

## Assimilation of groundwater level and cosmic-ray neutron sensor soil moisture measurements into integrated terrestrial system models for better predictions

Fang Li

Energie & Umwelt / Energy & Environment

Band / Volume 650

ISBN 978-3-95806-796-7





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

# **Assimilation of groundwater level and cosmic-ray neutron sensor soil moisture measurements into integrated terrestrial system models for better predictionst**

Fang Li

Schriften des Forschungszentrums Jülich  
Reihe Energie & Umwelt / Energy & Environment

Band / Volume 650

---

ISSN 1866-1793

ISBN 978-3-95806-796-7

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](mailto:zb-publikation@fz-juelich.de)  
[www.fz-juelich.de/zb](http://www.fz-juelich.de/zb)

Umschlaggestaltung: Grafische Medien, Forschungszentrum Jülich GmbH

Druck: Grafische Medien, Forschungszentrum Jülich GmbH

Copyright: Forschungszentrum Jülich 2024

Schriften des Forschungszentrums Jülich  
Reihe Energie & Umwelt / Energy & Environment, Band / Volume 650

D 82 (Diss. RWTH Aachen University, 2024)

ISSN 1866-1793  
ISBN 978-3-95806-796-7

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



This is an Open Access publication distributed under the terms of the [Creative Commons Attribution License 4.0](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Contents

Contents .....	i
List of Figures .....	iv
List of Tables .....	xi
List of Acronyms .....	xv
Abstract .....	1
Zusammenfassung .....	3
Chapter 1: Introduction .....	5
Chapter 2: Theory, methods and materials .....	15
2.1 Integrated terrestrial systems model TSMP .....	15
2.1.1 Land surface model Community Land Model (CLM), version 3.5 .....	15
2.1.2 Subsurface hydrological model ParFlow .....	18
2.1.3 Coupling interface OASIS-MCT .....	20
2.2 Data assimilation .....	20
2.2.1 Ensemble Kalman Filter (EnKF) .....	20
2.2.2 Localized Ensemble Kalman Filter (LEnKF) .....	23
2.2.3 TSMP-PDAF .....	24
2.3 Study area and hydrological measurements .....	26
Chapter 3: Water table depth assimilation in integrated terrestrial system models at the larger catchment scale .....	29
3.1 Introduction .....	29
3.2 Materials and methods .....	31
3.2.1 Study area and data .....	31
3.2.2 Model description (TSMP) .....	35
3.2.3 LEnKF methodology .....	38
3.2.4 Assimilation methodology .....	41
3.3 Experimental setup .....	42
3.3.1 Ensemble generation and simulations .....	42
3.3.2 Selection criteria for assimilated sites .....	44
3.3.3 Evaluation of model performance .....	45

3.4 Results and discussion.....	47
3.4.1 Water table depth .....	47
3.4.2 Soil moisture.....	53
3.4.3 Discussion.....	55
3.5 Conclusions.....	57
Chapter 4: Can a sparse network of cosmic ray neutron sensors improve soil moisture and evapotranspiration estimation at the larger catchment scale? .....	58
4.1 Introduction.....	58
4.2 Materials and methods.....	60
4.2.1 Study area.....	60
4.2.2 Terrestrial System Modeling Platform (TSMP).....	61
4.2.3 Data .....	63
4.2.4 Data assimilation methodology.....	66
4.3 Model and experiment setup .....	69
4.3.1 TSMP-PDAF setup.....	69
4.3.2 Ensemble generation.....	70
4.3.3 Setup of the DA Experiments .....	71
4.3.4 Evaluation of model performance.....	72
4.4 Results.....	74
4.4.1 Soil moisture data assimilation general results .....	74
4.4.2 Jackknife simulations.....	76
4.4.3 Temporal evolution of parameter estimates and parameter verification.....	78
4.4.4 Evapotranspiration and discharge .....	81
4.4.5 Discussion.....	85
4.5 Conclusions.....	89
Appendix A .....	90
Chapter 5: A new approach for joint assimilation of cosmic-ray neutron soil moisture and groundwater level data into an integrated terrestrial model.....	100
5.1 Introduction.....	100

5.2 Materials and methods.....	104
5.2.1 Study area .....	104
5.2.2 Terrestrial System Modeling Platform (TSMP).....	105
5.2.3 Model input data and measurements.....	107
5.2.4 Data assimilation: localized EnKF.....	110
5.3 Model and Experiment Setup.....	113
5.3.1 Ensemble generation.....	113
5.3.2 Setup of data assimilation experiments.....	115
5.3.3 Evaluation of model performance.....	116
5.4 Results.....	117
5.4.1 Univariate data assimilation of soil moisture.....	117
5.4.2 Univariate data assimilation of groundwater level.....	119
5.4.3 Multivariate data assimilation of groundwater level and soil moisture.....	122
5.4.4 Influence of $K_s$ updates on simulations.....	129
5.5 Discussions.....	131
5.5.1 Strengths and limitations of new multivariate data assimilation approach .....	131
5.5.2 Uncertainties and potential improvement .....	133
5.6 Conclusions.....	135
Appendix B .....	136
Chapter 6: Summary and outlook .....	151
Bibliography .....	159
Acknowledgements.....	171

## List of Figures

- Figure 2.1 Configuration of the CLM subgrid hierarchy, adapted from Oleson et al. (2004). Each grid cell is composed of multiple landunits, snow/soil columns, and Plant Functional Types (PFTs). The first subgrid level is the landunit, which can be glacier, wetland, vegetated, lake, or urban. The second subgrid level is the column, which can be soil or snow within a single landunit. The third subgrid level is the PFT level, where up to 4 of 15 possible PFTs that differ in physiology and structure can coexist on a single column..... 16
- Figure 2.2 Topography of the Rur catchment (a) and locations of the hydrological stations (b), including groundwater wells, cosmic-ray neutron sensors, and eddy covariance stations..... 27
- Figure 3.1 Map of the Rur catchment and locations of the 13 cosmic-ray neutron sensors (CRNS) (black points) and groundwater measurement sites (red points) in the year 2018. The Rur catchment is situated in western Germany..... 32
- Figure 3.2 Sand (a) and clay (b) content (%) for the Rur catchment derived from the BK50 soil map. .... 33
- Figure 3.3 Hydraulic conductivity of the aquifer material for the Rur catchment. .... 33
- Figure 3.4 Illustration of the link between groundwater level observation and data to be assimilated (revised from Zhang et al. (2018)). The blue colour indicates the groundwater level at layer  $i-1$ . The red layers (from layer  $i$  to the bottom layer) are saturated and are incorporated as groundwater observations, and converted to pressure heads assuming hydrostatic conditions. .... 41
- Figure 3.5 Spatial autocorrelation functions of measured and simulated (open loop) groundwater table depth for the year 2018 (a), the 26<sup>th</sup> of February 2018 (b), and the 24<sup>th</sup> of September 2018 (c). Filled squares or circles indicate autocorrelation coefficients are significantly different from zero ( $p < 0.05$ ). When the number of comparison pairs was smaller than 1000, the number of comparison pairs is indicated next to the marker. .... 47
- Figure 3.6 The locations of the assimilated groundwater sites (dots) in the Rur catchment together with the average groundwater table depth (a); time series of root mean square error (RMSE) of water table depth at measurement locations for the year 2018 for the open loop and data assimilation runs (b); the histogram of the water table depth errors

at the measurement locations for the year 2018 from the open loop and data assimilation runs (10 km (c), 5 km (d), 2.5 km (e)).....	48
Figure 3.7 Water table depth time series for 10 assimilation sites: observations (Obs, red), ensemble mean of open loop (OL, blue) and ensemble mean of data assimilation run with 10 km localization radius (DA, green) for the year 2018.....	50
Figure 3.8 Difference in average water table depth between data assimilation and open loop runs (data assimilation - open loop) for different localization radii (10 km (a), 5 km (b), 2.5 km (c)) on the 31 <sup>st</sup> of December 2018 for the Rur catchment; and difference in standard deviation for data assimilation and open loop runs for different localization radii (10 km (d), 5 km (e), 2.5 km (f)) on the 31 <sup>st</sup> of December 2018 for the Rur catchment.....	51
Figure 3.9 Time series of RMSE of groundwater table depth for the open loop (OL) and data assimilation (DA) runs (10 km (a, d, f), 5 km (b,e), 2.5 km (c) localization radius) at verification locations which were 0~2.5 km, 2.5~5 km and 5~10 km away from assimilated observations.....	52
Figure 3.10 Soil moisture time series from cosmic-ray neutron sensors (CRNS) (red), ensemble mean of open loop (OL, grey), and ensemble mean of data assimilation with 10 km assimilation radius (DA 10 km, blue) for the year 2018.....	54
Figure 4.1 Map of the Rur catchment with the altitude above sea level and the locations of the cosmic-ray neutron sensors, eddy covariance stations and discharge station. The Rur catchment is situated in western Germany.....	61
Figure 4.2 Sand (a) and clay content (b) for the Rur catchment derived from the BK50 soil map. ....	64
Figure 4.3 Hydraulic conductivity of the bedrock for the Rur catchment.....	64
Figure 4.4 Schematic overview of the assimilation of soil water content from CRNS with PDAF into TSMP (CLM-ParFlow). The flows represented by the red dashed line are outside TSMP-PDAF, including the weighting calculation and the comparison of CRNS soil moisture with the simulations.....	69
Figure 4.5 Examples of the simulated soil moisture distribution over the Rur catchment on the 22 <sup>nd</sup> of July in 2016 and 2018. Subplots a) and d) are from the open loop, b) and e) are from data assimilation with state update, and c) and f) are from joint state-	

parameter update simulations.....	76
Figure 4.6 Temporal evolution of simulated soil moisture from the open loop mean (OL, blue) and jackknife simulation mean (DA, green), together with the observed soil moisture from the CRNS (red), for 2016 (a, b) and 2018 (c, d) at the CRNS sites. Simulated soil moisture was vertically weighted using the revised method by Schrön et al. (2017). .....	78
Figure 4.7 Estimates of averaged saturated hydraulic conductivity ( $\log_{10}K_s$ ) from data assimilation experiments with joint state-parameter updating during the periods of 2016 and 2018 at CRNS locations. The input value of $K_s$ is indicated at the first time step.....	79
Figure 4.8 Ensemble averaged $\log_{10}K_s$ fields of the soil at 2 cm depth: (a) prior field; (b) DA with joint state-parameter updates at the end of 2016; (c) DA with joint state-parameter updates at the end of 2018. The black asterisk is the location of the CRNS sites.....	80
Figure 4.9 Temporal evolution of simulated evapotranspiration from open loop (OL, blue), data assimilation (joint state-parameter updates, DA, green), and the observed evapotranspiration (red) at the sites Rollesbroich, Wüstebach, and Selhausen for the assimilation periods of 2016 (first row) and 2018 (second row). Monthly Leaf Area Index (LAI) for the plant functional types at the sites Rollesbroich (grassland), Wüstebach (needle leaf forest), and Selhausen (cropland), as well as the available daily LAI measurements from 2016 and 2018 (third row).....	83
Figure 4.10 Annual evapotranspiration from open loop (a, c) and data assimilation runs (joint state-parameter updates) (b, d) over the Rur catchment during the assimilation periods. ....	84
Figure 4.11 Temporal evolution of simulated discharge from the open loop mean (OL, blue circle) and joint state-parameter assimilation mean (DA, green circle), together with the observed discharge (black circle) for 2016 and 2018 at the Erkensruhr-Einruhr in situ station. The temporal evolution of simulated soil moisture from the open loop mean (OL, blue line) and joint state-parameter assimilation mean (DA, green line), together with the observations (red dot) for 2016 and 2018 at the Wüstebach site.	85
Figure A.1 Temporal evolution of mean simulated soil moisture from the open loop (OL, blue), joint state-parameter estimation (DA, green), together with observed soil moisture	

from CRNS (red), for the year 2016 at the CRNS sites. Simulated soil moisture was vertically weighted using the revised method. ....	93
Figure A.2 Temporal evolution of mean simulated soil moisture from the open loop (OL, blue), joint state-parameter estimation (DA, green), together with observed soil moisture from CRNS (red), for the year 2018 at the CRNS sites. Simulated soil moisture was vertically weighted using the revised method. ....	94
Figure A.3 Soil moisture scatter plots for CRNS observations versus ensemble mean soil moisture from the open loop run (OL, blue) and ensemble mean soil moisture from joint state-parameter estimation (DA, red) for 2016. ....	95
Figure A.4 Soil moisture scatter plots for CRNS observations versus ensemble mean soil moisture from the open loop run (OL, blue) and ensemble mean soil moisture from joint state-parameter estimation (DA, red) for 2018. ....	96
Figure A.5 Temporal evolution of mean simulated soil moisture from the open loop run (OL, blue), jackknife simulations (DA, green), together with the observed soil moisture from CRNS (red) for the year 2016 at the CRNS sites. Simulated soil moisture was vertically weighted using the revised method. ....	97
Figure A.6 Temporal evolution of mean simulated soil moisture from the open loop run (OL, blue), jackknife simulations (DA, green), together with the observed soil moisture from CRNS (red) for the year 2018 at the CRNS sites. Simulated soil moisture was vertically weighted using the revised method. ....	98
Figure A.7 Examples of the spatial correlations of soil moisture between CRNS sites and other grid cells over the Rur catchment, for the open loop run. Subplots a) and d) are from Gevenich on the 29 <sup>th</sup> of June in 2016 and 2018, b) and e) are from Heinsberg on the 2 <sup>nd</sup> of August in 2016 and 2018, and c) and f) are from Schöenseiffen on the 28 <sup>th</sup> of July in 2016 and 2018. The black asterisk is the location of the CRNS sites.....	99
Figure 5.1 Topography of the Rur catchment (a) and locations of the hydrological stations (b), including groundwater wells, cosmic-ray neutron sensors, and eddy covariance stations.....	105
Figure 5.2 Sand (a) and clay (b) content, hydraulic conductivity (c) of the bedrock for the Rur catchment.....	108
Figure 5.3 Schematic overview of the assimilation of soil moisture from CRNS and	

groundwater (pressure head) with PDAF into TSMP (CLM-ParFlow).  $\theta^f$  and  $\theta^a$  are the forecasted and analyzed soil moisture in the unsaturated zone;  $h^f$  and  $h^a$  are the forecasted and analyzed pressure heads in the saturated zone. The observed pressure heads are derived from the measured groundwater levels..... 113

Figure 5.4 Subplots (a) to (c) are the differences (SM\_DA - OL) of annual SM (0-80 cm), ET, and GWL in 2018. Subplots (d) to (f) are the differences (SM\_DA\_PAR - OL) of annual SM (0-80 cm), ET, and GWL in 2018. The locations of the CRNS stations are indicated by the black pentagrams..... 119

Figure 5.5 Subplots (a) to (c) are the differences (GWL\_DA - OL) of annual GWL, SM (0-80 cm), and ET in 2018. Subplots (d) to (f) are the differences (GWL\_DA\_PAR - OL) of annual GWL, SM (0-80 cm), and ET in 2018. The locations of the assimilated groundwater sites are indicated by the black circles. .... 122

Figure 5.6 Average ubRMSE of GWL (at different distances from the assimilated groundwater sites) and SM for each univariate and multivariate assimilation experiment. The left axis corresponds to GWL and the right axis to SM..... 125

Figure 5.7 Differences in annual GWL, SM, and ET between various multivariate DA experiments and OL run in 2018: (a, d, g) FC\_DA\_PAR; (b, e, h) WC\_DA\_PAR; (c, f, i) WC\_DA\_r\_PAR. The red pentagrams and black circles indicate the locations of the CRNS stations and the assimilated groundwater sites, respectively..... 126

Figure 5.8 Temporal evolution of SM at the CRNS station Kall and GWL at a groundwater site for OL and different DA experiments in 2018: (a, b) SM\_DA and SM\_DA\_PAR; (c, d) GWL\_DA and GWL\_DA\_PAR; (e, f) FC\_DA and FC\_DA\_PAR; (g, h) WC\_DA and WC\_DA\_PAR; (i, j) WC\_DA\_r and WC\_DA\_r\_PAR..... 128

Figure B.1 Temporal evolution of simulated soil moisture from the OL (blue), SM\_DA (green), and SM\_DA\_PAR (black) experiments, together with observed soil moisture from CRNS (Obs, red) for the year 2016 at the CRNS sites. Simulated soil moisture was vertically weighted. .... 139

Figure B.2 Temporal evolution of simulated soil moisture from the OL (blue), SM\_DA (green), and SM\_DA\_PAR (black) experiments, together with observed soil moisture from CRNS (Obs, red) for the year 2017 at the CRNS sites. Simulated soil moisture was vertically weighted. .... 140

Figure B.3 Temporal evolution of simulated soil moisture from the OL (blue), SM\_DA (green),

and SM_DA_PAR (black) experiments, together with observed soil moisture from CRNS (Obs, red) for the year 2018 at the CRNS sites. Simulated soil moisture was vertically weighted. ....	141
Figure B.4 Subplots (a) to (c) are the differences (SM_DA - OL) of annual SM (0-80 cm), ET, and GWL in 2016. Subplots (d) to (f) are the differences (SM_DA_PAR - OL) of annual SM (0-80 cm), ET, and GWL in 2016. The locations of the CRNS stations are indicated by the black pentagrams.....	142
Figure B.5 Subplots (a) to (c) are the differences (SM_DA - OL) of annual SM (0-80 cm), ET, and GWL in 2017. Subplots (d) to (f) are the differences (SM_DA_PAR - OL) of annual SM (0-80 cm), ET, and GWL in 2017. The locations of the CRNS stations are indicated by the black pentagrams.....	143
Figure B.6 Temporal evolution of simulated groundwater level from the OL (blue), GWL_DA (green), and GWL_DA_PAR (black) experiments, together with observed groundwater level (Obs, red) at the 12 selected assimilated groundwater sites in 2018. ....	144
Figure B.7 Subplots (a) to (c) are the differences (GWL_DA - OL) of annual GWL, SM (0-80 cm), and ET in 2016. Subplots (d) to (f) are the differences (GWL_DA_PAR - OL) of annual GWL, SM (0-80 cm), and ET in 2016. The locations of the assimilated groundwater sites are indicated by the black circles. ....	145
Figure B.8 Subplots (a) to (c) are the differences (GWL_DA - OL) of annual GWL, SM (0-80 cm), and ET in 2017. Subplots (d) to (f) are the differences (GWL_DA_PAR - OL) of annual GWL, SM (0-80 cm), and ET in 2017. The locations of the assimilated groundwater sites are indicated by the black circles. ....	146
Figure B.9 The differences of annual GWL, SM and ET for OL and different DA experiments in 2016: FC_DA_PAR (a, d, g); WC_DA_PAR (b, e, h); WC_DA_r_PAR (c, f, i). The red pentagrams and black circles indicate the locations of the CRNS stations and the assimilated groundwater sites, respectively. ....	147
Figure B.10 The differences of annual GWL, SM and ET for OL and different DA experiments in 2017: FC_DA_PAR (a, d, g); WC_DA_PAR (b, e, h); WC_DA_r_PAR (c, f, i). The red pentagrams and black circles indicate the locations of the CRNS stations and the assimilated groundwater sites, respectively. ....	148
Figure B.11 Differences in ensemble averaged $\log K_s$ between OL and different DA experiments	

in 2016: (a) and (d) SM\_DA\_PAR; (b) and (e) GWL\_DA\_PAR; (c) and (f) WC\_DA\_r\_PAR. The first row is from 2 cm depth, and the second row is from 10 m depth. The red pentagrams and black circles indicate the locations of CRNS stations and the assimilated groundwater sites, respectively. .... 149

Figure B.12 Differences in ensemble averaged  $\log K_s$  between OL and different DA experiments in 2017: (a) and (d) SM\_DA\_PAR; (b) and (e) GWL\_DA\_PAR; (c) and (f) WC\_DA\_r\_PAR. The first row is from 2 cm depth, and the second row is from 10 m depth. The red pentagrams and black circles indicate the locations of CRNS stations and the assimilated groundwater sites, respectively. .... 150

## List of Tables

Table 2.1 CRNS sites used in this study, including key site characteristics. ....	27
Table 3.1 CRNS sites with geographical information .....	34
Table 3.2 The listed cross-correlations give the cross-correlations between the perturbations for the different atmospheric variables, following the order as indicated in the left column of the table. ....	43
Table 3.3 Perturbation of saturated hydraulic conductivities for different subsurface layers..	44
Table 3.4 The time averaged RMSE of the water table depth at the verification locations for the open loop (OL) and data assimilation (DA) runs (10 sites, 10 km, 5 km, 2.5 km localization radius). ....	52
Table 3.5 Comparison metrics for the soil moisture from CRNS compared to open loop (OL) and data assimilation runs (DA10 and DA5 are for 10 km and 5 km localization radius, respectively) for the year 2018. ....	53
Table 4.1 CRNS sites used in this study, including key site characteristics. ....	65
Table 4.2 The listed cross-correlations give the cross-correlations between the perturbations for the different atmospheric variables, following the order as indicated in the left column of the table. ....	70
Table 4.3 List of conducted simulation experiments: open loop (OL), data assimilation with state update (State) or joint state and parameter update (Joint), jackknife evaluation runs (Jackknife), and verification experiments in 2017 using the updated saturated hydraulic conductivity ( $K_s$ ) from joint assimilation experiments of 2016 and 2018 (Updated $K_s$ from 2016 and Updated $K_s$ from 2018). ....	72
Table 4.4 Error statistics for open loop (OL), data assimilation with state updates (State), joint state-parameter updates (Joint), and jackknife simulations with joint state-parameter updates (Jackknife) for the assimilation periods of 2016 and 2018. The indicators were averaged over all sites with CRNS soil moisture observations. Site-specific indicators are provided in Appendix Tables A.1 to A.4. ....	75
Table 4.5 Comparison of measured and simulated soil moisture for the year 2017 (evaluation period, no assimilation). The updated parameters used for verification were from the assimilation period (2016 and 2018). The error statistics were averaged over all CRNS sites.....	81

Table 4.6 Comparison of measured and simulated evapotranspiration (monthly) and discharge (monthly) from open loop (OL) and data assimilation runs with joint state-parameter updates (DA) for two assimilation periods (2016 and 2018). .....	82
Table A.1 Comparison of CRNS soil moisture measurements and simulated soil moisture from open loop (OL) and data assimilation with joint state and parameter updating (DA) for the year 2016.....	90
Table A.2 Comparison of CRNS soil moisture measurements and simulated soil moisture from open loop (OL) and data assimilation with joint state and parameter updating (DA) for the year 2018.....	90
Table A.3 Comparison of CRNS soil moisture measurements and simulated soil moisture from open loop (OL) and jackknife simulations (DA) for the year 2016.....	91
Table A.4 Comparison of CRNS soil moisture measurements and simulated soil moisture from open loop (OL) and jackknife simulations (DA) for the year 2018.....	91
Table A.5 Root mean square error (RMSE) for open loop (OL), data assimilation with state updates (State), joint state-parameter updates (Joint), and jackknife simulations with joint state-parameter updates (Jackknife) for the assimilation periods of 2016 and 2018. The seasonal indicator was averaged over all sites with CRNS soil moisture observations.....	92
Table A.6 Comparison of daily measured evapotranspiration and simulated evapotranspiration from open loop (OL), data assimilation with state updates (State), and joint state-parameter updates (Joint) for two assimilation periods (2016 and 2018).....	92
Table 5.1 CRNS sites used in this study, including key site characteristics. The mean air temperature and mean annual precipitation were obtained from the ECMWF climate reanalysis data product ERA5-Land (Muñoz Sabater, 2021), resulting in identical precipitation and air temperature values for some of the sites. ....	109
Table 5.2 Statistics for the perturbed atmospheric variables, including cross correlations between the atmospheric variables in the last column. The order of the variables in the last column is as indicated in the left column of the table.....	114
Table 5.3 List of experiments conducted. GWL and SM denote groundwater level and soil moisture observations, respectively. $h$ is pressure head, $\theta$ is soil moisture, and $K_s$ is saturated hydraulic conductivity. The subscripts <i>sat</i> and <i>unsat</i> denote the saturated	

and unsaturated zones, respectively. The experiments FC\_DA and FC\_DA\_PAR used the same fully coupled method as Hung et al. (2022). ..... 115

Table 5.4 Annual ubRMSE for estimated soil moisture (SM, cm<sup>3</sup>/cm<sup>3</sup>), evapotranspiration (ET, mm/day), and groundwater level (GWL, m) from 2016 to 2018 for OL, SM\_DA, and SM\_DA\_PAR experiments..... 117

Table 5.5 Annual ubRMSE for estimated groundwater level (GWL, m), soil moisture (SM, cm<sup>3</sup>/cm<sup>3</sup>), and evapotranspiration (ET, mm/day) from 2016 to 2018 for OL, GWL\_DA, and GWL\_DA\_PAR experiments. *Note:* Distance 0 indicates the groundwater assimilation locations, and 0-0.5 km, 0.5-2.5 km, and 2.5-5 km indicate groundwater validation locations at different distances from the groundwater assimilation sites..... 120

Table 5.6 Annual ubRMSE for estimated groundwater level (GWL, m), soil moisture (SM, cm<sup>3</sup>/cm<sup>3</sup>), and evapotranspiration (ET, mm/day) from 2016 to 2018 for OL and multivariate assimilation experiments. *Note:* Distance 0 indicates the groundwater assimilation locations, and 0-0.5 km, 0.5-2.5 km, and 2.5-5 km indicate groundwater validation locations at different distances from the groundwater assimilation sites. .... 123

Table 5.7 Averaged statistical metrics for the estimated groundwater level (GWL), soil moisture (SM), and evapotranspiration (ET) for all the validation experiments from 2016 to 2018. *Note:* Distance 0 indicates the groundwater assimilation locations, and 0-0.5 km, 0.5-2.5 km, and 2.5-5 km indicate groundwater validation locations at different distances from the groundwater assimilation sites. .... 129

Table B.1 RMSE and R for estimated SM and ET from 2016 to 2018 for OL and univariate data assimilation experiments. .... 136

Table B.2 Annual RMSE for estimated GWL (m) from 2016 to 2018 for OL and univariate data assimilation experiments. .... 137

Table B.3 Annual RMSE for estimated GWL (m) from 2016 to 2018 for OL and multivariate data assimilation experiments. *Note:* Distance 0 indicates the groundwater assimilation locations, and 0-0.5 km, 0.5-2.5 km, and 2.5-5 km indicate groundwater validation locations at different distances from the groundwater assimilation sites. .... 137

Table B.4 RMSE and R for estimated SM and ET from 2016 to 2018 for OL and multivariate

data assimilation experiments..... 138

## List of Acronyms

ANN	Artificial neural networks
BATS	Biosphere-Atmosphere Transfer Scheme
CATHY	CATchment HYdrology
CLM	Community Land Model
COSMIC	COsmic-ray Soil Moisture Interaction Code
COSMO	Consortium for Small-Scale Modeling
CRNS	Cosmic-ray neutron sensors
DA	Data assimilation
DEM	Digital Elevation Model
EC	Eddy covariance
EnKF	Ensemble Kalman Filter
En-Var	Ensemble Variational
ESDB	European Soil Database
ET	Evapotranspiration
ETKF	Ensemble Transform Kalman filter
FEFLOW	Finite Element subsurface FLOW system
GNSS	Global Navigation Satellite System
GRACE	Gravity Recovery and Climate Experiment

GRACE-FO	GRACE Follow-On
GWL	Groundwater Level
$K_s$	Saturated hydraulic conductivity
LAI	Leaf area index
LEnKF	Localized Ensemble Kalman Filter
LiDAR	Light Detection and Ranging
LSM	Land Surface Model
InSAR	Interferometric Synthetic Aperture Radar
ISMON	Irish Soil Moisture Observation Network
MAE	Mean absolute error
MODFLOW	Modular three-dimensional finite-difference groundwater flow model
NSE	Nash-Sutcliffe Model Efficiency
NWP	Numerical weather prediction
OL	Open loop
OASIS-MCT	Ocean Atmosphere Sea Ice Soil coupling Model Coupling Toolkit
ParFlow	Parallel Watershed Flow Model
PDAF	Parallel Data Assimilation Framework
PFT	Plant Function Type
R	Correlation coefficient

RMSE	Root Mean Square Error
RS	Remote sensing
RZSM	Root zone soil moisture
SIB	Simple Biosphere Model
SM	Soil moisture
SMAP	Soil Moisture Active Passive
SMOS	Soil Moisture Ocean Salinity
SNAP	Sentinel Application Platform
TERENO	TERrestrial Environmental Observatories
TSMP	Terrestrial System Modeling Platform
ubRMSD	Unbiased root mean square difference
UAV	Unmanned aerial vehicle
USGS	United States Geological Survey
WTD	Water Table Depth



## Abstract

Groundwater and soil moisture (SM) play a crucial role in the hydrological cycle, and therefore the dynamics of these two variables need to be accurately quantified on spatial and temporal scales. In situ observation networks can provide direct and accurate information on groundwater level (GWL) and SM. However, observations from observation networks are not sufficient to fully represent the Earth's hydrological system without the help of models. Integrated models such as the Terrestrial System Modelling Platform (TSMP) can simulate the hydrological system from the subsurface to the atmosphere and accurately capture the full terrestrial hydrological cycle. Current model estimates of GWL and SM are highly uncertain due to data limitations and model uncertainties. The main sources of uncertainty are related to atmospheric forcings, model structural errors, and uncertain parameterization. Data assimilation (DA) can merge numerical models with observations, resulting in a correction of hydrological states and fluxes and improved parameter estimates.

Different sources of uncertainty may lead to unsatisfactory simulations of groundwater hydrodynamics with hydrological models. The goal of first study is to investigate the impact of assimilating groundwater data into TSMP for improving hydrological modelling in a real-world case. Daily groundwater table depth (WTD) measurements from the year 2018 for the Rur catchment in Germany were assimilated by the Localized Ensemble Kalman Filter (LEnKF) into TSMP. The LEnKF is used with a localization radius so that the assimilated measurements only update model states in a limited radius around the measurements, in order to avoid unphysical updates related to spurious correlations. Due to the mismatch between groundwater measurements and the coarse model resolution (500 m), the measurements need careful screening before DA. Based on the spatial autocorrelation of the WTD deduced from the measurements, three different filter localization radii (2.5 km, 5 km and 10 km) were evaluated for assimilation. The bias in the simulated water table and the root mean square error (RMSE) are reduced after DA, compared with runs without DA (i.e., open loop (OL) runs). The best results at the assimilated locations are obtained for a localization radius of 10km, with an 81% reduction of RMSE at the measurement locations, and slightly smaller RMSE reductions for the 5 km and 2.5 km radius. The validation with independent WTD data showed the best results for a localization radius of 10 km, but groundwater table characterization could only be improved for sites less than 2.5 km from measurement locations. In case of a localization radius of 10km the RMSE-reduction was 30% for those nearby sites. Simulated soil moisture was validated against soil moisture measured by cosmic-ray neutron sensors (CRNS), but no RMSE reduction was observed for DA-runs compared to OL-run. However, in both cases, the correlation between measured and simulated soil moisture content was high (between 0.70 and 0.89, except for the Wüstebach site).

