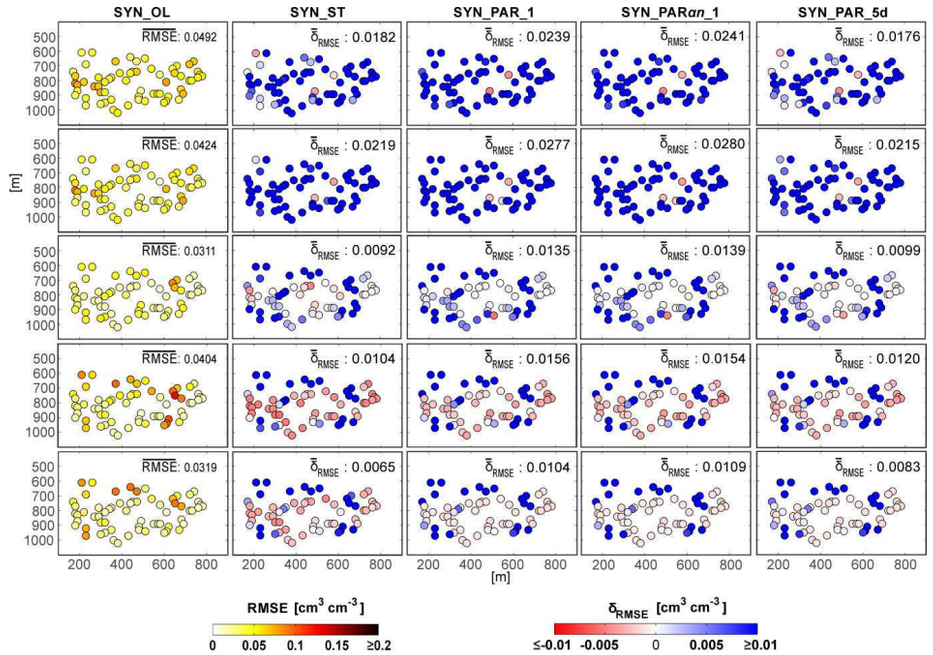

Figure 5.11: RMSE of SWC at individual locations for the open loop runs (left column) and changes in RMSE (increase implies improvement and decrease implies impoverishment) for four data assimilation scenarios (four columns on the right) with synthetic observations at 5, 20, 50, 130, and 190 cm depth.

5.4.2.2 Parameter Estimation

Figure 5.12 illustrates the temporal evolution of the hydraulic conductivity ($\log_{10}K_s$) estimates for the synthetic experiments. In the synthetic scenario, the bedrock layer $\log_{10}K_s$ is reproduced well, despite relative large differences between initial average $\log_{10}K_s$ and the synthetic truth. Results for the other soil horizons for the synthetic case vary among the different data assimilation experiments. While SYN_PAR_1 overestimates the hydraulic conductivity for horizon 2 (20 cm) and horizon 3 (50 cm), SYN_PAR_5d and SYN_PAR*an*_1 overestimate the conductivity for horizon 1 (5 cm). However, in most of these cases the deviations are very small (≤ 0.05 $\log_{10}Ks$). Only SYN_PAR*an*_1 horizon 3 shows larger deviations (ca. 0.10 $\log_{10}Ks$). The parameter estimation for this scenario is also affected by temporal instability for horizon 1 and 4 during the first month of the estimation, due to a too small ensemble spread. The parameter updates (Figure 5.13) for the synthetic scenarios are very similar to each other. The parameter updates for SYN_PAR_1 and SYN_PAR*an*_1 are only slightly stronger than for SYN_PAR_5d. Similar to the real-world scenarios, the synthetic experiments show high contrasts in updates between the sensor locations and also larger updates for the bedrock layer.

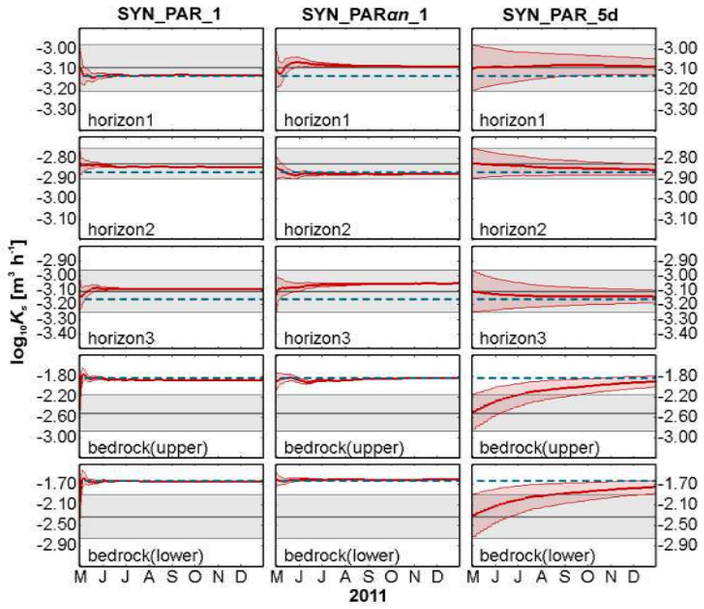

Figure 5.12: Ensemble mean soil hydraulic conductivity and standard deviation for synthetic experiments (red) against the open loop parameter set (grey) for different subsurface horizons of the Rollesbroich catchment. The dashed line in blue indicates the synthetic truth. Notice that the used scales on the vertical axis differ for the different horizons.

Figure 5.14 shows the spatial distribution of the differences between the ensemble hydraulic conductivity estimated for different data assimilation scenarios and the reference hydraulic conductivity. This gives further insight in the change of spatial patterns of the hydraulic conductivity during the assimilation period. For the scenarios SYN_PAR_1, SYN_PARαn_1, and SYN_PAR_5d in general a strong reduction of the RMSE is found. A large improvement is found at the sensor locations of horizons 1-3, but also within the bedrock layers where no measurements were available. The spatial distribution of RMSE reduction differs between SYN_PAR_1 and SYN_PAR_5d. Horizon 1 of SYN_PAR_1 is on average closer to the reference than its SYN_PAR_5d counterpart, but shows some locations with relatively high RMSE (0.4-0.5 $\log_{10}K_s$) in the center and the west of the study site. This is similar for horizon 3. In contrast, the RMSE reduction is smoother and less spiky for the SYN_PAR_5d scenario also in comparison with SYN_PARαn_1, which shows many similarities with the spatial patterns of SYN_PAR_1. The spatial RMSE distribution for the bedrock layers also differs among the simulation scenarios. Local maxima of RMSE are found mainly in the north and center of the study site. From these findings it can be concluded that the different spatial patterns of the estimated hydraulic conductivity are strongly associated with the update strategy resp. the model scenarios.

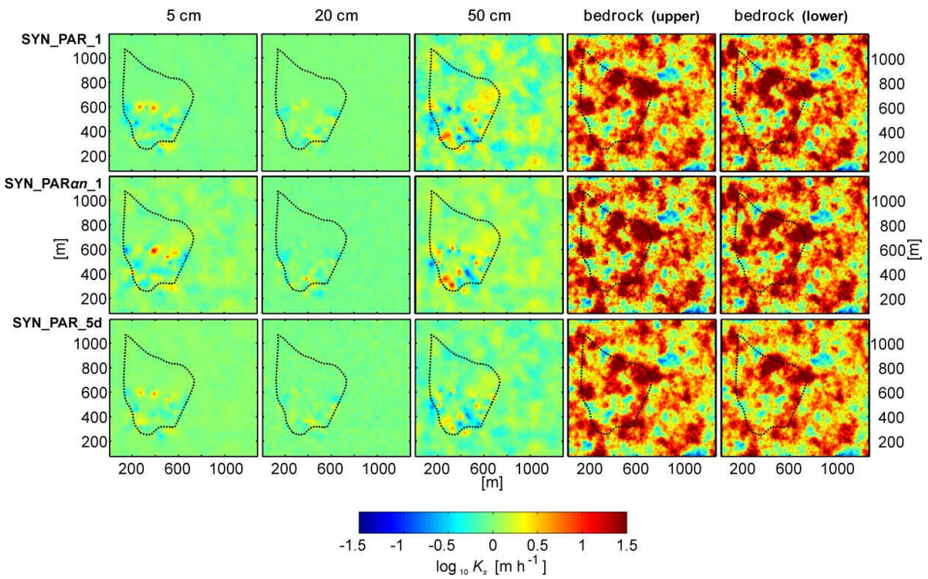

Figure 5.13: The ensemble averaged perturbation applied to the hydraulic conductivity field ($\log_{10}K_S$) during data assimilation for different synthetic simulation scenarios: Indicated are results for the soil layers at 5, 20, and 50 cm depth as well as the upper and lower bedrock layer. The dashed line marks the test site borders within the model domain.

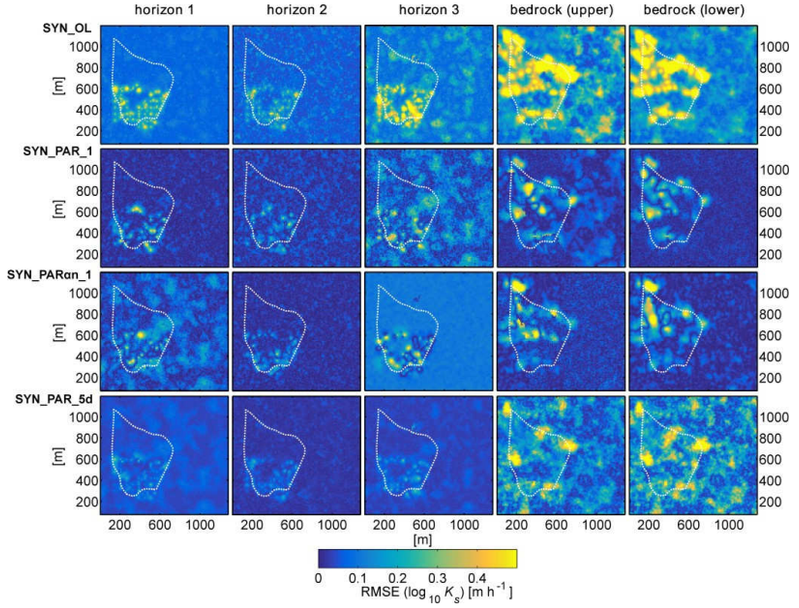

Figure 5.14: RMSE of the ensemble average hydraulic conductivity compared to the reference hydraulic conductivity ($\log_{10}K_S$) for the open loop and DA experiments and for different soil horizons. The study site is indicated by a white dashed line.

5.4.2.3 Discharge

In terms of ensemble spread and discharge fluctuation the results of the synthetic experiments show similarity to the results of their real-world counterparts. SYN_ST shows strong discharge fluctuations until the high flow period associated with a high ensemble spread. In some realizations the Kieselbach fell dry during assimilation, in particular during the period May-June 2011. This leads to an underestimation of discharge which is also indicated by a negative relative bias (-22.94 %) (Table 5.7). The NSE (0.44) is higher and the RMSE smaller (8.88 $m^3 h^{-1}$) compared to the open loop simulations. This is probably related to the strong reduction of discharge peaks for which these measures are most sensitive. In contrast, SYN_PAR_1 shows limited discharge fluctuations and low ensemble spread. The performance for SYN_PAR_1 is much better than for the open loop run or state updating only (NSE: 0.61; PBIAS: 12.16 %; RMSE: 7.34 $m^3 h^{-1}$), indicating a good adaptation of the simulations to the synthetic reference. The discharge SYN_PAR*αn*_1 has similar characteristics, whereas the discharge of SYN_PAR_5d can be characterized as intermediate between SYN_PAR_1 and SYN_ST. With an increasing number of parameter updates during assimilation the discharge fluctuations and ensemble spread decrease, with a better adjustment of the discharge to the synthetic reference.

Table 5.7: Discharge model performance at the catchment outlet for different update scenarios with synthetic observations for 2011. The model performance is indicated by mean NSE, standard deviation of NSE, average bias and root mean square error (RMSE).

	SYN_OL	SYN_ST	SYN_PAR_1	SYN_PAR*αn*_1	SYN_PAR_5d
NSE [-]	-0.04	0.44	0.61	0.46	0.59
PBIAS %	78.42	-22.94	12.16	12.38	-5.92
RMSE [$m^3 h^{-1}$]	14.03	8.88	7.34	8.72	7.60

5.4.2.4 Evaporation and Water Balance

In the synthetic experiments the reproduction of ET is already almost perfect in the open loop run with very small bias and RMSE and very high NSE (0.99). The data assimilation runs with exception of SYN_PAR*α*n_1 give an additional small further improvement.

Similar to the real-world case the water balance gaps for the synthetic experiments are related to discharge differences (Table 5.8). The discharge of the open loop simulations is higher than the reference truth in the assimilation period (+24 %). This gap is reduced by all DA-scenarios. All DA-scenarios are closer to the reference truth but show now a negative bias (-17 - -9 %). This is 62-68 % of the total open loop discharge which indicates that the EnKF extracts a significant amount of water in the assimilation period through the adaptation of the pressure heads.

Table 5.8: Observed and simulated yearly water balance components: precipitation (P), evapotranspiration (ET) and surface runoff (Q) of the Rollesbroich catchment for different synthetic simulation scenarios.

Scenario	P [mm]	ET [mm]			Q [mm]			P-ET-Q [mm]	
	Sum	Sum	Max	Min	Sum	Max	Min	Mean	%
SYN_OL	953	488	513	437	546	1118	165	-81	-8.5
Reference Truth	953	487	n/a	n/a	406	n/a	n/a	60	6.3
SYN_ST	953	483	496	462	336	625	160	134	14.1
SYN_PAR_1	953	484	496	463	371	594	205	98	10.3
SYN_PAR*an*_1	953	484	492	470	371	587	206	113	11.9
SYN_PAR_5d	953	484	496	462	353	599	170	116	12.2

5.5 Discussion

Our findings indicate that the EnKF in combination with fully coupled land surface-subsurface models and SWC data from a dense observation network is able to improve the characterization of the hydrology at the hillslope scale. For the real-world case joint updating of model states and hydraulic conductivity was more effective in updating the SWC than state update alone. Compared to open loop simulations the RMSE averaged over all SWC sensor locations in 5, 20 and 50 cm depth is reduced by 14 % for the scenarios with joint state-parameter updating and damping (REAL_PAR_5d), but only 6 % for state updating alone. A larger impact for parameter updating could have been expected.