CRNS fill the gap between locally measured in situ SM and remotely sensed (RS) SM by providing accurate SM estimation at the field scale. This is promising for improving hydrologic model predictions, as CRNS can provide valuable information on SM in the root zone at the typical scale of a model grid cell. In a second study of this PhD-work, SM measurements from a network of 12 CRNS in the Rur catchment (Germany) were assimilated into TSMP to investigate its potential for improving SM, evapotranspiration (ET) and river discharge characterization and estimating soil hydraulic parameters at the larger catchment scale. DA experiments (with and without parameter estimation) were conducted in both a wet year (2016) and a dry year (2018) with the Ensemble Kalman Filter (EnKF), and later verified with an independent year (2017) without DA. The results show that SM characterization was significantly improved at measurement locations (with up to 60% RMSE reduction), and that joint state-parameter estimation improved SM simulation more than state estimation alone (more than 15% additional RMSE reduction). Jackknife experiments showed that SM at verification locations had lower and different improvements in the wet and dry years (an RMSE

reduction of 40% in 2016 and 16% in 2018). In addition, SM assimilation was found to improve ET characterization to a much lesser extent, with a 15% RMSE reduction of monthly ET in the wet year and 9% in the dry year.

In a third study, we tested different approaches for joint DA of observed SM data from CRNS and GWL data into the TSMP model. DA experiments (with and without parameter estimation) were conducted with LEnKF for the Rur catchment in Germany for the years 2016-2018, followed by cross-validation (if parameters were estimated) in independent years. Univariate SM assimilation reduced the RMSE of SM over the assimilation locations by more than 50%. Univariate GWL assimilation reduced the monthly RMSE of GWL at assimilation locations by 70%. Within 5 km of the assimilated sites, GWL estimation was still obviously improved, with RMSE reductions 2-50%. However, the univariate assimilation of GWL degraded the characterization of SM, and the univariate assimilation of SM also diminished the simulation of GWL. A new multivariate DA approach that assimilates GWL and SM separately is proposed. GWL data are assimilated and used to estimate the interface between the unsaturated and saturated zones, and update the states (and possibly parameters) of the saturated zone. SM measurements are assimilated to update states of the unsaturated zone. In addition, observation specific localization is proposed. With multivariate DA, at the assimilation locations the estimates of variables (GWL, SM, and ET) are close to those in univariate assimilation. However, there were more than 15% RMSE reductions for GWL at 2.5~5 km validation locations compared to univariate assimilation. In addition, only SM assimilation (univariate or multivariate) improves very slightly ET estimates, with an overall RMSE reduction of 3%. Parameter updating reduced the RMSE of variable estimates by up to 17% compared to updating states alone.

This work was carried out for the Rur catchment (2354 km<sup>2</sup>), which has a well-established monitoring infrastructure and considerable regional diversity in climate, soil types, and land use. In contrast to previously reported small-domain tests, which were primarily conducted in synthetic experiments or oversimplified real-world cases, the assimilation of real-world data at the larger catchment scale faces additional complexities and challenges. The effectiveness of DA can be limited by the uneven distribution of monitoring stations, coarse model resolution, and model structure errors. This thesis, using EnKF and its variants, proposes specific strategies for the assimilation of GWL and SM (from CRNS) separately or jointly in the integrated terrestrial model TSMP. Overall, the results of this thesis provide insights for improving the characterization of multiple variables and parameters (GWL, SM, ET, and saturated hydraulic conductivity ( $K_s$ )) by DA. Possible promising approaches for future improvement of DA performance in coupled models are: (i) improving the accuracy of terrestrial system modeling, including the addition of an atmospheric model, the inclusion of more detailed agro-ecological processes into land surface models and increasing model resolution; (ii) attempting to assimilate data from more diverse sources such as RS, unmanned aerial vehicles (UAVs), and small satellites to address the sparse distribution of in situ observations; and (iii) exploring advanced DA algorithms, potentially EnKF variants or hybrid methods with machine learning (ML) integration, to address the issues and challenges of multivariate assimilation at large scales.

## Zusammenfassung

Grundwasser und Bodenfeuchte (SM) sind von entscheidender Bedeutung für den Wasserkreislauf. Daher muss die Dynamik dieser beiden Variablen auf räumlicher und zeitlicher Ebene genau quantifiziert werden. In situ-Monitoringnetze können direkte und genaue Informationen über den Grundwasserspiegel (GWL) und die Bodenfeuchte liefern, reichen jedoch nicht aus, um das hydrologische System der Erde vollständig zu erfassen. Dies ist nur möglich, indem auf Modelle zurückgegriffen wird. Integrierte Modelle wie die Terrestrial System Modelling Platform (TSMP) ermöglichen eine umfassende Simulation des hydrologischen Kreislaufs der Erde, welcher vom tiefen Untergrund bis zur oberen Atmosphäre reicht. Die aktuellen Modellgeschätzungen von GWL und SM sind jedoch aufgrund von Datenbeschränkungen und Modellfehlern mit erheblicher Unsicherheit behaftet. Die Hauptursachen für diese Unsicherheiten sind atmosphärische Einflüssen, strukturelle Modellfehler und ungenaue Parametrisierung. Durch Datenassimilation (DA) können numerische Modelle mit Beobachtungen kombiniert werden, was zu einer Korrektur der hydrologischen Zustände und Flüsse sowie zu verbesserten Parameterschätzungen führt.

Unterschiedliche Unsicherheitsquellen können zu unbefriedigenden Simulationen der Grundwasserhydrodynamik mit hydrologischen Modellen führen. Ziel dieser Studie ist es, die Auswirkungen der Assimilierung von Grundwasserdaten in die TSMP zur Verbesserung der hydrologischen Modellierung in einem realen Fall zu untersuchen. Tägliche Grundwasserstandsmessungen (WTD) aus dem Jahr 2018 für das Einzugsgebiet der Rur in Deutschland wurden mit dem Localized Ensemble Kalman Filter (LEnKF) in die TSMP assimiliert. Der LEnKF wird mit einem Lokalisierungsradius verwendet. Das bedeutet, dass die assimilierten Messungen nur Modellzustände in einem begrenzten Radius um die Messungen aktualisieren, um unphysikalische Aktualisierungen im Zusammenhang mit unerwünschten Korrelationen zu vermeiden. Aufgrund der Diskrepanz zwischen den Grundwassermessungen und der groben Modellauflösung (500 m) müssen die Messungen vor der DA sorgfältig geprüft werden. Basierend auf der räumlichen Autokorrelation des WTD, die aus den Messungen abgeleitet wurde, wurden drei verschiedene Filter-Lokalisierungsradien (2,5 km, 5 km und 10 km) für die Assimilation evaluiert. Der Fehler des simulierten Grundwasserspiegels und die Wurzel der mittleren Fehlerquadratsumme (RMSE) sind nach der DA, im Vergleich zu Modellläufen ohne DA (bspw. Modellläufen mit offener Schleife (OL)), reduziert. Die besten Ergebnisse an den assimilierten Standorten werden mit einem Lokalisierungsradius von 10 km erzielt, wobei eine 81%ige Verringerung des RMSE an den Messstandorten festgestellt wird. Für die Lokalisierungsradien von 5 km und 2,5 km fallen die RMSE-Reduzierungen etwas geringer aus. Die Validierung mit unabhängigen WTD-Daten zeigte die besten Ergebnisse für einen Lokalisierungsradius von 10 km, aber die Charakterisierung des Grundwasserspiegels konnte nur für Standorte verbessert werden, die weniger als 2,5 km von den Messstandorten entfernt sind. Bei einem Lokalisierungsradius von 10 km betrug die RMSE-Reduzierung für die nahe gelegenen Standorte 30 %. Die Validierung der simulierten Bodenfeuchte erfolgte anhand der gemessenen Bodenfeuchte, welche mit kosmischen Neutronensensoren (CRNS) erfasst wurde. Eine Reduktion des RMSE für DA-Läufe im Vergleich zu OL-Läufen konnte jedoch nicht festgestellt werden. In beiden Fällen war die Korrelation zwischen gemessener und simulierter Bodenfeuchte jedoch hoch (zwischen 0,70 und 0,89, außer für den Standort Wüstebach).

CRNS schließen die Lücke zwischen lokal gemessenem In situ-SM und fernerkundetem SM, indem sie eine genaue SM-Schätzung auf der Feldskala liefern. Dies ist vielversprechend für die Verbesserung von hydrologischen Modellvorhersagen, da CRNS wertvolle Informationen über den SM in der Wurzelzone auf der typischen Skala einer Modellgitterzelle liefern können. In einer zweiten Studie im Rahmen dieser Doktorarbeit wurden SM-Messungen aus einem Netz von 12 CRNS im Einzugsgebiet der Rur (Deutschland) in die TSMP integriert, um dessen Potenzial zur Verbesserung der Charakterisierung von SM, Evapotranspiration (ET) und Flussabfluss sowie zur Schätzung bodenhydraulischer Parameter auf der größeren Einzugsgebietsskala zu untersuchen. DA-Experimente (mit und ohne Parameterschätzung) wurden sowohl in einem feuchten Jahr (2016) als auch in einem trockenen Jahr (2018) mit dem Ensemble Kalman Filter (EnKF) durchgeführt und später mit einem durchschnittlichen Jahr (2017) ohne DA validiert. Die Ergebnisse zeigen, dass die

SM-Charakterisierung an Messstandorten signifikant verbessert wurde (mit einer Verringerung des RMSE um bis zu 60 %) und dass die gemeinsame Schätzung der Zustandsparameter die SM-Simulation stärker verbesserte als die Zustandsschätzung allein (mehr als 15 % zusätzliche RMSE-Reduzierung). Jackknife-Experimente zeigten, dass SM an Validierungsstandorten geringere und unterschiedliche Verbesserungen in nassen und trockenen Jahren aufwies (eine RMSE-Reduzierung von 40 % im Jahr 2016 und 16 % im Jahr 2018). Darüber hinaus wurde festgestellt, dass die SM-Assimilation die ET-Charakterisierung in einem viel geringeren Ausmaß verbessert, mit einer RMSE-Reduzierung der monatlichen ET um 15 % in einem feuchten Jahr und um 9 % in einem trockenen Jahr.

In einer dritten Studie wurden verschiedene Ansätze für die gemeinsame DA von beobachteten SM-Daten aus CRNS und GWL-Daten in das TSMP-Modell getestet. DA-Experimente (mit und ohne Parameterschätzung) wurden mit dem LENKF für das Rur-Einzugsgebiet in Deutschland für die Jahre 2016-2018 durchgeführt, gefolgt von einer Kreuzvalidierung (bei Parameterschätzung) in unabhängigen Jahren. Die univariate SM-Assimilation reduzierte den RMSE von SM an den Assimilationsstandorten um mehr als 50 %. Die univariate GWL-Assimilation reduzierte den monatlichen RMSE der GWL an den Assimilationsstandorten um 70 %. Im Umkreis von 5 km um die assimilierten Standorte wurde die GWL-Schätzung mit einer Verringerung des RMSE um 2-50 % deutlich verbessert. Allerdings verschlechterte die univariate Assimilation von GWL die Charakterisierung von SM, und die univariate Assimilation von SM verschlechterte die Simulation von GWL. Ein neuer multivariater DA-Ansatz wird vorgeschlagen, bei dem GWL und SM getrennt assimiliert werden. GWL-Daten werden assimiliert und genutzt, um die Grenzfläche zwischen der ungesättigten und der gesättigten Zone zu schätzen und die Zustände (und möglicherweise Parameter) der gesättigten Zone zu aktualisieren. SM-Messungen werden assimiliert, um die Zustände der ungesättigten Zone zu aktualisieren. Zusätzlich wird eine beobachtungsspezifische Lokalisierung vorgeschlagen. Mit der multivariaten DA liegen die Schätzungen der Variablen (GWL, SM und ET) an den Assimilationsorten nahe an denen der univariaten Assimilation. An Validierungsorten in einem Umkreis von 2,5-5 km konnte jedoch im Vergleich zur univariaten Assimilation eine Verringerung des RMSE für GWL um mehr als 15 % erzielt werden. Darüber hinaus verbessert nur die SM-Assimilation (univariat oder multivariat) die ET-Schätzungen geringfügig, mit einer RMSE-Reduzierung von insgesamt 3 %. Die Parameteraktualisierung reduzierte den RMSE der Variablenschätzungen um bis zu 17 % im Vergleich zur alleinigen Zustandsaktualisierung.

Diese Arbeit wurde für das Einzugsgebiet der Rur (2354 km<sup>2</sup>) durchgeführt, das über eine gut etablierte Monitoring-Infrastruktur und eine beträchtliche regionale Vielfalt in Bezug auf Klima, Bodenarten und Landnutzung verfügt. Im Gegensatz zu bisherigen Untersuchungen in kleineren Domänen, die hauptsächlich in synthetischen Experimenten oder vereinfachten realen Fällen durchgeführt wurden, steht die Assimilation von realen Daten auf der Ebene eines größeren Einzugsgebiets vor zusätzlichen Komplexitäten und Herausforderungen. Die Effektivität der DA kann durch die ungleichmäßige Verteilung von Messstationen, die grobe Modellauflösung und Modellstrukturfehler eingeschränkt werden. In dieser Arbeit werden spezifische Strategien für die Assimilation von GWL und SM (aus CRNS) unter Verwendung von ENKF und seinen Varianten getrennt oder gemeinsam in das integrierte terrestrische Modell TSMP vorgeschlagen. Insgesamt liefern die Ergebnisse dieser Arbeit Erkenntnisse zur Verbesserung der Charakterisierung mehrerer Variablen und Parameter (GWL, SM, ET und  $K_s$ ) durch DA. Mögliche vielversprechende Ansätze für eine zukünftige Verbesserung der DA-Leistung in gekoppelten Modellen sind: (i) Verbesserung der Genauigkeit der Modellierung des terrestrischen Systems, einschließlich der Hinzufügung eines atmosphärischen Modells, der Einbeziehung detaillierterer agrarökologischer Prozesse in Landoberflächenmodelle und der Erhöhung der Modellauflösung; (ii) der Versuch, Daten aus vielfältigeren Quellen wie Fernerkundung (RS), unbemannten Luftfahrzeugen (UAVs) und Kleinsatelliten zu assimilieren, um der spärlichen Verteilung von In situ-Beobachtungen entgegenzuwirken; und (iii) die Erforschung fortschrittlicher DA-Algorithmen, möglicherweise ENKF-Varianten oder hybride Methoden mit Integration von maschinellem Lernen (ML), um die Probleme und Herausforderungen der multivariaten Assimilation in großen Maßstäben anzugehen.

## Chapter 1: Introduction

Groundwater is an important source of fresh water for human life and agricultural and industrial production. Changes in groundwater storage can have significant impacts on agricultural irrigation supplies, industrial production costs and infrastructure security, water quality, and the biodiversity of groundwater-dependent ecosystems. Therefore, a good knowledge of hydrological conditions such as groundwater levels is essential for socio-economic development, including water resource management and environmental protection. GWL or WTD are an important hydrological variable that can range from 0 m in wetlands to hundreds of meters in arid regions. Shallow groundwater is critical in terrestrial ecosystems because it interacts with the unsaturated zone and can directly provide water for transpiration and soil evaporation. Changes in groundwater levels and soil moisture directly affect the availability of groundwater resources, plant growth, land-atmosphere interactions, surface runoff and groundwater recharge processes, linkages between surface water bodies, groundwater systems, and terrestrial ecosystems, and energy balance, and understanding their variations can help in assessing water resources and predicting droughts or floods. Therefore, accurate information on the spatiotemporal variability of groundwater levels and soil moisture content is essential for a better understanding of hydrological processes and water and energy cycles in terrestrial systems (Vereecken et al., 2022).

Groundwater level measurements provide an accurate assessment of local groundwater dynamics, derived either from direct measurements of water table depth or indirect measurements of water pressure converted to GWL data (Freeze and Cherry, 1979). Traditional instruments and electronic survey techniques enable point-scale GWL measurements, utilizing tools such as steel tapes, electronic measuring tapes, piezometers, digital water level recorders, pressure transducers, and data loggers (Masood et al., 2022). Manual methods like the steel tape (Peralta et al., 1983) and electronic measuring tapes (Chevalking et al., 2008) employ a water level indicator to assess GWL, while piezometers measure GWL by detecting hydraulic head (Yin et al., 2017; Sattari et al., 2018). The frequency of monitoring GWL with different instruments varies greatly. Automated instruments, including digital water level recorders and pressure transducers connected to data loggers, continuously record water depth, facilitating frequent monitoring at daily or shorter intervals (Husain et al., 1998; Kalbus et al., 2006). However, despite their advantages, in situ GWL measurements are expensive and time-consuming, especially during the installation phase, which requires drilling a well or piezometer, limiting the feasibility of monitoring large areas. Consequently, the number of measurements available in a given area is often sparse, and the spatial representativeness of in situ GWL data

can be constrained by hydrogeological heterogeneity. To complement point-scale measurements, geophysical techniques and RS have emerged as valuable tools for large-scale monitoring of groundwater changes. Geophysical methods, such as electrical resistivity (Garcia-Menendez et al., 2018), seismic techniques (Rosid and Keping, 2005), and gravity measurements (Sabri et al., 2020), provide indirect approaches (non-invasive) for regional-scale groundwater level monitoring by detecting variations in subsurface properties (resistivity, seismic wave velocity, and density) but often require complex modeling or integration with borehole data for accurate groundwater estimates (Masood et al., 2022). Available RS techniques for groundwater monitoring encompass gravity-based measurements, surface deformation analysis, radar altimetry, and LiDAR (Adams et al., 2022). Gravity-based approaches, such as the Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On (GRACE-FO), detect groundwater storage changes by measuring mass variations, offering global-scale data with limited spatial resolution (300 km) and potential processing errors (Richey et al., 2015; Landerer et al., 2020). Surface deformation analysis techniques, including Interferometric Synthetic Aperture Radar (InSAR) and Global Navigation Satellite System (GNSS), along with radar altimetry and LiDAR, detect high-precision surface deformations such as subsidence or uplift over large areas to indirectly estimate changes in groundwater storage (Abdalati et al., 2010; Hwang et al., 2016). However, the effectiveness of these methods often depends on supplementary geological data and they are susceptible to interference from terrain and vegetation (Chen et al., 2017; Smith and Majumdar, 2020; White et al., 2022). While these technologies provide broad coverage and rapid data collection, challenges remain, including addressing spatial heterogeneity, integrating different datasets, and downscaling coarse-resolution RS data for local groundwater monitoring. Therefore, RS technologies cannot fully substitute in situ methods for groundwater level monitoring, and achieving comprehensive groundwater monitoring at regional or global scales remains a challenge (Condon et al., 2021).

SM can be estimated at different scales using various techniques and approaches; it can be measured directly in situ or estimated indirectly from RS observations. The resolution of RS-derived products is often coarse both spatially and temporally, and the data are subject to various errors, especially in areas with dense vegetation cover (Bauer-Marschallinger et al., 2019; Kim et al., 2020). In addition, RS products only provide SM for the upper surface soil layer. In contrast, in situ measurements can provide more accurate data for deeper soil layers, but only at the point scale. To bridge the gap between in situ measurements and RS products, the CRNS has been introduced as a new technique to monitor SM (Zreda et al., 2008). The CRNS measurement is promising because it can provide non-invasive and real-time SM at the field

scale. The CRNS detects epithermal neutrons produced by cosmic radiation, which can be used to measure SM because the detected neutron count rate is inversely correlated with the amount of hydrogen in the soil (Zreda et al., 2008). The CRNS probe measures integral SM centered on the detector, with the measurement range inversely related to soil water content (Zreda et al., 2008; Köhli et al., 2015). The footprint of the CRNS can extend to an area of up to 18 hectares with a measurement depth of up to 80 centimeters, corresponding to the upper root zone (Zreda et al., 2008; Bogena et al., 2015; Köhli et al., 2015). Hence, the spatial extent of CRNS is the desired application scale of land-surface processes, as it corresponds to the desired grid cell size of a high-resolution land surface model (Crow et al., 2012; Ajami et al., 2014). In addition, over a wide range of land uses and climatic conditions, the CRNS has demonstrated excellent reliability (Bogena et al., 2013; Heidbüchel et al., 2016; Vather et al., 2020). Over the past decade, the continuous development of CRNS technology has promoted its application in hydrological modeling (Schattan et al., 2020; Dimitrova-Petrova et al., 2021), land surface modeling (Han et al., 2015; Roland Baatz et al., 2017), satellite product validation (Montzka et al., 2017; Zhao et al., 2021), ecohydrological (e.g., snow, precipitation, and vegetation) monitoring (Schattan et al., 2017; Jakobi et al., 2018; Bogena et al., 2020), and agricultural management (Han et al., 2016; Li et al., 2019). Because of the advantages offered by CRNS, CRNS networks with multiple stationary sensors have been established in several countries, including Europe (Bogena et al., 2022), the United Kingdom (Evans et al., 2016), Australia (Hawdon et al., 2014), and the United States (Zreda et al., 2012). These distributed CRNS networks are intended to support environmental monitoring at larger scales, which can also benefit diverse hydrological applications.

However, measuring GWL and SM over large areas is uneconomical and impractical. Considering the limited data availability, the most common method to estimate GWL and SM is to use physical models to simulate these variables, validate the simulation results with measured data, evaluate the effectiveness of the model, and finally use the hydrological information from the model for extensive regions as well as for future prediction. In this work, we mainly focus on the two key variables in hydrological modeling: GWL for the saturated zone and SM for the unsaturated zone.

Groundwater modeling is a way to study the dynamics of groundwater systems and predict the spatial and temporal distribution of groundwater levels (de Marsily, 1986; Baalousha, 2008). The most important step in the groundwater numerical modeling process is the conceptualization of the groundwater model (Bear and Verruijt, 1987; Anderson et al., 2015). It provides a basic understanding of the groundwater system and also describes the

characterization of hydrogeology (Bear and Verruijt, 1987; Enemark et al., 2019; El-Rawy et al., 2022). Then, through a set of mathematical equations with initial and boundary conditions, groundwater models can represent the conceptual understanding of the groundwater system (Kumar, 2004). Early numerical groundwater flow models were based on simplified two-dimensional flow equations and idealized geological conditions (Darcy, 1856; V., 1965; Bear, 1975). Later, with the development of groundwater models based on three-dimensional equations (Hubbert and Rubey, 1959), groundwater systems can be more accurately described and represented, such as the United States Geological Survey (USGS) MODFLOW (Modular three-dimensional finite-difference groundwater flow model) (McDonald and Harbaugh, 1988), CATHY (Catchment Hydrology) model (Paniconi and Wood, 1993; Paniconi and Mario, 1994), ParFlow (Ashby and Falgout, 1996; Jones and Woodward, 2001), and FEFLOW (Finite Element subsurface FLOW system) (Diersch, 2014).

Land surface models (LSMs) provide an effective approach to simulating SM in terrestrial systems. They are numerical models designed to solve the coupled fluxes of water, energy, and carbon between the land surface and atmosphere (Fisher and Koven, 2020). LSMs were originally developed as parts of atmospheric and climate models to investigate the influence of the land surface on meteorological processes (Fisher and Koven, 2020). Early models introduced the concept of integrating soil and vegetation to represent energy and water exchanges between the land surface and the atmosphere (Blyth et al., 2021), such as the Biosphere-Atmosphere Transfer Scheme (BATS) and the Simple Biosphere Model (SIB). Over time, LSMs have evolved, and numerous processes affecting the terrestrial system have been added to LSMs (Fisher and Koven, 2020). Different LSMs have different parameterizations for modeling land surface processes, and the focus of each model differs, with some focusing more on the water cycle and others focusing more on land use and carbon (Blyth et al., 2021). Some processes have also evolved from simple representations to more complex processes. For example, representations of soil hydrology in LSMs have evolved from a simple “bucket” representation (Manabe, 1969) to the 1D Richards equation (Bonan, 1996; Cox et al., 1999) to 3D variably saturated flow models (Bisht & Riley, 2019). Some studies have suggested the need to improve the representation of hydrologic processes in LSMs to better simulate atmosphere-land surface interactions (Blyth et al., 2021). Coupling hydrologic models and LSMs can improve the representation of hydrologic processes in the soil-vegetation-atmosphere continuum.

In most LSMs, the dynamic interactions between root zone soil moisture (RZSM) and shallow groundwater are either neglected or simplified (Chen and Hu, 2004; Kollet and

Maxwell, 2008). This results in a poor understanding of the influence of groundwater on the land surface and atmospheric processes. However, groundwater plays a critical role in both the water and energy cycles of the terrestrial system (Condon et al., 2021). Shallow groundwater affects root zone SM, and surface runoff is more likely to occur when the groundwater table is shallow (Zhang, 2018). As a result, shallow groundwater also indirectly affects evapotranspiration. Therefore, quantifying the temporal dynamics of groundwater levels across spatial and temporal scales and investigating the effects of its variability on other hydrological variables is of high interest.

Considering that groundwater flow models can simulate subsurface water fluxes, attempts have been made to integrate a groundwater model and a LSM in recent years (Maxwell and Miller, 2005; Rihani et al., 2010; Sulis et al., 2010; Bizhanimanzar et al., 2019). The coupled model allows consideration of the dynamic interactions between groundwater, surface water, and land surface processes, leading to improved simulation accuracy and, consequently, more accurate predictions (Kollet and Maxwell, 2008; Rihani et al., 2010; Shrestha et al., 2014; Keune et al., 2016; Sridhar et al., 2018; Zhao et al., 2021). Yeh and Eltahir (2005) integrated a lumped unconfined aquifer model with a LSM, and the results of the model showed that simulated hydrologic variables (soil saturation and WTD) and fluxes (evaporation, runoff, and groundwater recharge) were in good agreement with observations. Graham and Butts (2005) developed the integrated hydrological model MIKE-SHE, which allows each hydrological process to be represented at different spatial and temporal scales and levels of complexity. Maxwell and Miller (2005) coupled a LSM (Common Land Model) and a variably saturated groundwater model (ParFlow) as a single column model and showed that the SM and groundwater table depths simulated by the model were consistent with observations. In addition, comparisons of hydrologic variables between the uncoupled and coupled models showed the need for improved representation of groundwater in LSMs.

Shrestha et al. (2014) developed the TSMP model, which can simulate hydrological and biogeophysical processes in terrestrial systems from the atmosphere to the land surface and subsurface in a fully coupled fashion. Surface water-groundwater interactions are also considered by the model. To understand the effects of temporal and spatial variations of GWL on terrestrial ecosystems and the interactions between groundwater and root zone SM, integrated models such as TSMP are well suited to physically simulate the groundwater-soil-vegetation-atmosphere system. The applicability of this model has been demonstrated in many studies (e.g., Shrestha et al., 2015; Keune et al., 2016; Shrestha et al., 2018; Furusho-Percot et al., 2019; Zhao et al., 2021; Naz et al., 2023).

However, LSMs or hydrologic models tend to have many parameters, including parameters that are difficult to know, such as large scale information on soil hydraulic properties and aquifer properties (Vereecken et al., 1990; Wen et al., 2002; Vereecken et al., 2008). In addition, coupled models generally have more model parameters than stand-alone models. These parameters contribute to model uncertainty and affect its accuracy. Modeling uncertainties can also arise from model forcings, model structure, and initial conditions (Freeze, 1975; Roland Baatz et al., 2017). To quantify and reduce uncertainties in model predictions, DA can be used to correct model predictions with observations and improve the accuracy of simulations (De Lannoy et al., 2014). Furthermore, DA incorporates all sources of uncertainty and can provide uncertainty estimations for parameters. The most widely used DA technique in hydrology is the EnKF (Evensen, 1994; 2003). As a sequential filtering algorithm, the state probability density is propagated forward using an ensemble of model realizations taking into account information from measurements (McLaughlin, 2002). EnKF has been shown to be effective for solving nonlinear hydrological problems with many unknowns (Camporese et al., 2009b; Schöniger et al., 2012). In addition, EnKF can combine state and parameter estimation using an augmented state vector approach (Chen and Zhang, 2006; Hendricks Franssen and Kinzelbach, 2008), which is beneficial for characterizing and optimizing model parameters and can provide better model predictions (Brandhorst et al., 2017). The not-too-large ensemble size and the efficient parallel computation allow EnKF to run at a relatively low computational cost (Hendricks Franssen and Kinzelbach, 2008; Yin et al., 2015; Kurtz et al., 2016). This advantage has led to the widespread use of EnKF, which can be combined with different types of models and observations.

Several studies assimilated hydraulic head data into groundwater flow models to estimate hydraulic head and hydraulic conductivity. However, due to a scarcity of real groundwater observations, GWL (or hydraulic head) assimilation studies were mainly conducted with synthetic data. Chen and Zhang (2006) assimilated synthetic pressure head data into transient flow models, demonstrating the ability of EnKF to estimate hydraulic conductivity fields. Hendricks Franssen and Kinzelbach (2008) conducted synthetic studies using a two-dimensional transient groundwater flow model to explore effective ways to reduce the filter inbreeding problem when using the EnKF for parameter estimation. Tong et al. (2011) employed LEnKF in a two-dimensional synthetic transient groundwater flow model and found that EnKF with localization can solve the problem of filter divergence and capture the heterogeneous hydraulic conductivity field reliably with higher efficiency and a smaller ensemble size than EnKF. Panzeri et al. (2013, 2014) proposed EnKF variants (direct solution of nonlocal (integrodifferential) stochastic ensemble moment equations or direct computation

of stochastic ensemble moment equations governing the space-time evolution of ensemble means and covariances of hydraulic heads and fluxes) for two-dimensional transient groundwater flow models to address issues arising in groundwater DA, including heavy computation and filter inbreeding, and the validity of this approach was verified in a field application (Panzeri et al., 2015). The previous experiments demonstrated that EnKF provides an efficient approach to obtaining satisfactory estimations of hydraulic head and hydraulic conductivity fields with groundwater head data (Chen and Zhang, 2006; Hendricks Franssen et al., 2011; Kurtz et al., 2012).