In contrast to their real-world counterparts, the synthetic scenarios exhibit a more significant improvement of SWC after data assimilation. An excellent performance with low bias particularly is found at 5, 20, and 130 cm depth (NSE: >0.9; BIAS: 0.001-0.019 cm^3 cm^{-3}). Compared to open loop simulations the RMSE averaged over all SWC verification locations (including all layers) is reduced between 40 % (SYN_ST) and 55 % (SYN_PARαn_1). These results can be called satisfying, especially because the evaluations in these synthetic runs are made now at verification locations not used in the assimilation.

The fact that for the synthetic case simulation results are better than for real-world case (with 40% lower layer-averaged RMSE of SWC for the open loop run) indicates that structural model errors play a role for the SWC estimation. Preferential flow through macropores, which was already observed at the nearby Wüstebach site (Wiekenkamp et al., 2016) and other studies investigating in-situ SWC observations (e.g. Martini et al., 2016; Polteradnev et al., 2015) is potentially under-represented in the model simulations. This might explain the high bias at 50 cm depth for the real-world simulations having a large influence on the infiltration of a grassland with an extensive rooted top soil layer (Weiler and Naef, 2003). This issue was already indicated by Gebler et al. (2017) for a slightly different ParFlow-CLM setup for the Rollesbroich study site. Furthermore, the representation of the sapprolite and the drainage system in the subsurface might be suboptimal in our simulation model and could be better represented with a dual porosity approach (e.g. Frey et al., 2012; Frey et al., 2016). Moreover, an unknown additional storage in the bedrock fractures potentially has influence on the baseflow of the Kieselbach (Hale et al., 2016). The representation of the drainage system in the model might be a cause for the model structural error, although no direct evidence in the spatial error distribution of SWC was found for this.

The finding that data assimilation has a relatively small impact on the SWC-characterization for the real-world case, and also much smaller than for the synthetic case, points to the fact that data assimilation in combination with integrated physically based terrestrial systems models can strongly improve SWC- and parameter characterization, also at the hillslope scale, but that for real-world cases additional complications arise. The structural model errors which we referred to before

might also have inhibited a higher efficiency of the data assimilation as parameter updates might have compensated for model structural errors. This points to the fact that, maybe especially for physically based models, the representation of all processes which affect the soil water distribution is very critical. The representation of the drainage network or the role of the upper weathered bedrock layer are potential associated with significant model structural errors.

One further potential explanation is the fact that all soil hydraulic parameters are uncertain, but only hydraulic conductivity is estimated. For a relatively simple 1D subsurface flow model Li and Ren (2011) already demonstrated that the update of multiple van Genuchten parameters (i.e., hydraulic conductivity, α and n) is more effective than updating hydraulic conductivity alone. This issue has been analyzed by the SYN_PARαn_1 and SYN_PARαn_5_256 scenarios with perfect knowledge of van Genuchten α and n. Although overall the SYN_PARαn scenarios give smaller errors, the improvement with respect to their counterparts with uncertain α and n is small. This indicates that model simulations could be further improved by an estimation of these parameters, but that their impact at least in this study is relatively small.

Other limitations of the data assimilation experiments are potentially associated with ensemble size and filter inbreeding. Whereas the larger ensemble size of 256 gives slightly better results than the ensemble size of 128, the overall effect of the increased ensemble size is small. Nevertheless, it is found that in the data assimilation experiments, especially in case of daily updating, the ensemble variance decreases too fast and is too low. As a consequence, too small weights are assigned to the measurements which have too little influence to update model predictions and model parameters (e.g. Houtekamer and Mitchel, 1998; Evensen, 2004; Zhang et al., 2007; Hendricks Franssen and Kinzelbach, 2008). This might explain the erroneous estimation of soil hydraulic conductivity for the SYN_PARαn_1 scenario for the horizons 1 and 3 as well as the instabilities during parameter estimation for this scenario. The comparison of the estimated spatial hydraulic conductivity distribution (SYN_PAR_1 and SYN_PAR_5d) with the synthetic truth indicates that longer update intervals and damping led to a spatially more smooth estimation of hydraulic conductivity and a slower decrease of ensemble spread. Rasmussen et al. (2015) stressed that large ensemble sizes in combination with localization could further improve the performance of data assimilation in combination with integrated terrestrial models. The application of localization reduces the influence of physical not meaningful spurious correlations (Houtekamer and Mitchel, 2001; Anderson et al., 2007). It could therefore be that a more sophisticated data assimilation strategy including localization could improve the results, but this option was not yet available and implemented in our TerrSysMP-PDAF framework.

Discharge was in general not well reproduced in the experiments. This shows that characterizing SWC is not the main factor for improving discharge estimation. Nevertheless, for the best performing scenarios in the synthetic case the NSE improved from -0.04 to around 0.61, which is

still a considerable improvement related to assimilation of SWC-data and parameter estimation. For this small hillslope scale where precipitation uncertainty does not play such a major role, maybe still a higher impact of SWC-assimilation could have been expected, especially for the real-world case where the improvement of discharge estimation related to data assimilation and parameter estimation was smaller. Since subsurface SWC and overland flow in ParFlow are directly linked via pressure head for the individual grid cells, the update of the SWC and therefore pressure head has a large impact on the discharge simulation. The extraction of water by subsurface state updates reduced discharge performance particularly during dry periods. Also the representation of the Kieselbach in the model is critical (Gebler et al., 2017). Given the small spatial extent of the channel (0.3-1.0 m) the interaction between surface water and groundwater is highly sensitive to the vertical pressure gradient which controls the re-infiltration of groundwater into the channel. Although the channel bed of the Kieselbach was carefully derived from a 1 × 1 m grid small deviations might lead to a model bias. The assimilation of discharge data, combined with the updating of Manning's roughness coefficient as well as other van Genuchten soil hydraulic properties could further improve discharge estimation. Baatz et al. (2017) demonstrated the benefit of estimating Manning's roughness coefficients on the discharge simulation of a 2D regional scale synthetic ParFlow model.

The assimilation of SWC data had no significant effect on the reproduction of the annual actual ET in the synthetic simulations. Information on SWC and/or hydraulic conductivity does not contribute in those cases to improve ET-characterization. This is also illustrated by the limited ensemble spread of the ET-predictions for the open loop simulations. Simulation results are therefore close to the expected ones. However, more impact of SWC-assimilation for drier conditions with limited water supply can be expected.

In summary, the potential of assimilating SWC-observations at the hillslope scale to improve the characterization of spatially distributed SWC and discharge has been demonstrated for the synthetic case, which mimicked the real-world case with all its complexities including 3D fully distributed fields of saturated hydraulic conductivity and Mualem-van Genuchten parameters α and n. The RMSE of SWC at verification locations decreased up to 55% for the data assimilation scenarios including parameter estimation and the NSE for discharge estimation was ca. 0.61 coming from -0.04 for the open loop. However, for the equivalent real-world case the improvements induced by data assimilation were much smaller. This can be attributed mainly to model structural errors, which were already subject of investigation in the past, but could not be clearly identified (Gebler et al., 2017). In addition, also the fact that only saturated hydraulic conductivity was estimated, and some setting in the data assimilation might have given some contribution here.

5.6 Conclusions

This study investigated the assimilation of high resolution SWC data with the Ensemble Kalman Filter (EnKF) in the fully coupled land surface-subsurface model TerrSysMP. The assimilation experiments were performed for the small headwater grassland catchment Rollesbroich in the Eifel (Germany), or alternatively, for a synthetic reality which mimics this catchment. Data assimilation experiments were made with an ensemble of 128 (256) model realizations at 10×10 m lateral resolution and a variable vertical resolution (0.025-0.575 m) which resulted in a problem size of 0.3 million unknowns. The individual realizations were setup with a fully heterogeneous subsurface with geostatistical simulations of van Genuchten soil hydraulic properties. The following data assimilation experiments were performed: (i) daily state updating; (ii) daily joint state-parameter updating; (iii) daily state updating combined with parameter updating each 5 days and a damping factor. In addition, for the synthetic test case also experiments were performed where Mualem-van Genuchten parameters α and n were deterministic and only saturated hydraulic conductivity uncertain. In all experiments SWC data from 61 sensor network locations and three depths (5, 20, and 50 cm) were assimilated.

The results for the real-world case showed that the EnKF in combination with joint updating of model states and hydraulic conductivity is more efficient in updating SWC than state updating alone. For joint state-parameter updating the average RMSE for the overall SWC sensor locations decreased by 14 % for the real-world scenario with joint state-parameter updating and damping while state updating reduced the RMSE only by 6 %. The improvement of the SWC characterization for the real-world case was limited. The uppermost layer in 5 cm depth exhibited performance impoverishment indicated by strong RMSE-increases. The characterization of discharge was only marginally improved although some discharge improvement was found if both states and parameters were updated with DA. In particular, the systematic bias up to 65 % between real-world model simulations and measurement data was not significantly improved. This is probably related to an underrepresentation of preferential flow, too little drainage by either the fissured bedrock or the on-site drainage system in the model.

Considerable improvement was found for the synthetic case, which showed on average 40-55 % RMSE reduction for the verification locations and for different data assimilation scenarios. Also the discharge estimation was improved now with a NSE of -0.04 for the open loop and a NSE of ca. 0.61 for the data assimilation scenarios which involved parameter estimation. It was found that estimated saturated hydraulic conductivity was much closer to the reference values (after joint state-parameter updating).

In order to explain the very different results for the real-world case and the synthetic case it is important to point to the very different RMSE for the open loop runs for the real-world case and the synthetic case. The RMSE for the open loop run, calculated over the 61 measurement locations and

upper three layers, was 53% lower for the synthetic case than for the real-world case. This can be related to model structural errors, and these same model structural errors might have inhibited the positive impact of data assimilation as updated parameters just might have compensated for model structural errors. It was also investigated whether results could be better if besides saturated hydraulic conductivity also Mualem-van Genuchten parameters α and n were estimated. This was tested for the synthetic case by assuming that α and n were perfectly known and only saturated hydraulic conductivity was uncertain. Results only improved very slightly compared to the scenarios where α and n were unknown and point to the fact that for this case the uncertain α and n were not a main additional limitation to improve SWC characterization by data assimilation.

In summary, the potential of assimilating SWC-observations at the hillslope scale to improve the characterization of spatially distributed SWC and discharge has been demonstrated for the synthetic case, which mimicked the real-world case with all its complexities including 3D fully distributed fields of saturated hydraulic conductivity and Mualem-van Genuchten parameters α and n. The fact that for the real-world case the improvements induced by data assimilation were much smaller can be attributed mainly to model structural errors, but also other data assimilation setups will have played some role. These results might be site-specific, and a similar approach could give better results at another hillslope site. However, it might also point to the importance of small scale processes which are difficult to capture, even with a physically based integrated terrestrial systems model. Process-based integrated models in combination with DA can be a valuable tool for identifying structural model errors. It can help to detect the importance of including additional natural processes (e.g. macropore flow) as well as anthropogenic structures (e.g. drainage system) in the modelling process to further improve the integrated terrestrial modelling approach.

Chapter 6 Summary and Outlook

Understanding the soil-vegetation-atmosphere continuum is essential to improve hydrological model predictions. In particular the spatio-temporal variability of the soil water content (SWC) and its feedbacks on other terrestrial compartments are of high interest in many geoscientific fields, since these patterns influence for example the rainfall-runoff response and the partitioning of the net radiation into latent and sensible heat fluxes while interacting with the vegetation cover. The characterization and prediction of the highly variable spatial and temporal patterns of this key variable and its controlling factors are therefore of great importance. Within this context, this PhD thesis explores the degree of model complexity that is necessary to adequately represent heterogeneous subsurface processes, as well as the benefit of merging soil moisture data with a fully-coupled land surface-subsurface terrestrial model. This includes also an uncertainty analysis of model forcing (i.e. precipitation) and evaluation data (actual evapotranspiration). For this purpose, the fully-coupled land surface-subsurface terrestrial model ParFlow-CLM (part of TerrSysMP framework) was applied to a 38 ha grassland headwater catchment located in the Eifel (Germany). ParFlow-CLM takes lateral subsurface flow, overland flow, as well as topography into account. Detailed long-term data for model setup, calibration, and evaluation for this approach were provided by the TERENO infrastructure initiative (Zacharias et al., 2011), the North Rhine-Westphalian State Environment Agency (LUA NRW), and the Transregional Collaborative Research Center 32 (TR32). This combination of process orientated model and extensive observation data was expected to contribute to a better understanding of the complex processes of the energy and water cycle at the hillslope, the elementary unit for the runoff generation process.