SM assimilation has been shown in many studies to be an effective way to estimate SM dynamics (e.g., Reichle et al., 2002; De Lannoy et al., 2007; Reichle et al., 2008; Brocca et al., 2010; Montzka et al., 2011; Hain et al., 2012; Han et al., 2014; Dong et al., 2016; Zhou et al., 2020). The simulated SM data can either be in situ measurements (Jiang et al., 2021; Fu et al., 2023) or RS retrievals, such as Soil Moisture Active Passive (SMAP) (Reichle et al., 2017; Liming He et al., 2021; Seo et al., 2021) and Soil Moisture Ocean Salinity (SMOS) (De Lannoy and Reichle, 2016; Lievens et al., 2016; Román-Cascón et al., 2017). For SM assimilation from CRNS, some studies with LSMs demonstrated the effectiveness of DA in improving SM simulations. The COsmic-ray Soil Moisture Interaction Code (COSMIC) model developed by Shuttleworth et al. (2013) allows a quick and easy conversion of neutron counts (from CRNS observations) to SM values. COSMIC has been used as an observation operator in several studies to assimilate neutron counts into LSMs to improve SM prediction (Shuttleworth et al., 2013; Rosolem et al., 2014; Han et al., 2015; Han et al., 2016; Roland Baatz et al., 2017; Patil et al., 2021). For example, in a study by Roland Baatz et al. (2017), assimilation of SM from a CRNS network was found to improve catchment-scale SM modeling, and joint state and parameter estimation was found to outperform state estimation alone. Several other studies have also shown that joint state parameter updating has a positive effect on state estimation when assimilating SM in land surface models (Pauwels et al., 2009; Shi et al., 2014; Han et al., 2015; Shi et al., 2015; Zhang et al., 2017).

However, the majority of DA studies focused on a single compartment of the terrestrial system, either the land surface or the subsurface, and often the assimilation was performed in synthetic experiments or oversimplified real-world cases. With the advent of coupled models, multivariate assimilation of different data types in coupled models has also been investigated. The understanding of various measured data can be greatly improved by using multivariate assimilation in integrated hydrological models (Zhang et al., 2016). Several studies have reported the application of multivariate assimilation in integrated hydrological models.

Camporese et al. (2009a) assimilated synthetic observations of pressure head and streamflow into the coupled surface-subsurface model CATHY (CATchment HYdrology) with EnKF, and the results showed that streamflow prediction can be improved by assimilating pressure head and streamflow, either individually or simultaneously. Botto et al. (2018) assimilated pressure head, SM, and subsurface outflow with EnKF into the CATHY model on an artificial hillslope. Their results demonstrate the efficiency of EnKF in correcting states and parameters even with strong nonlinearities, but they also showed that multisource data assimilation can lead to the degradation of model predictions for other variables. Kurtz et al. (2014) assimilated hydraulic head and groundwater temperature data into the coupled subsurface flow and heat transport model SPRING to update uncertain subsurface hydraulic parameters (i.e., hydraulic conductivities and leakage coefficients) and found that joint assimilation of the two observations with updating of uncertain hydraulic parameters provided the best characterization of model states and hydraulic properties. Shi et al. (2014) performed synthetic multivariate data assimilation to estimate multiple parameters in a coupled physically based land surface hydrologic model (Flux-PIHM) with EnKF. The observations included discharge, water table depth, SM, land surface temperature, sensible and latent heat fluxes, and transpiration. The results showed the great potential of EnKF to estimate hydrological model parameters using multivariate observations. The capability of EnKF was further confirmed at a field site with multivariate measurements (Shi et al., 2015). Zhang et al. (2016) jointly assimilated SM and groundwater head measurements into the integrated hydrological model MIKE-SHE with the ensemble transform Kalman filter (ETKF), and the results showed that DA can improve the model performance. They also showed the need to use localization (distance and variable) when assimilating both groundwater head and SM. Zhang et al. (2016) subtracted the average difference between observations and model simulations from the original data when comparing in situ head measurements with predictions.

Some assimilation experiments have also been reported for the integrated model TSMP. Zhang et al. (2018) compared five data assimilation methods for assimilating GWL data via EnKF to improve root zone SM estimation with TSMP over a synthetic domain (4 soil grid cells). The synthetic experiments showed that the joint assimilation of GWL and SM gave the best result for predicting root zone SM. Hung et al. (2022) investigated the assimilation of SM and GWL in TSMP by constructing a virtual reality that mimicked the Neckar catchment. The DA experiments showed clear improvements in SM or GWL characterization when assimilating the corresponding observations, and the benefits of DA were observed in the vicinity of the measurement locations. The joint assimilation of SM and GWL gave a better estimation of the model states when using the weakly coupled scheme (only saturated subsurface states are

updated) than the fully coupled scenario (all subsurface states are updated).

Although many DA platforms and applications have been reported, to the best of our knowledge, assimilation experiments with CRNS soil moisture data have only been conducted with land surface models. The published studies on SM and groundwater assimilation reported mainly tests on small domains in synthetic experiments or oversimplified real-world cases. There is still a lack of studies with integrated land surface-subsurface models to investigate the assimilation of SM measurements from CRNS and real groundwater data at a larger catchment scale. Therefore, this is still an emerging research topic. To explore the information provided by both groundwater measurements and SM data through DA, integrated models such as TSMP that can account for the interactions between root zone SM and groundwater are promising. The DA framework TSMP-PDAF was used to perform the DA experiments in this work.

The objectives of this PhD work are to:

(1) investigate whether the assimilation of groundwater data into the integrated model TSMP at the larger catchment scale for a real-world case is able to provide better GWL characterization than the model run without DA, and identify the main limitations and complications of groundwater assimilation in practice. This is a novel contribution, as the assimilation of real groundwater measurements in integrated land-surface-subsurface models at a larger catchment scale has not been done before.

(2) investigate the effectiveness of a CRNS network to estimate SM with the integrated model TSMP under different hydrological conditions and whether the assimilation of CRNS soil moisture data can also improve ET estimates. This work is also innovative because it is the first time that SM from a CRNS network is assimilated into an integrated land surface-subsurface model at the large catchment scale.

(3) formulate an improved methodology for joint assimilation of GWL and SM from CRNS in TSMP under real-world conditions, investigating the differences between assimilating groundwater data and SM observations separately and jointly, and exploring the interactions between root zone SM, GWL and other hydrological processes (e.g., ET).

Chapter 2 provides a brief overview of the TSMP integrated model and the PDAF data assimilation system used in this work. In addition, this chapter presents a basic introduction to the DA methods of EnKF and LEnKF, as well as the study area and its hydrological measurements.

In Chapter 3, GWL assimilation in the integrated terrestrial system model TSMP for a real-

world case is demonstrated and analyzed.

In Chapter 4, SM measurements from a network of 12 CRNS in the Rur catchment were assimilated into TSMP to investigate its potential for improving SM and ET characterization and estimation of soil hydraulic parameters at the larger catchment scale.

In Chapter 5, a new multivariate assimilation approach is proposed for the joint assimilation of groundwater and SM in integrated terrestrial system models. The univariate and multivariate DA experiments (with and without parameter estimation) were conducted with LEnKF for the years 2016-2018.

Finally, Chapter 6 provides discussions, conclusions, and an outlook for future research.

## **Chapter 2: Theory, methods and materials**

### **2.1 Integrated terrestrial systems model TSMP**

The TSMP model is a fully integrated biogeophysical terrestrial system model that can simulate the transport processes of water and energy between the atmosphere, land surface, and subsurface in a physically based representation (Gasper et al., 2014; Shrestha et al., 2014). The TSMP includes three models that are already well established: an atmospheric model (Consortium for Small Scale Modeling, COSMO) (Baldauf et al., 2011), a land surface model (Community Land Model (CLM, version 3.5)) (Oleson et al., 2004; Oleson et al., 2008) from the National Center for Atmospheric Research, and the 3D variably saturated groundwater flow model ParFlow (Kollet and Maxwell, 2006) for the subsurface. These three models are two-way coupled by the Ocean Atmosphere Sea Ice Soil coupling Model Coupling Toolkit (OASIS-MCT, version 3) (Valcke, 2013). The OASIS-MCT coupler is used to exchange variables and fluxes between the different sub-models. The TSMP modeling platform can be run with different combinations of component models, such as a fully coupled configuration (COSMO-CLM-ParFlow), partially coupled configurations (COSMO-CLM or CLM-ParFlow), or a single model (COSMO/CLM/ParFlow). In this Ph.D. work, the partially coupled configuration with the land surface model CLM and the subsurface model ParFlow is used.

#### **2.1.1 Land surface model Community Land Model (CLM), version 3.5**

The land surface model CLM simulates the biophysical processes in the terrestrial system, including solar and longwave radiation interactions with the vegetation canopy and soil, momentum and turbulent fluxes from the canopy and soil, canopy hydrology (e.g., interception processes), soil hydrology, and stomatal physiology and photosynthesis (Oleson et al., 2007). The mass and energy balance components in CLM include soil evaporation, evaporation from intercepted water, plant transpiration, infiltration of water in the soil, sensible and ground heat fluxes, and freeze-thaw processes (Oleson et al., 2004; Oleson et al., 2008).

A nested subgrid hierarchy is used to represent the spatial heterogeneity of the land surface in CLM (Oleson et al., 2008). Each grid cell is divided into different types of land units (glacier, lake, wetland, urban, and vegetated), where each land unit can have a different number of snow/soil columns (see Fig. 2.1). In CLM, the soil column and snow column are discretized into 10 and 5 vertical layers, respectively (Oleson et al., 2004; Oleson et al., 2008). In addition, each column can have multiple Plant Functional Types (PFTs) (Bonan et al., 2002; Oleson et al., 2008). Each PFT is characterized by distinct plant physiological parameters that should

capture the biogeophysical and biogeochemical differences between the different vegetation types (Oleson et al., 2004; Oleson et al., 2008).

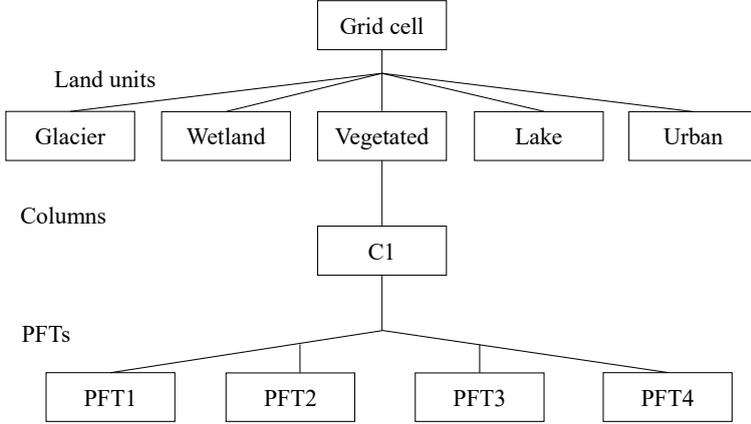


Figure 2.1 Configuration of the CLM subgrid hierarchy, adapted from Oleson et al. (2004). Each grid cell is composed of multiple landunits, snow/soil columns, and Plant Functional Types (PFTs). The first subgrid level is the landunit, which can be glacier, wetland, vegetated, lake, or urban. The second subgrid level is the column, which can be soil or snow within a single landunit. The third subgrid level is the PFT level, where up to 4 of 15 possible PFTs that differ in physiology and structure can coexist on a single column.

The primary function of the CLM in the TSMP is to calculate evapotranspiration from the soil and vegetation. The CLM calculates evaporation and transpiration, considering both vegetated and non-vegetated surfaces (Oleson et al., 2007). For the non-vegetated surface (bare soil), the evaporation  $E_g$  [M/L<sup>2</sup>/T] from the ground is calculated as follows:

$$E_g = -\frac{\rho_{atm}(q_{atm}-q_g)}{\gamma_{aw}} \quad (2.1)$$

where  $\rho_{atm}$  [M/L<sup>3</sup>] is the air density,  $q_{atm}$  [M/M] is the atmospheric specific humidity,  $q_g$  [M/M] is the soil surface specific humidity, and  $\gamma_{aw}$  [T/L] is the aerodynamic resistance to water

vapor transfer.  $q_g$  is proportional to the saturation specific humidity:

$$q_g = \alpha q_{sat}^{T_g} \quad (2.2)$$

and  $q_{sat}^{T_g}$  [M/M] is the saturated specific humidity given the ground surface temperature  $T_g$  [Q]. The factor  $\alpha$  [-] is a combined value for soil and snow:

$$\alpha = \alpha_{soi,1}(1 - f_{sno}) + f_{sno} \quad (2.3)$$

where  $f_{sno}$  [-] is the fraction of snow coverage,  $\alpha_{soi,1}$  [-] refers to the surface soil layer and is a function of the surface soil layer water matrix potential  $\psi_1$ [L]:

$$\alpha_{soi,1} = \exp\left(\frac{\psi_1 g}{1 \times 10^3 R_{wv} T_g}\right) \quad (2.4)$$

where  $R_{wv}$  [L<sup>2</sup>/T<sup>2</sup>/Q] is the gas constant for water vapor,  $g$  [L/T<sup>2</sup>] is the gravitational acceleration, and  $\psi_1$  [L] is calculated as:

$$\psi_1 = \psi_{sat,1} s_1^{-B_1} \text{ and } \psi_1 \geq -1 \times 10^8 \quad (2.5)$$

where  $\psi_{sat,1}$  [L] is the saturated matric potential for the surface soil layer,  $B_1$  [-] is the Clapp and Hornberger parameter (Clapp and Hornberger, 1978), and  $s_1$  [-] is the wetness of the top soil layer with respect to saturation:

$$s_1 = \frac{1}{\Delta z_1 \theta_{sat,1}} \left[ \frac{w_{liq,1}}{\rho_{liq}} + \frac{w_{ice,1}}{\rho_{ice}} \right] \text{ and } 0.01 \leq s_1 \leq 1.0 \quad (2.6)$$

where  $\Delta z_1$  [L] is the thickness of the top soil layer,  $\theta_{\text{sat},1}$  [ $\text{L}^3/\text{L}^3$ ] is the saturated soil moisture of the top soil layer (i.e., porosity),  $w_{\text{liq},1}$  [ $\text{M}/\text{L}^2$ ] and  $w_{\text{ice},1}$  [ $\text{M}/\text{L}^2$ ] are the mass of liquid water and ice of the top soil layer,  $\rho_{\text{liq}}$  [ $\text{M}/\text{L}^3$ ] and  $\rho_{\text{ice}}$  [ $\text{M}/\text{L}^3$ ] are the density of liquid water and ice.

For the vegetated surface, the evapotranspiration flux  $E$  [ $\text{M}/\text{L}^2/\text{T}$ ] includes the water vapor flux from the vegetation  $E_v$  and the ground  $E_g$ :

$$E = E_v + E_g \quad (2.7)$$

$$E_v = - \frac{\rho_{\text{atm}}(q_s - q_{\text{sat}}^{T_v})}{r_{\text{total}}} \quad (2.8)$$

$$E_g = - \frac{\rho_{\text{atm}}(q_s - q_g)}{r'_{\text{aw}}} \quad (2.9)$$

where  $q_{\text{sat}}^{T_v}$  [ $\text{M}/\text{M}$ ] is the saturated specific humidity at vegetation temperature  $T_v$  [Q],  $r'_{\text{aw}}$  [ $\text{T}/\text{L}$ ] is the aerodynamic resistance to water vapor transfer between the ground and canopy air,  $r_{\text{total}}$  [ $\text{T}/\text{L}$ ] is the aerodynamic resistance to water vapor transfer from the canopy to the canopy air, and  $q_s$  [ $\text{M}/\text{M}$ ] is the canopy specific humidity:

$$q_s = \frac{c_a^w q_{\text{atm}} + c_g^w q_g + c_v^w q_{\text{sat}}^{T_v}}{c_a^w + c_g^w + c_v^w} \quad (2.10)$$

where  $c_a^w$  [ $\text{L}/\text{T}$ ],  $c_g^w$  [ $\text{L}/\text{T}$ ] and  $c_v^w$  [ $\text{L}/\text{T}$ ] are water vapor conductances from the canopy air to the atmosphere, the leaf to the canopy air, and the ground to the canopy air, respectively.

### 2.1.2 Subsurface hydrological model ParFlow

ParFlow is used to simulate surface runoff and groundwater flow in TSMP. The soil hydrologic processes of the CLM are replaced by ParFlow in the TSMP (Ashby and Falgout,

1996; Jones and Woodward, 2001; Kollet and Maxwell, 2006; Maxwell, 2013). ParFlow is improved from the parallel three-dimensional variably saturated subsurface flow code ParFlow with a two-dimensional overland flow simulator (Ashby and Falgout, 1996; Kollet and Maxwell, 2006). ParFlow solves the kinematic wave equation (Lighthill and Whitham, 1955) for overland flow and the 3D Richards equation (Richards, 1931) for groundwater flow in the unsaturated and saturated zones using the Newton-Krylov nonlinear solver (Jones and Woodward, 2001).

In ParFlow, the three-dimensional Richards' equation can be written as follows (Maxwell, 2013):

$$S_s S_w \frac{\partial h}{\partial t} + \Phi \frac{\partial S_w(h)}{\partial t} = \nabla \cdot \mathbf{q} + q_r(x, z), \quad (2.11)$$

and

$$\mathbf{q} = \Phi S_w(h) \mathbf{v} = -K_s(x) k_r(h) \nabla(h + z), \quad (2.12)$$

where  $S_s$  is the specific storage [ $L^{-1}$ ];  $S_w$  is the relative saturation;  $h$  is the pressure head [ $L$ ];  $t$  is time [ $T$ ];  $\Phi$  is the porosity;  $\mathbf{q}$  is the specific volumetric (Darcy) flux [ $LT^{-1}$ ];  $q_r$  is a general source/sink term representing transpiration, wells, and other fluxes [ $LT^{-1}$ ];  $x$  is the length along the  $x$ -axis specified as horizontal direction [ $L$ ];  $z$  is the elevation along the  $z$ -axis specified as upward to be positive [ $L$ ];  $\mathbf{v}$  is the subsurface flow velocity [ $LT^{-1}$ ];  $K_s(x)$  is the saturated hydraulic conductivity tensor [ $LT^{-1}$ ];  $k_r$  is the relative permeability.

Input soil hydraulic parameters for ParFlow include saturated hydraulic conductivity, porosity, and Mualem-van Genuchten parameters (van Genuchten, 1980; Kollet and Maxwell, 2008). The van Genuchten formulation (van Genuchten, 1980) is used to estimate the pressure head from SM data:

$$\psi = - \frac{\left[ \left( \frac{\theta_s - \theta_r}{\theta(\psi) - \theta_r} \right)^{\frac{1}{m}} - 1 \right]^{\frac{1}{n}}}{\alpha}, \quad (2.13)$$

where  $\Psi$  is the subsurface pressure head [L];  $\theta_s$  is the porosity;  $\theta_r$  is the residual soil moisture content;  $\alpha$  is a measure of the first moment of the pore size density function [L<sup>-1</sup>];  $n$  is an inverse measure of the second moment of the pore size density function; and  $m=1-1/n$ .

In addition, a terrain-following grid transformation with variable vertical discretization has been implemented in ParFlow by Maxwell (2013), which can solve groundwater problems with high topographic gradients. Furthermore, ParFlow has been developed for parallel computing systems and has been shown in many studies to be effective in solving large-scale problems at high resolution in highly heterogeneous media and under variably saturated conditions (Jones and Woodward, 2001; Kollet and Maxwell, 2006; 2008; Maxwell, 2013).

### **2.1.3 Coupling interface OASIS-MCT**

The OASIS-MCT coupler is used to control the exchange of fluxes and state variables between the three component models in TSMP (Valcke, 2013; Shrestha et al., 2014). In addition, OASIS-MCT uses time integration/averaging and spatial interpolation operations to keep scales constant when fluxes are at different spatial and temporal scales (Sulis et al., 2015). The exchange of data between the atmosphere and the land surface is ignored because only the CLM and ParFlow models were used in this work, and the atmospheric model was replaced by the reanalysis forcing data. CLM provides ParFlow with the upper boundary condition, which is net infiltration or exfiltration. Net infiltration includes precipitation, interception, total evaporation, and total transpiration (Kurtz et al., 2016; Zhang et al., 2018). In turn, ParFlow provides the calculated pressure and saturation values of the top ten subsurface layers to CLM.

## **2.2 Data assimilation**

### **2.2.1 Ensemble Kalman Filter (EnKF)**

The Ensemble Kalman Filter, a sequential filtering algorithm, can assimilate the measurement data and update the model states based on the optimal combination of the ensemble of model predictions and the measurements (Evensen, 1994; 2003). EnKF is a widely used DA technique and has been shown to be effective for nonlinear systems and high-dimensional problems (Camporese et al., 2009b; Schöniger et al., 2012).

Data assimilation consists of two primary steps: the forecast step and the analysis step. In the forecast step, the state estimate is based only on past data, whereas in the analysis step, information from current measurements and from a prior short-term forecast (based on past

data) is used to produce a current state estimate, and the probability density of the state is propagated forward (McLaughlin, 2002). This estimate is then used to initialize the next short-term forecast, which is then used in the next analysis, and so on (Hunt et al., 2007). The EnKF alternates sequentially between model forecast and filter analysis steps (also called filter updating), either state updates alone or joint state parameter updates.

The efficiency of the filter depends on the accurate determination of the forecast error covariance from the ensemble, and the sources of forecast errors are mainly uncertain initial conditions, forcing data, model parameters, and model equations (Turner et al., 2008). To achieve the ideal assimilation effect, the ensemble generation should ensure that errors from different sources are taken into account.

For each ensemble member  $j$  at time step  $i$ , the state vector  $\mathbf{x}_{j,i}$  in the forecast step is updated by the model predictions and is given by:

$$\mathbf{x}_{j,i} = \mathcal{M}(\mathbf{x}_{j,i-1}, \mathbf{q}_{j,i}, \mathbf{p}_{j,i}) \quad (2.14)$$

where  $j$  is the ensemble member,  $\mathbf{x}_{j,i}$  is the model forecast state vector at time step  $i$  (pressure head or soil moisture in this work),  $\mathcal{M}$  is the model (TSMP in this work),  $\mathbf{x}_{j,i-1}$  is the earlier model state vector at time step  $i-1$ ,  $\mathbf{q}_{j,i}$  is the vector with (perturbed) model forcings, and  $\mathbf{p}_{j,i}$  is the model perturbation vector with parameters. Model forecasts are updated according to:

$$\mathbf{x}_{j,i}^a = \mathbf{x}_{j,i}^f + \mathbf{K}_i(\mathbf{y}_{j,i} - \mathbf{H}_i\mathbf{x}_{j,i}^f) \quad (2.15)$$

where  $\mathbf{y}_{j,i}$  is the vector of (perturbed) observations (size is  $m$ ), and the superscripts  $a$  and  $f$  refer to the updated state vector (the analysis) and the model-predicted state vector, respectively. The observation operator  $\mathbf{H}_i$  is used to map the model forecasts into the observation space, which is assumed to be linear, and  $\mathbf{K}_i$  denotes the Kalman gain, which is calculated as:

$$\mathbf{K}_i = \mathbf{P}_i\mathbf{H}_i^T(\mathbf{H}_i\mathbf{P}_i\mathbf{H}_i^T + \mathbf{R}_i)^{-1} \quad (2.16)$$

where  $\mathbf{R}_i$  is the measurement error covariance matrix defined based on the expected measurement error of the assimilated data.  $\mathbf{P}_i$  is the model covariance matrix computed from the forecasted ensemble of model simulations at time step  $i$  according to:

$$\mathbf{P}_i = \frac{\sum_{j=1}^N (\mathbf{x}_{j,i}^f - \bar{\mathbf{x}}^f)(\mathbf{x}_{j,i}^f - \bar{\mathbf{x}}^f)^T}{N-1} \quad (2.17)$$

where  $\bar{\mathbf{x}}^f$  is a vector of ensemble means for the model states at time step  $i$ .  $N$  is the number of ensemble members.

The updated states are then finally given by:

$$\mathbf{x}_{j,i}^a = \mathbf{x}_{j,i}^f + \mathbf{P}_i \mathbf{H}_i^T (\mathbf{H}_i \mathbf{P}_i \mathbf{H}_i^T + \mathbf{R}_i)^{-1} (\mathbf{y}_{j,i} - \mathbf{H}_i \mathbf{x}_{j,i}^f) \quad (2.18)$$

The augmented state vector approach of Hendricks Franssen et al. (2008) is used to update both states ( $\boldsymbol{\psi}$ ) and parameters ( $\mathbf{Y}$ ) jointly.

$$\mathbf{x}_{j,i}^f = \begin{pmatrix} \boldsymbol{\psi}_{j,i} \\ \mathbf{Y}_{j,i} \end{pmatrix} \quad (2.19)$$

The dimension of the augmented state vector  $\mathbf{x}$  for realization  $j$  is expanded to  $2n$  (where all model grid cell sizes are  $n$ ). Furthermore, the covariance matrix ( $2n \times n$ ), Kalman gain ( $2n \times m$ ), and mapping operator ( $m \times 2n$ ) are all adjusted to include the unknown parameters. There are now two blocks of these vectors and matrices:

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{\psi\psi} \\ \mathbf{P}_{\psi Y} \end{bmatrix}$$

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_{\psi\psi} \\ \mathbf{K}_{\psi Y} \end{bmatrix} \quad (2.20)$$

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_\psi \\ \mathbf{H}_Y \end{bmatrix}$$

where the subscript  $\psi$  denotes the component related to the states, the subscript  $Y$  represents the component associated with the parameters, the subscript  $\psi\psi$  indicates evaluations (e.g., covariances) between two states at two locations, and the subscript  $\psi Y$  signifies evaluations between one state at a location and one parameter at a location.

The iterative application of the EnKF analysis may lead to filter inbreeding, i.e., underestimation of the ensemble covariance, due to the limited ensemble size in practical applications (Hendricks Franssen and Kinzelbach, 2008). Therefore, a damping factor ( $\alpha$ ) can be used during the analysis step to reduce the modification of the forecast with the Kalman gain and to limit the intensity of the perturbation of the states (pressure head) or parameters ( $\log_{10} K_s$ ) (Gebler et al., 2019; Hung et al., 2022). This leads to the subsequent updating equation for joint estimation of both state and parameter:

$$\mathbf{x}_{j,i}^a = \mathbf{x}_{j,i}^f + \boldsymbol{\alpha}^T \mathbf{P}_i \mathbf{H}_i^T (\mathbf{H}_i \mathbf{P}_i \mathbf{H}_i^T + \mathbf{R}_i)^{-1} (\mathbf{y}_{j,i} - \mathbf{H}_i \mathbf{x}_{j,i}^f) \quad (2.21)$$

where  $\boldsymbol{\alpha}^T$  is a vector of damping factors, which can have values between 0 and 1, for updating states or parameters.

### 2.2.2 Localized Ensemble Kalman Filter (LEnKF)

The effectiveness of DA with EnKF is often compromised by inaccuracies in the estimated model error covariances and the presence of spurious correlations between spatially distant grid cells (Houtekamer and Mitchell, 1998; Houtekamer and Mitchell, 2001). Houtekamer and Mitchell (1998) proposed a localization approach to restrict the filter update to the vicinity of the observation locations to avoid filter divergence caused by spurious correlations. Therefore, in this study, LEnKF, a variant of the sequential DA algorithm EnKF, is also used. Based on the localization of error covariances proposed by Houtekamer and Mitchell (2001), in the evaluation of the Kalman gain in Equation 2.16,  $\mathbf{P}_i$  is replaced by  $\rho \circ \mathbf{P}_i$ , where  $\rho \circ \mathbf{P}_i$  represents the Schur product of the correlation matrix  $\rho$  and the covariance matrix  $\mathbf{P}_i$ , where  $\rho$  is a correlation matrix containing the correlations between the grid cells (which are set to zero for

grid cell combinations that are separated beyond a certain threshold). Matrices  $\rho$  and  $\mathbf{P}_i$  have the same dimensions, so that the LEnKF analysis scheme can be expressed as:

$$\mathbf{x}_{j,i}^a = \mathbf{x}_{j,i}^f + [\rho_1 \circ \mathbf{P}_i \mathbf{H}_i^T] [\rho_2 \circ (\mathbf{H}_i \mathbf{P}_i \mathbf{H}_i^T) + \mathbf{R}_i]^{-1} (\mathbf{y}_{j,i} - \mathbf{H}_i \mathbf{x}_{j,i}^f), \quad (2.22)$$

Here  $\rho$  is determined using a piecewise fifth-order function, as given by Gaspari and Cohn (1999). The correlation  $\omega$  between a grid point and an observation, i.e., an element in  $\rho$ , can be approximated as (Hu et al., 2012):

$$\omega(l, e) = \begin{cases} 1 - \frac{1}{4} \left(\frac{e}{l}\right)^5 + \frac{1}{2} \left(\frac{e}{l}\right)^4 + \frac{5}{8} \left(\frac{e}{l}\right)^3 - \frac{5}{3} \left(\frac{e}{l}\right)^2, & 0 \leq e \leq l \\ \frac{1}{12} \left(\frac{e}{l}\right)^5 - \frac{1}{2} \left(\frac{e}{l}\right)^4 + \frac{5}{8} \left(\frac{e}{l}\right)^3 + \frac{5}{3} \left(\frac{e}{l}\right)^2 - 5 \left(\frac{e}{l}\right) + 4 - \frac{2}{3} \left(\frac{e}{l}\right)^{-1}, & l < e \leq 2l \\ 0, & e > 2l \end{cases} \quad (2.23)$$

where  $l$  is the defined localization radius and  $e$  is the Euclidean distance between an analyzed grid point and an observation location. The correlation  $\omega$  is distance-dependent and varies between 1 at observation locations and 0 at distances greater than twice the influence radius  $l$ .

### 2.2.3 TSMP-PDAF

The TSMP-PDAF framework was developed by Kurtz et al. (2016), which couples the TSMP integrated model with the Parallel Data Assimilation Framework (PDAF) (Nerger et al., 2005). By developing a parallel DA system, PDAF facilitates the implementation of parallel assimilation systems with numerical models for more efficient DA. In this framework, a fully parallel modular environment is provided to perform DA for the partially or fully coupled model components. Technical details on the coupling of PDAF to TSMP can be found in Kurtz et al. (2016). The current version of TSMP-PDAF allows assimilation of SM (in situ or RS), pressure measurements (river stage or groundwater levels), and water storage, either separately or jointly. States can be updated separately or jointly with parameters, and the estimated hydraulic parameters can be saturated hydraulic conductivity, porosity, Mulaem-van Genuchten parameters, and Manning's roughness coefficient.