In order to investigate the uncertainty of the model input and evaluation data, the accuracy of precipitation and actual evapotranspiration measurements was evaluated. The assessment of the accuracy of both major water balance components is important. In particular deviations of precipitation potentially lead to a high model bias in a non-linear terrestrial model system. With the help of the extensive measurements of the TERENO and the TR32 infrastructure both, input and evaluation data can be cross-checked to detect and correct probable errors in these data. Therefore, for a one year period (2012) the estimates of precipitation and evapotranspiration by a set of six weighable lysimeters were compared with precipitation data from a standard tipping bucket and evapotranspiration observations from an eddy covariance station installed at the study site. The simultaneous estimation of precipitation and evapotranspiration with lysimeters is a relatively new approach since the technical prerequisites (i.e. high temporal resolution and weighing precision) as well as adequate noise reduction algorithms have been developed only very recently. It is assumed that lysimeters, which are well embedded into the (vegetated) surroundings, have the potential for very precise estimates of the water balance. Eddy covariance data were corrected for the energy

balance deficit using a Bowen ratio method according to Kessomkiat et al. (2013). The energy balance deficit was ca. 16% of the annual net radiation. The comparison of lysimeter and eddy covariance data revealed a good correspondence in terms of evapotranspiration, in particular during summer. This indicates that the correction method (i.e. adding the energy balance gap to the turbulent fluxes according to the evaporative fraction) was adequate for the Rollesbroich grassland. Both devices also fit well to the reference evapotranspiration calculated with the full-form Penman-Monteith equation, indicating that the evapotranspiration of 2012 was energy limited and not water limited. Major differences in evapotranspiration between all three methods were related to differences in field maintenance (i.e. grass harvest). Our study revealed a relationship between the differences in grass length (between the lysimeter and the EC-field) and the differences in evapotranspiration between lysimeter, EC and the calculated Penman-Monteith evapotranspiration. Different harvesting dates for the lysimeter site and the field behind the EC station led to periods with large differences in grass lengths which induced evapotranspiration differences up to 2.1 mm d^{-1}. The comparison of precipitation measured by the lysimeters and a classical tipping bucket allowed a deeper insight into the seasonality of the precipitation measurement errors and the role of different weather conditions on the differences. In general, the lysimeter showed a 16 % (145 mm) higher precipitation than the tipping bucket for 2012. This difference was partly related to a small but significant contribution of dew and rime (17 % of the difference) as well as to fog and drizzling rain (5.5 %). With the help of a surveillance system, 8 % of the difference could be assigned to light and moderate snowfall, under those conditions the precipitation differences between the lysimeter and tipping bucket were larger. However, the camera system also revealed that in case the lysimeters were fully covered with snow, the precipitation estimates by lysimeters were unreliable. This situation was associated with 36 % of the precipitation difference, also in this case with more precipitation for the lysimeter than for the tipping bucket. Nevertheless, in summary it was concluded that lysimeters were able to increase the precision of water balance estimates including contributions from dew, fog and rime. The results further indicate that common correction methods (e.g. Richter, 1995) can be used to correct the precipitation estimates in case expensive lysimeter measurements are not available. The precision of lysimeter precipitation and evapotranspiration estimates can also still be further improved in the future. Specifically, winter weather conditions with a high accumulated snow cover and frozen soil conditions cause ponding of surface water at the lysimeter surface and are still associated with high uncertainty and remain a challenge. More accurate precipitation estimates under these conditions require intensive and frequent maintenance (e.g. frequent manual separation of lysimeter snow layer and surroundings) and additional cross-comparison with snow hydrological measurements (e.g. snow pillows, snow tube sampling). These insights subsequently can be used for the calibration and evaluation of novel snow precipitation measurement techniques, for example neutron counters (Desilets et al., 2010, Schattan et al., 2017), which might be an alternative to classical approaches. Another open question

associated with the accuracy of weighable lysimeters is the representativeness of the thermal regimes between the lysimeter and its direct surroundings. Deviations due to differences in the thermal regime potentially have a strong impact on the accuracy of the measured evapotranspiration. They depend among others on individual design, enclosure material (e.g. metal, plastic), lysimeter setup in the field as well as vegetation cover and are hence very specific for the individual lysimeters and study sites. This should be investigated in more detail also by controlled laboratory experiments.

The second study of this thesis was dedicated to the degree of model complexity that is necessary to adequately represent the heterogeneous processes at the hydrological hillslope scale. This insight into small scale process heterogeneity becomes important, since hydrological modeling more and more evolves towards high-resolution process orientated models and should allow a better understanding of the runoff generation at a hillslope. However, fully coupled subsurface-surface models have rarely been applied for complex subsurface structures and fully heterogeneous fields of various soil hydraulic parameters. Therefore, we performed 32 stochastic ParFlow-CLM simulations for four different complexity levels of soil hydraulic properties (homogeneous, homogeneous layers, homogeneous layers differing between different soil units, and fully heterogeneous subsurface generated by stochastic simulations). Additionally, soil hydraulic parameters were derived with two different approaches. In one approach, soil hydraulic properties were estimated with a pedotransfer function from measured soil texture (ROS). In the second approach, soil hydraulic properties were inversely estimated in 1D simulations (SCE) at each measurement location. The simulations for the eight different scenarios (four complexity levels and both ROS and SCE) were subsequently evaluated with a dense network of soil moisture sensors, as well as evapotranspiration and discharge data. The results showed that topography and the heterogeneous distribution of soil hydraulic parameters are the main controlling factors of the variability of soil water content (SWC) at the study site. Topography in combination with homogeneous soil hydraulic properties could only explain a small proportion (25-40 %) of the observed spatial SWC variability. On the contrary, the spatial SWC variability could be mainly explained with a heterogeneous distribution of soil hydraulic parameters. This indicates that the distribution of soil hydraulic properties was more important to explain the overall SWC variability than topography at the study site. Results also differed between the ROS and SCE-scenarios. In particular for the intermediate complexity levels with homogeneous layers and homogeneous soil units the SCE and ROS scenarios showed significant differences. Major differences between both setups were also found in the temporal SWC dynamics. The SCE-setup outperformed the ROS-setup in particular during dry periods. This is mainly related to the difference in the inverse air entry pressure (α) for both parameter setups which had a strong impact on the drying and wetting behavior of the subsurface. This illustrates the uncertainty related to the estimation of soil hydraulic parameters with a pedotransfer function and also the difficulty to upscale local values of soil

hydraulic properties to the larger scale. The 1D soil hydraulic properties of the SCE-setup by contrast showed an acceptable performance although lateral flow processes were not considered. This indicates that lateral flow processes had minor impact on the SWC specifically for the upper soil layers. However, this finding is site specific and might vary with the influence of lateral flow processes (e.g. depending on the topographic gradient). The inverse air entry pressure also affected the annual evapotranspiration estimate of ROS and SCE. With a lower α, the SWC for the ROS-setup was consistently higher than SWC for the SCE-setup in particular during dry periods. Thus, also evapotranspiration for the ROS-setup was 15- 25 mm/year higher than for the SCE-setup. Despite the good seasonal performance found for the simulated evapotranspiration of both setups, short-term ET dynamics were not represented well. This also affected the simulated discharge, which in general showed a poor model performance. Whereas peak discharge events in the winter months were overestimated, low flow in summer was underestimated by the model simulations. This probably can be explained by a lack of information on bedrock characteristics and its interaction with the on-site drainage system. Furthermore, discharge is very sensitive to preferential flow and soil hydraulic model parameters, which were uncalibrated in this study. Other parameter configurations potentially better represent infiltration and preferential flow processes than the uncalibrated model stochastic realizations.

The third study explored the potential of data assimilation and in particular the EnKF to improve the characterization of SWC as well as discharge and evapotranspiration with information from a dense network of soil moisture observations in combination with a high-resolution Parflow-CLM model (10 m lateral resolution, 0.3 million unknowns) at the hillslope scale. This is the first study which investigates the benefit of conditioning a fully-coupled surface-subsurface model using 3D heterogeneous soil hydraulic properties with real-world long-term soil moisture network data. Specifically it was investigated, how different scenarios of updating model states and parameters (i.e. hydraulic conductivity) affect the transient model simulations and whether an approach which is commonly followed in 2D or 3D-studies, namely solely updating the hydraulic conductivity, is sufficient to improve model predictions which include also other spatial heterogeneous and uncertain soil hydraulic parameters (e.g. inverse air entry). An ensemble of 128 (256) stochastic model realizations was generated with soil hydraulic parameters and precipitation differing between the individual model runs. The basic model setup of the upper soil horizons was kept similar to the SCE setup of the second study with exception of the range of soil hydraulic properties that was changed in particular for the bedrock horizons to explore a larger parameter space. Subsequently, a larger range of soil hydraulic conductivity was assumed for the parameter samples in order to compensate for the fact that hydraulic conductivity could not be laterally increased during data assimilation. Different variants of model state updates with or without parameter updates were tested in the experiments including also different update intervals for states and parameters as well as damping of the parameter update. Model simulations were not only

carried out for the Rollesbroich site but also for a synthetic case which mimicks the Rollesbroich site. The comparison of results for the real case and the synthetic case give more insight into the role of model conceptual errors which might be indicated by systematic worse performance of the real-world experiments. The results for the real-world case showed that the EnKF in combination with joint updating of model states and hydraulic conductivity is more efficient in updating SWC than state updating alone. For joint state-parameter updating the average RMSE for the overall SWC sensor locations decreased by 14 % for the real-world scenario with joint state-parameter updating and damping. while state updating reduced the RMSE only by 6 %. Particularly, the lower soil layers in 20 and 50 cm depth improved in terms of layer-averaged RMSE (8-28 % reduction). However the improvement of the real-world case was limited. In particular, the systematic bias up to 65 % between real-world model simulations and measurement data was not significantly improved. More improvement was found for the synthetic case, which showed on average 40-55 % RMSE reduction for the verification locations in the individual scenarios. After the joint state-parameter update the saturated hydraulic conductivity was much closer to the synthetic reference value. Also the discharge estimation was improved now with a NSE of -0.04 for the open loop and a NSE of ca. 0.61 for the data assimilation scenarios which involved parameter estimation. This demonstrates the potential of assimilating SWC-observations at the hillslope scale to improve the characterization of spatially distributed SWC and discharge for the synthetic case, which mimicked the real-world case with all its complexities including 3D fully distributed fields of saturated hydraulic conductivity and Mualem-van Genuchten parameters α and n. The fact that for the real-world case the improvements induced by data assimilation were much smaller can be attributed to model structural errors which might have inhibited the positive impact of data assimilation. Further studies should explore whether the role of model structural errors was related to site specific characteristics (i.e. on-site drainage system), or whether small scale processes which is not captured by the modelling approach (i.e. preferential flow, drainage by the fissured bedrock) play a more dominant role.

From these studies it can be concluded that high-resolution fully-coupled subsurface-surface models at the hillslope scale are a valuable tool to better understand the basic hydrological processes. These models better represent the rainfall-runoff response and will improve the model predictions of the hydrological cycle on the long term. However, on the way to a predictive application of these high or “hyper-resolution” (Wood et al., 2011) models with a lateral resolution equal to or beyond 100 × 100 m still many challenges remain. An important limitation that comes along with parameter estimation of fully integrated terrestrial models at the hillslope scale is indicated by the third study of this thesis which has shown that the adjustment of one specific parameter (i.e. hydraulic conductivity) might not be enough for a significant improvement of the model predictions. This is contradictory to many previous studies with simplified or synthetic model setups for the vadose zone that showed a good model performance if only hydraulic

conductivity was constrained. However, for more complex model setups, i.e. models setups with heterogeneous fields of soil hydraulic properties, the parameter space has to be explored including additional soil hydraulic properties. On this account, future DA studies with complex models should also aim at estimating simultaneous multiple soil hydraulic parameters. It can be expected that complex terrestrial model systems will benefit from the inclusion of multiple observation types (e.g. soil moisture, discharge, evapotranspiration) to be able to estimate the various spatially heterogeneous fields of soil hydraulic parameters. Further improvements can be also achieved by increasing the ensemble size, including localization and tune the ensemble spread by introducing for example additional sources of uncertainty.

Another remaining challenge is model structural error and the interference of it with the model parameterization. The second study of this PhD work illustrated the uncertainty related to the parameterization of the heterogeneous subsurface for a fully-coupled land-surface subsurface model at the hillslope scale. Within this context the impact of structural model errors caused by a misrepresentation of local (subsurface) processes (i.e. preferential flow, drainage by the fissured bedrock) has to be taken better into account for future high-resolution model studies. With the presence of structural model errors, parameter estimates as can be made by the Ensemble Kalman Filter are affected by those model structural errors and can therefore be erroneous. In addition, the impact of such model structural errors might also depend on parameter values. For example, the impact of preferential flow, not included in the model, might vary, since it interacts with the hydraulic properties of the different overlying soil layers. For a soil with high hydraulic conductivity (e.g. sand) preferential flow processes might be of minor importance for the model performance, while for a soil with low hydraulic conductivity (e.g. silty loam) the influence of preferential flow can massively alter the model performance. As a consequence, model structural errors or unknown processes also have a strong effect on the estimation of model states and parameters by inverse modelling or sequential data assimilation. An improved representation of the significant processes is therefore important for these model-data fusion techniques.

The experiments in this PhD study also demonstrated that despite a well-equipped study site which provided a broad basis of high-resolution data, still important information is missing which affects the simulation outcomes. This includes for example information about underlying bedrock structures as well as common small-scale anthropogenic infrastructure (e.g. the on-site drainage system). These features are site-specific and of importance for high resolution models since their impact will not be averaged out at some specific model resolution level. This illustrates that processes which significantly contribute to the uncertainty in runoff generation at a specific (unknown) scale might be easily overlooked. As a consequence, the specific requirements of fully-coupled high-resolution land surface-subsurface models have to be taken better into account for the operation and planning of monitoring campaigns. In fact, high-resolution models can be used to

identify significant information gaps in existing measurement infrastructure. In turn, the measurement data of an infrastructure campaign can be used to unravel structural model deficiencies and contribute to the model development.

134

Acknowledgements

First, I would like to thank my supervisor Professor Dr. Harrie-Jan Hendricks Franssen for his intensive supervision during my doctoral studies introducing me to the field of terrestrial modelling and data assimilation. Without his constructive feedback und the scientific discussions all along the track this work would never have been finished.

Furthermore, I thank the Helmholtz Association for providing financial support within the Water Network. I also would like to express my gratitude to all scientists and technicians accounting for the realization and maintenance of field observations or computer facilities such as the TERENO initiative, the Transregional Collaborative Research Centre 32 (TR32) and the Jülich Supercomputing Centre (JSC). I also thank the Centre for High-Performance Scientific Computing in Terrestrial Systems: HPSC TerrSys, Geoverbund ABC/J for scientific guidance.

I am very grateful for Professor Dr. Valentijn Pauwels and the water engineering group for the support, the hospitality, and the great experience during my research stay at Monash University in Melbourne, Australia.

This work would not have been possible without the colleagues from the Institute of Bio- and Geosciences at the Jülich Research Center. Many thanks to Professor Dr. Harry Vereecken, the head of the institute, for providing me support, suggestions and feedback; Professor Dr. Stefan Kollet and Dr. Klaus Görgen who supported me in questions of Parflow and TerrSysMP; Dr. Wolfgang Kurtz for his intensive TerrSysMP-PDAF support; Dr. Thomas Pütz for introducing me to the world of the weighable lysimeters; Thomas Schuster, Horst Hardelauf, Pramod Kumbar, Guowei He, and Mukund Pondkule for their IT support; and Qu Wei, Ansgar Weuthen, Bernd Schilling, and Marius Schmidt for providing extensive datasets and always having an open door for questions.

Special thanks go to my fellow PhD students and friends at IBG-3, who made work and free time so pleasant, in particular the lunch group(s) Anja, Katrin, Laura, Markus, Michael, Kathrina (other lunchgroup), Asta, Magdalena, Wei, and the IBG-3's institute band Michael, Maria, Ralf, Anne, Stocki, Wolfgang, and Alex.

Finally, I thank my family, especially my parents for their support and encouragement. And, last but not least, I want to thank my wife Kathrina and my little son Philipp for their love, understanding, and patience whenever I need it.