The TSMP-PDAF framework has been applied to assimilate different hydrologic variables (SM or groundwater) at different scales (e.g., hillslope, catchment, and continental scales) (Kurtz et al., 2016; D. Baatz et al., 2017; Zhang et al., 2018; Gebler et al., 2019; Naz et al., 2019; Naz et al., 2020; Hung et al., 2022; Brandhorst and Neuweiler, 2023). Kurtz et al. (2016) assimilated SM into the TSMP model (CLM-ParFlow) in a virtual catchment, and the effects of DA were demonstrated in the simple synthetic DA, and the model states and parameters were corrected to the reference values. Their experiments also showed that the TSMP-PDAF framework can run efficiently with parallel computations and large problem sizes, and thus holds promise for simulating and predicting states and fluxes of the terrestrial system over large spatial scales at high resolution. D. Baatz et al. (2017) proposed a catchment tomography approach and assimilated stream water level observations into TSMP (ParFlow stand-alone)-PDAF in a synthetic catchment. The successful estimation of the spatial distribution of Manning's coefficient in their experiments proved the effectiveness of the joint state parameter estimation in the TSMP-PDAF framework. Zhang et al. (2018) found that the joint assimilation of GWL and SM data has great potential to improve the SM characterization of the root zone, and the assimilation was performed via TSMP (CLM-ParFlow)-PDAF in synthetic experiments with only four grid cells. Naz et al. (2019) used the TSMP-PDAF framework to assimilate the satellite SM into the stand-alone CLM 3.5 model over Europe. The results of their assimilation experiment showed an overall improvement in SM and runoff, demonstrating the potential of assimilating remotely sensed SM to improve continental-scale SM and runoff simulations at high resolution. Gebler et al. (2019) assimilated SM data into TSMP in a synthetic test case and concluded that EnKF can significantly improve SM characterization. However, they also found that DA experiments yielded relatively poor performance for the real-world case of Rollesbroich, a small catchment in the Eifel mountains in western Germany. They argue that the different performance for the real-world case and the synthetic case may be related to model structure errors. Hung et al. (2022) assimilated SM and GWL at the large catchment scale in a synthetic study that mimicked the Neckar catchment in Southwestern Germany with the TSMP (CLM-ParFlow)-PDAF model. In their experiments, SM and pressure were updated together or separately, and hydraulic conductivity was updated together with the states in some experiments. For the SM update, the CRNS data were vertically weighted according to the scheme from Franz et al. (2012). The synthetic experiment showed that the assimilation of SM and GWL was able to correct the predictions of SM and GWL towards the reference values. Joint state parameter updating performed better than updating states alone, although the improvement was local and limited. Brandhorst and Neuweiler (2023) investigated the effect of parameter updates on SM estimation with a synthetic three-dimensional hydrologic hillslope

model. The results showed that the best SM estimates were obtained by performing a joint update of porosity, van Genuchten parameters, and (optionally) saturated hydraulic conductivity.

Previous studies have investigated the assimilation of SM and GWL measurements with TSMP (CLM-ParFlow)-PDAF and demonstrated the effectiveness and reliability of the assimilation system in this framework. However, they mainly focused on synthetic experiments, small catchments, stand-alone models, or greatly simplified the representation of spatial heterogeneity in the real catchment. In this work, we assimilated SM from a network of 12 CRNS sites and real groundwater measurements into the integrated model TSMP using the DA methods EnKF and LEnKF.

### **2.3 Study area and hydrological measurements**

The study area for the assimilation experiments conducted for this dissertation is the Rur catchment (see Fig. 2.2). With a total area of 2354 km<sup>2</sup>, it is mainly located in western Germany, with smaller parts in Belgium and the Netherlands. The Rur catchment was chosen as the study domain because of its spatial variability in climate, altitude, land use, soil properties, and geology. More importantly, the Rur catchment is part of the TERENO (TERrestrial Environmental Observatories, <https://www.tereno.net/>) observing platform and is equipped with high-density monitoring infrastructure, including CRNS soil moisture and GWL measurement networks, eddy covariance stations, and river discharge stations (Bogena et al., 2018). Therefore, the extensive observations can provide valuable information for DA in integrated terrestrial models. The elevation of the terrain in the Rur catchment generally decreases from 690 to 15 m a.s.l. from the south to the north, so that the river flows from the south to the north. Corresponding to the topography from north to south, mean annual air temperature decreases from 10 to 7 °C, annual precipitation increases from 650 to 1300 mm, and annual potential ET decreases from 850 to 450 mm (Montzka et al., 2008; Bogena et al., 2018). The main types of land use are arable agriculture in the north (mainly maize and wheat) and grassland, and coniferous and deciduous forests in the south (Waldhoff and Lusse, 2015; Roland Baatz et al., 2017). Other important land uses include open-cast lignite quarries and urban areas (Shukla et al., 2023). The permeability of the upper unconfined aquifer and the groundwater recharge in the southern low-mountain range are relatively low due to the prevailing consolidated bedrock, while they are relatively high in the northern lowland region due to the presence of unconsolidated bedrock (Bogena et al., 2018).

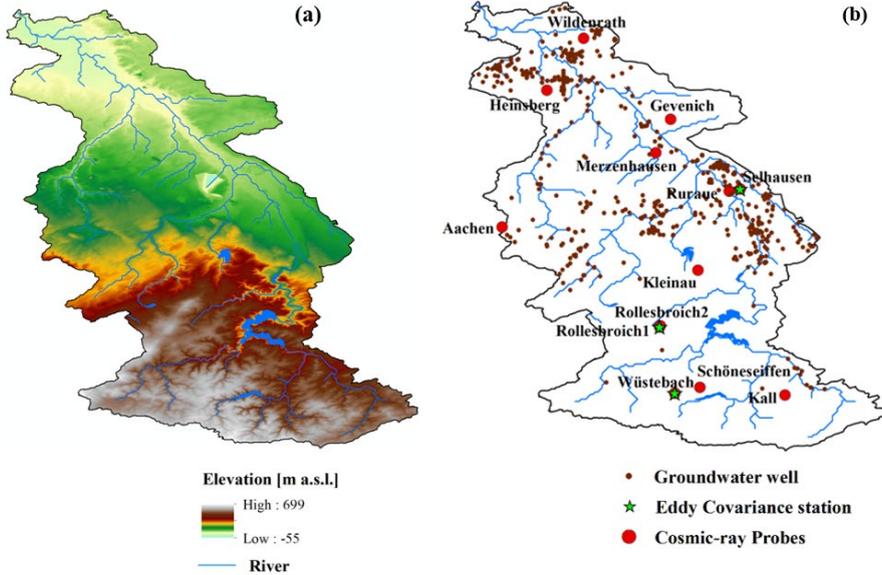


Figure 2.2 Topography of the Rur catchment (a) and locations of the hydrological stations (b), including groundwater wells, cosmic-ray neutron sensors, and eddy covariance stations.

In the Rur catchment, the CRNS network consists of 13 relatively evenly distributed cosmic ray neutron stations (Table 2.1), covering the relevant land use types and allowing temporal dynamic monitoring of field-scale SM over the entire Rur catchment (Baatz et al., 2014; Bogena et al., 2018; Bogena et al., 2022). Either a CRS-1000 or CRS-2000/B cosmic-ray neutron probe (Hydroinnova LLC) is used in the CRNS (Bogena et al., 2018). The raw measured neutron intensity detected by the CRNS probe must be corrected and can then be converted to SM values (Bogena et al., 2013; R. Baatz et al., 2015; Bogena et al., 2022). The near-real-time SM data provided by CRNS are well suited for catchment-scale DA and hydrological modeling verification.

Table 2.1 CRNS sites used in this study, including key site characteristics.

Name	Latitude (degr)	Longitude (degr)	Altitude (m)	Mean annual precipitation (mm y <sup>-1</sup> )	Mean air temperature (°C)	Land use
Merzenhausen	50.93	6.30	91	718	10.3	crop
Rollesbroich1	50.62	6.30	515	1018	7	grassland

Rollesbroich2	50.62	6.31	506	1018	7	grassland
Gevenich	50.99	6.32	107	718	10.3	crop
Ruraue	50.86	6.43	100	718	10.3	grassland
Wildenrath	51.13	6.17	72	722	10.3	needleleaf
Wüstebach	50.51	6.33	605	1401	7	spruce
Heinsberg	51.04	6.10	58	722	10.3	crop
Kall	50.50	6.53	505	857	8	grassland
Selhausen	50.87	6.45	101	718	10.3	crop
Schöneseiffen	50.52	6.38	611	870	7	grassland
Kleinhau	50.72	6.37	374	614	9	grassland
Aachen	50.80	6.03	232	865	10.3	crop

There are hundreds of groundwater wells in the Rur catchment that are used to monitor groundwater levels, with the number of observation wells varying slightly from year to year. The observation frequency also varies for each well and can be on a daily, weekly, or monthly basis. Groundwater wells are mainly concentrated in the northern part of the catchment, and there are many sites distributed along the river. The measured groundwater data can be accessed via the monitoring network Geoportal NRW ([www.geoportal.nrw](http://www.geoportal.nrw)).

Three eddy covariance (EC) stations (Rollesbroich, Wüstebach, and Selhausen) in the Rur catchment continuously monitor evapotranspiration. The EC measurements were performed with a sonic anemometer (CSAT3, Campbell Scientific Inc., Logan, USA) to measure the 3D wind components, an open-path gas analyzer (Li7500, LI-COR Inc., Lincoln, USA) to determine H<sub>2</sub>O and CO<sub>2</sub> concentrations in the air, and an air temperature and humidity sensor (HMP45C, Vaisala Inc., Helsinki, Finland). The EC data require conversion and uncertainty estimation before being used for model verification (Mauder et al., 2013).

## **Chapter 3: Water table depth assimilation in integrated terrestrial system models at the larger catchment scale**

\*adapted from: Li, F., Kurtz, W., Hung, C. P., Vereecken, H., and Hendricks Franssen, H.-J.: Water table depth assimilation in integrated terrestrial system models at the larger catchment scale, *Frontiers in Water*, 5, 10.3389/frwa.2023.1150999, 2023.

### **3.1 Introduction**

As a widespread and highly used resource, groundwater provides globally 50% of the drinking water, with higher values for inhabitants of dry regions, and 2.5 billion people depend entirely on groundwater resources for their basic daily water needs (UNWWAP, 2015). Either groundwater level or water table depth is a significant variable related to groundwater and can vary between 0 m in wetland areas to depth of hundreds of meters from the land surface in arid regions. Shallow groundwater is crucial in terrestrial ecosystems as it strongly influences the soil water content in the root zone and thus exerts an important control on water and energy fluxes between the subsurface, land surface, and the atmosphere (Koster et al., 2004; Vereecken et al., 2016). To understand the influence of temporal and spatial variations of groundwater level on terrestrial ecosystems, models like the integrated TSMP model (Shrestha et al., 2014) are used, which can model the groundwater-soil-vegetation-atmosphere system in a physical manner.

However, the accuracy of modeling is often affected by uncertain model forcings, parameters, and initial conditions (Freeze, 1975; D. Baatz et al., 2017). Especially for groundwater systems, the extreme spatial heterogeneity of hydraulic parameters is challenging (de Marsily, 1986). To quantify and reduce the uncertainties of model predictions, DA can be used to correct model predictions with observations and improve the estimation of hydrological variables (Reichle et al., 2002; Reichle et al., 2008). One of the most commonly used DA algorithms is the EnKF (Evensen, 1994; 2003). EnKF uses a Monte Carlo approach to forecast model error statistics (Evensen, 1994). DA with EnKF is often affected by a poor quality of the estimated model error covariances and spurious correlations between grid cells which are separated far in space. Therefore, the LEnKF approach (Houtekamer and Mitchell, 2001) is used in this study, which can improve the effectiveness of the EnKF analysis (Hu et al., 2012).

Past studies have proven the effectiveness of DA in improving real-time hydrological modeling and forecasting (e.g. Han et al., 2014; Zhang et al., 2018; Yu et al., 2020), and some

studies investigated the assimilation of groundwater data with hydrological models. Camporese et al. (2009b) assimilated both pressure head and SM data with the EnKF, and the results showed that assimilation of either pressure head or soil moisture can improve the characterization of subsurface states in the vicinity of the measurement locations. Camporese et al. (2009a) assimilated synthetic observations of pressure head and streamflow for a v-tilted catchment, and the results suggested that streamflow prediction can be improved by assimilation of pressure head and streamflow, either individually or simultaneously. Kurtz et al. (2014) assimilated jointly piezometric heads and groundwater temperatures with EnKF to update uncertain hydraulic subsurface parameters (i.e., hydraulic conductivities and leakage coefficients) for an area near the river Limmat in Switzerland, and found that the joint assimilation of the two kinds of data with updating of uncertain hydraulic parameters gives the best characterization. Zhang et al. (2016) assimilated soil moisture and groundwater head measurements with the MIKE SHE hydrological model for catchments of different complexities and using different assimilation settings (observation types, ensemble sizes, and localization schemes) and found that the ensemble transform Kalman filter (ETKF) method improved the model performance compared to the OL run. But the average difference between observations and model simulations was subtracted from the original data when comparing in-situ head measurements with predictions. The proposed scheme by Zhang et al. (2016) with both distance localization and variable localization was shown to be more robust than only using one localization scheme and provided better results. However, these experiments on groundwater assimilation have only been conducted by hydrological models in synthetic experiments or over-simplified real-world cases. No studies demonstrated the potential of assimilating real groundwater observations into integrated terrestrial system models to improve groundwater estimates at the regional scale. This is therefore still an emerging research topic.

The integrated model TSMP, which is composed of an atmospheric, land surface, and subsurface model, was used in this work, in combination with the Parallel Data Assimilation Framework (PDAF) (Nerger et al., 2005; Kurtz et al., 2016). TSMP has been applied in a series of studies (e.g., Shrestha et al., 2015; Keune et al., 2016; Furusho-Percot et al., 2019). The combination of PDAF and TSMP has been used for the assimilation of different hydrological variables (e.g., SM and groundwater) at different scales (e.g., hillslope, catchment, and continental scale) (Kurtz et al., 2016; Zhang et al., 2018; Gebler et al., 2019; Naz et al., 2019; Naz et al., 2020; Hung et al., 2022). Zhang et al. (2018) demonstrated in synthetic experiments with only four grid cells that the joint assimilation of GWL and SM data has great potential to improve root zone SM characterization. Hung et al. (2022) assimilated groundwater levels at the large catchment scale in a synthetic study that mimicked the Neckar catchment, and the

results showed that GWL assimilation can lead to a large scale improved characterization of groundwater levels, also between groundwater wells, but the impact of GWL assimilation on other compartments of the terrestrial system was limited, except for the deep vadose zone. However, there is still a lack of studies with integrated land surface-subsurface models to investigate assimilation of real groundwater measurements at the larger catchment scale. The objective of this study was to investigate whether assimilating groundwater data into the integrated terrestrial systems model TSMP at the larger catchment scale for a real-world case is able to achieve a better characterization of groundwater levels (and other terrestrial system states and fluxes) than an OL run and identify the main limitations and complications in practice. Furthermore, SM measured by cosmic-ray neutron sensors was used to verify the model simulation accuracy and evaluate whether assimilating WTD data can improve SM characterization. This is a novel contribution, as the assimilation of groundwater measurements in integrated land surface-subsurface models at the larger catchment scale with real data has not been carried out before.

## **3.2 Materials and methods**

### **3.2.1 Study area and data**

The simulation domain is the Rur catchment (2354 km<sup>2</sup>) which is situated in western Germany and includes a small part of Belgium and the Netherlands. The Digital Elevation Model (DEM) for the area was acquired from SRTM 90m Version 4 (Jarvis et al., 2008) and is shown in Fig 3.1. The altitude ranges from 15m to 690 m a.s.l., decreasing from south to north, and the Rur river flows from the Eifel hills in the south to the northern flat terrain. From the northern to the southern part of the catchment, long-term average annual precipitation ranges from 650mm to 1300mm, mean annual air temperature decreases from 10 to 7°C and mean annual potential evapotranspiration varies between 850mm and 450mm (Montzka et al., 2008; Bogen et al., 2018). The land use types were taken from the CRC/TR32 Database (Waldhoff and Lussem, 2015) and are mainly agriculture (corn, sugarbeet and wheat in the north), grassland, and coniferous and deciduous forest (southern mountainous areas).

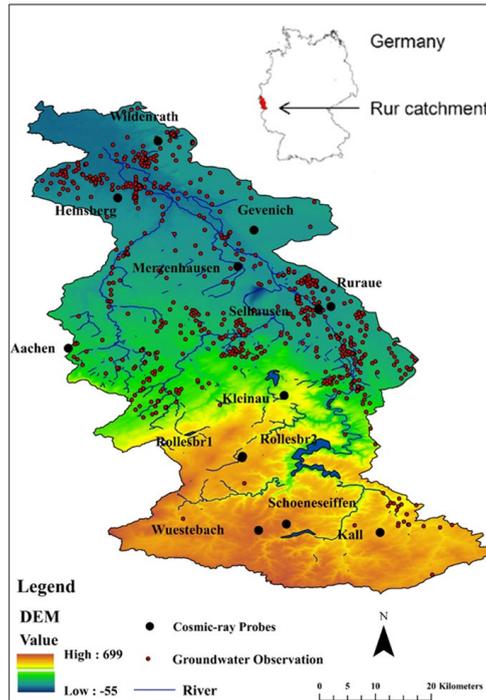


Figure 3.1 Map of the Rur catchment and locations of the 13 cosmic-ray neutron sensors (CRNS) (black points) and groundwater measurement sites (red points) in the year 2018. The Rur catchment is situated in western Germany.

The high-resolution (1:50,000) regional soil map BK50 (Geologischer Dienst NRW, 2009) (see Fig 3.2) and European Soil Database (ESDB) (Pano, 2006) were used to obtain the soil characteristics and to calculate the soil hydraulic properties. Bulk density was obtained from ESDB.

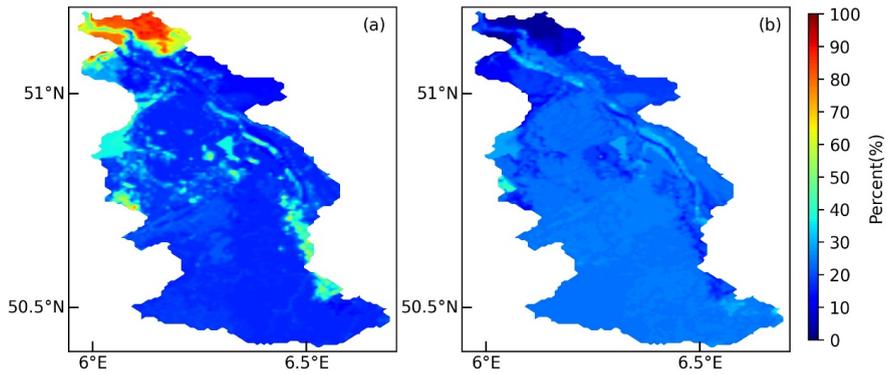


Figure 3.2 Sand (a) and clay (b) content (%) for the Rur catchment derived from the BK50 soil map.

Based on the thickness of the BK50 soil layers, we treat the layers below the soil layers as aquifer layers. The upper aquifer hydraulic conductivity (see Fig 3.3) was obtained from the Information System Hydrogeological Map of North Rhine-Westphalia with a resolution of 1:100,000 (<https://www.opengeodata.nrw.de/produkte/geologie/geologie/HK/ISHK100/>). The permeability of the aquifer is based on different classes of rock types.

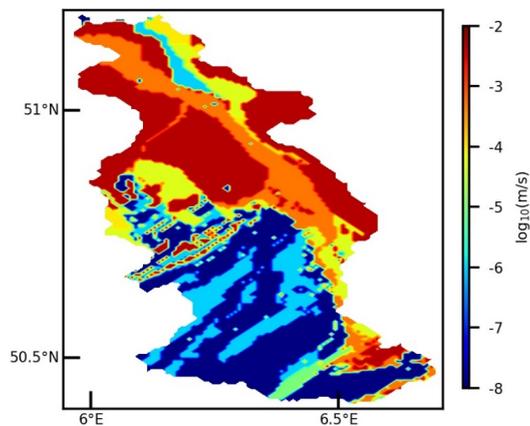


Figure 3.3 Hydraulic conductivity of the aquifer material for the Rur catchment.

The high-resolution reanalysis dataset COSMO-REA6 developed with the numerical weather prediction (NWP) model COSMO (Baldauf et al., 2011) is used as atmospheric forcing in this work (Bollmeyer et al., 2015; Wahl et al., 2017). Currently, the reanalysis covers the period 1995-2019 at a high spatial resolution of  $0.055^\circ$  (6 km) and is continuously extended by the German Weather Service (Deutscher Wetterdienst; DWD). The forcing data include precipitation, air temperature, air pressure, wind velocity, specific humidity, incoming shortwave radiation, and incoming longwave radiation.

The measured WTD data from the monitoring network Geoportal NRW ([www.geoportal.nrw](http://www.geoportal.nrw)) were used for assimilation and some as independent verification data for the model simulations. For the year 2018, there were 865 sites located in shallow and deep aquifers of the Rur catchment that monitored the WTD (see Fig 3.1), and most measurement sites are distributed along the river. The observation frequency varies for each site, and can be daily, weekly, or monthly. In 2018, there were 575 sites with WTD between 0 and 20 m. We only used the sites with WTD between 0 and 20 m to be sure that only measurements from the upper aquifer were included for assimilation or verification, as our model only considered the upper 20 m.

CRNS is a precise method to measure SM at the field scale (Zreda et al., 2008; Baatz et al., 2014; Köhli et al., 2015). The Rur catchment CRNS network comprises 13 CRNS stations (see Table 3.1) (CRS1000, HydroInnova LLC, 2009) (Roland Baatz et al., 2017; Bogen et al., 2022) and these observation sites are relatively evenly distributed over the study area (see Fig 3.1). CRNS measures the fast neutron intensity, and the measured number of neutron counts shows an inverse correlation with SM content. Fast neutrons originate from the collisions of secondary cosmic particles from outer space with terrestrial atoms. Fast neutrons are moderated most effectively by hydrogen, since the mass of the neutron is similar to the mass of a nucleus of the hydrogen atom. Thus, the corresponding fast neutron intensity measured by CRNS strongly depends on the amount of hydrogen within the CRNS footprint, allowing for a continuous non-invasive SM estimate at the field scale (Roland Baatz et al., 2017). The horizontal footprint of this measurement matches the 500 m horizontal model resolution quite well. It can measure SM until 83 cm depth under very dry conditions, and to 15 cm depth under very wet soil conditions (Köhli et al., 2015; Schrön et al., 2017).

Table 3.1 CRNS sites with geographical information

Name	Latitude	Longitude	Altitude (m a.s.l.)	Mean annual precipitation (mm y <sup>-1</sup> )	Mean air temperature (°C)	Land Use
Merzenhausen	50.930	6.297	91	718	10.3	crop
Rollesbroich1	50.622	6.304	515	1018	7	grassland
Rollesbroich2	50.624	6.305	506	1018	7	grassland
Gevenich	50.989	6.324	107	718	10.3	crop
Ruraue	50.862	6.427	100	718	10.3	grassland
Wildenrath	51.133	6.169	72	722	10.3	needleleaf
Wuestebach	50.505	6.331	605	1401	7	spruce
Heinsberg	51.041	6.104	58	722	10.3	grassland, crop
Kall	50.501	6.526	505	857	8	grassland
Selhausen	50.866	6.447	101	718	10.3	crop
Schoeneseiffen	50.515	6.376	611	870	7	grassland
Kleinhau	50.722	6.372	374	614	9	grassland
Aachen	50.799	6.025	232	865	10.3	crop

### 3.2.2 Model description (TSMP)

The coupled terrestrial system model used consists of three compartments integrated under the framework TSMP, the 3D variably saturated groundwater flow model ParFlow for the subsurface (Kollet and Maxwell, 2006), the land surface model CLM version 3.5 (Community Land Model) from the National Center for Atmospheric Research (Oleson et al., 2004; Oleson et al., 2008), and the numerical weather prediction model COSMO (Consortium for Small Scale Modelling) (Baldauf et al., 2011). These three models are two-way coupled by the Ocean Atmosphere Sea Ice Soil coupling Model Coupling Toolkit (OASIS-MCT, version 3) (Valcke, 2013). The integrated modeling platform TSMP can run with different combinations of the component models (Shrestha et al., 2014). In this study, CLM-ParFlow was used without COSMO.

#### 3.2.2.1 Land surface model CLM

The biophysical processes simulated by CLM3.5 include solar and longwave radiation interactions with vegetation canopy and soil, momentum and turbulent fluxes from canopy and soil, canopy hydrology (e.g., interception processes), soil hydrology, and stomatal physiology and photosynthesis (Oleson et al., 2007). The mass and energy balance solved by CLM include soil evaporation, evaporation from intercepted water, transpiration from plants, infiltration of water in the soil, sensible and ground heat fluxes, and freeze-thaw processes (Oleson et al., 2004; Oleson et al., 2008).

The nested subgrid hierarchy is used to represent spatial land surface heterogeneity (Oleson et al., 2008). Each grid cell is divided into a variety of land units (glacier, lake, wetland, urban, and vegetated), where each land unit can have a different number of snow/soil columns, and each column can have multiple plant functional types (PFTs) (Bonan et al., 2002; Oleson et al., 2008). In CLM, the soil column and snow column are discretized into 10 and 5 vertical layers, respectively (Oleson et al., 2004; Oleson et al., 2008). Each PFT is characterized by distinct plant physiological parameters, which could capture the biogeophysical and biogeochemical differences between the different vegetation types (Oleson et al., 2004; Oleson et al., 2008).

### 3.2.2.2 Subsurface hydrological model ParFlow

In TSMP, the soil hydrology of CLM is substituted by the soil hydrology of ParFlow (Kollet and Maxwell, 2008) and also surface runoff and groundwater flow are calculated by ParFlow. ParFlow is a three-dimensional variably saturated groundwater flow model improved with a two-dimensional overland flow simulator (Ashby and Falgout, 1996; Kollet and Maxwell, 2006). It combines the kinematic wave equation (Lighthill and Whitham, 1955) and the 3D Richards' equation (Richards, 1931) to describe the dynamic coupling of surface-subsurface flow under overland flow boundary conditions (Kollet and Maxwell, 2006). In ParFlow, the three-dimensional Richards' equation can be written as follows (Maxwell, 2013):

$$S_s S_w \frac{\partial h}{\partial t} + \Phi \frac{\partial S_w(h)}{\partial t} = \nabla \cdot \mathbf{q} + q_r(x, z), \quad (3.1)$$

and

$$\mathbf{q} = \Phi S_w(h) \mathbf{v} = -K_s(x) k_r(h) \nabla(h + z), \quad (3.2)$$

where  $S_s$  is the specific storage [ $L^{-1}$ ];  $S_w$  is the relative saturation;  $h$  is the pressure-head [L];  $t$  is time [T];  $\Phi$  is the porosity;  $\mathbf{q}$  is the specific volumetric (Darcy) flux [ $LT^{-1}$ ];  $q_r$  is a general source/sink term that represents transpiration, wells, and other fluxes [ $LT^{-1}$ ];  $x$  is the length along the  $x$ -axis specified as horizontal direction [L];  $z$  is the elevation along the  $z$ -axis

specified as upward to be positive [L];  $\mathbf{v}$  is the subsurface flow velocity [LT<sup>-1</sup>];  $K_s(x)$  is the saturated hydraulic conductivity tensor [LT<sup>-1</sup>];  $k_r$  is the relative permeability.

ParFlow requires input soil hydraulic parameters like saturated hydraulic conductivity, porosity, and Mualem-van Genuchten parameters (van Genuchten, 1980; Kollet and Maxwell, 2008). For our study, saturated hydraulic conductivities of soil layers were calculated from sand, silt and clay contents and bulk density using the software rosettav3 H3, which is based on an artificial neural network analysis coupled with a bootstrap resampling method (Schaap et al., 2001; Zhang and Schaap, 2017). To keep hydraulic consistency between CLM and ParFlow, porosity ( $\theta_s$ ) for both models is calculated on the basis of the sand fraction via the following pedotransfer function in CLM (Oleson et al., 2004):

$$\theta_s = 0.489 - 0.00126(\text{sand}\%), \quad (3.3)$$

The van Genuchten formulation (van Genuchten, 1980) is employed to evaluate the pressure head from soil moisture data:

$$\Psi = - \frac{\left[ \left( \frac{\theta_s - \theta_r}{\theta(\Psi) - \theta_r} \right)^{\frac{1}{m}} - 1 \right]^{\frac{1}{n}}}{\alpha}, \quad (3.4)$$

where  $\Psi$  is subsurface pressure head [L];  $\theta_s$  is porosity;  $\theta_r$  is residual soil moisture content;  $\alpha$  is a measure of the first moment of the pore size density function [L<sup>-1</sup>];  $n$  is an inverse measure of the second moment of the pore size density function; and  $m=1-1/n$ .

The Newton Krylov solution technique is applied in ParFlow and acts as a nonlinear solver (Jones and Woodward, 2001). The coupled partial differential equations for subsurface flow and surface water flow are solved by the Newton-Krylov method with multigrid preconditioning, which is good at handling subsurface flow problems at large-scales in highly heterogeneous media and under variably saturated conditions (Kollet and Maxwell, 2006; 2008; Maxwell, 2013). A prominent advantage of ParFlow is that it was designed for parallel computer systems, so that it can efficiently compute large-scale problems at high resolution, which has been demonstrated in many studies (Jones and Woodward, 2001; Kollet and Maxwell,

2006; 2008).

### 3.2.2.3 Coupling interface OASIS-MCT

The external coupler OASIS-MCT (Valcke, 2013) is used to couple CLM and ParFlow, and control the exchange of fluxes between the different component models (Shrestha et al., 2014). When the fluxes correspond to different spatial and temporal scales, OASIS-MCT uses time integration/averaging and spatial interpolation operators to keep the scales consistent (Sulis et al., 2015). In TSMP, ParFlow provides CLM with the upper 10 subsurface layers' pressure and saturation, and in turn, CLM provides ParFlow with the upper boundary condition, which is net infiltration or exfiltration. The net infiltration includes precipitation, interception, total evaporation, and total transpiration (Zhang et al., 2018).

### 3.2.3 LEnKF methodology

DA consists of a forecast and an analysis step. For the forecast step, the state estimation is only based on past data (McLaughlin, 2002). For the analysis step, the information from current measurements and from a prior short-term forecast (which is based on past data) is used to produce a current state estimate. Then, the estimate will be used to initialize the next short-term forecast, which is subsequently used in the next analysis, and so on (Hunt et al., 2007). The EnKF sequentially performs a model forecast and a filter analysis. The efficiency of the filter relies on the accurate determination of the forecast error covariance from the ensemble, and the main sources of forecast errors are initial conditions, forcing data, model parameters, and model equations (Turner et al., 2008). Perturbation approaches can take these error sources for the ensemble generation into account.

For each ensemble member  $j$  at time step  $i$ , the state vector  $\mathbf{x}_{j,i}$  in the forecast step is updated by model predictions and is given by:

$$\mathbf{x}_{j,i} = \mathcal{M}(\mathbf{x}_{j,i-1}, \mathbf{q}_{j,i}, \mathbf{p}_{j,i}), \quad (3.5)$$

where  $j$  is the ensemble member,  $\mathbf{x}_{j,i}$  is the model forecast state vector at time step  $i$  (pressure head in our study),  $\mathcal{M}$  is the model TSMP,  $\mathbf{x}_{j,i-1}$  is the earlier model analysis with state

vector at time step  $i-1$ ,  $\mathbf{q}_{j,i}$  is the vector with (perturbed) model forcings (perturbed forcings are precipitation, incoming shortwave radiation, incoming longwave radiation, and air temperature in this study) and  $\mathbf{p}_{j,i}$  denotes the model perturbation vector with parameters (porosity and saturated hydraulic conductivity in this study). In summary, in this work, the ensemble of model realizations is generated by different initial conditions, forcings, and parameters. Model forecasts are updated according to:

$$\mathbf{x}_{j,i}^a = \mathbf{x}_{j,i}^f + \mathbf{K}_i (\mathbf{y}_{j,i} - \mathbf{H}_i \mathbf{x}_{j,i}^f), \quad (3.6)$$

where  $\mathbf{y}_{j,i}$  is the vector with (perturbed) observations, and the superscripts  $a$  and  $f$  refer to the updated state vector (the analysis) and the model predicted state vector, respectively. The observation operator  $\mathbf{H}_i$  is used to map model forecasts into the observation space, which is here assumed to be linear, and  $\mathbf{K}_i$  denotes the Kalman gain that is calculated as:

$$\mathbf{K}_i = \mathbf{P}_i \mathbf{H}_i^T (\mathbf{H}_i \mathbf{P}_i \mathbf{H}_i^T + \mathbf{R}_i)^{-1}, \quad (3.7)$$

where  $\mathbf{R}_i$  is the measurement error covariance matrix, and  $\mathbf{P}_i$  is the model covariance matrix, which is calculated from the forecasted ensemble of model simulations at time step  $i$  according to:

$$\mathbf{P}_i = \frac{\sum_{j=1}^N (\mathbf{x}_{j,i}^f - \bar{\mathbf{x}}^f) (\mathbf{x}_{j,i}^f - \bar{\mathbf{x}}^f)^T}{N-1}, \quad (3.8)$$

where  $\bar{\mathbf{x}}^f$  is a vector with ensemble mean values for the model states at time step  $i$ .  $N$  is the number of ensemble members.

The estimation of the covariances with a limited ensemble is affected by strong sampling fluctuations, and the estimated covariances might be affected by spurious correlations (Houtekamer and Mitchell, 1998; Houtekamer and Mitchell, 2001). Houtekamer and Mitchell (1998) suggested a localization approach to remove spurious correlations to avoid filter

divergence, limiting the updates to the surroundings of observations. Based on the localization of the error covariances proposed by Houtekamer and Mitchell (2001), in the evaluation of the Kalman gain in equation 3.7,  $\mathbf{P}_i$  is replaced by  $\rho \circ \mathbf{P}_i$ ,  $\rho \circ \mathbf{P}_i$  represents the Schur product of the correlation matrix  $\rho$  and covariance matrix  $\mathbf{P}_i$ , where  $\rho$  is a correlation matrix containing correlations between the grid cells (which are set to zero for grid cell combinations that are separated beyond a certain threshold). And the  $\rho$  and  $\mathbf{P}_i$  should have the same dimensions, so the LEnKF analysis scheme can be expressed as:

$$\mathbf{x}_{j,i}^a = \mathbf{x}_{j,i}^f + [\rho_1 \circ \mathbf{P}_i \mathbf{H}_i^T] [\rho_2 \circ (\mathbf{H}_i \mathbf{P}_i \mathbf{H}_i^T) + \mathbf{R}_i]^{-1} (\mathbf{y}_{j,i} - \mathbf{H}_i \mathbf{x}_{j,i}^f), \quad (3.9)$$

Here  $\rho$  is determined by using a fifth-order piecewise function, as given by Gaspari and Cohn (1999). The correlation  $\omega$  between a grid point and an observation, i.e., an element in  $\rho$ , can be approximated as (Hu et al., 2012):

$$\omega(l, e) = \begin{cases} 1 - \frac{1}{4} \left(\frac{e}{l}\right)^5 + \frac{1}{2} \left(\frac{e}{l}\right)^4 + \frac{5}{8} \left(\frac{e}{l}\right)^3 - \frac{5}{3} \left(\frac{e}{l}\right)^2, & 0 \leq e \leq l \\ \frac{1}{12} \left(\frac{e}{l}\right)^5 - \frac{1}{2} \left(\frac{e}{l}\right)^4 + \frac{5}{8} \left(\frac{e}{l}\right)^3 + \frac{5}{3} \left(\frac{e}{l}\right)^2 - 5 \left(\frac{e}{l}\right) + 4 - \frac{2}{3} \left(\frac{e}{l}\right)^{-1}, & l < e \leq 2l \\ 0, & e > 2l \end{cases} \quad (3.10)$$

where  $l$  is the defined localization radius and  $e$  is the Euclidean distance between an analyzed grid point and an observation location. The correlations  $\omega$  are distance-dependent and vary between 1 at observation locations and 0 at distances greater than twice the influence radius  $l$ . Only the observations located within the localization radius from an analyzed grid point can contribute to the analysis for this grid point (Hu et al., 2012). The cutoff radius can filter out small and noisy correlations associated with remote observations (Houtekamer and Mitchell, 2001). A larger radius may contain more spurious correlation, resulting in less effective assimilation. In contrast, a radius that is too small limits the influence of observations too much to update neighboring grid cells. Therefore, determining an appropriate assimilation localization radius is crucial.

### 3.2.4 Assimilation methodology

To be able to assimilate WTD measurements into TSMP, WTD data need to be transferred into pressure accordingly (see Fig. 3.4). At locations with WTD measurements, the pressure head in the saturated zone is calculated from the measured WTD assuming a hydrostatic pressure distribution, according to Zhang et al. (2018):

$$\Psi_i = D_i - WTD^{obs}, \quad (3.11)$$

Where  $\Psi_i$  is the pressure head at the  $i^{\text{th}}$  soil layer [L],  $D_i$  is the depth from land surface to the  $i^{\text{th}}$  soil layer [L], and  $WTD^{obs}$  is the observed WTD [L].

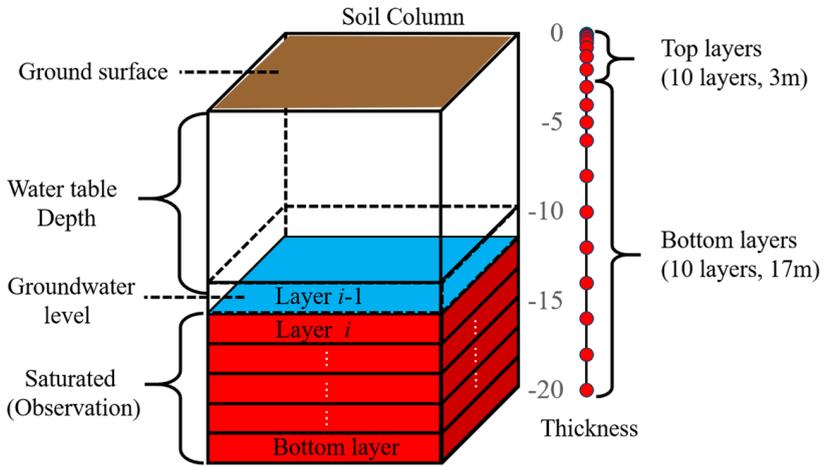


Figure 3.4 Illustration of the link between groundwater level observation and data to be assimilated (revised from Zhang et al. (2018)). The blue colour indicates the groundwater level at layer  $i-1$ . The red layers (from layer  $i$  to the bottom layer) are saturated and are incorporated as groundwater observations, and converted to pressure heads assuming hydrostatic conditions.

In our study, in order to ensure stability and avoid the occurrence of anomalous pressure values in the unsaturated zone related to updating pressure in the DA step, a weakly coupled approach was followed, which implies that only pressure in saturated layers is updated during

assimilation. Hung et al. (2022) found that the weakly coupled approach outperformed the fully coupled approach for assimilating WTD measurements in TSMP. In the OL run, the vertical division between the unsaturated and saturated zones will differ among ensemble members. But as stated in Zhang et al. (2018), every grid cell should be updated consistently in DA, so the definition of the state vector should be the same for all ensemble members. The saturated and unsaturated zones are defined by the deepest WTD among the ensemble members, following Zhang et al. (2018). In the analysis step, only the pressure head values for the defined saturated zone will be directly updated via LEnKF.

### **3.3 Experimental setup**

#### **3.3.1 Ensemble generation and simulations**

The simulation domain was discretized with a horizontal spatial resolution of 500 m. The study domain has a vertical extension of 20 m, which is discretized into 20 soil and aquifer layers with variable thicknesses. The thicknesses of the 10 uppermost layers increase exponentially with depth and extend to a total of 3 m. The deeper ten subsurface layers have thicknesses of 1 m (three layers) or 2 m (seven layers).

It is expected that the assimilation performance improves with increasing ensemble size (number of realizations), as found, for example, in studies with groundwater flow models (Chen and Zhang, 2006). An increasing ensemble size also implies higher computational costs. Hendricks Franssen and Kinzelbach (2008) indicated that 100 realizations should be sufficient for real-time groundwater flow modeling problems with state updating only. For combining state and parameter estimation, the ensemble size needs to be larger. As a compromise between accuracy and available compute time and data storage, we established an ensemble with 128 members for WTD assimilation in this work.

Meteorological forcings, hydraulic conductivities, and porosity were perturbed to generate the ensemble. Four atmospheric variables were perturbed: precipitation, incoming shortwave radiation, incoming longwave radiation, and air temperature. The meteorological forcings were perturbed without spatial correlation, while temporal correlations were induced by a first-order autoregressive model (Reichle et al., 2010; Han et al., 2015). Since the four meteorological variables are correlated, random values were drawn from a multivariate normal distribution. The statistics of the perturbed atmospheric variables are summarized in Table 3.2. The temporal correlations and standard deviations of the perturbations were chosen based on previous catchment-scale and regional-scale DA experiments (Reichle et al., 2010; Han et al., 2013; Han

et al., 2015; Roland Baatz et al., 2017).

Table 3.2 The listed cross-correlations give the cross-correlations between the perturbations for the different atmospheric variables, following the order as indicated in the left column of the table.

Variables	Noise	Standard deviation	Time correlation scale	Cross correlation
Precipitation	Multiplicative	0.3	24 h	[ 1.0, -0.8, 0.5, 0.0,
Shortwave radiation	Multiplicative	0.2	24 h	-0.8, 1.0, -0.5, 0.4,
Longwave radiation	Additive	20 W m <sup>-2</sup>	24 h	0.5, -0.5, 1.0, 0.4,
Air temperature	Additive	1 K	24 h	0.0, 0.4, 0.4, 1.0]

Precipitation and shortwave radiation were multiplied by lognormally distributed noise (Han et al., 2013). A direct back transformation would induce a bias (resulting typically in a larger mean precipitation and larger mean incoming shortwave radiation), and therefore a correction is applied (Yamamoto, 2007):

$$Z_{j,i}^* = K \times Z_i \exp(x_{j,i}) = \frac{Z_i \exp(x_{j,i})}{\frac{\sum_{m=1}^{365} \sum_{n=1}^N Z_n \exp(x_{n,m})}{(N \times 365)}} , \quad (3.12)$$

Where  $Z_{j,i}^*$  is the bias corrected perturbed variable of ensemble member  $j$  at day  $i$ ,  $Z_i$  is the original variable at day  $i$ ,  $K$  is the corrective factor, and  $x_{j,i}$  is the random perturbation of ensemble member  $j$  at day  $i$  from the multivariate normal distribution.  $N$  is the number of ensemble members (128 in this study).

Not only atmospheric forcings, but also hydraulic conductivity was perturbed in this study as the uncertainty of this parameter is in general large with an important effect on the groundwater flow prediction. We use different  $K_s$  data for the upper soil layers and lower aquifer layers. Hence, the  $K_s$  values of the soil and aquifer layers were perturbed separately (Table 3.3). The  $K_s$  values for soil layers were perturbed by perturbing the sand and clay contents first, and then applying the Rosetta pedotransfer functions (Schaap et al., 2001; Zhang and Schaap, 2017) to obtain the perturbed  $K_s$ . Sand and clay content were perturbed by calculating a field of spatially correlated perturbation values with geostatistical simulation and mean zero. A spherical variogram model was used, with nugget 0, sill of 50%<sup>2</sup>, and range 25km. In order to

avoid unphysical values for the soil textures, the sum of the sand and clay content were constrained between 0% and 100%. The  $K_s$  of the bottom aquifer layers were perturbed for each hydrogeological unit by taking a value from a univariate uniform distribution with values between -0.5 and 0.5 and adding this to the mean  $K_s$  of the unit ( $\log_{10}(\text{m/s})$ ). The porosity for the upper soil layers was determined according to Eq. 3.3 and on the basis of the perturbed sand contents, while for the bottom aquifer layers, the constant value of 0.15 was used, without perturbation.

Table 3.3 Perturbation of saturated hydraulic conductivities for different subsurface layers.

Layers	Texture classes	Perturbation	Generation
Soil layers	Sand content	Simple kriging (sill=50% <sup>2</sup> , range=25 km, and nugget=0)	Rosettav3 (Zhang and Schaap, 2017)
	Clay content		
	Silt content	100- sand(%)-clay(%)	
Aquifer layers	Hydraulic conductivity (in $\log_{10}(\text{m/s})$ unit) was perturbed by uniform distribution (-0.5, 0.5)		

It is known from previous studies that the spin-up for the model TSMP significantly influences the simulated WTD. The 100 year spin-up for 128 ensemble members departed from a WTD of 0 m, and was forced by 30-year average recharge values (derived from gridded data of precipitation and actual evapotranspiration provided by the German Weather Service) as an upper boundary condition. Next, an exit spin-up was done by running CLM-ParFlow for additional 10 years, using meteorological forcings from the year 2017 for all ten years. The conditions at the end of the spin-up were used to initialize the DA experiments for the year 2018. The model time step is set to hourly.

### 3.3.2 Selection criteria for assimilated sites

There is a spatial mismatch between the point-scale groundwater measurements and the TSMP grid cell size of 500 m. In order to compare the measured WTD data with the model simulated values, each groundwater observation site was assigned to the nearest grid cell center. It is therefore common to have several measurement sites located in the same grid cell. We kept the groundwater measurement site which had the median value of all measurement sites in the grid cell for the year 2018, while the rest of the measurement sites were excluded from assimilation.

In addition, due to the relatively coarse model resolution, some measurement sites were located in model river grid cells. If all soil layers for each ensemble member in a grid cell were saturated for the complete year, the grid cell was considered to be a river grid cell. The river grid cells were eliminated from the analysis as groundwater measurements are not informative for the pressure values in river grid cells. Grid cells directly next to rivers were also excluded from the DA procedure, as these grid cells were also saturated most of the time and sometimes became part of the river. In this study, within the localization radius, we assimilated only observations from one site, with measurement values that are median values considering all sites in the localization radius.

In our study, the impact of the localization radius on the assimilation results was investigated, and three different localization radii were considered for assimilating groundwater measurements: 10 km, 5 km and 2.5 km. According to the assimilation site selection criteria, three different localization radii resulted in 10 groundwater sites being selected for each of the DA experiments. Also, we used the groundwater data from unassimilated locations to investigate whether the localized assimilation could also improve the groundwater estimation at locations without assimilating data.

### 3.3.3 Evaluation of model performance

The root mean square error (RMSE) and bias (BIAS) were calculated to evaluate the performance of the WTD assimilation. The RMSE of WTD at each time step is calculated as:

$$RMSE_i = \sqrt{\frac{\sum_{n=1}^M (WTD_{n,i}^{sim} - WTD_{n,i}^{obs})^2}{M}}, \quad (3.13)$$

where  $M$  is the total number of observation sites,  $WTD_{n,i}^{sim}$  is the ensemble average groundwater table depth of the grid cell where the observation site is located at the time step  $i$  (either from an OL run or a DA run), and  $WTD_{n,i}^{obs}$  is the observed WTD at the  $n^{th}$  site and time step  $i$ .

The bias is also specified to quantify systematic differences between simulated and measured WTD:

$$BIAS_{n,i} = WTD_{n,i}^{sim} - WTD_{n,i}^{obs}, \quad (3.14)$$

In addition, simulation results were also compared with measured SM content by CRNS. We follow the approach by Schrön et al. (2017), where weighted SM content from the simulations was compared with CRNS measurements. The indicators, including BIAS, correlation coefficient (R), RMSE and unbiased root mean square difference (ubRMSD), are used to evaluate simulated SM compared with the CRNS measurements. For each CRNS site, the above indicators were calculated individually and aggregated over time.

$$BIAS = \frac{\sum_{i=1}^T (\theta_i^{sim} - \theta_i^{obs})}{T}, \quad (3.15)$$

$$R = \frac{\sum_{i=1}^T (\theta_i^{obs} - \overline{\theta^{obs}}) (\theta_i^{sim} - \overline{\theta^{sim}})}{\sqrt{\sum_{i=1}^T (\theta_i^{obs} - \overline{\theta^{obs}})^2 \sum_{i=1}^T (\theta_i^{sim} - \overline{\theta^{sim}})^2}}, \quad (3.16)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^T (\theta_i^{sim} - \theta_i^{obs})^2}{T}}, \quad (3.17)$$

$$ubRMSD = \sqrt{\frac{\sum_{i=1}^T ((\theta_i^{obs} - \overline{\theta^{obs}}) - (\theta_i^{sim} - \overline{\theta^{sim}}))^2}{T-1}}, \quad (3.18)$$

where  $T$  is the total number of time steps,  $\theta_i^{sim}$  is the simulated (either from an OL run or a DA run) ensemble average SM content at the time step  $i$ , and  $\theta_i^{obs}$  is the observed SM by CRNS at the time step  $i$ . The overbars in equations 3.16 and 3.18 indicate the temporal mean over the entire time period.

### 3.4 Results and discussion

#### 3.4.1 Water table depth

##### 3.4.1.1 Spatial autocorrelation analysis

For localized assimilation, the selection of the appropriate radius of localization is important. The localization radius should not be too small in order not to neglect positive correlations, and it should not be too long so that areas with spurious correlations are excluded. Therefore, we calculated the spatial autocorrelation of groundwater level measurements and simulated groundwater tables in the OL run (see Fig 3.5) to identify the appropriate radius. The spatial autocorrelations for different distance classes (0-0.5 km, 0.5-2.5 km, 2.5-5 km, 5-10 km, 10-20 km, 20-30 km, 30-40 km, 40-50 km, 50-60 km and 60-70 km) were determined. The spatial correlation functions for the measured and WTD are quite close, implying that the model represents quite well the spatial correlation of the real groundwater levels in the Rur catchment. The largest differences are found for shorter distances, where the model autocorrelation is higher than the measured autocorrelation.

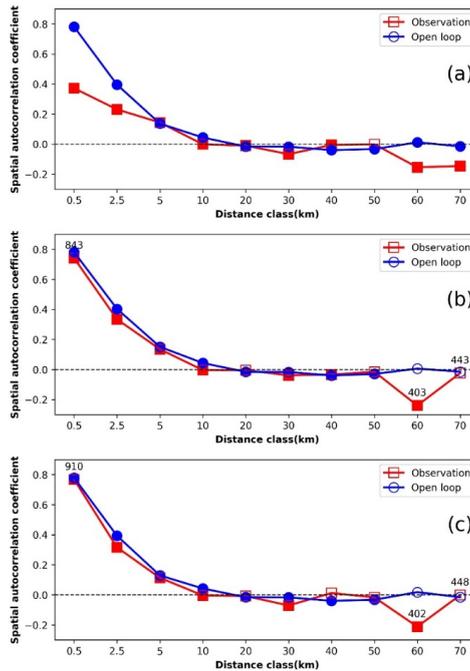


Figure 3.5 Spatial autocorrelation functions of measured and simulated (open loop)

groundwater table depth for the year 2018 (a), the 26<sup>th</sup> of February 2018 (b), and the 24<sup>th</sup> of September 2018 (c). Filled squares or circles indicate autocorrelation coefficients are significantly different from zero ( $p < 0.05$ ). When the number of comparison pairs was smaller than 1000, the number of comparison pairs is indicated next to the marker.

### 3.4.1.2 Different localization radius assimilation strategies

Based on the spatial autocorrelation analysis, the localization radius could be up to 10 km. We tested 10 km, 5 km and 2.5 km as localization radii, including 10 groundwater measurement sites for the assimilation.

For all scenarios, the RMSEs of WTD after 1 year of assimilation were lower than those of OL at the measurement locations (see Fig 3.6). The histogram of WTD errors at measurement locations also illustrates the improved WTD characterization after DA, compared with the OL. It can also be observed that the OL results on average in WTD values were larger than the measured ones, implying that the simulated WTDs were deeper than the measured ones. DA resulted in a reduction of the bias, and the peak of the histogram is closer to zero than OL. Thus, in all cases, LEnKF strongly reduced the bias and RMSE of WTD, compared to the scenario without assimilation of groundwater data, and the simulation improvement is best when using 10 km or 5 km localization radius, and slightly worse for 2.5 km radius.

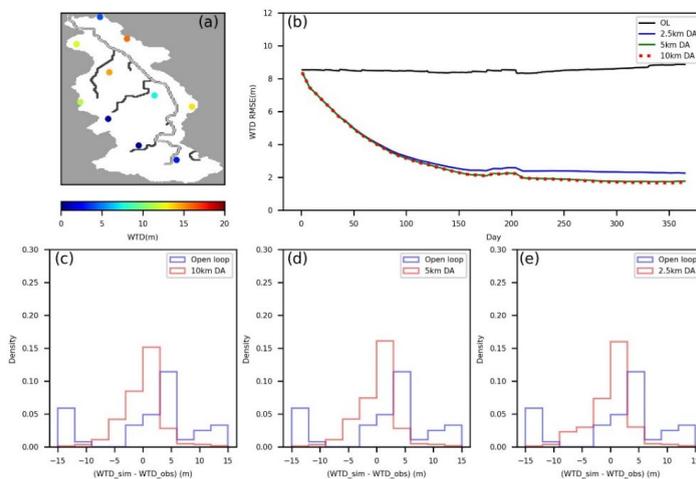


Figure 3.6 The locations of the assimilated groundwater sites (dots) in the Rur catchment

together with the average groundwater table depth (a); time series of root mean square error (RMSE) of water table depth at measurement locations for the year 2018 for the open loop and data assimilation runs (b); the histogram of the water table depth errors at the measurement locations for the year 2018 from the open loop and data assimilation runs (10 km (c), 5 km (d), 2.5 km (e)).

Since the groundwater assimilation results for the three radii were similar at the assimilation sites, and the best results were obtained for the 10 km radius, only the simulated WTD from the 10 km localization radius DA run is shown in Fig 3.7, and compared with the WTD from the OL run and the measurements. The changes in groundwater levels during assimilation show that once assimilation starts, the WTD gets closer to the measurements.

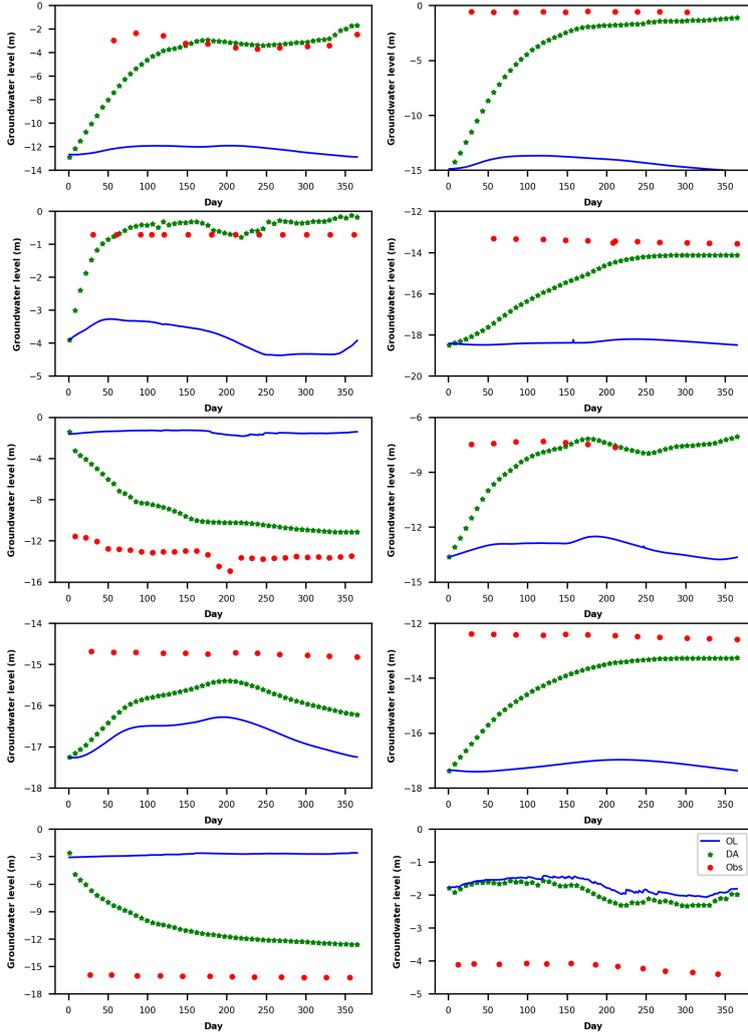


Figure 3.7 Water table depth time series for 10 assimilation sites: observations (Obs, red), ensemble mean of open loop (OL, blue) and ensemble mean of data assimilation run with 10 km localization radius (DA, green) for the year 2018.

The impact of the DA is similar for the different localization radii, with the same regions affected by increases or decreases in WTD (see in Fig 3.8). The main difference is that for a larger localization radius, the area updated by assimilation is larger, with a stronger reduction of ensemble standard deviations. However, for some areas, the ensemble standard deviation

was larger for the DA run than for the OL run. This occurred when the measurements deviated strongly from the ensemble of OL runs. With DA, some ensemble members became closer to the observations, but others were not, resulting in an increase in the ensemble dispersion.

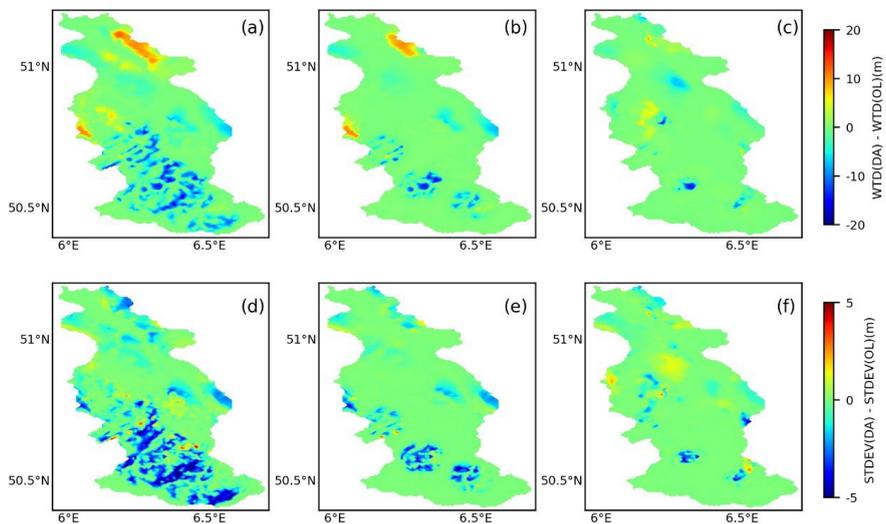


Figure 3.8 Difference in average water table depth between data assimilation and open loop runs (data assimilation - open loop) for different localization radii (10 km (a), 5 km (b), 2.5 km (c)) on the 31<sup>st</sup> of December 2018 for the Rur catchment; and difference in standard deviation for data assimilation and open loop runs for different localization radii (10 km (d), 5 km (e), 2.5 km (f)) on the 31<sup>st</sup> of December 2018 for the Rur catchment.

### 3.4.1.3 Data assimilation verification

To explore the impact of the assimilation of WTD measurements, we also evaluated the WTD characterization at verification locations (555 sites in total) which were not included in the assimilation, for the three different localization radii. We only show results for verification locations within the localization radius and only if enough measurement data were available for assimilation at a given time step (see Fig 3.9). Table 3.4 shows the RMSE for the OL and DA simulations, averaged for the period of one year. At verification locations, the RMSE of the WTD also decreased, especially closer to the assimilation location, with verification locations separated less than 2.5 km from assimilation locations. DA could improve the groundwater

simulation around measurement locations, which is consistent with the results by Hung et al. (2022).

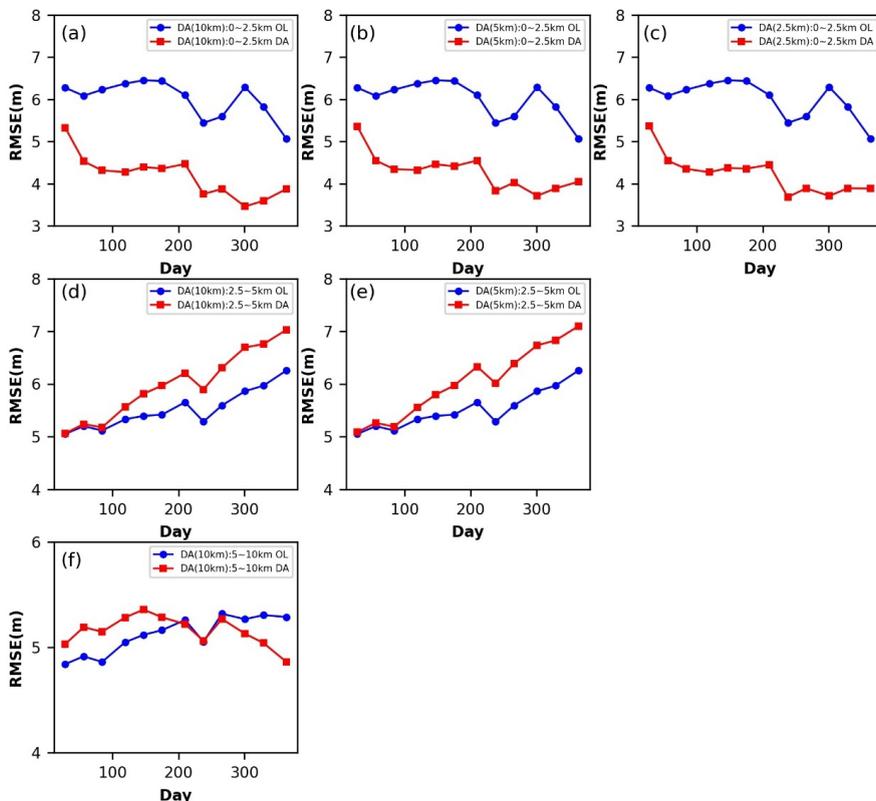


Figure 3.9 Time series of RMSE of groundwater table depth for the open loop (OL) and data assimilation (DA) runs (10 km (a, d, f), 5 km (b,e), 2.5 km (c) localization radius) at verification locations which were 0~2.5 km, 2.5~5 km and 5~10 km away from assimilated observations.

Table 3.4 The time averaged RMSE of the water table depth at the verification locations for the open loop (OL) and data assimilation (DA) runs (10 sites, 10 km, 5 km, 2.5 km localization radius).