136

Bibliography

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E., Rasmussen, J., 1986. An introduction to the European Hydrological System — Systeme Hydrologique Europeen, “SHE”, 2: Structure of a physically-based, distributed modelling system. Journal of Hydrology, 87(1–2): 61-77.

Akaike, H., 1974. New Look at Statistical-Model Identification. IEEE Transactions on Automatic Control, 19(6): 716-723.

Alfieri, J.G., Kustas, W.P., Prueger, J.H., Hipps, L.E., Evett, S.R., Basara, J.B., Neale, C.M.U., French, A.N., Colaizzi, P., Agam, N., Cosh, M.H., Chavez, J.L., Howell, T.A., 2012. On the discrepancy between eddy covariance and lysimetry-based surface flux measurements under strongly advective conditions. Advances in Water Resources, 50(0): 62-78.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. FAO, Rome, 300(9): 1-15.

Allen, R.G., 2000. Using the FAO-56 dual crop coefficient method over an irrigated region as part of an evapotranspiration intercomparison study. Journal of Hydrology, 229(1-2): 27-41.

Allen, R.G., Pruitt, W.O., Wright, J.L., Howell, T.A., Ventura, F., Snyder, R., Itenfisu, D., Steduto, P., Berengena, J., Yrisarry, J.B., Smith, M., Pereira, L.S., Raes, D., Perrier, A., Alves, I., Walter, I., Elliott, R., 2006. A recommendation on standardized surface resistance for hourly calculation of reference ETo by the FAO56 Penman-Monteith method. Agricultural Water Management, 81(1–2): 1-22.

Anderson, J.L., 2001. An Ensemble Adjustment Kalman Filter for Data Assimilation. Monthly Weather Review, 129(12): 2884-2903.

Anderson, J. L., 2007. Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. Physica D: Nonlinear Phenomena, 230(1-2), 99-111.

Andreadis, K.M., Lettenmaier, D.P., 2006. Assimilating remotely sensed snow observations into a macroscale hydrology model. Advances in Water Resources, 29(6): 872-886.

Annan, J., Hargreaves, J., 2004. Efficient parameter estimation for a highly chaotic system. Tellus A, 56(5): 520-526.

Ashby, S.F., Falgout, R.D., 1996. A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations. Nuclear science and engineering, 124(1): 145-159.

Atchley, A., Maxwell, R., 2011. Influences of subsurface heterogeneity and vegetation cover on soil moisture, surface temperature and evapotranspiration at hillslope scales. Hydrogeology Journal, 19(2): 289-305.

Baatz, D., Kurtz, W., Hendricks Franssen, H. J., Vereecken, H., and Kollet, S. J., 2017. Catchment tomography - An approach for spatial parameter estimation, Advances in Water Resources, 107, 147-159.

Bailey, R.T., Baù, D., 2012. Estimating geostatistical parameters and spatially-variable hydraulic conductivity within a catchment system using an ensemble smoother. Hydrol. Earth Syst. Sci., 16(2): 287-304.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., Reinhardt, T., 2011. Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. Monthly Weather Review, 139(12): 3887-3905.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., 2001. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities. Bulletin of the American Meteorological Society, 82(11): 2415-2434.

Bárdossy, A., Das, T., 2008. Influence of rainfall observation network on model calibration and application. Hydrol Earth Syst Sc, 12(1): 77-89.

Betts, A. K., Ball, J. H., Beljaars, A. C. M., Miller, M. J., Viterbo, P. A., 1996. The landatmosphere interaction: A review based on observational and global modeling perspectives, J. Geophys. Res., 101(D3), 7209-7225.

Beven, K., Binley, A., 1992. The future of distributed models: Model calibration and uncertainty prediction. Hydrological Processes, 6(3): 279-298.

Beven, K., 1992. Future of Distributed Modeling - Special Issue of Hydrological Processes. Hydrological Processes, 6(3): 253-254.

Beven, K., 1996. A Discussion of Distributed Hydrological Modelling. In: Abbott, M., Refsgaard, J. (Eds.), Distributed Hydrological Modelling. Water Science and Technology Library. Springer Netherlands, 22: 255-278.

Bierkens, M.F., Van den Hurk, B.J., 2007. Groundwater convergence as a possible mechanism for multi-year persistence in rainfall. Geophys Res Lett, 34(2):GL028396.

Binley, A., Beven, K., 1992. Three-dimensional modelling of hillslope hydrology. Hydrological Processes, 6(3): 347-359.

Bishop, C.H., Etherton, B.J., Majumdar, S.J., 2001. Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. Monthly Weather Review, 129(3): 420-436.

Bixio, A., Gambolati, G., Paniconi, C., Putti, M., Shestopalov, V., Bublias, V., Bohuslavsky, A., Kasteltseva, N., Rudenko, Y., 2002. Modeling groundwater-surface water interactions including effects of morphogenetic depressions in the Chernobyl exclusion zone. Environmental Geology, 42(2): 162-177.

Blöschl, G., Sivapalan, M., 1995. Scale Issues in Hydrological Modeling - a Review. Hydrological Processes, 9(3-4): 251-290.

Bogena, H., 2003. Die Grundwasserneubildung in Nordrhein-Westfalen. URL: http://wwwzb.fz-juelich.de/contentenrichment/inhaltsverzeichnisse/2010/verlag/umwelt37_i.PDF. Schriften des Forschungszentrums Jülich. Reihe Umwelt, Forschungszentrum, Zentralbibliothek, Jülich, 148 pp.

Bonan, G.B., Oleson, K.W., Vertenstein, M., Levis, S., Zeng, X.B., Dai, Y.J., Dickinson, R.E., Yang, Z.L., 2002. The land surface climatology of the community land model coupled to the NCAR community climate model. Journal of Climate, 15(22): 3123-3149.

Bonetti, S., Manoli, G., Domec, J.-C., Putti, M., Marani, M., Katul, G.G., 2015. The influence of water table depth and the free atmospheric state on convective rainfall predisposition. Water Resources Research, 51(4): 2283-2297.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., Hasenauer, S., 2010. Improving runoff prediction through the assimilation of the ASCAT soil moisture product. Hydrol. Earth Syst. Sci., 14(10): 1881-1893.

Brown, R.N., Percivalle, C., Narkiewicz, S., DeCuollo, S., 2010. Relative Rooting Depths of Native Grasses and Amenity Grasses with Potential for Use on Roadsides in New England. HortScience, 45(3): 393-400.

Brutsaert, W., 2010. Hydrology: An introduction / Wilfried Brutsaert. Univ. Press, Cambridge, XI, 605 pp.

Burgess, T.M., Webster, R., 1980. Optimal Interpolation and Isarithmic Mapping of Soil Properties. Journal of Soil Science, 31(2): 333-341.

Burgers, G., Jan van Leeuwen, P., Evensen, G., 1998. Analysis Scheme in the Ensemble Kalman Filter. Monthly Weather Review, 126(6): 1719-1724.

Butts, M., Drews, M., Larsen, M.A.D., Lerer, S., Rasmussen, S.H., Grooss, J., Overgaard, J., Refsgaard, J.C., Christensen, O.B., Christensen, J.H., 2014. Embedding complex hydrology in the regional climate system – Dynamic coupling across different modelling domains. Advances in Water Resources, 74: 166-184.

Camporese, M., Paniconi, C., Putti, M., Salandin, P., 2009. Ensemble Kalman filter data assimilation for a process-based catchment scale model of surface and subsurface flow. Water Resources Research, 45(10):WR007031.

Camporese, M., Paniconi, C., Putti, M., Orlandini, S., 2010. Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data. Water Resources Research, 46(2): W02512.

Carsel, R.F., Parrish, R.S., 1988. Developing joint probability distributions of soil water retention characteristics. Water Resources Research, 24(5): 755-769.

Chaudhuri, A., Hendricks Franssen, H.-J., Sekhar, M., 2018. Iterative filter based estimation of fully 3D heterogeneous fields of permeability and Mualem-van Genuchten parameters. Advances in Water Resources, 122: 340-354.

Chávez, J., Howell, T., Copeland, K., 2009. Evaluating eddy covariance cotton ET measurements in an advective environment with large weighing lysimeters. Irrig Sci, 28(1): 35-50.

Chen, F., Crow, W.T., Starks, P.J., Moriasi, D.N., 2011. Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture. Advances in Water Resources, 34(4): 526-536.

Chen, H., Yang, D., Hong, Y., Gourley, J.J., Zhang, Y., 2013. Hydrological data assimilation with the Ensemble Square-Root-Filter: Use of streamflow observations to update model states for real-time flash flood forecasting. Advances in Water Resources, 59: 209-220.

Chen, Y., Oliver, D.S., 2010. Cross-covariances and localization for EnKF in multiphase flow data assimilation. Computational Geosciences, 14(4): 579-601.

Chen, Y., Zhang, D., 2006. Data assimilation for transient flow in geologic formations via ensemble Kalman filter. Advances in Water Resources, 29(8): 1107-1122.

Chen, X., Hu, Q., 2004. Groundwater influences on soil moisture and surface evaporation. Journal of Hydrology, 297(1-4): 285-300.

Choi, M., Jacobs, J.M., 2007. Soil moisture variability of root zone profiles within SMEX02 remote sensing footprints. Advances in Water Resources, 30(4): 883-896.

Christiaens, K., Feyen, J., 2000. The influence of different methods to derive soil hydraulic properties on the uncertainty of various model outputs of a distributed hydrological model. Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, 25(7): 679-683.

Christiaens, K., Feyen, J., 2001. Analysis of uncertainties associated with different methods to determine soil hydraulic properties and their propagation in the distributed hydrological MIKE SHE model. Journal of Hydrology, 246(1–4): 63-81.

Chvíla, B., Sevruk, B., Ondrás, M., 2005. The wind-induced loss of thunderstorm precipitation measurements. Atmospheric Research, 77(1–4): 29-38.

Ciach, G.J., 2003. Local Random Errors in Tipping-Bucket Rain Gauge Measurements. Journal of Atmospheric and Oceanic Technology, 20(5): 752-759.

Cornelissen, T., Diekkrüger, B., Bogena, H.R., 2014. Significance of scale and lower boundary condition in the 3D simulation of hydrological processes and soil moisture variability in a forested headwater catchment. Journal of Hydrology, 516(0): 140-153.

Cornelissen, T., Diekkrüger, B., Bogena, H., 2016. Using High-Resolution Data to Test Parameter Sensitivity of the Distributed Hydrological Model HydroGeoSphere. Water, 8(5): w805020.

Crow, W.T., Wood, E.F., 2003. The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using Ensemble Kalman filtering: a case study based on ESTAR measurements during SGP97. Advances in Water Resources, 26(2): 137-149.

Crow, W.T., Loon, E.V., 2006. Impact of Incorrect Model Error Assumptions on the Sequential Assimilation of Remotely Sensed Surface Soil Moisture. Journal of Hydrometeorology, 7(3): 421-432.

Darcy, H., 1856. Les fontaines publiques de la ville de Dijon: Exposition et application des principes a suivre et des formules a employer dans les questions de distribution d'eau; ouvrage terminé par un appendice relatif aux fournitures d'eau de plusieurs villes au filtrage des eaux et a la fabrication des tuyaux de fonte, de plomb, de tole et de bitume. Atlas. Victor Dalmont, Libraire des Corps imperiaux des ponts et chaussées et des mines.

De Lannoy, G.J.M., Reichle, R.H., Houser, P.R., Pauwels, V.R.N., Verhoest, N.E.C., 2007. Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter. Water Resources Research, 43(9): WR005449.

Desilets, D., Zreda, M., Ferré, T., 2010. Nature's neutron probe: Land surface hydrology at an elusive scale with cosmic rays. Water Resources Research, 46(11): WR008726.

Deutscher Wetterdienst (DWD), 1993. Richtlinie für automatische Klimastationen, Offenbach am Main.

Devegowda, D., Arroyo-Negrete, E., Datta-Gupta, A., 2010. Flow relevant covariance localization during dynamic data assimilation using EnKF. Advances in water resources, 33(2): 129-145.

Dickinson, R.E., Henderson-Sellers, A., Kennedy, P.J., 1993. Biosphere-Atmosphere Transfer Scheme (BATS) version le as coupled to the NCAR community climate model. Version 1e as Coupled to the NCAR Community Climate Model. National Center for Atmospheric Research (NCAR) Technical Note United States.

Ding, R., Kang, S., Li, F., Zhang, Y., Tong, L., Sun, Q., 2010. Evaluating eddy covariance method by large-scale weighing lysimeter in a maize field of northwest China. Agricultural Water Management, 98(1): 87-95.

Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resources Research, 28(4): 1015-1031.

Erdal, D., Neuweiler, I., Wollschläger, U., 2014. Using a bias aware EnKF to account for unresolved structure in an unsaturated zone model. Water Resources Research, 50(1): 132-147.

Erdal, D., Rahman, M.A., Neuweiler, I., 2015. The importance of state transformations when using the ensemble Kalman filter for unsaturated flow modeling: Dealing with strong nonlinearities. Advances in Water Resources, 86: 354-365.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research: Oceans, 99(C5): 10143-10162.

Evensen, G., 2004. Sampling strategies and square root analysis schemes for the EnKF. Ocean Dynam, 54(6): 539-560.

Evensen, G. (2009). Data assimilation: the ensemble Kalman filter. Springer Science & Business Media, 272 pp.

Evett, S.R., Schwartz, R.C., Howell, T.A., Baumhardt, R.L., Copeland, K.S., 2012. Can weighing lysimeter ET represent surrounding field ET well enough to test flux station measurements of daily and sub-daily ET? Advances in Water Resources, 50: 79-90.

Famiglietti, J.S., Rudnicki, J.W., Rodell, M., 1998. Variability in surface moisture content along a hillslope transect: Rattlesnake Hill, Texas. Journal of Hydrology, 210(1-4): 259-281.

Famiglietti, J.S., Devereaux, J. A., Laymon, C. A., Tsegaye, T., Houser, P. R., Jackson, T. J., Graham, S. T., Rodell, M., van Oevelen, P. J., 1999. Ground-based investigation of soil moisture variability within remote sensing footprints During the Southern Great Plains 1997 (SGP97) Hydrology Experiment. Water Resources Research, 35(6): 1839-1851.

Fang, Z., Bogena, H., Kollet, S., Koch, J., Vereecken, H., 2015. Spatio-temporal validation of long-term 3D hydrological simulations of a forested catchment using empirical orthogonal functions and wavelet coherence analysis. Journal of Hydrology, 529: 1754-1767.

Fatichi, S., Katul, G.G., Ivanov, V.Y., Pappas, C., Paschalis, A., Consolo, A., Kim, J., Burlando, P., 2015. Abiotic and biotic controls of soil moisture spatiotemporal variability and the occurrence of hysteresis. Water Resources Research, 51(5): 3505-3524.

Fatichi, S., Vivoni, E.R., Ogden, F.L., Ivanov, V.Y., Mirus, B., Gochis, D., Downer, C.W., Camporese, M., Davison, J.H., Ebel, B., Jones, N., Kim, J., Mascaro, G., Niswonger, R., Restrepo, P., Rigon, R., Shen, C., Sulis, M., Tarboton, D., 2016. An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. Journal of Hydrology, 537: 45-60.

Feldhake, C.M., Boyer, D.G., 1986. Effect of soil temperature on evapotranspiration by C3 and C4 grasses. Agricultural and Forest Meteorology, 37(4): 309-318.

Ferguson, I.M., Maxwell, R.M., 2010. Role of groundwater in watershed response and land surface feedbacks under climate change. Water Resources Research, 46(10): W00F02.

Finnigan, J., 2004. The footprint concept in complex terrain. Agricultural and Forest Meteorology, 127(3–4): 117-129.

Foken, T., Wichura, B., 1996. Tools for quality assessment of surface-based flux measurements. Agricultural and Forest Meteorology, 78(1-2): 83-105.

Foken, T., 2008. The energy balance closure problem: An overview. Ecological Applications, 18(6): 1351-1367.

Foken, T., Aubinet, M., Finnigan, J.J., Leclerc, M.Y., Mauder, M., Paw U, K.T., 2011. Results Of A Panel Discussion About The Energy Balance Closure Correction For Trace Gases. Bulletin of the American Meteorological Society, 92(4): 13-18.

Foufoula-Georgiou, E., Ganti, V., Dietrich, W.E., 2010. A nonlocal theory of sediment transport on hillslopes. J Geophys Res-Earth, 115(2): F001280.

Frey, S. K., Rudolph, D. L., Lapen, D. R., Ball Coelho, B. R., 2012. Viscosity dependent dual-permeability modeling of liquid manure movement in layered, macroporous, tile drained soil. Water Resources Research, 48(6).

Frey, S. K., Hwang, H. T., Park, Y. J., Hussain, S. I., Gottschall, N., Edwards, M., Lapen, D. R., 2016. Dual permeability modeling of tile drain management influences on hydrologic and nutrient transport characteristics in macroporous soil. Journal of Hydrology, 535, 392-406.

Gasper, F., Goergen, K., Kollet, S., Shrestha, P., Sulis, M., Rihani, J., Geimer, M., 2014. Implementation and scaling of the fully coupled Terrestrial Systems Modeling Platform (TerrSysMP) in a massively parallel supercomputing environment – a case study on JUQUEEN (IBM Blue Gene/Q). Geosci. Model Dev. Discuss., 7(3): 3545-3573.

Gatel, L., Lauvernet, C., Carluer, N., Paniconi, C., 2016. Effect of surface and subsurface heterogeneity on the hydrological response of a grassed buffer zone. Journal of Hydrology, 542: 637-647.

Gebler, S., Hendricks Franssen, H.J., Pütz, T., Post, H., Schmidt, M., Vereecken, H., 2015. Actual evapotranspiration and precipitation measured by lysimeters: a comparison with eddy covariance and tipping bucket. Hydrol. Earth Syst. Sci., 19(5): 2145-2161.

Gebler, S., Hendricks Franssen, H.J., Kollet, S.J., Qu, W., Vereecken, H., 2017. High resolution modelling of soil moisture patterns with TerrSysMP: A comparison with sensor network data. Journal of Hydrology, 547, 309-331.

Gentine, P., Troy, T.J., Lintner, B.R., Findell, K.L., 2012. Scaling in Surface Hydrology: Progress and Challenges. Journal of Contemporary Water Research & Education, 147(1): 28-40.

Gharamti, M.E., Kadoura, A., Valstar, J., Sun, S., Hoteit, I., 2014. Constraining a compositional flow model with flow-chemical data using an ensemble-based Kalman filter. Water Resources Research, 50(3): 2444-2467.

Ghent, D., Kaduk, J., Remedios, J., Ardö, J., Balzter, H., 2010. Assimilation of land surface temperature into the land surface model JULES with an ensemble Kalman filter. Journal of Geophysical Research: Atmospheres, 115(19): D014382.

Gómez-Hernández, J.J., Journel, A., 1993. Joint Sequential Simulation of MultiGaussian Fields. In: Soares, A. (Ed.), Geostatistics Tróia '92. Quantitative Geology and Geostatistics. Springer Netherlands: 85-94.

Gómez-Hernández, J.J., Wen, X.-H., 1998. To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. Advances in Water Resources, 21(1): 47-61.

Gómez-Plaza, A., Martínez-Mena, M., Albaladejo, J., Castillo, V.M., 2001. Factors regulating spatial distribution of soil water content in small semiarid catchments. Journal of Hydrology, 253(1–4): 211-226.

Goodison, B., Louie, P., Yang, D., 1997. The WMO solid precipitation measurement intercomparison. World Meteorological Organization-Publications-WMO TD: 65-70.

Graf, A., Bogena, H.R., Drüe, C., Hardelauf, H., Pütz, T., Heinemann, G., Vereecken, H., 2014. Spatiotemporal relations between water budget components and soil water content in a forested tributary catchment. Water Resources Research, 50(6): 4837-4857.

Graham, D.N., Butts, M.B., 2005. Flexible, integrated watershed modelling with MIKE SHE. Watershed models, 849336090: 245-272.

Grayson, R.B., Western, A.W., Chiew, F.H.S., Blöschl, G., 1997. Preferred states in spatial soil moisture patterns: Local and nonlocal controls. Water Resources Research, 33(12): 2897-2908.

Greifeneder, F., Notarnicola, C., Bertoldi, G., Niedrist, G., Wagner, W., 2016. From Point to Pixel Scale: An Upscaling Approach for In Situ Soil Moisture Measurements. Vadose Zone Journal, 15(6): vzj2015.03.0048.

Groisman, P.Y., Legates, D.R., Groisman, P.Y., Legates, D.R., 1994. The Accuracy of United States Precipitation Data. Bulletin of the American Meteorological Society, 75(2): 215-227.

Guay, C., Nastev, M., Paniconi, C., Sulis, M., 2013. Comparison of two modeling approaches for groundwater–surface water interactions. Hydrological Processes, 27(16): 2258-2270.

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1998. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resources Research, 34(4): 751-763.

Gutmann, E. D., and E. E. Small, 2007: A comparison of land surface model soil hydraulic properties estimated by inverse modeling and pedotransfer functions. Water Resources Research, 43(5): 1-13.

Guo, L., Chen, J., Lin, H., 2014. Subsurface lateral preferential flow network revealed by time-lapse ground-penetrating radar in a hillslope. Water Resources Research, 50(12): 9127-9147.

Hale, V.C., McDonnell, J.J., Stewart, M.K., Solomon, D.K., Doolitte, J., Ice, G.G., Pack, R.T., 2016. Effect of bedrock permeability on stream base flow mean transit time scaling relationships: 2. Process study of storage and release. Water Resources Research, 52(2): 1375-1397.

Han, X., Hendricks Franssen, H.-J., Li, X., Zhang, Y., Montzka, C., Vereecken, H., 2013. Joint Assimilation of Surface Temperature and L-Band Microwave Brightness Temperature in Land Data Assimilation. Vadose Zone Journal, 12(3): vzj2012.0072.

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika, 57(1): 97-109.

Heathman, G.C., Starks, P.J., Ahuja, L.R., Jackson, T.J., 2003. Assimilation of surface soil moisture to estimate profile soil water content. Journal of Hydrology, 279(1–4): 1-17.

Heathman, G.C., Larose, M., Cosh, M.H., Bindlish, R., 2009. Surface and profile soil moisture spatio-temporal analysis during an excessive rainfall period in the Southern Great Plains, USA. CATENA, 78(2): 159-169.

Hendricks Franssen, H.J., Kinzelbach, W., 2008. Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem. Water Resources Research, 44(9): WR006505.

Hendricks Franssen, H.-J., Stöckli, R., Lehner, I., Rotenberg, E., Seneviratne, S., 2010. Energy balance closure of eddy-covariance data: a multisite analysis for European FLUXNET stations. Agricultural and Forest Meteorology, 150(12): 1553-1567.

Hendricks Franssen, H.J., Kaiser, H.P., Kuhlmann, U., Bauser, G., Stauffer, F., Müller, R., Kinzelbach, W., 2011. Operational real-time modeling with ensemble Kalman filter of variably saturated subsurface flow including stream-aquifer interaction and parameter updating. Water Resources Research, 47(2): WR009480.

Herbst, M., Diekkrüger, B., 2003. Modelling the spatial variability of soil moisture in a micro-scale catchment and comparison with field data using geostatistics. Physics and Chemistry of the Earth, Parts A/B/C, 28(6–7): 239-245.

Herbst, M., Diekkrüger, B., Vanderborght, J., 2006. Numerical experiments on the sensitivity of runoff generation to the spatial variation of soil hydraulic properties. Journal of Hydrology, 326(1-4): 43-58.

Heuvelink, G.B.M., Webster, R., 2001. Modelling soil variation: past, present, and future. Geoderma, 100(3–4): 269-301.

Hohenbrink, T. L., and G. Lischeid, 2014: Texture-depending performance of an in situ method assessing deep seepage. Journal of Hydrology, 511, 61-71.

Hohenbrink, T.L., Lischeid, G., Schindler, U., Hufnagel, J., 2016. Disentangling the Effects of Land Management and Soil Heterogeneity on Soil Moisture Dynamics. Vadose Zone Journal, 15(1): vzj2015.07.0107.

Hu, W., Shao, M., Han, F., Reichardt, K., Tan, J., 2010. Watershed scale temporal stability of soil water content. Geoderma, 158(3–4): 181-198.

Hu, Z., Islam, S., 1998. Effects of subgrid-scale heterogeneity of soil wetness and temperature on grid-scale evaporation and its parameterization. Int J Climatol, 18(1): 49-63.

Huang, W., Zhang, C., Xue, X., Chen, L., 2012. A Data Acquisition System Based on Outlier Detection Method for Weighing Lysimeters. In: Li, D., Chen, Y. (Eds.), Computer and Computing Technologies in Agriculture V. IFIP Advances in Information and Communication Technology. Springer Berlin Heidelberg: 471-478.

Hurvich, C.M., Tsai, C.-L., 1989. Regression and time series model selection in small samples. Biometrika, 76(2): 297-307.

Houser, P.R., Shuttleworth, W.J., Famiglietti, J.S., Gupta, H.V., Syed, K.H., Goodrich, D.C., 1998. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. Water Resources Research, 34(12): 3405-3420.

Houtekamer, P.L., Mitchell, H.L., 1998. Data Assimilation Using an Ensemble Kalman Filter Technique. Monthly Weather Review, 126(3): 796-811.

Imukova, K., Ingwersen, J., Hevart, M., Streck, T., 2016. Energy balance closure on a winter wheat stand: comparing the eddy covariance technique with the soil water balance method. Biogeosciences, 13(1): 63-75.

Ines, A.V.M., Das, N.N., Hansen, J.W., Njoku, E.G., 2013. Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction. Remote Sensing of Environment, 138: 149-164.

Ingwersen, J., Steffens, K., Högy, P., Warrach-Sagi, K., Zhunusbayeva, D., Poltoradnev, M., Gäbler, R., Wizemann, H.D., Fangmeier, A., Wulfmeyer, V., Streck, T., 2011. Comparison of Noah simulations with eddy covariance and soil water measurements at a winter wheat stand. Agricultural and Forest Meteorology, 151(3): 345-355.

Ivanov, V.Y., Bras, R.L., Vivoni, E.R., 2008. Vegetation-hydrology dynamics in complex terrain of semiarid areas: 1. A mechanistic approach to modeling dynamic feedbacks. Water Resources Research, 44(3): WR005588.

Ivanov, V.Y., Fatichi, S., Jenerette, G.D., Espeleta, J.F., Troch, P.A., Huxman, T.E., 2010. Hysteresis of soil moisture spatial heterogeneity and the "homogenizing" effect of vegetation. Water Resources Research, 46(9): WR008611.

Jacobs, A.F.G., Heusinkveld, B.G., Wichink Kruit, R.J., Berkowicz, S.M., 2006. Contribution of dew to the water budget of a grassland area in the Netherlands. Water Resources Research, 42(3): WR004055.

Jones, J.E., Woodward, C.S., 2001. Newton–Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems. Advances in Water Resources, 24(7): 763-774.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Journal of basic Engineering, 82(1): 35-45.

Kampf, S.K., Burges, S.J., 2010. Quantifying the water balance in a planar hillslope plot: Effects of measurement errors on flow prediction. Journal of Hydrology, 380(1): 191-202.

Kessomkiat, W., Franssen, H.-J.H., Graf, A., Vereecken, H., 2013. Estimating random errors of eddy covariance data: An extended two-tower approach. Agricultural and Forest Meteorology, 171–172(0): 203-219.

Kerrou, J., Renard, P., Franssen, H.-J.H., Lunati, I., 2008. Issues in characterizing heterogeneity and connectivity in non-multiGaussian media. Advances in Water Resources, 31(1): 147-159.

Koch, J., Jensen, K.H., Stisen, S., 2015. Toward a true spatial model evaluation in distributed hydrological modeling: Kappa statistics, Fuzzy theory, and EOF-analysis benchmarked by the human perception and evaluated against a modeling case study. Water Resources Research, 51(2): 1225-1246.

Kohnke, H., Davidson, J.M., Dreibelbis, F.R., 1940. A survey and discussion of lysimeters and a bibliography on their construction and performance. U. S. Govt. print. off., Washington, D.C., 68 pp.

Kollet, S.J., Maxwell, R.M., 2006. Integrated surface-groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. Advances in Water Resources, 29(7): 945-958.

Kollet, S.J., Maxwell, R.M., 2008. Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model. Water Resour. Res., 44(2): W02402.

Kollet, S.J., 2009. Influence of soil heterogeneity on evapotranspiration under shallow water table conditions: transient, stochastic simulations. Environ Res Lett, 4(3): 035007.