Experiment	Horizontal distance 0-2.5 km		Horizontal distance 2.5-5 km		Horizontal distance 5-10 km	
	OL	DA	OL	DA	OL	DA

DA (10 km)	6.03	4.22	5.52	6.01	5.12	5.16
DA (5 km)	6.03	4.31	5.52	6.05	/	/
DA (2.5 km)	6.03	4.26	/	/	/	/

### 3.4.2 Soil moisture

The impact of WTD assimilation was also evaluated with in situ SM measurements from CRNS networks. Simulated SM in OL and DA runs (for 10 km and 5 km localization radius) was compared with CRNS measurements. The OL results indicate that simulated SM contents have similar temporal variations as measured SM contents (see Fig 3.10). Assimilation of WTD measurements did not result in an obvious improvement for SM estimation (see Table 3.5). Hung et al. (2022) also found that assimilating groundwater table data only slightly improved SM characterization with RMSE reductions between 1% and 6%, and the improvements were limited to a relatively small area around observation locations. This is related to the fact that SM is only indirectly updated by the propagation of the pressure below the groundwater table. Therefore, when the groundwater table is deep, the impact of WTD assimilation on the upper SM is small.

Table 3.5 Comparison metrics for the soil moisture from CRNS compared to open loop (OL) and data assimilation runs (DA10 and DA5 are for 10 km and 5 km localization radius, respectively) for the year 2018.

Site	BIAS (cm <sup>3</sup> /cm <sup>3</sup> )			R			RMSE (cm <sup>3</sup> /cm <sup>3</sup> )			ubRMSD (cm <sup>3</sup> /cm <sup>3</sup> )		
	OL	DA10	DA5	OL	DA10	DA5	OL	DA10	DA5	OL	DA10	DA5
Merzenhausen	-0.04	-0.04	-0.04	0.74	0.74	0.74	0.06	0.06	0.06	0.05	0.05	0.05
Rollesbroich1	0.05	0.07	0.07	0.89	0.88	0.88	0.06	0.08	0.08	0.04	0.04	0.04
Rollesbroich2	0.05	0.07	0.07	0.87	0.85	0.85	0.06	0.08	0.08	0.05	0.05	0.05
Gevenich	-0.02	-0.02	-0.02	0.76	0.76	0.76	0.06	0.06	0.06	0.06	0.06	0.06
Ruraue	-0.04	-0.04	-0.04	0.76	0.75	0.75	0.07	0.07	0.07	0.06	0.06	0.06
Wildenrath	-0.04	-0.04	-0.04	0.76	0.76	0.76	0.05	0.05	0.05	0.04	0.04	0.04
Wuestebach	-0.13	-0.12	-0.13	0.53	0.54	0.53	0.13	0.13	0.13	0.06	0.06	0.06
Heinsberg	-0.03	-0.03	-0.03	0.79	0.79	0.79	0.06	0.06	0.06	0.05	0.05	0.05
Kall	0.02	-0.00	-0.00	0.82	0.77	0.77	0.05	0.05	0.05	0.05	0.06	0.06
Selhausen	-0.06	-0.06	-0.06	0.70	0.69	0.70	0.09	0.08	0.09	0.07	0.07	0.07
Schoeneseiffen	-0.08	-0.04	-0.08	0.81	0.72	0.81	0.09	0.07	0.09	0.05	0.06	0.05
Kleinau	-0.03	0.00	-0.03	0.81	0.71	0.81	0.06	0.06	0.06	0.05	0.07	0.05
Aachen	-0.14	-0.14	-0.14	0.80	0.80	0.80	0.14	0.14	0.14	0.05	0.05	0.05
Average	-0.04	-0.03	-0.04	0.77	0.75	0.77	0.08	0.08	0.08	0.05	0.06	0.05

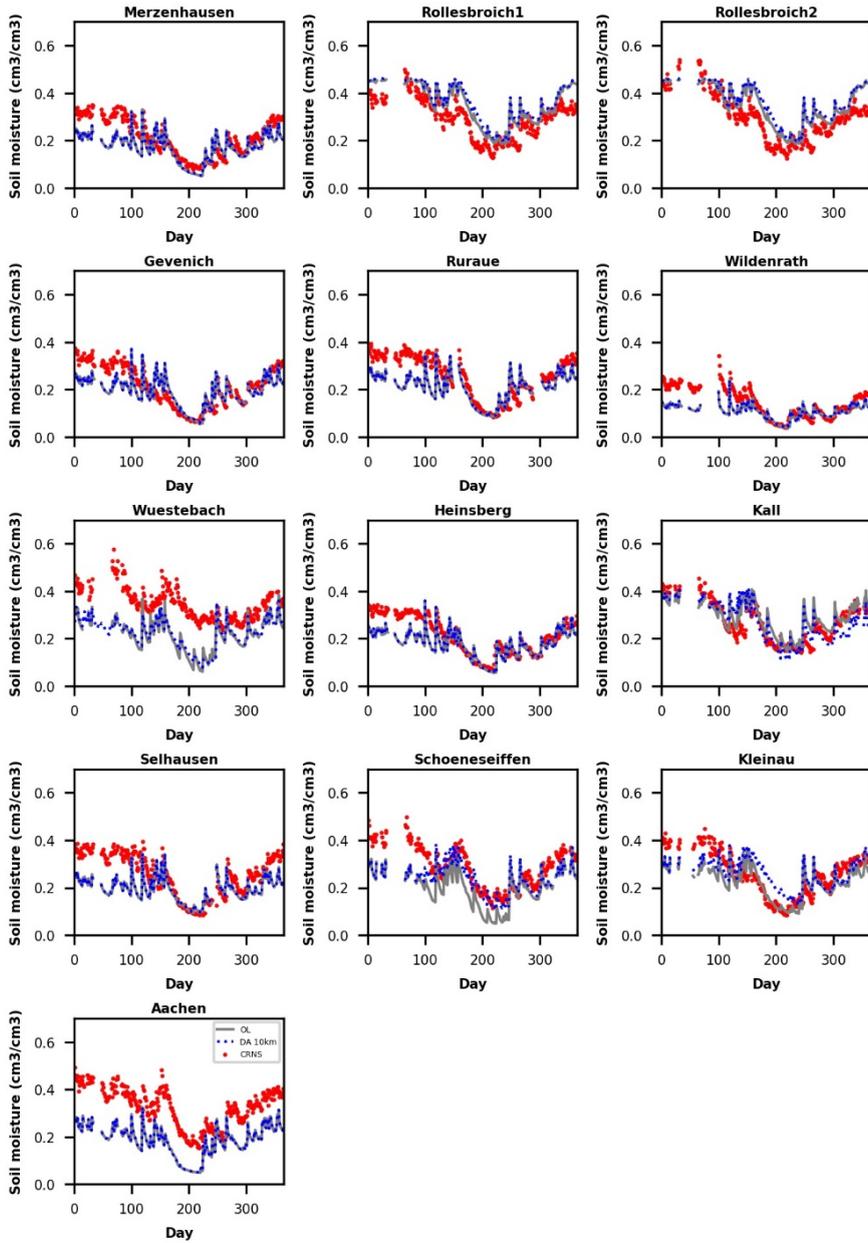


Figure 3.10 Soil moisture time series from cosmic-ray neutron sensors (CRNS) (red), ensemble mean of open loop (OL, grey), and ensemble mean of data assimilation with 10 km assimilation radius (DA 10 km, blue) for the year 2018.

### 3.4.3 Discussion

In all DA experiments, the estimation of the WTD improved, and also close to observation sites an improved groundwater characterization was found. This shows that for real-world cases, the localized EnKF could merge the integrated model TSMP with data to more accurately simulate the groundwater table.

There are some caveats regarding the use of in situ groundwater observations to do assimilation and validate model estimates. Since the spatial representativeness of model and measurements are different, it is non-trivial to assimilate the in situ groundwater measurements into the integrated model and to evaluate the coarse resolution model results against in situ measurements. In our study, the model has a grid resolution of 500 m, while the groundwater measurements are obtained from points. Many observation sites were located in the same grid cell and were not included in the assimilation in this work. In future work, these measurements should be assimilated by modifying the measurement operator. However, this will not resolve all issues regarding scale mismatches. As the coarser model resolution flattens the topography, and therefore the gradients for surface and subsurface water flow, a systematic bias in the simulated groundwater table can be expected and is also observed in this study. In theory, for data assimilation, we should not have systematic differences between simulated and measured values, and prior bias correction would be a strategy to consider. In practice, we normally have to deal with systematic biases in data assimilation and if the ensemble spread is large enough, the model states can still be corrected towards the measurements. Removing the systematic bias in simulated groundwater levels with TSMP is not trivial as it depends on the model resolution. An extensive effort is needed to remove the systematic bias, which is a research question in itself and beyond the scope of this study. We argue that in the future better results can be obtained if a higher model resolution of 100 m instead of 500 m is used so that groundwater bodies related to narrow valleys can be better represented.

In addition, the model TSMP in this study only considers a vertical depth up to 20 m, and only one upper unconfined aquifer is better modelled. However, the real situation is much more complicated, as typically multiple unconfined and confined aquifer layers exist. As our model only models the 20 m subsurface, measurements relating to deeper aquifers were also excluded. In future work, an extension of the vertical depth could provide more realistic simulations, but for this, it would be important to have more detailed 3D geological information.

The spatial autocorrelation analysis indicates that groundwater levels were correlated for

separation distances up to 10 km. However, groundwater level characterization was only improved in a smaller area for locations separated by less than 2.5 km from measurement sites. Hung et al. (2022) used a dense observation network in a synthetic experiment that closely mimicked the Neckar catchment of southwestern Germany, and proved that GWL assimilation could improve GWL estimation between the measurement locations. They found that the improvement of the GWL simulation decreased with increasing horizontal distance, and improvements in GWL simulations could extend to 8 km away from the observations for a localization radius of 12 km. Our results illustrate that for a real-world application, the improvement is more limited, which will be related to model structural errors like inadequate grid resolution and missing information on pumping activities.

Theoretically, only grid cells within the localization radius can be updated in the analysis step (Houtekamer and Mitchell, 2001). However, as the assimilation proceeds over time, updates around measurement locations can laterally propagate through the working of the physical equations, and this effect could be particularly strong in the saturated zone given the importance of lateral flow in the saturated zone. In the assimilation experiments with 10 km and 5 km localization radius, there were no obvious improvements in the characterizations of SM content by TSMP. Though the groundwater bias was corrected after assimilation, SM does not change significantly with the change in deep groundwater tables. Also, Hung et al. (2022) discovered topography variations and lateral groundwater flow greatly influence groundwater levels, making soil moisture data probably less informative for groundwater levels, which also supports the findings of this study. Hung et al. (2022) found a slight improvement for SM characterization related to GWL assimilation, which was not found in this study. The worse performance in this real-world study might be related to model structural errors, as Hung et al. (2022) simulated a catchment of similar complexity, but in a synthetic version that mimicked that catchment. A further reason might be the limited number of groundwater assimilation and SM validation sites used in this study. We assimilated only GWL data, but SM was not measured at the same locations, and SM verification locations were separated from the groundwater monitoring sites.

Improved results can be expected for more ensemble members and/or a higher spatial resolution, which was not feasible in this work, as only one single DA experiment with 128 ensemble members required 73,728 core hours (the spin-up not included) and 1.75 TB of computer storage for one year of simulation at a daily timescale.

Furthermore, although updating saturated hydraulic conductivities in Hung et al. (2022) only marginally improved the simulation of subsurface states, compared with only state

updating, we will explore in future work the role of this important parameter in groundwater modelling. This should be evaluated for simulations at higher spatial resolution and larger ensemble sizes.

### 3.5 Conclusions

The localized Ensemble Kalman Filter was used to assimilate GWL measurements into the integrated terrestrial system model TSMP for the ~2000 km<sup>2</sup> Rur catchment. This is the first application of the assimilation of observed WTD data into the integrated land surface-subsurface model TSMP for a real-world case. Earlier work focused on a synthetic case, mimicking the Neckar catchment in southwest Germany. For the Rur catchment, 128 ensemble members were generated by perturbing four atmospheric forcing variables, saturated hydraulic conductivities and porosity. The perturbed ensemble was used as input in the TSMP-PDAF data assimilation framework and assimilation experiments were done for different localization radii (10 km, 5 km and 2.5 km). The performance of WTD assimilation was assessed by comparing results from OL and DA experiments, and using groundwater observations and SM measurements from cosmic ray neutron sensors as verification data. The main findings are:

1. The WTD simulated by the integrated model TSMP could be improved by localized EnKF, with more than 75% RMSE reduction at the assimilated locations for 3 different localization radii. The positive impact of assimilation is limited to the vicinity of the assimilated locations. The localized WTD assimilation is greatly affected by the unevenly distributed groundwater observations.

2. Simulated SM generally reproduced the observed temporal fluctuations of soil water content, but SM characterization was not improved after WTD assimilation. This can be related to the fact that only the saturated zone was directly updated via assimilation (and the unsaturated zone only indirectly), and the presence of model structural errors like a relatively coarse grid resolution of 500 m and missing information on groundwater pumping activities, for example.

3. Systematic differences between simulated and measured WTD might be related to the too coarse model resolution and model structural errors. Future work should focus on DA with integrated land surface-subsurface models at a higher spatial resolution and with more ensemble members, which would allow parameter estimation. In addition, the measurement operator needs to be considered for multiple GWL observations in a grid cell.

## **Chapter 4: Can a sparse network of cosmic ray neutron sensors improve soil moisture and evapotranspiration estimation at the larger catchment scale?**

\*adapted from: Li, F., Bogena, H. R., Bayat, B., Kurtz, W., and Hendricks Franssen, H.-J.: Can a sparse network of cosmic ray neutron sensors improve soil moisture and evapotranspiration estimation at the larger catchment scale? *Water Resources Research*, 10.1029/2023WR035056, 2023.

### **4.1 Introduction**

Soil moisture has a significant influence on water and energy fluxes between the subsurface, land surface, and the atmosphere (Chen and Hu, 2004). Accurate information on the spatio-temporal variability of SM is crucial to better understand the role of SM in terrestrial systems (Vereecken et al., 2022). SM can be estimated at various scales, indirectly with RS observations and directly by in situ measurements using electromagnetic techniques being the most commonly used methods. The resolution of RS-derived products is often coarse, spatially or temporally, and the data are subjected to various errors, especially in areas with dense vegetation coverage (Bauer-Marschallinger et al., 2019; Kim et al., 2020). On the other hand, in situ measurements are point-scale measurements and, therefore, do not provide area coverage (Hailong He et al., 2021).

Over the last decade, the CRNS has been introduced as an alternative method, providing real-time SM estimation and bridging the gap between in situ measurements and RS products. The CRNS footprint covers up to 18 ha with a measurement depth of up to 80 cm (Zreda et al., 2008; Bogena et al., 2015; Köhli et al., 2015). The continuous development of CRNS technology has enabled SM monitoring under a variety of climatic conditions, which has promoted its application in hydrological modeling (Roland Baatz et al., 2017), satellite product validation (Zhao et al., 2021), extreme weather event (drought and flood risk) assessment (Bogena et al., 2022), ecohydrological (e.g., snow, precipitation, and vegetation) monitoring (Bogena et al., 2020), and agricultural management (Li et al., 2019). The advantages of CRNS have made it increasingly attractive, and large-scale CRNS networks have been established in Europe, the USA, Australia, and India for large-scale SM monitoring with high temporal resolution, which can also benefit the multifaceted hydrological applications mentioned above.

However, it is neither economical nor feasible to measure area-wide SM over large areas

using CRNS. Therefore, it is important to establish a scientific and economical observation network that can cover all major land use types and climatic zones in the study area while ensuring that the sensor coverage is representative of SM patterns across the region. Then hydrological or land surface models can be combined with observations to effectively monitor SM at larger scales, including unobserved locations, and validate the model's performance.

Nevertheless, modeling accuracy is often limited by uncertainties arising mainly from model forcings, parameters, and initial conditions (Freeze, 1975; Roland Baatz et al., 2017). DA is a technique to combine different information sources to update or correct the model predictions and improve the simulations (De Lannoy et al., 2014). The Ensemble Kalman filter (Evensen, 1994; 2003), a sequential filtering algorithm, is the most widely used DA technique and has been proven effective for nonlinear systems and high-dimensional problems (Camporese et al., 2009b; Schöniger et al., 2012). The process of DA can be summarized in two steps: a forecast step and an analysis step. For the forecast step, the state estimation is only based on past data, while for the analysis step, the probability density of the state is propagated forward, considering the information from current measurements (McLaughlin, 2002).

Shuttleworth et al. (2013) developed the forward COsmic-ray Soil Moisture Interaction Code (COSMIC) model to enable rapid conversion of neutron counts to SM values. Since then, COSMIC has been used as an observation operator in several studies for assimilating neutron counts into land surface models to improve SM prediction (Shuttleworth et al., 2013; Rosolem et al., 2014; Han et al., 2015; Han et al., 2016; Roland Baatz et al., 2017; Patil et al., 2021). For example, Roland Baatz et al. (2017) found that catchment-scale SM prediction can be improved by assimilating SM from a CRNS network and that joint estimation of state and parameters performs better than state estimation alone. To date, however, such assimilation experiments with CRNS data have been conducted only with land surface models that do not adequately describe lateral water movement and groundwater-land surface interactions (Kollet and Maxwell, 2008). Zhao et al. (2021) compared the CRNS data with simulated SM using both the land surface model Community Land Model (CLM, version 3.5) and a coupled land surface-subsurface model (CLM-ParFlow). They found that the coupled model simulations showed less bias and reproduced better SM dynamics than the CLM stand-alone, demonstrating the importance of considering lateral subsurface flow in subsurface hydrological simulations. Therefore, there is still strong interest in applying DA with coupled land surface-subsurface models to exploit the full potential of CRNS data.

In this work, the integrated Terrestrial System Modelling Platform (Shrestha et al., 2014; Kurtz et al., 2016) is used, which is a coupled atmosphere-land surface-subsurface model with

the Parallel Data Assimilation Framework (PDAF). The integrated model TSMP has been utilized in a number of studies (Shrestha et al., 2015; Keune et al., 2016; Shrestha et al., 2018b; Furusho-Percot et al., 2019). Previous studies investigating the assimilation of SM measurements with TSMP-PDAF focused on synthetic experiments, small catchments, or greatly simplified the representation of spatial heterogeneity in the real catchment (Zhang et al., 2018; Gebler et al., 2019; Hung et al., 2022). Recently, Hung et al. (2022) adopted the conventional vertical weighting calculation of CRNS data from Franz et al. (2012) for the assimilation of SM data in a virtual reality experiment with TSMP and discovered that DA improved the vertical SM profile characterization and SM estimation for the surrounding grid cells.

In this study, SM from a distributed network of 12 CRNS in the Rur catchment was assimilated into TSMP to investigate whether the sensor density is sufficient to represent the SM for the whole catchment. So far, to the best of our knowledge, this is the first study to assimilate SM from such a high-density CRNS monitoring network into the integrated model. Our work, for the first time, explores the information content of CRNS observations through fully coupled TSMP and DA techniques. It unlocks the full potential of CRNS to characterize the SM and ET across a relatively large catchment. The main objectives of this study are to investigate: (i) how effective a CRNS network can be in improving SM characterization with fully integrated terrestrial models such as TSMP at the catchment scale; (ii) whether the assimilation of CRNS soil moisture data can result in better prediction of ET and discharge; and (iii) how DA performance can vary between years with different hydrological conditions (wet versus dry).

## **4.2 Materials and methods**

### **4.2.1 Study area**

The Rur catchment covers an area of 2354 km<sup>2</sup> and is located in western Germany, including small portions of Belgium and the Netherlands. Fig 4.1 shows the Digital Elevation Model (DEM) for the Rur catchment obtained from SRTM 90m Version 4 (Jarvis et al., 2008). The Rur River flows from the Eifel low-mountain range in the south with a maximum altitude of 690 m a.s.l. to the northern lowlands with a minimum altitude of 15 m a.s.l. The land use types in the Rur catchment are arable agriculture in the north (mainly maize and wheat) and grassland, coniferous and deciduous forests in the south (Waldhoff and Lussem, 2015; Roland Baatz et al., 2017). From the northern to the southern part of the catchment, long-term average

annual precipitation ranges from 650 to 1300 mm, the mean annual air temperature decreases from 10 to 7°C, and the mean annual potential evapotranspiration ranges from 450 to 850mm (Montzka et al., 2008; Bogena et al., 2018). The mean river discharge in the upper catchment (controlled by the in situ station Erkensruhr-Einruhr (see Fig 4.1)) was about 0.26 m<sup>3</sup>/s from 2013 to 2022.

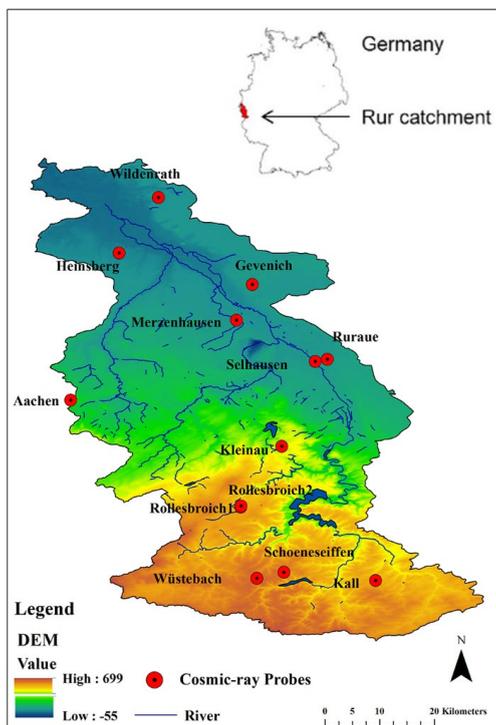


Figure 4.1 Map of the Rur catchment with the altitude above sea level and the locations of the cosmic-ray neutron sensors, eddy covariance stations and discharge station. The Rur catchment is situated in western Germany.

#### 4.2.2 Terrestrial System Modeling Platform (TSMP)

TSMP is a modular coupled biogeophysical terrestrial systems model consisting of atmospheric, surface, and subsurface models (Shrestha et al., 2014). The three component models that make up TSMP are the numerical weather prediction model COSMO (Consortium

for Small Scale Modelling) (Baldauf et al., 2011), the Community Land Model CLM 3.5 (Oleson et al., 2004; Oleson et al., 2008) from the National Center for Atmospheric Research, and the 3D variably saturated groundwater flow model ParFlow (Kollet and Maxwell, 2006) for the subsurface. These three models are two-way coupled by the Ocean Atmosphere Sea Ice Soil coupling Model Coupling Toolkit (OASIS-MCT, version 3) (Valcke, 2013). The OASIS-MCT coupler is included in the model platform and is used for the exchange of variables and fluxes between different sub-models. In this work, only the land surface model CLM 3.5 and the subsurface model ParFlow were used.

The biophysical processes simulated by the land surface model CLM 3.5 include energy and water exchange between the land and atmosphere, snow accumulation and melting, energy and water transport in the soil, and stomatal physiology and photosynthesis (Oleson et al., 2004; Oleson et al., 2008). Spatial land surface heterogeneity is represented by the nested subgrid hierarchy in CLM (Oleson et al., 2008). Each grid cell is divided into different types of land units (glacier, lake, wetland, urban, and vegetated), and each land unit in the grid cell can have a different number of snow/soil columns, and each column can have multiple plant functional types (PFTs) with different plant physiological parameters (Bonan et al., 2002; Oleson et al., 2008), for example, leaf area index (LAI). The input LAI used in this study was taken from previous studies, in which the study area included the Rur domain (Sulis et al., 2015; Sulis et al., 2018). The primary function of the CLM in the TSMP is to calculate evapotranspiration from the ground and vegetation. See Appendix A for details on the computing of evapotranspiration by CLM.

In the coupled model TSMP, the hydrological processes of the CLM are replaced by ParFlow (Ashby and Falgout, 1996; Jones and Woodward, 2001; Kollet and Maxwell, 2006; Maxwell, 2013). ParFlow solves the 3D Richards equation (Richards, 1931) for groundwater flow in the unsaturated and saturated zones and the kinematic wave equation (Lighthill and Whitham, 1955) for overland flow. The coupled partial differential equations for subsurface flow and surface water flow are solved by the Newton-Krylov nonlinear solver (Jones and Woodward, 2001). Moreover, ParFlow was created for parallel computing systems and can effectively solve large-scale problems at high resolution, which has been proven in numerous studies (Jones and Woodward, 2001; Kollet and Maxwell, 2006; 2008; Hung et al., 2022). In addition, ParFlow employs a terrain-following grid transformation with variable vertical discretization, which can resolve groundwater problems with high topographic gradients and reduce the computation time (Maxwell, 2013).

The coupler OASIS-MCT controls the exchange of fluxes and state variables between

CLM and ParFlow, ensuring that the spatial and temporal scales of the fluxes exchanged by the different components remain consistent (Valcke, 2013; Shrestha et al., 2014). In the coupled model TSMP, ParFlow provides the pressure and saturation of the upper ten subsurface layers to CLM, while in turn, CLM provides the upper boundary conditions, i.e., net infiltration or exfiltration, to ParFlow. The net infiltration includes precipitation, interception, total evaporation, and total transpiration (Zhang et al., 2018). More comprehensive information about the implementation of the coupler in TSMP and its operation is presented by Kurtz et al. (2016).

### **4.2.3 Data**

#### **4.2.3.1 Atmospheric forcing**

The high-resolution atmospheric reanalysis dataset COSMO-REA6 (0.055° (6 km)) is used as forcing data for the land surface model CLM (Bollmeyer et al., 2015; Wahl et al., 2017). The reanalysis dataset was developed by the German Meteorological Service (Deutscher Wetterdienst; DWD) based on the numerical weather prediction (NWP) model COSMO (Baldauf et al., 2011), covering the period 1995-2020, and is continuously being extended. Forcing data include precipitation, air temperature, air pressure, wind velocity, specific humidity, incoming shortwave radiation, and incoming longwave radiation. In addition, to maintain consistency with the atmospheric forcings, daily air pressure and air humidity from COSMO-REA6 were used to calculate the weighting of SM based on the revised approach from Schrön et al. (2017). The coupled model CLM-ParFlow of the Rur domain has a horizontal spatial resolution of 500 m for the land surface and a total depth of 100 m for the subsurface.

#### **4.2.3.2 Soil data**

The high-resolution regional soil map BK50 (Geologischer Dienst NRW, 2009) at a scale of 1:50,000 (<https://www.opengeodata.nrw.de/produkte/geologie/boden/BK/ISBK50/>; last access: 7 July 2023) and the European Soil Database (ESDB) (Pano, 2006) were utilized to obtain the soil texture and compute its hydrological parameters. Sand and clay contents (see Fig 4.2) were derived from BK50, and bulk density was obtained from ESDB.

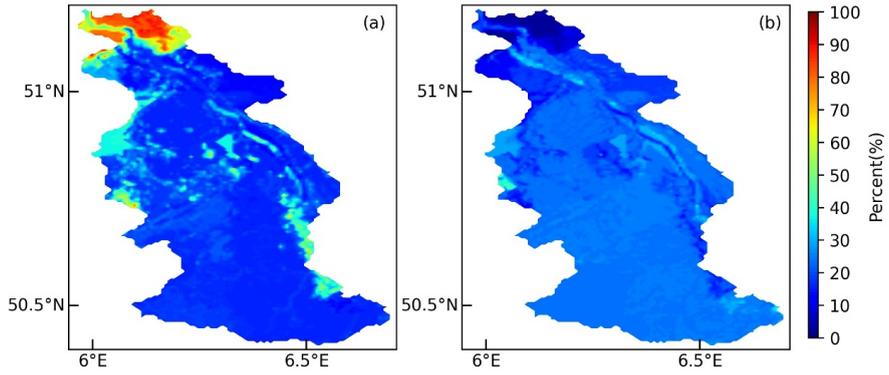


Figure 4.2 Sand (a) and clay content (b) for the Rur catchment derived from the BK50 soil map.

The aquifer permeability for the layers below the soil layers was taken from the 100 m resolution regional hydraulic conductivity ( $K_s$ ) map (Fig 4.3) from the North Rhine-Westphalia Geological Survey database.

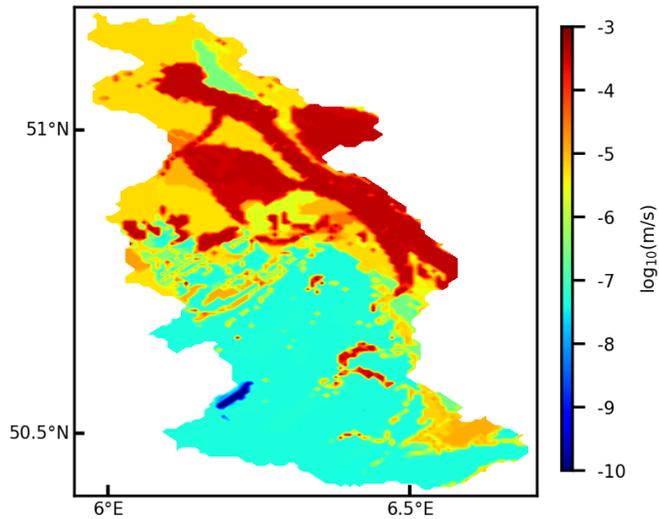


Figure 4.3 Hydraulic conductivity of the bedrock for the Rur catchment.

#### 4.2.3.3 CRNS and flux data

The CRNS detects epithermal neutrons produced by cosmic radiation, which can be used to measure SM because the detected neutron count rate is inversely correlated with the amount of hydrogen in the soil (Zreda et al., 2008). The CRNS soil moisture data were obtained from the “Dataset of COSMOS-Europe: A European network of Cosmic-Ray Neutron Soil Moisture Sensors” (Bogena et al., 2022). The raw neutron count data were measured by the CRNS stations and then transformed into SM values with harmonized correction and processing by Bogena et al. (2022). There are 13 CRNS stations (Andreasen et al., 2017; Roland Baatz et al., 2017; Bogena et al., 2022) relatively evenly distributed over the domain (Fig 4.1), and detailed information is presented in Table 4.1 for all sites. In this work, the SM measured by CRNS is used for DA and as independent verification data for jackknife simulations. The CRNS stations Rollesbroich1 and Rollesbroich2 are regarded as one site since they are too close and located in the same model grid cell, and the average values for the two sites were used in this study. This is, therefore, equivalent to having 12 CRNS sites for final assimilation.

The observed flux data of three eddy covariance (EC) stations (Rollesbroich, Wüstebach, and Selhausen) and discharge data of the Erkensruhr-Einruhr in situ station from TERENO (TERrestrial ENvironmental Observatories; <https://www.tereno.net/>; last access: 7 July 2023) were used to verify the evapotranspiration and discharge simulations. The runoff station in the upstream catchment was chosen for validation because the downstream catchment is highly influenced by water management activities (e.g., water reservoirs, wastewater treatment plant discharges, opencast lignite mining) (Bogena et al., 2005b). EC measurements have been taken with a sonic anemometer (CSAT3, Campbell Scientific Inc., Logan, USA) to measure the 3D wind components, an open-path gas analyzer (Li7500, LI-COR Inc., Lincoln, USA) to determine the H<sub>2</sub>O and CO<sub>2</sub> concentrations in the air, and an air temperature and humidity sensor (HMP45C, Vaisala Inc., Helsinki, Finland). Conversions to fluxes, including uncertainty information, are based on Mauder et al. (2013). The daily EC data were gap-filled by grass reference evapotranspiration calculated from the FAO Penman-Monteith equation (Allan et al., 1998). The non-closure of the energy balance of the EC data was not corrected. For further information on EC measurements and processing, the reader is referred to Bogena et al. (2018).