Kollet, S.J., Maxwell, R.M., Woodward, C.S., Smith, S., Vanderborght, J., Vereecken, H., Simmer, C., 2010. Proof of concept of regional scale hydrologic simulations at hydrologic resolution utilizing massively parallel computer resources. Water Resources Research, 46: W04201.

Kormann, R., Meixner, F.X., 2001. An analytical footprint model for non-neutral stratification. Bound-Lay Meteorol, 99(2): 207-224.

Korres, W., Koyama, C.N., Fiener, P., Schneider, K., 2010. Analysis of surface soil moisture patterns in agricultural landscapes using Empirical Orthogonal Functions. Hydrol Earth Syst Sc, 14(5): 751-764.

Koyama, C.N., Korres, W., Fiener, P., Schneider, K., 2010. Variability of Surface Soil Moisture Observed from Multitemporal C-Band Synthetic Aperture Radar and Field Data. Vadose Zone Journal, 9(4): 1014-1024.

Kumar, M., 2009: Toward a hydrologic modeling system. Ph.D. thesis, The Pennsylvania State University, 251 pp.

Kurtz, W., Hendricks Franssen, H.-J., Vereecken, H., 2012. Identification of time-variant river bed properties with the ensemble Kalman filter. Water Resources Research, 48(10): W10534.

Kurtz, W., Hendricks Franssen, H.-J., Kaiser, H.-P., Vereecken, H., 2014. Joint assimilation of piezometric heads and groundwater temperatures for improved modeling of river-aquifer interactions. Water Resources Research, 50(2): 1665-1688.

Kurtz, W., He, G., Kollet, S.J., Maxwell, R.M., Vereecken, H., Hendricks Franssen, H.J., 2016. TerrSysMP–PDAF (version 1.0): a modular high-performance data assimilation framework for an integrated land surface–subsurface model. Geosci. Model Dev., 9(4): 1341-1360.

Legates, D.R., DeLiberty, T.L., 1993. Precipiation Measurement biases in the United States. JAWRA Journal of the American Water Resources Association, 29(5): 855-861.

Legates, D.R., Mahmood, R., Levia, D.F., DeLiberty, T.L., Quiring, S.M., Houser, C., Nelson, F.E., 2011. Soil moisture: A central and unifying theme in physical geography. Progress in Physical Geography, 35(1): 65-86.

Li, C., Ren, L., 2011. Estimation of Unsaturated Soil Hydraulic Parameters Using the Ensemble Kalman Filter. Vadose Zone Journal, 10(4): 1205-1227.

Li, S., Kang, S., Zhang, L., Li, F., Zhu, Z., Zhang, B., 2008. A comparison of three methods for determining vineyard evapotranspiration in the arid desert regions of northwest China. Hydrological Processes, 22(23): 4554-4564.

Lievens, H., Tomer, S. K., Al Bitar, A., De Lannoy, G. J., Drusch, M., Dumedah, G., Hendricks Franssen, H.-J., Kerr, Y.H., Martens, B., Pan, M., Roundy, J.K., Vereecken, H. , Walker, J.P., Wood, E.F., Verhoest, N.E.C., Pauwels, V.R.N., 2015. SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia. Remote Sensing of Environment, 168, 146-162.

Lievens, H., De Lannoy, G. J. M ., Al Bitar, A., Drusch, M., Dumedah, G., Hendricks Franssen, H.-J., Kerr, Y. H., Tomer, S. K., Martens, B., Merlin, O., Pan, M., Roundy, J. K., Vereecken, H., Walker, J. P., Wood, E. F., Verhoest, N. E. C., Pauwels, V. R. N., 2016. Assimilation of SMOS soil moisture and brightness temperature products into a land surface model. In: Remote Sensing of Environment, 180: 292-304.

Liu, H., Lin, H., 2015. Frequency and Control of Subsurface Preferential Flow: From Pedon to Catchment Scales. Soil Sci Soc Am J, 79(2): 362-377.

Liu, Y., Gupta, H.V., Sorooshian, S., Bastidas, L.A., Shuttleworth, W.J., 2005. Constraining Land Surface and Atmospheric Parameters of a Locally Coupled Model Using Observational Data. Journal of Hydrometeorology, 6(2): 156-172.

Liu, Y., Gupta, H.V., 2007. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. Water Resources Research, 43(7): WR005756.

López-Urrea, R., Olalla, F.M.d.S., Fabeiro, C., Moratalla, A., 2006. An evaluation of two hourly reference evapotranspiration equations for semiarid conditions. Agricultural Water Management, 86(3): 277-282.

Manfreda, S., McCabe, M.F., Fiorentino, M., Rodríguez-Iturbe, I., Wood, E.F., 2007. Scaling characteristics of spatial patterns of soil moisture from distributed modelling. Advances in Water Resources, 30(10): 2145-2150.

Martínez, G., Pachepsky, Y.A., Vereecken, H., 2014. Temporal stability of soil water content as affected by climate and soil hydraulic properties: a simulation study. Hydrological Processes, 28(4): 1899-1915.

Martínez García, G., Pachepsky, Y.A., Vereecken, H., 2014. Effect of soil hydraulic properties on the relationship between the spatial mean and variability of soil moisture. Journal of Hydrology, 516: 154-160.

Martini, E., Wollschläger, U., Kögler, S., Behrens, T., Dietrich, P., Reinstorf, F., Schmidt, C., Weiler, M. Werban, U., Zacharias, S., 2015. Spatial and temporal dynamics of hillslope-scale soil moisture patterns: Characteristic states and transition mechanisms. Vadose Zone Journal, 14(4): vzj2014.10.0150.

Mascaro, G., Vivoni, E.R., Méndez-Barroso, L.A., 2015. Hyperresolution hydrologic modeling in a regional watershed and its interpretation using empirical orthogonal functions. Advances in Water Resources, 83: 190-206.

Mascaro, G., Vivoni, E., 2016. On the observed hysteresis in field-scale soil moisture variability and its physical controls. Environ Res Lett, 11(8): 084008.

Mauder, M., Foken, T., 2011. Documentation and instruction manual of the eddy-covariance software package TK3.

Mauder, M., Cuntz, M., Druee, C., Graf, A., Rebmann, C., Schmid, H.P., Schmidt, M., Steinbrecher, R., 2013. A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements. Agricultural and Forest Meteorology, 169: 122-135.

Maxwell, R.M., Miller, N.L., 2005. Development of a Coupled Land Surface and Groundwater Model. Journal of Hydrometeorology, 6(3): 233-247.

Maxwell, R.M., Chow, F.K., Kollet, S.J., 2007. The groundwater–land-surface–atmosphere connection: Soil moisture effects on the atmospheric boundary layer in fully-coupled simulations. Advances in Water Resources, 30(12): 2447-2466.

Maxwell, R.M., Kollet, S.J., 2008. Quantifying the effects of three-dimensional subsurface heterogeneity on Hortonian runoff processes using a coupled numerical, stochastic approach. Advances in Water Resources, 31(5): 807-817.

Maxwell, R.M., Lundquist, J.K., Mirocha, J.D., Smith, S.G., Woodward, C.S., Tompson, A.F.B., 2011. Development of a Coupled Groundwater–Atmosphere Model. Monthly Weather Review, 139(1): 96-116.

Maxwell, R.M., 2013. A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling. Advances in Water Resources, 53: 109-117.

Maxwell, R.M., Condon, L.E., 2016. Connections between groundwater flow and transpiration partitioning. Science, 353(6297): 377-380.

McDonnell, J.J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M.L., Selker, J., Weiler, M., 2007. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. Water Resources Research, 43(7): WR005467.

Meissner, R., Seeger, J., Rupp, H., Seyfarth, M., Borg, H., 2007. Measurement of dew, fog, and rime with a high-precision gravitation lysimeter. Journal of Plant Nutrition and Soil Science, 170(3): 335-344.

Michelson, D.B., 2004. Systematic correction of precipitation gauge observations using analyzed meteorological variables. Journal of Hydrology, 290(3–4): 161-177.

Meyerhoff, S.B., Maxwell, R.M., 2011. Quantifying the effects of subsurface heterogeneity on hillslope runoff using a stochastic approach. Hydrogeology Journal, 19(8): 1515-1530.

Miguez-Macho, G., Fan, Y., 2012. The role of groundwater in the Amazon water cycle: 2. Influence on seasonal soil moisture and evapotranspiration. Journal of Geophysical Research: Atmospheres, 117(D15): D017540.

Montzka, C., Moradkhani, H., Weihermüller, L., Franssen, H.-J.H., Canty, M., Vereecken, H., 2011. Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter. Journal of Hydrology, 399(3–4): 410-421.

Montzka, C., Grant, J. P., Moradkhani, H., Franssen, H.-J. H., Weihermüller, L., Drusch, M., & Vereecken, H., 2013. Estimation of radiative transfer parameters from L-band passive microwave brightness temperatures using advanced data assimilation. Vadose Zone Journal, 12(3): vzj2012.0040.

Moore, C.J., 1986. Freuquency response corrections for eddy-correlation systems. Bound-Lay Meteorol, 37(1-2): 17-35.

Moriasi, D.N. et al., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 50(3): 885-900.

Nan, T., Wu, J., 2011. Groundwater parameter estimation using the ensemble Kalman filter with localization. Hydrogeology Journal, 19(3): 547-561.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. Journal of Hydrology, 10(3): 282-290.

Nerger, L., Hiller, W., 2013. Software for ensemble-based data assimilation systems—Implementation strategies and scalability. Computers & Geosciences, 55: 110-118.

Nešpor, V., Sevruk, B., 1999. Estimation of Wind-Induced Error of Rainfall Gauge Measurements Using a Numerical Simulation. Journal of Atmospheric and Oceanic Technology, 16(4): 450-464.

Neteler, M., Bowman, M.H., Landa, M., Metz, M., 2012. GRASS GIS: A multi-purpose open source GIS. Environ Modell Softw, 31: 124-130.

Niu, G.-Y., Paniconi, C., Troch, P.A., Scott, R.L., Durcik, M., Zeng, X., Huxman, T., Goodrich, D.C., 2014. An integrated modelling framework of catchment-scale ecohydrological processes: 1. Model description and tests over an energy-limited watershed. Ecohydrology, 7(2): 427-439.

Nolz, R., Kammerer, G., Cepuder, P., 2013. Interpretation of lysimeter weighing data affected by wind. Journal of Plant Nutrition and Soil Science, 176(2): 200-208.

Oleson, K., Bonan, G.B., Levis, S., Thornton, P., Vertenstein, M., Yang, Z., Dai, Y., Bosilovich, M., Dirmeyer, P., Hoffman, F., Houser, P., Niu, G., Zeng, X., 2004. Technical Description of the Community Land Model (CLM). NCAR Technical Note NCAR/TN-461+STR.

Oleson, K.W., Niu, G.Y., Yang, Z.L., Lawrence, D.M., Thornton, P.E., Lawrence, P.J., Stockli, R., Dickinson, R.E., Bonan, G.B., Levis, S., Dai, A., Qian, T., 2008. Improvements to the Community Land Model and their impact on the hydrological cycle. J Geophys Res-Biogeo, 113: G000563.

Oudin, L., Perrin, C., Mathevet, T., Andréassian, V., Michel, C., 2006. Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models. Journal of Hydrology, 320(1): 62-83.

Pan, F., Peters-Lidard, C.D., 2008. On the relationship between mean and variance of soil moisture fields. J Am Water Resour As, 44(1): 235-242.

Panday, S., Huyakorn, P.S., 2004. A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. Advances in Water Resources, 27(4): 361-382.

Paschalis, A., Fatichi, S., Molnar, P., Rimkus, S., Burlando, P., 2014. On the effects of small scale space–time variability of rainfall on basin flood response. Journal of Hydrology, 514: 313-327.

Pasetto, D., Niu, G.-Y., Pangle, L., Paniconi, C., Putti, M., Troch, P.A., 2015. Impact of sensor failure on the observability of flow dynamics at the Biosphere 2 LEO hillslopes. Advances in Water Resources, 86: 327-339.

Pauwels, V.R.N., Hoeben, R., Verhoest, N.E.C., De Troch, F.P., 2001. The importance of the spatial patterns of remotely sensed soil moisture in the improvement of discharge predictions for small-scale basins through data assimilation. Journal of Hydrology, 251(1–2): 88-102.

Pauwels, V.R.N., Hoeben, R., Verhoest, N.E.C., De Troch, F.P., Troch, P.A., 2002. Improvement of TOPLATS-based discharge predictions through assimilation of ERS-based remotely sensed soil moisture values. Hydrological Processes, 16(5): 995-1013.

Pauwels, V.R.N., De Lannoy, G.J.M., 2009. Ensemble-based assimilation of discharge into rainfall-runoff models: A comparison of approaches to mapping observational information to state space. Water Resources Research, 45(8): W08428.

Peters, A., Nehls, T., Schonsky, H., Wessolek, G., 2014. Separating precipitation and evapotranspiration from noise – a new filter routine for high-resolution lysimeter data. Hydrol. Earth Syst. Sci., 18(3): 1189-1198.

Plauborg, F., Iversen, B.V., Lærke, P.E., 2005. In situ comparison of three dielectric soil moisture sensors in drip irrigated sandy soils. Vadose Zone Journal, 4(4): 1037-1047.

Poltoradnev, M., Ingwersen, J., Streck, T., 2016. Spatial and Temporal Variability of Soil Water Content in Two Regions of Southwest Germany during a Three-Year Observation Period. Vadose Zone Journal, 15(6): vzj2015.11.0143.

Pruitt, W., Lourence, F.J., 1985. Experiences in lysimetry for ET and surface drag measurements. Advances in Evapotranspiration, ASAE, 51- 69.

Qu, W., Bogena, H.R., Huisman, J.A., Vereecken, H., 2013. Calibration of a Novel Low-Cost Soil Water Content Sensor Based on a Ring Oscillator, Vadose Zone Journal., 12(2): vzj2012.0139.

Qu, W., Bogena, H.R., Huisman, J.A., Martinez, G., Pachepsky, Y.A., Vereecken, H., 2014. Effects of Soil Hydraulic Properties on the Spatial Variability of Soil Water Content: Evidence from Sensor Network Data and Inverse Modeling. Vadose Zone Journal, 13(12): vzj2014.07.0099.

Qu, W., Bogena, H. R., Huisman, J. A., Vanderborght, J., Schuh, M., Priesack, E., Vereecken, H., 2015, Predicting subgrid variability of soil water content from basic soil information. Geophys. Res. Lett., 42: 789–796.