Table 4.1 CRNS sites used in this study, including key site characteristics.

Name	Latitude (°C)	Longitude (°C)	Altitude (m)	Mean annual precipitation (mm y <sup>-1</sup> )	Mean air temperature (°C)	Landuse
Merzenhausen	50.930	6.297	91	718	10.3	crop
Rollesbroich1	50.622	6.304	515	1018	7	grassland
Rollesbroich2	50.624	6.305	506	1018	7	grassland
Gevenich	50.989	6.324	107	718	10.3	crop
Ruraue	50.862	6.427	100	718	10.3	grassland
Wildenrath	51.133	6.169	72	722	10.3	needleleaf
Wüstebach	50.505	6.331	605	1401	7	spruce
Heinsberg	51.041	6.104	58	722	10.3	crop
Kall	50.501	6.526	505	857	8	grassland
Selhausen	50.866	6.447	101	718	10.3	crop
Schönseiffen	50.515	6.376	611	870	7	grassland
Kleinhau	50.722	6.372	374	614	9	grassland
Aachen	50.799	6.025	232	865	10.3	crop

#### 4.2.4 Data assimilation methodology

The EnKF was used in this work to assimilate SM measured by CRNS into the coupled model TSMP. The EnKF sequentially alternates model prediction and filter updating steps (also called filter analysis), either state updates alone or joint state-parameter updates. The effectiveness of the filter depends on the accurate determination of the forecast error covariance from the ensemble, and the sources of forecast errors are mainly uncertain initial conditions, forcing data, and model equations (Turner et al., 2008). To ensure that errors from various sources are taken into consideration to improve assimilation results, perturbation is used to create an ensemble that takes into account the different error sources. In this work, the ensemble of model realizations takes into account the uncertainty of model forcings (including precipitation, incoming shortwave radiation, incoming longwave radiation, and air temperature), parameters (including saturated hydraulic conductivity and porosity), and initial conditions (from spin-up).

For each ensemble member  $i$  at time step  $t$ , the SM state vector  $\mathbf{x}_{i,t}$  is updated by the model prediction. The forecast step is given by:

$$\mathbf{x}_{t,i} = f(\mathbf{x}_{t-1,i}, \mathbf{Q}_{t,i}, \mathbf{P}_{t,i}) \quad (4.1)$$

where  $i$  is the ensemble member,  $\mathbf{x}_{t,i}$  is the model forecast state vector at time step  $t$ ,  $f$  is the model TSMP,  $\mathbf{x}_{t-1,i}$  is the earlier model analysis state vector at time step  $t-1$ ,  $\mathbf{q}_{t,i}$  is the vector with (perturbed) model forcings and  $\mathbf{p}_{t,i}$  denotes the model perturbation vector with parameters. Model forecasts are updated according to:

$$\mathbf{x}_{t,i}^a = \mathbf{x}_{t,i}^f + \mathbf{K}_t (\mathbf{y}_{t,i} - \mathbf{H}_t \mathbf{x}_{t,i}^f) \quad (4.2)$$

where  $\mathbf{y}_{t,i}$  is the vector with (perturbed) observations, and the superscripts  $a$  and  $f$  refer to the updated state vector (the analysis) and the model predicted state vector, respectively. The observation operator  $\mathbf{H}_t$  is used to map model forecasts into the observation space, which is assumed to be linear, and  $\mathbf{K}_t$  denotes the Kalman gain that is calculated as:

$$\mathbf{K}_t = \mathbf{P}_t \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_t \mathbf{H}_t^T + \mathbf{R}_t)^{-1} \quad (4.3)$$

where  $\mathbf{P}_t$  is the model covariance matrix, which is calculated from the forecasted ensemble of model simulations at time step  $t$  according to:

$$\mathbf{P}_t = \frac{\sum_{i=1}^N (\mathbf{x}_{t,i}^f - \bar{\mathbf{x}}^f) (\mathbf{x}_{t,i}^f - \bar{\mathbf{x}}^f)^T}{N-1} \quad (4.4)$$

where  $\bar{\mathbf{x}}^f$  is a vector with ensemble mean values for the model states at time step  $t$ .  $\mathbf{R}_t$  is the measurement error covariance matrix, which is defined based on the expected measurement error of the CRNS soil moisture data ( $0.03 \text{ cm}^3/\text{cm}^3$ ).  $N$  is the number of ensemble members.

The updated states are then finally given by:

$$\mathbf{x}_{t,i}^a = \mathbf{x}_{t,i}^f + \mathbf{P}_t \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_t \mathbf{H}_t^T + \mathbf{R}_t)^{-1} (\mathbf{y}_{t,i} - \mathbf{H}_t \mathbf{x}_{t,i}^f) \quad (4.5)$$

In this work, the EnKF is also used to update the most sensitive parameter (saturated

hydraulic conductivities) in ParFlow. The other parameters were not updated because Brandhorst and Neuweiler (2023) found that updating multiple parameters for the unsaturated zone is prone to causing numerical instabilities, even in synthetic studies. The augmented state vector for updating both states and parameters is then extended and defined as follows:

$$\mathbf{x}_{t,i}^f = \begin{pmatrix} \Psi_{t,i} \\ \mathbf{y}_{t,i} \end{pmatrix} \quad (4.6)$$

where  $\mathbf{x}$  is the augmented state vector, including pressure heads ( $\Psi$ ) (m) and the logarithm of hydraulic conductivities ( $\mathbf{Y} = \log_{10}K_s$  (m/s)).

A damping factor ( $\boldsymbol{\alpha}$ ) is used when both states and parameters are updated, so as to reduce filter inbreeding (Hendricks Franssen and Kinzelbach, 2008; Hung et al., 2022). Filter inbreeding refers to the underestimation of the ensemble variance that occurs after the EnKF analysis is applied repeatedly, which happens when the ensemble size is small (Hendricks Franssen and Kinzelbach, 2008). The damping factor could reduce the modification of the forecast with the Kalman gain and limit the intensity of the perturbation of the parameter ( $\log_{10}K_s$ ) (Gebler et al., 2019). This results in the following updating equation for the joint state-parameter estimation:

$$\mathbf{x}_{t,i}^a = \mathbf{x}_{t,i}^f + \boldsymbol{\alpha}^T \mathbf{P}_t \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_t \mathbf{H}_t^T + \mathbf{R}_t)^{-1} (\mathbf{y}_{t,i} - \mathbf{H}_t \mathbf{x}_{t,i}^f) \quad (4.7)$$

where  $\boldsymbol{\alpha}^T$  is a vector with damping factors, of which 1 is for updating states and values between 0 and 1 are for updating parameters.

The DA updates states (and possibly also parameters) at all grid cells via the calculated model covariances, which give the covariances between all grid cells. Thus, during the analysis step, the states or parameters of the unassimilated locations are also updated, and the update is influenced by the correlation with the states or parameters of the assimilated locations.

### 4.3 Model and experiment setup

#### 4.3.1 TSMP-PDAF setup

The operation mode of TSMP-PDAF, when applied with the assimilation of CRNS soil moisture, is schematically illustrated in Fig 4.4. Before assimilation, the measurement depth needs to be determined. In order to determine a reasonable penetration depth for the CRNS observations for the corresponding model grid (500 m), a mean value for three distances to the CRNS station (2, 25 and 85 m) was calculated based on the revised method of Schrön et al. (2017). This calculation is necessary because the penetration depth depends on the distance to the CRNS station. And then the CRNS soil moisture observation is specified by PDAF for the soil layers until the measurement depth.

The states and parameters of each TSMP realization run are collected by PDAF after a predefined assimilation interval (Gebler et al., 2019). By assimilating SM observations, either model states or both model states and parameters are updated and passed back to the TSMP realizations, and then the updated states and parameters are used in the next prediction step, which is subsequently used in the next analysis, and so on. After the update has been made, the average weighted SM from the simulations is compared with the observed SM from CRNS.

We took the soil layers above 80 cm into account when calculating the weights to be assigned to the different soil layers for their contribution to the CRNS measurement signal. We calculated the weights for 1 mm thick soil layers and integrated the values to calculate the normalized weights for each model soil layer. Readers are referred to Schrön et al. (2017) for a more detailed description of weighting calculations.

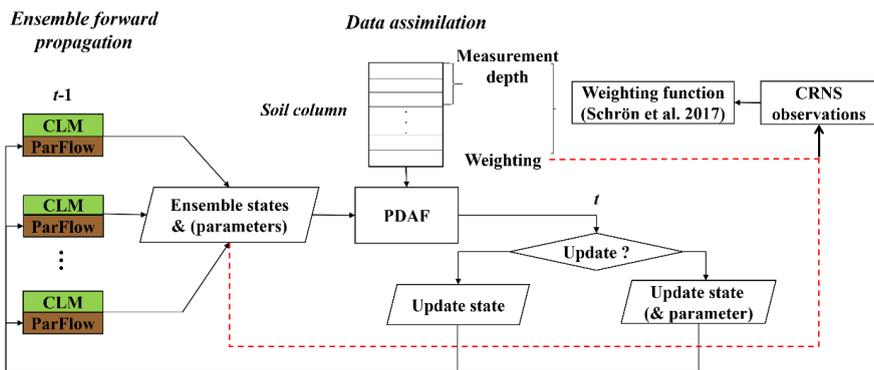


Figure 4.4 Schematic overview of the assimilation of soil water content from CRNS with PDAF

into TSMP (CLM-ParFlow). The flows represented by the red dashed line are outside TSMP-PDAF, including the weighting calculation and the comparison of CRNS soil moisture with the simulations.

### 4.3.2 Ensemble generation

The soil moisture DA experiments employ the EnKF with a total of 128 ensemble members. Each ensemble member was perturbed, with perturbations for meteorological forcings (precipitation, incoming shortwave radiation, longwave radiation, and air temperature), hydraulic conductivity, and porosity to account for uncertainties. The perturbed values were drawn from a multivariate normal distribution, considering the temporal correlation of the four meteorological variables, which were induced by a first-order autoregressive model (Reichle et al., 2010; Han et al., 2015). Table 4.2 summarizes the atmospheric forcing perturbations. The temporal correlations and standard deviations of the perturbations were assigned on the basis of earlier catchment-scale and regional-scale DA studies (Reichle et al., 2010; Han et al., 2013; Han et al., 2015; Roland Baatz et al., 2017). To ensure mass and energy balance, the perturbed precipitation and shortwave radiation are multiplied by the corrected lognormally distributed noises (Yamamoto, 2007; Han et al., 2013).

Table 4.2 The listed cross-correlations give the cross-correlations between the perturbations for the different atmospheric variables, following the order as indicated in the left column of the table.

Variables	Noise	Standard deviation	Time correlation scale	Cross correlation
Precipitation	Multiplicative	0.3	24 h	[ 1.0, -0.8, 0.5, 0.0,
Shortwave radiation	Multiplicative	0.2	24 h	-0.8, 1.0, -0.5, 0.4,
Longwave radiation	Additive	20 W m <sup>-2</sup>	24 h	0.5, -0.5, 1.0, 0.4,
Air temperature	Additive	1 K	24 h	0.0, 0.4, 0.4, 1.0]

The vertical profile of the model has a vertical extension of 100 m and is discretized into 25 layers with varying thicknesses. The upper ten layers extend to 3 m, which coincides with the CLM soil layers, and the lower layers are treated as aquifer layers. The  $K_s$  and porosity of the soil and bedrock layers were perturbed separately.

For the soil layers, soil texture (sand and clay contents) was perturbed by geostatistical

simulation. A spherical variogram model was adopted to generate a spatially correlated random field with zero mean, variance  $50\%^2$  and correlation length of 12.5 km, and the generated perturbation field was added to the original soil texture as derived from the soil map. The sand and clay contents were perturbed separately. The percentages for sand and clay were limited to a range between 0% and 100% to prevent unphysical values for soil texture. The final silt contents were calculated from the perturbed sand and clay contents. Subsequently, the Rosetta pedotransfer functions (Schaap et al., 2001; Zhang and Schaap, 2017) were employed to calculate the perturbed  $K_s$  and porosity based on the perturbed soil texture.

For the bedrock layers, the original  $K_s$  values are from the hydrogeological map (Fig 4.3). For each ensemble member, the  $\log_{10}K_s$  of all the bedrock layers were perturbed by additive random values (same values for each ensemble member) from a univariate uniform distribution with values between -0.5 and 0.5. The porosity for the lower bedrock layers was set to a constant value of 0.15.

### 4.3.3 Setup of the DA Experiments

After generating the ensemble, spin-up simulations for the ensemble members were performed in order to achieve a dynamic groundwater equilibrium. The multi-year average water table depth derived by Bogaen et al. (2005b) was used as the initial condition, and the 30-year average recharge values (derived from gridded German Meteorological Service data on precipitation and actual evapotranspiration) were used as the upper boundary condition for the ParFlow model. The spin-up simulations for ParFlow were conducted for 100 years. Next, the final conditions from ParFlow's spin-up were used to continue the spin-up for TSMP, including both CLM and ParFlow. This was done for a period of 5 years, using atmospheric forcings from the year 2015 (for the DA experiments in the wet year 2016) or the year 2017 (for the DA experiments in the dry year 2018) as input.

The CRNS soil moisture data were assimilated into the model TSMP by PDAF. In the DA experiments, the states were updated daily by DA, and saturated hydraulic conductivity was updated every three days. Those are the optimal updating frequencies found after conducting different assimilation experiments. When jointly updating states and parameters, a damping factor of 0.1 was employed to limit the intensity of the hydraulic conductivity perturbation (Hung et al., 2022) and reduce the possibility of filter inbreeding (Hendricks Franssen and Kinzelbach, 2008). The river grid cells were masked during assimilation analysis to avoid instabilities. The year 2017 is used as an independent evaluation period for the DA experiments

of 2016 and 2018. For the year 2017, the ensemble model ran with the updated parameters from 2016 and 2018, but without assimilation.

In addition to the DA experiments, jackknife simulations were also carried out to assess the effect of the CRNS assimilation on SM simulation at unassimilated locations in the model domain. When performing a jackknife simulation, 11 sites were used for assimilation (jointly updating states and parameters with a damping factor of 0.1), and the remaining one site was used for evaluation, so there were 12 jackknife experiments for each assimilation year (2016 and 2018). Table 4.3 lists all the experiments conducted.

Table 4.3 List of conducted simulation experiments: open loop (OL), data assimilation with state update (State) or joint state and parameter update (Joint), jackknife evaluation runs (Jackknife), and verification experiments in 2017 using the updated saturated hydraulic conductivity ( $K_s$ ) from joint assimilation experiments of 2016 and 2018 (Updated  $K_s$  from 2016 and Updated  $K_s$  from 2018).

Year	Experiment	Update state	Update parameter
2016	OL	-	-
	State	+	-
	Joint	+	+
	Jackknife	+	+
2018	OL	-	-
	State	+	-
	Joint	+	+
	Jackknife	+	+
2017	OL	-	-
	Updated $K_s$ from 2016	-	-
	Updated $K_s$ from 2018	-	-

#### 4.3.4 Evaluation of model performance

The simulated SM results were evaluated with the following statistical metrics: bias (BIAS), mean absolute error (MAE), correlation coefficient (R), RMSE, and unbiased root mean square difference (ubRMSD):

$$BIAS = \sum_{t=1}^n (SM_t^{sim} - SM_t^{obs}) \quad (4.8)$$

$$MAE = \frac{\sum_{t=1}^n (|SM_t^{sim} - SM_t^{obs}|)}{n} \quad (4.9)$$

$$R = \frac{\sum_{t=1}^n (SM_t^{obs} - \overline{SM^{obs}})(SM_t^{sim} - \overline{SM^{sim}})}{\sqrt{\sum_{t=1}^n (SM_t^{obs} - \overline{SM^{obs}})^2 \sum_{t=1}^n (SM_t^{sim} - \overline{SM^{sim}})^2}} \quad (4.10)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (SM_t^{sim} - SM_t^{obs})^2}{n}} \quad (4.11)$$

$$ubRMSD = \sqrt{\frac{\sum_{t=1}^n ((SM_t^{obs} - \overline{SM^{obs}}) - (SM_t^{sim} - \overline{SM^{sim}}))^2}{n-1}} \quad (4.12)$$

where  $n$  is the total number of time steps,  $SM_t^{sim}$  the simulated ensemble average soil moisture content at the time step  $t$  (either from an OL or DA run), and  $SM_t^{obs}$  the observed soil moisture by CRNS at the time step  $t$ . The overbar in equations 10, 12 and 13 indicates the temporal mean over the study period.

The above performance measures were also used to evaluate the effect of CRNS soil moisture assimilation on evapotranspiration and discharge characteristics simply by replacing SM with evapotranspiration and discharge in the equations.

The Nash-Sutcliffe efficiency (NSE) index was also used to evaluate the simulations of discharge. The NSE was calculated according:

$$NSE = 1 - \frac{\sum_{t=1}^n (Q_t^{sim} - \overline{Q^{sim}})^2}{\sum_{t=1}^n (Q_t^{obs} - \overline{Q^{obs}})^2} \quad (4.13)$$

where  $Q_t^{sim}$  is the simulated ensemble average discharge at the time step  $t$  (either from an open loop (OL) or DA run), and  $Q_t^{obs}$  is the observed discharge at the time step  $t$ . The NSE range is between  $-\infty$  and 1. The closer to 1, the more accurate the model is.

## 4.4 Results

### 4.4.1 Soil moisture data assimilation general results

Table 4.4 summarizes the performance in terms of error statistics for OL and different DA experiments (state updates alone and joint state-parameter updates) for 2016 and 2018. For all the DA experiments, both for the wet year 2016 and the dry year 2018, and both for state updating alone and joint state-parameter updating, BIAS, MAE, RMSE, and ubRMSD were lower than for the OL run, and R was higher than for the OL run, indicating that simulated soil moisture was closer to the measurements after assimilation. Joint state-parameter estimation gave better results than state estimation alone, both in 2016 and 2018. The RMSE of SM decreased by 42.9% (2016) and 36.2% (2018) for state updating only, while for joint state-parameter updates, decreases were 59.7% (2016) and 52.2% (2018). The best assimilation results resulted in similar soil moisture RMSE values for 2016 and 2018, namely 0.031 and 0.033 cm<sup>3</sup>/cm<sup>3</sup>, respectively. The DA results for two different years illustrate that the effect of CRNS assimilation at the assimilated locations is consistent.

Fig A.1 and A.2 show the temporal courses of CRNS measured SM, and simulated SM from OL and joint state-parameter estimation in 2016 and 2018 at all the CRNS locations. The figures clearly show that simulated SM for all CRNS sites is closer to measurements after assimilation. The error statistics for all CRNS sites can be found in Tables A.1 and A.2. The RMSE is less than 0.035 cm<sup>3</sup>/cm<sup>3</sup> for most sites (except Wüstebach in 2016, and Rollesbroich and Aachen in 2018), which is within the acceptable range.

For Wüstebach, the RMSE was 0.059 cm<sup>3</sup>/cm<sup>3</sup> in 2016, which is still much larger than the measurement error. The scatter plot for Wüstebach in Fig A.3 indicates that modelled SM updates reach an upper plafond, which is defined by the maximum possible porosity in the model, which is determined by the soil texture and pedotransfer function. However, the remarkably high porosity of the forest soil is due to the very high content of organic material in the topsoil, so the real porosity of Wüstebach cannot be represented in the model by the soil texture alone (Strebel et al., 2022). Therefore, the upper porosity limit inhibits further improvement of SM characterization. For Rollesbroich, the RMSE was 0.048 cm<sup>3</sup>/cm<sup>3</sup> in 2018, and the poor performance is also limited by the porosity, as can be seen in the scatter plot. As for Aachen, the RMSE was 0.038 cm<sup>3</sup>/cm<sup>3</sup> in 2018, and SM simulation might have been negatively impacted by irrigation, which was not accounted for in the model (land use is crop and 2018 was a dry year).

Table 4.4 Error statistics for open loop (OL), data assimilation with state updates (State), joint state-parameter updates (Joint), and jackknife simulations with joint state-parameter updates (Jackknife) for the assimilation periods of 2016 and 2018. The indicators were averaged over all sites with CRNS soil moisture observations. Site-specific indicators are provided in Appendix Tables A.1 to A.4.

Year	Simulation	BIAS (cm <sup>3</sup> /cm <sup>3</sup> )	MAE (cm <sup>3</sup> /cm <sup>3</sup> )	R	RMSE (cm <sup>3</sup> /cm <sup>3</sup> )	ubRMSD (cm <sup>3</sup> /cm <sup>3</sup> )
<b>2016</b>	OL	-0.051	0.062	0.795	0.077	0.058
	State	-0.024	0.034	0.918	0.044	0.037
	Joint	-0.004	0.023	0.942	0.031	0.031
	Jackknife	-0.012	0.036	0.879	0.046	0.045
<b>2018</b>	OL	0.005	0.054	0.739	0.069	0.069
	State	-0.008	0.032	0.895	0.044	0.043
	Joint	0.001	0.024	0.944	0.033	0.033
	Jackknife	0.008	0.046	0.816	0.058	0.058

To better understand the effect of the DA during different hydrological conditions, the RMSEs of the SM were also calculated over different seasons (see Appendix Table A.5). For DA experiments (both 2016 and 2018), the seasonal SM simulations were obviously improved after assimilation, and joint state-parameter estimation resulted in better performance compared to state update alone. The reduction in RMSE showed small differences across the four seasons in both 2016 and 2018, all ranging from 50% to 60%, suggesting the robustness of DA performance at assimilation locations under different hydrological conditions.

Examples of SM spatial distributions (vertically averaged) for the OL and joint state-parameter estimation runs are shown in Fig 4.5. The simulated SM for the whole catchment is corrected by DA for both state updates alone and joint state and parameter updates. However, the difference between state assimilation and joint assimilation is small, indicating that parameter update influence is limited. For 2016, a comparison between simulated values from OL and measurements at the CRNS sites revealed that the OL simulation was too dry. Data assimilation corrected the simulations, and in some too dry parts of the catchment, such as grid cells near the river, soil water content increased by assimilation. On the contrary, in 2018, the OL overestimated SM content, and the DA corrected SM toward lower values.

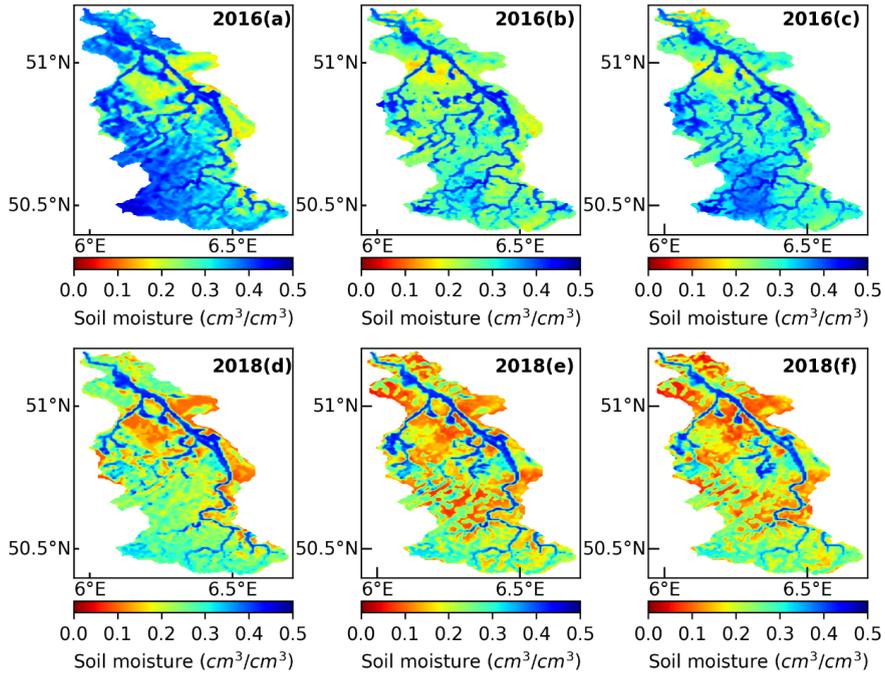


Figure 4.5 Examples of the simulated soil moisture distribution over the Rur catchment on the 22<sup>nd</sup> of July in 2016 and 2018. Subplots a) and d) are from the open loop, b) and e) are from data assimilation with state update, and c) and f) are from joint state-parameter update simulations.

#### 4.4.2 Jackknife simulations

In order to investigate whether the limited CRNS stations could improve the simulated SM at locations beyond the CRNS stations over the Rur catchment, 12 jackknife simulations were performed for each year (2016 and 2018). The EnKF may enhance the spatial accuracy of the simulated SM, given the spatial correlation of atmospheric forcings, soil hydraulic parameters, and SM. The overall error statistics of 12 jackknife simulations for 2016 and 2018 are shown in Table 4.3. Overall, the jackknife runs reduced MAE, RMSE, and ubRMSD, and increased R compared to OL, demonstrating that the SM simulation at verification locations also improved. On average, the RMSE of the 12 jackknife runs for 2016 was  $0.046 \text{ cm}^3/\text{cm}^3$ , which is much lower than the RMSE for the OL run ( $0.077 \text{ cm}^3/\text{cm}^3$ ) and only a bit higher than when only the state was updated. For the year 2018, the jackknife simulations resulted in a smaller RMSE

reduction at the verification locations, with an average RMSE of  $0.058 \text{ cm}^3/\text{cm}^3$  ( $0.069 \text{ cm}^3/\text{cm}^3$  for the OL run).

For jackknife runs, the seasonal SM simulations showed quite different performances (see Appendix Table A.5), and the extent of RMSE reduction was positively correlated with the average measured SM; the higher the SM content, the larger the RMSE reduction. Therefore, the best performance occurred during the winter, when SM was at its maximum for the whole year (47.7% and 34.3% RMSE reduction for 2016 and 2018, respectively). The worst performance was found for dry soil water conditions, e.g., 32.3% RMSE reduction in autumn 2016. In 2018, the extreme dry conditions in the summer even led the RMSE to increase by 8.8%.

For each CRNS site, the jackknife simulation performed differently in 2016 and 2018. More detailed site statistics can be found in Tables A.3 and A.4. For 2016, all jackknife simulations resulted in an improved RMSE at the verification locations compared to the OL run. Assimilation could reduce RMSE by 70% at sites with a high RMSE in the OL run, such as Aachen (see Fig 4.6). In 2018, the RMSE for Aachen decreased by 36%, but RMSE reductions were smaller at other sites (for Gevenich, Heinsberg, and Schönseiffen, the RMSE even increased after assimilation). For Heinsberg, the RMSE for the DA-run in 2018 is higher ( $0.057 \text{ cm}^3/\text{cm}^3$ ) than for the OL-run ( $0.044 \text{ cm}^3/\text{cm}^3$ ), while the RMSE value decreased by 35% in 2016 (see Fig 4.6). Fig A.7 shows the spatial correlation of SM from OL between the CRNS locations (Gevenich, Heinsberg, and Schönseiffen) and other grid cells in the catchment on a specific day in the summer. The figure indicates that the spatial correlation around the CRNS locations was weaker in 2018 (dry) compared to 2016 (wet), which resulted in less accurate jackknife simulations in 2018 compared to 2016.

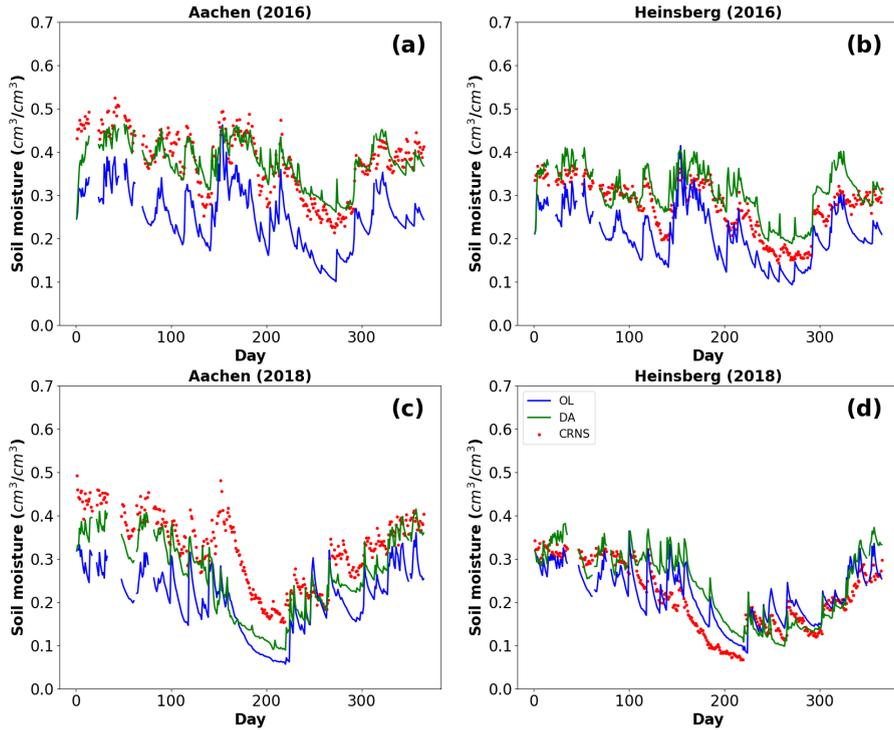


Figure 4.6 Temporal evolution of simulated soil moisture from the open loop mean (OL, blue) and jackknife simulation mean (DA, green), together with the observed soil moisture from the CRNS (red), for 2016 (a, b) and 2018 (c, d) at the CRNS sites. Simulated soil moisture was vertically weighted using the revised method by Schrön et al. (2017).

#### 4.4.3 Temporal evolution of parameter estimates and parameter verification

The temporal evolution of  $K_s$  estimates during the assimilation period (2016 and 2018) for the CRNS sites is shown in Fig 4.7. Once the assimilation began, the parameters varied considerably within short time intervals. For most sites, the updated  $K_s$  started to stabilize after about 100 days of assimilation. Compared to the initial input  $K_s$ , most sites showed a decreasing trend during assimilation, while only Rollesbroich in 2016 showed a slightly increasing trend. The changing values for  $K_s$  estimates for Merzenhausen, Gevenich, Ruraue, Heinsberg, Selhausen, and Kleinhau were remarkably consistent for the two distinct assimilation years.