Qu, W., Bogena, H.R., Huisman, J.A., Schmidt, M., Kunkel, R., Weuthen, A., Schiedung, H., Schilling, B., Sorg, J., Vereecken, H., 2016. The integrated water balance and soil data set of the Rollesbroich hydrological observatory. Earth Syst. Sci. Data, 8(2): 517-529.

Qu, Y., Duffy, C. J., 2007: A semidiscrete finite volume formulation for multiprocess watershed simulation. Water Resources Research,43: W08419.

Rahman, M., Sulis, M., Kollet, S.J., 2014. The concept of dual-boundary forcing in land surface-subsurface interactions of the terrestrial hydrologic and energy cycles. Water Resources Research, 50(11): 8531-8548.

Rana, G., Katerji, N., 2000. Measurement and estimation of actual evapotranspiration in the field under Mediterranean climate: a review. European Journal of Agronomy, 13(2): 125-153.

Rasmussen, J., Madsen, H., Jensen, K.H., Refsgaard, J.C., 2015. Data assimilation in integrated hydrological modeling using ensemble Kalman filtering: evaluating the effect of ensemble size and localization on filter performance. Hydrol. Earth Syst. Sci., 19(7): 2999-3013.

Rawls, W.J., Brakensiek, D.L., 1989. Estimation of Soil Water Retention and Hydraulic Properties. In: Morel-Seytoux, H.J. (Ed.), Unsaturated Flow in Hydrologic Modeling: Theory and Practice. Springer Netherlands, Dordrecht, 275-300.

Reichle, R.H., McLaughlin, D.B., Entekhabi, D., 2002. Hydrologic Data Assimilation with the Ensemble Kalman Filter. Monthly Weather Review, 130(1): 103-114.

Reichle, R.H., Kumar, S.V., Mahanama, S.P.P., Koster, R.D., Liu, Q., 2010. Assimilation of Satellite-Derived Skin Temperature Observations into Land Surface Models. Journal of Hydrometeorology, 11(5): 1103-1122.

Richards, L.A., 1931. Capillary conduction of liquids through porous mediums. Physics-J Gen Appl P, 1(1): 318-333.

Richter, D., 1995. Ergebnisse methodischer Untersuchungen zur Korrektur des systematischen Messfehlers des Hellmann-Niederschlagsmessers / von Dieter Richter. 93 S. Berichte des Deutschen Wetterdienstes ; 194. Selbstverl. des Dt. Wetterdienstes, Offenbach.

Ridler, M.E., van Velzen, N., Hummel, S., Sandholt, I., Falk, A.K., Heemink, A., Madsen, H., 2014. Data assimilation framework: Linking an open data assimilation library (OpenDA) to a widely adopted model interface (OpenMI). Environ Modell Softw, 57: 76-89.

Rihani, J.F., Maxwell, R.M., Chow, F.K., 2010. Coupling groundwater and land surface processes: Idealized simulations to identify effects of terrain and subsurface heterogeneity on land surface energy fluxes. Water Resources Research, 46: WR009111.

Robinson, D.A., Campbell, C.S., Hopmans, J.W., Hornbuckle, B.K., Jones, S.B., Knight, R., Ogden, F., Selker, J., Wendroth, O., 2008. Soil Moisture Measurement for Ecological and Hydrological Watershed-Scale Observatories. Vadose Zone Journal, 7(1): 358-389.

Rodriguez-Iturbe, I., Porporato, A., Laio, F., Ridolfi, L., 2001. Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress: I. Scope and general outline. Advances in Water Resources, 24(7): 695-705.

Rosenbaum, U., Bogena, H.R., Herbst, M., Huisman, J.A., Peterson, T.J., Weuthen, A., Western, A.W., Vereecken, H., 2012. Seasonal and event dynamics of spatial soil moisture patterns at the small catchment scale. Water Resources Research, 48(10): W10544.

Rötzer, K., Montzka, C., Bogena, H., Wagner, W., Kerr, Y.H., Kidd, R., Vereecken, H., 2014. Catchment scale validation of SMOS and ASCAT soil moisture products using hydrological modeling and temporal stability analysis. Journal of Hydrology, 519, Part A: 934-946.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry, 36(8): 1627-1639.

Schaap, M.G., Leij, F.J., van Genuchten, M.T., 2001. rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. Journal of Hydrology, 251(3–4): 163-176.

Schaefer, G.L., Cosh, M.H., Jackson, T.J., 2007. The USDA natural resources conservation service soil climate analysis network (SCAN). Journal of Atmospheric and Oceanic Technology, 24(12): 2073-2077.

Schattan, P., G. Baroni, S. E. Oswald, J. Schöber, C. Fey, C. Kormann, M. Huttenlau, and S. Achleitner, 2017. Continuous monitoring of snowpack dynamics in alpine terrain by aboveground neutron sensing, Water Resources Research., 53, 3615–3634.

Schiedung, H., 2015. Soil data Rollesbroich. CRC/TR32 Database (TR32DB). Accessed from http://tr32db.uni-koeln.de/data.php?dataID=1153 at 2015-09-29.

Schöniger, A., Nowak, W., Hendricks Franssen, H.J., 2012. Parameter estimation by ensemble Kalman filters with transformed data: Approach and application to hydraulic tomography. Water Resources Research, 48(4): W04502.

Schotanus, P., Nieuwstadt, F.T.M., Debruin, H.A.R., 1983. Temperatur-measurements with a sonic anemometer and its application to heat and moisture fluxes. Bound-Lay Meteorol, 26(1): 81-93.

Schrader, F., Durner, W., Fank, J., Gebler, S., Pütz, T., Hannes, M., Wollschläger, U., 2013. Estimating Precipitation and Actual Evapotranspiration from Precision Lysimeter Measurements. Procedia Environmental Sciences, 19(0): 543-552.

Schröter, I., Paasche, H., Dietrich, P., Wollschläger, U., 2015. Estimation of Catchment-Scale Soil Moisture Patterns Based on Terrain Data and Sparse TDR Measurements Using a Fuzzy C-Means Clustering Approach. Vadose Zone Journal, 14(11): vzj2015.01.0008.

Schume, H., Jost, G., Hager, H., 2004. Soil water depletion and recharge patterns in mixed and pure forest stands of European beech and Norway spruce. Journal of Hydrology, 289(1-4): 258-274.

Schume, H., Jost, G., Katzensteiner, K., 2003. Spatio-temporal analysis of the soil water content in a mixed Norway spruce (Picea abies (L.) Karst.)–European beech (Fagus sylvatica L.) stand. Geoderma, 112(3–4): 273-287.

Sciuto, G., Diekkrüger, B., 2010. Influence of soil heterogeneity and spatial discretization on catchment water balance modeling. Vadose Zone Journal, 9(4): 955-969.

Scott, R.L., 2010. Using watershed water balance to evaluate the accuracy of eddy covariance evaporation measurements for three semiarid ecosystems. Agricultural and Forest Meteorology, 150(2): 219-225.

Serreze, M.C., Clark, M.P., Armstrong, R.L., McGinnis, D.A., Pulwarty, R.S., 1999. Characteristics of the western United States snowpack from snowpack telemetry(SNOTEL) data. Water Resources Research, 35(7): 2145-2160.

Sevruk, B., 1981. Methodische Untersuchungen des systematischen Messfehlers der Hellmann-Regenmesser im Sommerhalbjahr in der Schweiz. Mitteilungen Nr. 52, Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie, Eidgenössische Technische Hochschule Zürich, 299 pp.

Sevruk, B., 1989. Wind induced measurement error for high-intensity rains. Proc. International Workshop on Precipitation Measurement, WMO Tech. Document 328, St. Moritz, Switzerland, 199–204.

Sevruk, B., 1996. Adjustment of tipping-bucket precipitation gauge measurements. Atmospheric Research, 42(1–4): 237-246.

Sevruk, B., Ondrás, M., Chvíla, B., 2009. The WMO precipitation measurement intercomparisons. Atmospheric Research, 92(3): 376-380.

Shi, Y., Davis, K.J., Zhang, F., Duffy, C.J., Yu, X., 2014. Parameter estimation of a physically based land surface hydrologic model using the ensemble Kalman filter: A synthetic experiment. Water Resources Research, 50(1): 706-724.

Shi, Y., Davis, K.J., Zhang, F., Duffy, C.J., Yu, X., 2015. Parameter estimation of a physically-based land surface hydrologic model using an ensemble Kalman filter: A multivariate real-data experiment. Advances in Water Resources, 83: 421-427.

Shrestha, P., Sulis, M., Masbou, M., Kollet, S., Simmer, C., 2014. A Scale-Consistent Terrestrial Systems Modeling Platform Based on COSMO, CLM, and ParFlow. Monthly Weather Review, 142(9): 3466-3483.

Shrestha, P., Sulis, M., Simmer, C., Kollet, S., 2015. Impacts of grid resolution on surface energy fluxes simulated with an integrated surface-groundwater flow model. Hydrol. Earth Syst. Sci., 19(10): 4317-4326.

Si, B.C., 2008. Spatial Scaling Analyses of Soil Physical Properties: A Review of Spectral and Wavelet Methods. Vadose Zone Journal, 7(2): vzj2007.0040.

Simmer, C., Thiele-Eich, I., Masbou, M., Amelung, W., Bogena, H., Crewell, S., Diekkrüger, B., Ewert, F., Franssen, H.-J.H., Huisman, J.A., Kemna, A., Klitzsch, N., Kollet, S., Langensiepen, M., Löhnert, U., Rahman, A.S.M.M., Rascher, U., Schneider, K., Schween, J., Shao, Y., Shrestha, P., Stiebler, M., Sulis, M., Vanderborght, J., Vereecken, H., Kruk, J.v.d., Waldhoff, G., Zerenner, T., 2015. Monitoring and Modeling the Terrestrial System from Pores to Catchments: The Transregional Collaborative Research Center on Patterns in the Soil–Vegetation–Atmosphere System. Bulletin of the American Meteorological Society, 96(10): 1765-1787.

Šimůnek, J., K. Huang, and M. Th. van Genuchten, 1995. The SWMS 3D code for simulating water flow and solute transport in three-dimensional variably saturated media. Version 1.0, Research Report No. 139, U.S. Salinity Laboratory, USDA, ARS, Riverside, California, 155.

Šimůnek, J., van Genuchten, M.T., Šejna, M., 2008. Development and applications of the HYDRUS and STANMOD software packages and related codes. Vadose Zone Journal, 7(2): 587-600.

Smirnova, T.G., Brown, J.M., Benjamin, S.G., 1997. Performance of Different Soil Model Configurations in Simulating Ground Surface Temperature and Surface Fluxes. Monthly Weather Review, 125(8): 1870-1884.

Stark, C.P., Foufoula-Georgiou, E., Ganti, V., 2009. A nonlocal theory of sediment buffering and bedrock channel evolution. J Geophys Res-Earth, 114: JF000981.

Stisen, S., Højberg, A.L., Troldborg, L., Refsgaard, J.C., Christensen, B.S.B., Olsen, M., Henriksen, H.J., 2012. On the importance of appropriate precipitation gauge catch correction for hydrological modelling at mid to high latitudes. Hydrol Earth Syst Sc, 16(11): 4157-4176.

Strangeways, I., 2010. A history of rain gauges. Weather, 65(5): 133-138.

Su, H., Yang, Z.-L., Dickinson, R.E., Wilson, C.R., Niu, G.-Y., 2010. Multisensor snow data assimilation at the continental scale: The value of Gravity Recovery and Climate Experiment terrestrial water storage information. Journal of Geophysical Research: Atmospheres, 115(D10): D013035.

Sulis, M., Paniconi, C., Camporese, M., 2011. Impact of grid resolution on the integrated and distributed response of a coupled surface-subsurface hydrological model for the des Anglais catchment, Quebec. Hydrological Processes, 25(12): 1853-1865.

Teuling, A.J., Troch, P.A., 2005. Improved understanding of soil moisture variability dynamics. Geophys Res Lett, 32(5): L05404.

Tian, W., Li, X., Cheng, G.D., Wang, X.S., Hu, B.X., 2012. Coupling a groundwater model with a land surface model to improve water and energy cycle simulation. Hydrol. Earth Syst. Sci., 16(12): 4707-4723.

Therrien, R., McLaren, R., Sudicky, E., Panday, S., 2010. HydroGeoSphere: A three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport. Groundwater Simulations Group, University of Waterloo, Waterloo, ON.

Tippett, M.K., Anderson, J.L., Bishop, C.H., Hamill, T.M., Whitaker, J.S., 2003. Ensemble Square Root Filters. Monthly Weather Review, 131(7): 1485-1490.

Torrence, C., Compo, G.P., 1998. A Practical Guide to Wavelet Analysis. Bulletin of the American Meteorological society 79(1): 61-78.

Tran, A.P., Vanclooster, M., Zupanski, M., Lambot, S., 2014. Joint estimation of soil moisture profile and hydraulic parameters by ground-penetrating radar data assimilation with maximum likelihood ensemble filter. Water resources research, 50(4): 3131-3146.

Twine, T.E., Kustas, W.P., Norman, J.M., Cook, D.R., Houser, P.R., Meyers, T.P., Prueger, J.H., Starks, P.J., Wesely, M.L., 2000. Correcting eddy-covariance flux underestimates over a grassland. Agricultural and Forest Meteorology, 103(3): 279-300.

Unold, G., Fank, J., 2008. Modular Design of Field Lysimeters for Specific Application Needs. Water Air Soil Pollut: Focus, 8(2): 233-242.

Valcke, S., 2013. The OASIS3 coupler: a European climate modelling community software. Geosci. Model Dev., 6(2): 373-388.

Valcke, S., Tony C., Coquart, L., 2013. OASIS3-MCT User Guide, OASIS3-MCT 2.0. CERFACS/CNRS SUC URA 1875.

van Dam, J.C., Feddes, R.A., 2000. Numerical simulation of infiltration, evaporation and shallow groundwater levels with the Richards equation. Journal of Hydrology, 233(1-4): 72-85.

VanderKwaak, J.E., Loague, K., 2001. Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model. Water Resources Research, 37(4): 999-1013.

van Genuchten, M.T., 1980. A Closed-form Equation For Predicting The Hydraulic Conductivity Of Unsaturated Soils. Soil Sci. Soc. Am. J., 44(5): 892-898.

Vaughan, P.J., Trout, T.J., Ayars, J.E., 2007. A processing method for weighing lysimeter data and comparison to micrometeorological ETo predictions. Agricultural Water Management, 88(1-3): 141-146.

Vaughan, P., Ayars, J., 2009. Noise Reduction Methods for Weighing Lysimeters. Journal of Irrigation and Drainage Engineering, 135(2): 235-240.

Vázquez, R.F., Willems, P., Feyen, J., 2008. Improving the predictions of a MIKE SHE catchment-scale application by using a multi-criteria approach. Hydrological Processes, 22(13): 2159-2179.

Vereecken, H., Kamai, T., Harter, T., Kasteel, R., Hopmans, J., Vanderborght, J., 2007. Explaining soil moisture variability as a function of mean soil moisture: A stochastic unsaturated flow perspective. Geophys Res Lett, 34(22): GL031813.

Vereecken, H., Huisman, J.A., Bogena, H., Vanderborght, J., Vrugt, J.A., Hopmans, J.W., 2008. On the value of soil moisture measurements in vadose zone hydrology: A review. Water Resources Research, 44(4): W00D06.

Vereecken, H., Schnepf, A., Hopmans, J.W., Javaux, M., Or, D., Roose, T., Vanderborght, J., Young, M.H., Amelung, W., Aitkenhead, M., Allison, S.D., Assouline, S., Baveye, P., Berli, M., Brüggemann, N., Finke, P., Flury, M., Gaiser, T., Govers, G., Ghezzehei, T., Hallett, P., Hendricks Franssen, H.J., Heppell, J., Horn, R., Huisman, J.A., Jacques, D., Jonard, F., Kollet, S., Lafolie, F., Lamorski, K., Leitner, D., McBratney, A., Minasny, B., Montzka, C., Nowak, W., Pachepsky, Y., Padarian, J., Romano, N., Roth, K., Rothfuss, Y., Rowe, E.C., Schwen, A., Šimůnek, J., Tiktak, A., Van Dam, J., van der Zee, S.E.A.T.M., Vogel, H.J., Vrugt, J.A., Wöhling, T., Young, I.M., 2016. Modeling Soil Processes: Review, Key Challenges, and New Perspectives. Vadose Zone Journal, 15(5): vzj2015.09.0131.

Vivoni, E.R., Rodríguez, J.C., Watts, C.J., 2010. On the spatiotemporal variability of soil moisture and evapotranspiration in a mountainous basin within the North American monsoon region. Water Resources Research, 46(2): W02509.

Vrugt, J.A., Gupta, H.V., Bouten, W., Sorooshian, S., 2003. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resources Research, 39(8): WR001642.

Vrugt, J.A., Diks, C.G.H., Gupta, H.V., Bouten, W., Verstraten, J.M., 2005. Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. Water Resources Research, 41(1): W01017.

Wagener, T., Gupta, H.V., 2005. Model identification for hydrological forecasting under uncertainty. Stochastic Environmental Research and Risk Assessment, 19(6): 378-387.

Walker, J.P., Willgoose, G.R., Kalma, J.D., 2004. In situ measurement of soil moisture: a comparison of techniques. Journal of Hydrology, 293(1): 85-99.

Wang, T., Franz, T.E., 2015. Field Observations of Regional Controls of Soil Hydraulic Properties on Soil Moisture Spatial Variability in Different Climate Zones. Vadose Zone Journal, 14(8): vzj2015.02.0032.

Wang, X., Hamill, T.M., Whitaker, J.S., Bishop, C.H., 2007. A Comparison of Hybrid Ensemble Transform Kalman Filter–Optimum Interpolation and Ensemble Square Root Filter Analysis Schemes. Monthly Weather Review, 135(3): 1055-1076.

Webb, E.K., Pearman, G.I., Leuning, R., 1980. Correction of flux measurements for density effects due to heat and water-vapor transfer. Quarterly Journal of the Royal Meteorological Society, 106(447): 85-100.

Wen, X.-H., Chen, W.H., 2007. Some Practical Issues on Real-Time Reservoir Model Updating Using Ensemble Kalman Filter. International Petroleum Technology Conference. International Petroleum Technology Conference.

Wegehenkel, M., Gerke Horst, H., 2013. Comparison of real evapotranspiration measured by weighing lysimeters with simulations based on the Penman formula and a crop growth model, Journal of Hydrology and Hydromechanics,61(2): johh-2013-0021.

Weiler, M., Naef, F., 2003. An experimental tracer study of the role of macropores in infiltration in grassland soils. Hydrological Processes, 17(2): 477-493.

Weill, S., Mazzia, A., Putti, M., Paniconi, C., 2011. Coupling water flow and solute transport into a physically-based surface-subsurface hydrological model. Advances in Water Resources, 34(1): 128-136.

Western, A.W., Bloschl, G., Grayson, R.B., 1998. Geostatistical characterisation of soil moisture patterns in the Tarrawarra a catchment. Journal of Hydrology, 205(1-2): 20-37.

Western, A.W., Grayson, R.B., Blöschl, G., 2002. Scaling of soil moisture: A hydrologic perspective. Annual Review of Earth and Planetary Sciences, 30(1): 149-180.

Western, A.W., Zhou, S.L., Grayson, R.B., McMahon, T.A., Bloschl, G., Wilson, D.J., 2004. Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes. Journal of Hydrology, 286(1-4): 113-134.

Whelan, B., McBratney, A., Minasny, B., 2002. Vesper 1.5–spatial prediction software for precision agriculture, Precision Agriculture, Proc. 6th Int. Conf. on Precision Agriculture, ASA/CSSA/SSSA, Madison, WI, United States.

Wiekenkamp, I., Huisman, J.A., Bogena, H.R., Lin, H.S., Vereecken, H., 2016. Spatial and temporal occurrence of preferential flow in a forested headwater catchment. Journal of Hydrology, 534: 139-149.

Wilczak, J.M., Oncley, S.P., Stage, S.A., 2001. Sonic anemometer tilt correction algorithms. Bound-Lay Meteorol, 99(1): 127-150.

Williams, J.L., Maxwell, R.M., 2011. Propagating Subsurface Uncertainty to the Atmosphere Using Fully Coupled Stochastic Simulations. Journal of Hydrometeorology, 12(4): 690-701.

Wilson, K.B., Hanson, P.J., Mulholland, P.J., Baldocchi, D.D., Wullschleger, S.D., 2001. A comparison of methods for determining forest evapotranspiration and its components: sap-flow, soil water budget, eddy covariance and catchment water balance. Agricultural and Forest Meteorology, 106(2): 153-168.

Wohlfahrt, G., Irschick, C., Thalinger, B., Hörtnagl, L., Obojes, N., Hammerle, A., 2010. Insights from Independent Evapotranspiration Estimates for Closing the Energy Balance: A Grassland Case Study. Vadose Zone Journal, 9(4): 1025-1033.

Wohlfahrt, G., Widmoser, P., 2013. Can an energy balance model provide additional constraints on how to close the energy imbalance? Agricultural and Forest Meteorology, 169: 85-91.

Wösten, J.H.M., Pachepsky, Y.A., Rawls, W.J., 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. Journal of Hydrology, 251(3–4): 123-150.

Wood, E.F., Roundy, J.K., Troy, T.J., van Beek, L.P.H., Bierkens, M.F.P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P.R., Kollet, S., Lehner, B., Lettenmaier, D.P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., Whitehead, P., 2011. Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. Water Resources Research, 47(5): W05301.

Wu, C., Margulis S.A., 2011. Feasibility of real-time soil state and flux characterization for wastewater reuse using an embedded sensor net-work data assimilation approach, Journal of Hydrology, 399(3–4), 313–325.

Wu, C., Margulis S.A., 2013. Real-Time Soil Moisture and Salinity Profile Estimation Using Assimilation of Embedded Sensor Datastreams. Vadose Zone Journal, 12(1): vzj2011.017.

Xie, X., Zhang, D., 2010. Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. Advances in Water Resources, 33(6): 678-690.

Xu, J., Shu, H., 2014. Assimilating MODIS-based albedo and snow cover fraction into the Common Land Model to improve snow depth simulation with direct insertion and deterministic ensemble Kalman filter methods. Journal of Geophysical Research: Atmospheres, 119(18): 10684-10701.

Yang, D., Goodison, B.E., Metcalfe, J.R., Golubev, V.S., Bates, R., Pangburn, T., Hanson, C.L., 1998. Accuracy of NWS 8" Standard Nonrecording Precipitation Gauge: Results and Application of WMO Intercomparison. Journal of Atmospheric and Oceanic Technology, 15(1): 54-68.

Yeh, P.J.F., Eltahir, E.A.B., 2005. Representation of Water Table Dynamics in a Land Surface Scheme. Part I: Model Development. Journal of Climate, 18(12): 1861-1880.

York, J.P., Person, M., Gutowski, W.J., Winter, T.C., 2002. Putting aquifers into atmospheric simulation models: An example from the Mill Creek Watershed, northeastern Kansas. Advances in Water Resources, 25(2): 221-238.

Zacharias, S., Bogena, H., Samaniego, L., Mauder, M., Fuss, R., Puetz, T., Frenzel, M., Schwank, M., Baessler, C., Butterbach-Bahl, K., Bens, O., Borg, E., Brauer, A., Dietrich, P., Hajnsek, I., Helle, G., Kiese, R., Kunstmann, H., Klotz, S., Munch, J.C., Papen, H., Priesack, E., Schmid, H.P., Steinbrecher, R., Rosenbaum, U., Teutsch, G., Vereecken, H., 2011. A Network of Terrestrial Environmental Observatories in Germany. Vadose Zone Journal, 10(3): 955-973.

Zeng, X., Barlage, M., Dickinson, R.E., Dai, Y., Wang, G., Oleson, K., 2005. Treatment of Undercanopy Turbulence in Land Models. Journal of Climate, 18(23): 5086-5094.

Zeng, X.B., Zhao, M., Dickinson, R.E., 1998. Intercomparison of bulk aerodynamic algorithms for the computation of sea surface fluxes using TOGA COARE and TAO data. Journal of Climate, 11(10): 2628-2644.

Zenker, T., 2003. Verdunstungswiderstände und Gras-Referenzverdunstung: Lysimeteruntersuchungen zum Penman-Monteith-Ansatz im Berliner Raum. PhD thesis, Berlin University of Technology, Germany.

Zhang, D., Lu, Z., Chen, Y., 2007. Dynamic Reservoir Data Assimilation With an Efficient, Dimension-Reduced Kalman Filter. Society of Petroleum Engineers.

Zhang, D., Madsen, H., Ridler, M.E., Refsgaard, J.C., Jensen, K.H., 2015. Impact of uncertainty description on assimilating hydraulic head in the MIKE SHE distributed hydrological model. Advances in Water Resources, 86: 400-413.

Zhang, H., Kurtz, W., Kollet, S., Vereecken, H., Hendricks Franssen, H.-J., 2018. Comparison of different assimilation methodologies of groundwater levels to improve predictions of root zone soil moisture with an integrated terrestrial system model. Advances in Water Resources, 111, 224-238.

Zhu, J., Mohanty, B.P., 2002. Spatial Averaging of van Genuchten Hydraulic Parameters for Steady-State Flow in Heterogeneous Soils. Vadose Zone Journal, 1(2): 261-272.

Zhou, H., Gómez-Hernández, J.J., Franssen, H.-J.H., Li, L., 2011. An approach to handling non-Gaussianity of parameters and state variables in ensemble Kalman filtering. Advances in Water Resources, 34(7): 844-864.

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment

Band / Volume 430
Der Einfluss von Wasserdampf auf den Sauerstofftransport in keramischen Hochtemperaturmembranen
F. Thaler (2018), ii, 93, XXXI pp
ISBN: 978-3-95806-340-2

Band / Volume 431
Analysis & modeling of metastable photovoltaic technologies: towards dynamic photovoltaic performance models
M. Görig (2018), 246 pp
ISBN: 978-3-95806-342-6

Band / Volume 432
Laser Treatment of Silicon Thin-Films for Photovoltaic Applications
C. Maurer (2018), vii, 165 pp
ISBN: 978-3-95806-347-1

Band / Volume 433
Mentalitäten und Verhaltensmuster im Kontext der Energiewende in NRW
K. Schürmann & D. Schumann (Hrsg.) (2018), 236 pp
ISBN: 978-3-95806-349-5

Band / Volume 434
Adhäsionsverhalten von wässrigen Nafion-Lösungen an dispersen Phasengrenzen
A. Schulz (2018), xii, 129 pp
ISBN: 978-3-95806-354-9

Band / Volume 435
Alterungs- und fehlertolerante optimale Betriebsführung eines Direktmethanol-Brennstoffzellensystems
R. Keller (2018), XX, 175 pp
ISBN: 978-3-95806-355-6

Band / Volume 436
Chamber study of biogenic volatile organic compounds: plant emission, oxidation products and their OH reactivity
Y. Zhujun (2018), ix, 139 pp
ISBN: 978-3-95806-356-3

Band / Volume 437
Characterization of High Temperature Polymer Electrolyte Fuel Cells
Y. Rahim (2018), iii, 162 pp
ISBN: 978-3-95806-359-4

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment

Band / Volume 438
Lattice Boltzmann Simulation in Components of Polymer Electrolyte Fuel Cell
J. Yu (2018), ii, 173 pp
ISBN: 978-3-95806-360-0

Band / Volume 439
Quantitative Luminescence Imaging of Solar Cells
V. Huhn (2018), 155 pp
ISBN: 978-3-95806-363-1

Band / Volume 440
Characterization of Phosphoric Acid Doped Polybenzimidazole Membranes
Y. Lin (2018), II, IV, 140 pp
ISBN: 978-3-95806-364-8

Band / Volume 441
Degradation Study of SOC Stacks with Impedance Spectroscopy
Y. Yan (2018), 135 pp
ISBN: 978-3-95806-367-9

Band / Volume 442
Future Grid Load of the Residential Building Sector
L. Kotzur (2018), xix, 213 pp
ISBN: 978-3-95806-370-9

Band / Volume 443
Yttriumoxid-Dünnschichten als Tritium-Permeationsbarriere
J. Engels (2018), 252 pp
ISBN: 978-3-95806-371-6

Band / Volume 444
Inverse conditioning of a high resolution integrated terrestrial model at the hillslope scale: the role of input data quality and model structural errors
S. Gebler (2018), xxii, 160 pp
ISBN: 978-3-95806-372-3

Energie & Umwelt / Energy & Environment
Band / Volume 444
ISBN 978-3-95806-372-3

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
Forschungszentrum