The years 2016 and 2018 resulted in very similar parameter sites, with differences smaller than  $0.10 \log_{10}$  (m/s) units at the end of the assimilation period for more than half of the sites

(Fig 4.7). Some sites like Gevenich and Kall showed only slight variations from the prior values, with  $K_s$  changes less than  $0.20 \log_{10} (\text{m/s})$  units, while for Ruraue and Wüstabach  $K_s$  changed more than  $0.45 \log_{10} (\text{m/s})$  units. Among all sites, Wildenrath has the largest absolute variation, with  $K_s$  varying more than  $10^{-5} \text{ m/s}$ , while Kall showed very small variations, with absolute  $K_s$  changes less than  $5 \cdot 10^{-7} \text{ m/s}$ . Temporally unstable and inconsistent parameter estimates imply that there may be multiple or seasonal optimal parameter values, so the fluctuations in  $K_s$  may be related to variations in atmospheric forcings. Some instability in the updated parameters could also be related to the compensation for other errors, e.g., errors in the inputs (from atmospheric forcings or soil hydraulic parameters) and model structural errors.

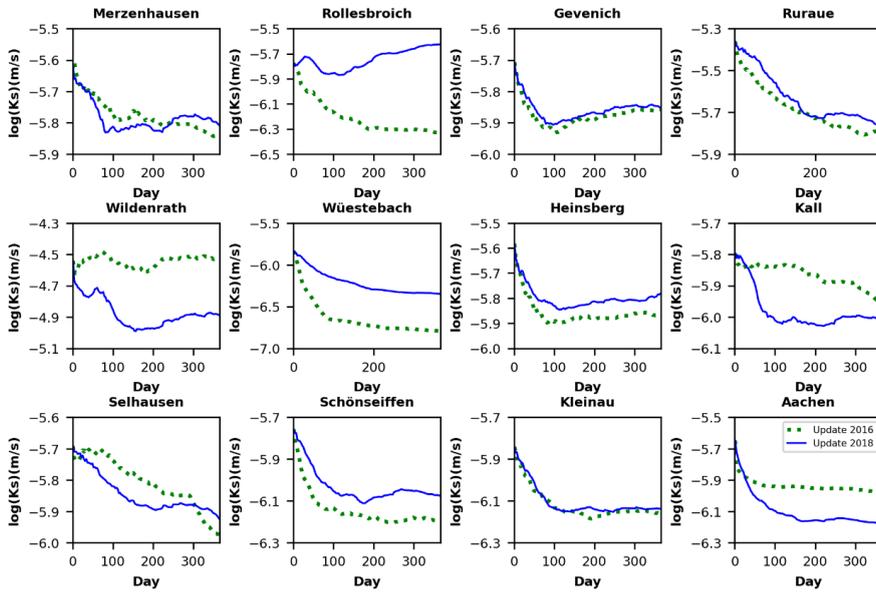


Figure 4.7 Estimates of averaged saturated hydraulic conductivity ( $\log_{10}K_s$ ) from data assimilation experiments with joint state-parameter updating during the periods of 2016 and 2018 at CRNS locations. The input value of  $K_s$  is indicated at the first time step.

Fig 4.8 depicts the prior and updated spatial ensemble mean of  $\log K_s$  at 2 cm depth (similar pattern for depths up to 80 cm), both for the years 2016 and 2018. Assimilation had a noticeable impact on  $K_s$ , particularly around the assimilated CRNS measurement locations, resulting in a

decrease in its overall value. The  $\log K_s$  changes for the simulation year 2016 were more noticeable than for 2018. The reason for the larger  $\log K_s$  updates in 2016 could be that the simulated SM content by the OL was lower than the observed values. 2018 was a particularly dry year, so the SM condition in 2018 was closer to that of OL than in 2016, resulting in a larger update of the  $\log K_s$  in 2016 than in 2018. Grids with larger distances to the CRNS sites show smaller  $\log K_s$  updates because of the weak correlations with the SM observations. Additionally, there are some grid cells with increased  $\log K_s$  after DA, suggesting that horizontal water redistribution, e.g., due to lateral groundwater flow or surface runoff, resulted in different  $\log K_s$  changes than at the CRNS sites.

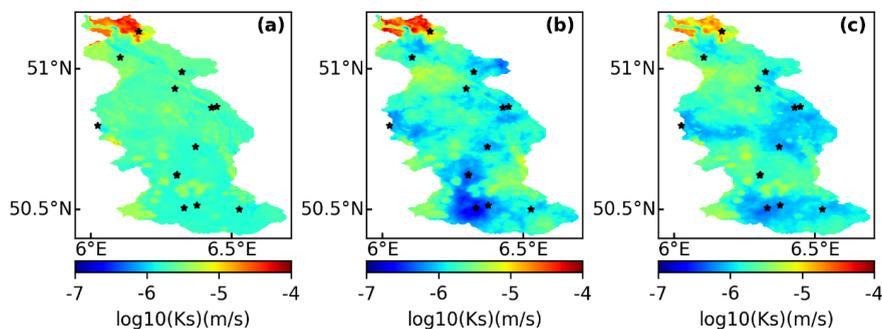


Figure 4.8 Ensemble averaged  $\log_{10}K_s$  fields of the soil at 2 cm depth: (a) prior field; (b) DA with joint state-parameter updates at the end of 2016; (c) DA with joint state-parameter updates at the end of 2018. The black asterisk is the location of the CRNS sites.

Simulations were made for the verification year 2017, using as input updated hydraulic parameters from either 2016 or 2018. For the verification year, reduced BIAS, MAE, RMSE, and ubRMSD, and a small increase in R compared to OL were found (see Table 4.5). Using the updated  $K_s$  from the 2016 simulation as input to the simulation for the year 2017 gave simulated SM contents that were closer to observations than when the updated  $K_s$  from the 2018 simulation were used as input. The updated parameter's verification for the year 2018 was less successful than for the year 2016, which may be due to the hydrological conditions during 2016 (average CRNS soil moisture  $0.31 \text{ cm}^3/\text{cm}^3$ ) being more similar to 2017 ( $0.29 \text{ cm}^3/\text{cm}^3$ ), as 2018 ( $0.26 \text{ cm}^3/\text{cm}^3$ ) was a dry year.

Table 4.5 Comparison of measured and simulated soil moisture for the year 2017 (evaluation period, no assimilation). The updated parameters used for verification were from the assimilation period (2016 and 2018). The error statistics were averaged over all CRNS sites.

Simulation	BIAS (cm <sup>3</sup> /cm <sup>3</sup> )	MAE (cm <sup>3</sup> /cm <sup>3</sup> )	R	RMSE (cm <sup>3</sup> /cm <sup>3</sup> )	ubRMSD (cm <sup>3</sup> /cm <sup>3</sup> )
OL	-0.019	0.051	0.733	0.066	0.063
Updated Ks from 2016	-0.007	0.047	0.769	0.060	0.059
Updated Ks from 2018	-0.012	0.049	0.760	0.061	0.060

#### 4.4.4 Evapotranspiration and discharge

The effect of soil moisture DA on ET modeling was also investigated. We used observed ET data from three EC stations for comparison with simulated values in order to examine the impact of CRNS soil moisture assimilation on ET simulations. Results of the DA experiments showed that SM states were significantly altered, and ET was also somewhat impacted by the different assimilation scenarios, depending on the simulation year (see Table A.6). Moreover, the joint state-parameter assimilation resulted in a better ET prediction than SM state updating alone.

The statistical performance measures BIAS, MAE, R, RMSE, and ubRMSD, comparing simulated ET (by OL and joint state-parameter updates) and EC data, are provided in Table 4.6. These statistical measures were computed on a monthly basis, as the parameter LAI in the CLM model is provided on a monthly scale. The joint state-parameter updates with CRNS soil moisture assimilation showed lower BIAS, MAE, and RMSE values than OL, except for Rollesbroich in 2018, demonstrating that ET simulation improved if SM simulation was improved by DA. However, the relative improvement in the characterization of ET is far smaller than for SM. The high correlation coefficients (larger than 0.95) for Rollesbroich and Wüstebach, either for OL or DA, indicate a good fit between simulated and measured ET, mainly because of the reproduction of the yearly cycle. For Selhausen, the correlation is lower (less than 0.85), which might be related to different crops being cultivated for the years 2016 and 2018 (winter barley in 2016 and winter wheat in 2018). CLM uses the same parameters for these crops, and for example, the harvest date is not well represented by the model. For Rollesbroich, the ET simulation in 2018 was worse after DA (compared to OL) because the overestimated SM in the OL run was corrected towards lower values, reducing ET, which further exacerbated the ET underestimation.

Table 4.6 Comparison of measured and simulated evapotranspiration (monthly) and discharge (monthly) from open loop (OL) and data assimilation runs with joint state-parameter updates (DA) for two assimilation periods (2016 and 2018).

Year	Site	BIAS (mm/month)		MAE (mm/month)		R		RMSE (mm/month)		ubRMSD (mm/month)	
		OL	DA	OL	DA	OL	DA	OL	DA	OL	DA
		2016	Rollesbroich	-10.17	-8.02	10.31	8.75	0.98	0.97	12.97	10.79
	Wüstebach	-17.93	-16.05	17.93	16.36	0.99	0.99	21.48	19.41	11.82	10.91
	Selhausen	-14.11	-7.27	15.77	12.37	0.87	0.85	18.53	14.63	12.02	12.69
2018	Rollesbroich	-12.77	-14.26	13.68	15.22	0.97	0.97	18.05	19.86	12.76	13.83
	Wüstebach	-11.46	-10.50	13.36	12.33	0.96	0.95	17.24	16.25	12.88	12.40
	Selhausen	-8.14	-3.87	16.33	14.02	0.80	0.81	20.66	18.73	18.99	18.32

Year	Site	BIAS (m <sup>3</sup> /s)		MAE (m <sup>3</sup> /s)		R		RMSE (m <sup>3</sup> /s)		NSE	
		OL	DA	OL	DA	OL	DA	OL	DA	OL	DA
		2016	Erkensruhr- Einruhr	-0.23	-0.23	0.32	0.31	0.85	0.86	0.48	0.46
2018	Erkensruhr- Einruhr	-0.15	-0.16	0.27	0.27	0.90	0.86	0.41	0.41	0.67	0.69

Fig 4.9 shows the monthly temporal ET variations for the OL run and the joint state-parameter experiment, compared to the EC data. The simulated ET tends to be closer to the observed values after SM assimilation, and the larger changes in ET simulation are observed during drier conditions, specifically in the summer, which is consistent with the results by Hung et al. (2022). For example, for Rollesbroich (2016) and Wüstebach (2018), the largest reductions in RMSE for ET occur in the summer, with 11.5% and 5.4%, respectively. However, for Wüstebach (2016), RMSE decreased most in the autumn, with 9.5%. For Selhausen, the largest RMSE reduction in ET occurred in the spring, for both 2016 and 2018, with reductions of 27.8% and 22.2%, respectively. In the winter, ET simulations hardly improved for Rollesbroich and Wüstebach. ET is limited by available energy under conditions of high SM, so SM changes have a minimal impact on ET. Therefore, the overall ET improvement for the entire year is limited.

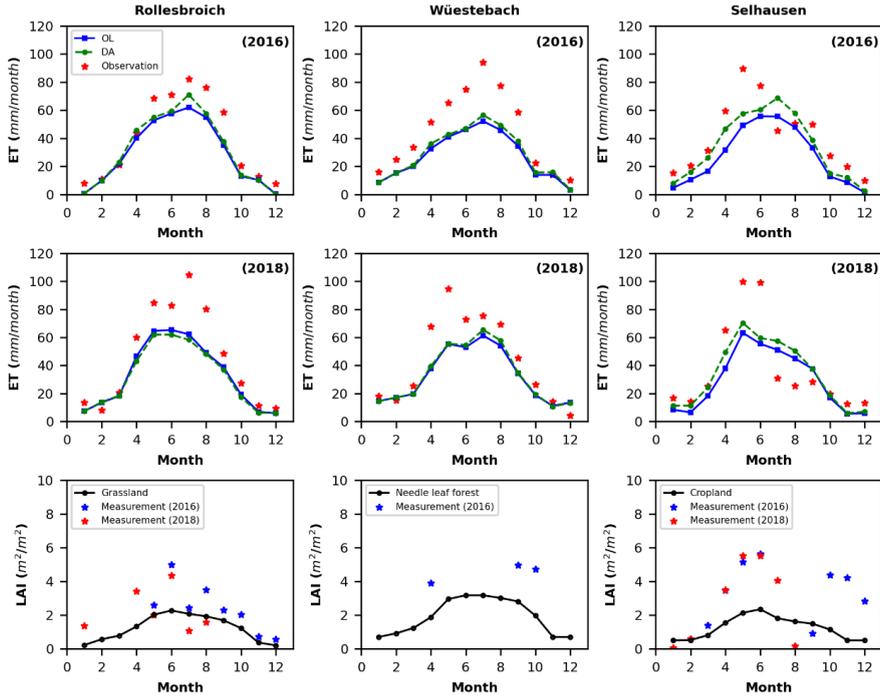


Figure 4.9 Temporal evolution of simulated evapotranspiration from open loop (OL, blue), data assimilation (joint state-parameter updates, DA, green), and the observed evapotranspiration (red) at the sites Rollesbroich, Wüstebach, and Selhausen for the assimilation periods of 2016 (first row) and 2018 (second row). Monthly Leaf Area Index (LAI) for the plant functional types at the sites Rollesbroich (grassland), Wüstebach (needle leaf forest), and Selhausen (cropland), as well as the available daily LAI measurements from 2016 and 2018 (third row).

The annual ET for 2016 and 2018 across the whole Rur catchment for OL and joint state-parameter updating are presented in Fig 4.10. Since the ET changes were minimal in the assimilation experiments, only results for joint state-parameter updating are shown. For the OL simulation, the ET in 2018 was greater than in 2016, mainly due to the significantly higher temperature and higher incoming shortwave radiation in 2018, and in spite of the drier conditions. DA did not much affect simulated ET in the southern part of the catchment, where ET was generally energy limited. In contrast, simulated ET in the northern part of the catchment with generally less precipitation was affected by DA, with ET increases of more than 50 mm yr<sup>-1</sup> for many grid cells in 2016, whereas ET was modified less by DA in 2018. This is related

to the larger update of soil hydraulic parameters in 2016 compared to 2018.

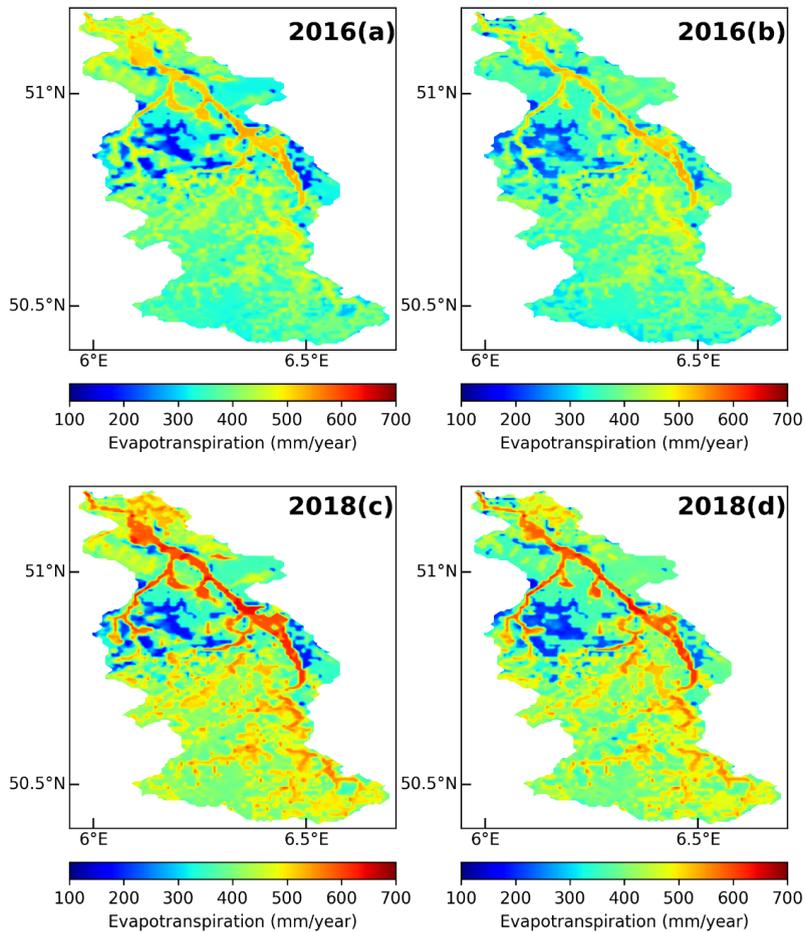


Figure 4.10 Annual evapotranspiration from open loop (a, c) and data assimilation runs (joint state-parameter updates) (b, d) over the Rur catchment during the assimilation periods.

Fig 4.11 compares the simulated river discharge from the OL and DA experiments to the discharge from the in situ station Erkensruhr-Einruhr (indicated in Fig 4.1). This sub-catchment was hardly affected by water management operations, so it was selected. For comparison purposes, the SM data of the CRNS station in Wüstebach are also shown since it is located in

the catchment area of the Erkensruhr-Einruhr station. For the sake of simplicity, only the results of the joint state-parameter update are shown, as the results for the other experiments are very similar. The coupled model performance for discharge simulation is satisfactory, as the NSEs of 2016 and 2018 were 0.62 and 0.64, respectively. The simulated discharge could capture the daily variations, including discharge peaks (see Fig 4.11). DA only slightly improves monthly discharge estimation compared to the OL (see Table 4.6), with an increase in NSE of about 0.02 for both 2016 and 2018, even though SM in Wüstebach was significantly improved by DA.

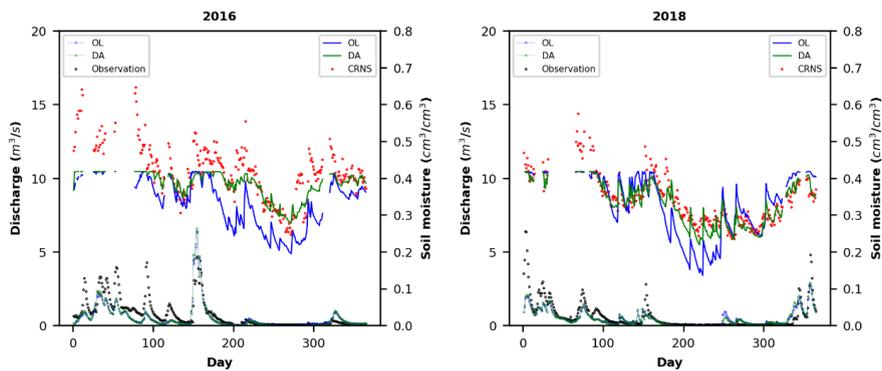


Figure 4.11 Temporal evolution of simulated discharge from the open loop mean (OL, blue circle) and joint state-parameter assimilation mean (DA, green circle), together with the observed discharge (black circle) for 2016 and 2018 at the Erkensruhr-Einruhr in situ station. The temporal evolution of simulated soil moisture from the open loop mean (OL, blue line) and joint state-parameter assimilation mean (DA, green line), together with the observations (red dot) for 2016 and 2018 at the Wüstebach site.

#### 4.4.5 Discussion

This study demonstrated that the assimilation of CRNS soil moisture data is beneficial and improves the integrated terrestrial system model simulations of SM over a real catchment, both for a wet and dry year. In addition, the jackknife simulations demonstrate the potential of the CRNS network to improve modelled SM at the catchment scale, but it performs differently in a wet and dry year. The improvement in the dry year is relatively small, due to the weaker spatial correlations in the dry year 2018, compared to the wet year 2016. The same perturbation methods were used for the two years and the spatial correlation length utilized for soil hydraulic

parameter perturbations ensured that each site used for validation was within the correlation length of the assimilated sites. However, the SM spatial correlation fluctuated under various soil hydrological conditions and was shown to be weakened under drought conditions. As a result, the drought in the summer of 2018 led to the worst validation performance of the seasonal SM simulation. The overall results from the jackknife experiments also indicate that the RMSE is much less reduced at the verification locations than at the assimilation locations.

It is very likely that a denser CRNS network may improve SM characterization, for example, related to better parameter estimates and compensate for variations in performance across years. To further investigate how dense an optimal measurement network of CRNS should be, one possible approach could involve conducting a synthetic study that tests varying numbers and locations of CRNS stations based on the model results established in this study. In addition, some denser CRNS observation networks are gradually being established, such as the new Irish Soil Moisture Observation Network (ISMON) (Finkele et al., 2022) and some field campaigns in which a large number of CRNS were operated together to explore the potential of a dense stationary CRNS network to monitor spatio-temporal SM dynamics at the catchment scale. For instance, a dense network of 24 CRNS was established in an area of only 1 km<sup>2</sup> in the pre-alpine Rott headwater catchment in Southern Germany (Fersch et al., 2020) and a network of 15 CRNS covering an area of 0.39 km<sup>2</sup> in the Wüstebach headwater catchment in Western Germany (Heistermann et al., 2022). In our study, we show the potential high-density CRNS networks have to correct for errors introduced by imperfect input data and spatial correlations, thus reducing the uncertainties in SM prediction. The establishment of the above-mentioned CRNS observation networks offers the opportunity to further investigate how the density of sensors influences SM assimilation.

Joint state-parameter estimating improved SM simulations, especially at measurement locations, but much less at verification locations. Hydraulic conductivity was only modified slightly and locally during joint state-parameter updating, and as a result, SM characterization only improved slightly. Better results could be achieved with a larger ensemble size. Here, DA experiments were performed with 128 ensemble members, but this ensemble size might still be too small. Hendricks Franssen and Kinzelbach (2008) suggested that 200-500 realizations are needed to achieve successful joint state-parameter estimation with groundwater hydrological models. A larger ensemble size, however, was not feasible in this work, given the needed compute time for a run with the high-resolution integrated model.

DA reduced differences between simulated and measured SM contents significantly, but the benefit of DA was not clear for the modelling of ET, and the findings are in line with the

synthetic study by Hung et al. (2022) with the TSMP-PDAF model. Similar findings were made by Ridler et al. (2014), who found that SM assimilation had little influence on flux estimation. Uncertain parameters and model structural errors are also possible reasons for the limited improvement in ET simulations after SM assimilation.

It is important to note that more studies have reported the underestimation of evapotranspiration (or latent heat flux) by CLM (Shrestha et al., 2014; Boas et al., 2021). The ET mismatch in our simulations was largely related to a systematic underestimation of ET (bias), which seems partly related to underestimated LAI values. The input LAI used in this study was taken from previous studies (Sulis et al., 2015; Sulis et al., 2018), but was found to be smaller than the measured LAI values at EC sites (see Fig 4.9). Notice that the measured LAI is on a daily basis, while the LAI in the model is defined on a monthly basis, but the systematic underestimation of the LAI in the model can nevertheless be observed.

The performance of the ET simulation is also affected by the use of a uniform set of parameters for crops, neglecting the fact that different crops in the region can have very different properties (e.g., LAI and stem area index). Sulis et al. (2015) incorporated crop-specific parameters in CLM 3.5 simulations, which resulted in improved simulations of land-atmosphere exchange fluxes compared to simulations using the generic crop type. Similarly, a more recent study by Boas et al. (2021) found that utilizing crop-specific parameters in the newer version 5.0 of CLM improved the representation of crop growth cycles and led to more accurate simulations of energy fluxes. The model CLM 5 shows a better characterization of ET than CLM 3.5 (Shrestha et al., 2018a), but it is not yet coupled to ParFlow, which is the reason why we did not use it in this work. In addition, other factors influence ET, like vegetation rooting depth and further vegetation characteristics (Li et al., 2020). The uncertainty of those parameters was not considered in this work, but in order to improve ET simulation in DA studies, their uncertainty should be considered in the future. Finally, mismatches between modelled and measured ET are probably also related to errors in the input of atmospheric forcings like incident radiation. A fully coupled atmosphere-land surface-subsurface model integrated into the DA framework may further improve the characterization of ET.

Therefore, it can be concluded that assimilating only SM is insufficient to significantly improve the simulation of ET and that parameter biases and model errors are more important for the ET simulation. Better results may be achieved by assimilating additional types of measurements, like LAI, and estimating further parameters, like vegetation parameters.

River discharge was used to investigate the effects of SM assimilation on lateral fluxes,

with slight improvements in discharge estimates. Our finding is consistent with previous synthetic SM assimilation experiments at the hillslope or larger catchment scale using the integrated model TSMP (Gebler et al., 2019; Hung et al., 2022). The limited improvement in discharge characterization may be attributed to the limited spread of discharge and to the fact that only the soil hydraulic parameter  $K_s$  was updated without large changes in the parameter values. On the other hand, the limited improvement might also be partly related to model structural errors (e.g., underrepresentation of preferential flows and representation of drainage) (Gebler et al., 2019). Furthermore, Baroni et al. (2017) found that river discharge in large catchments is only sensitive to the perturbation of long spatial structures and is not affected by small-scale soil variabilities. Therefore, with only SM and parameter  $K_s$  being updated in the integrated model, an improvement in the performance of discharge estimation is challenging. Possible improvements could be achieved by considering the uncertainties in other parameters. For instance, D. Baatz et al. (2017) found that the estimated Manning's roughness coefficients could improve the discharge simulation with TSMP in synthetic 2D experiments.

In this study, we did not directly assimilate the CRNS neutron intensity observations but used SM products derived from the CRNS observations. Next, it is planned to assimilate neutron count intensity directly with the COSMIC operator (Shuttleworth et al., 2013). In addition, although we used a 500 m resolution, which is already fine relative to the RS data, 500 m is still coarse compared to the footprint of CRNS data. Therefore, a higher resolution will be used in the future to include the calculation of the horizontal weighting of the CRNS observations.

Our study demonstrates the potential of a CRNS observational network to enhance SM estimation as well as other hydrological variables (evapotranspiration and discharge) at a larger catchment scale, suggesting promising prospects for the application of CRNS compared to traditional SM sensors or RS datasets. The footprint of CRNS covers areas with a diameter of 300-600 m and a depth of 15-70 cm, which is much larger than the measurement volume of conventional SM measurement methods such as point-type SM sensors. RS data can provide spatially continuous SM information, but typically only for the top soil (0-5 cm). Therefore, the SM data from CRNS better represent the scale of model grids and thus can provide more accurate parameters for hydrological models to simulate hydrological processes such as infiltration, evapotranspiration, and runoff. We have also shown that the assimilation of SM from CRNS can improve SM estimates in the vicinity as well as at distant locations from CRNS stations. This indicates that even a low number of CRNS can provide useful information for DA. For instance, Patil et al. (2021) demonstrated that assimilating SM from four CRNS

improved SM simulation in a 655 km<sup>2</sup> catchment. This suggests that the use of CRNS instead of point sensors could reduce the number of measuring stations, which in turn may reduce installation and maintenance costs.

#### 4.5 Conclusions

SM measurements from 12 CRNS distributed over the Rur catchment (~2400 km<sup>2</sup>) were assimilated into TSMP with EnKF. This is the first application of the assimilation of observed SM data from CRNS into an integrated land surface-subsurface model for a real-world case. To this end, 128 ensemble members were generated by considering uncertain atmospheric forcings and subsurface hydraulic parameters. DA experiments were conducted for a wet year (2016) and a dry one (2018), with state-only updates and joint updates of state and parameters. SM observations from CRNS, evapotranspiration from eddy covariance stations, and discharge from an in situ station were used as validation data to assess the impact of SM assimilation on SM and flux simulation. EnKF assimilation of SM from CRNS improves SM estimation at measurement sites strongly in both dry and wet years, with up to 60% RMSE reductions. Joint state-parameter estimation results in a slightly better SM simulation than state estimation alone, with an RMSE reduction of more than 15% compared to state estimation alone. Jackknife experiments show limited improvement in SM characterization at independent verification locations, and the verification performance is affected by hydrological conditions, showing worse performance in dry episodes, indicating that the measurement network (~1 site per 200 km<sup>2</sup>) is not dense enough. SM assimilation improved ET and discharge characterization to a much lesser degree than SM, indicating limited sensitivities of ET and discharge towards SM.

The DA experiments show that improving the characterization of states (spatially and temporally) in the integrated surface-subsurface model TSMP by assimilating SM from a distributed CRNS network at the catchment scale is challenging but also promising. Compared to RS and traditional point SM measurements, CRNS could provide larger-scale in situ SM data with high temporal resolution and deeper penetration depth. In combination with DA, a limited number of sensors makes it feasible to continuously and stably determine SM dynamics from the field to the catchment scale. Better SM information is important for improving our understanding of the processes in terrestrial water cycles and reducing the large uncertainties of hydrological fluxes during modeling. Large-scale networks of CRNS already exist in the USA, Europe, Australia and India, but the density of sensors is still not sufficient to completely represent SM patterns at the continental scale. As sensors become more affordable, higher observational coverage will become possible, enabling the acquisition of long-term SM datasets

to monitor climate change and support predictions.

To improve the characterization of the states and parameters of integrated land surface-subsurface simulation at large scales, in addition to assimilating SM from a denser CRNS measurement network, future work should focus on multivariate assimilation (e.g., joint assimilation with vegetation related data) and the estimation of further soil hydraulic and vegetation parameters with integrated terrestrial system models at a higher spatial resolution.

## Appendix A

Table A.1 Comparison of CRNS soil moisture measurements and simulated soil moisture from open loop (OL) and data assimilation with joint state and parameter updating (DA) for the year 2016.

Site	BIAS (cm <sup>3</sup> /cm <sup>3</sup> )		MAE (cm <sup>3</sup> /cm <sup>3</sup> )		R		RMSE (cm <sup>3</sup> /cm <sup>3</sup> )		ubRMSD (cm <sup>3</sup> /cm <sup>3</sup> )	
	OL	DA	OL	DA	OL	DA	OL	DA	OL	DA
Merzenhausen	-0.036	0.001	0.041	0.022	0.870	0.902	0.048	0.029	0.032	0.028
Rollesbroich	-0.010	-0.009	0.036	0.026	0.854	0.957	0.044	0.032	0.043	0.030
Gevenich	-0.050	0.016	0.055	0.023	0.879	0.949	0.063	0.030	0.038	0.025
Ruraue	-0.074	-0.019	0.075	0.025	0.779	0.936	0.084	0.030	0.041	0.023
Wildenrath	-0.012	-0.003	0.024	0.015	0.856	0.902	0.030	0.021	0.028	0.021
Wüstebach	-0.083	-0.033	0.084	0.042	0.731	0.797	0.097	0.059	0.051	0.049
Heinsberg	-0.064	0.003	0.065	0.017	0.854	0.928	0.072	0.023	0.033	0.023
Kall	0.014	0.000	0.031	0.020	0.886	0.946	0.037	0.024	0.034	0.024
Selhausen	-0.120	-0.011	0.121	0.026	0.852	0.906	0.127	0.035	0.040	0.033
Schönseiffen	-0.050	0.008	0.057	0.022	0.843	0.930	0.067	0.027	0.045	0.025
Kleinau	0.002	-0.005	0.027	0.023	0.915	0.947	0.035	0.028	0.034	0.028
Aachen	-0.126	-0.009	0.126	0.022	0.883	0.920	0.130	0.030	0.034	0.028

Table A.2 Comparison of CRNS soil moisture measurements and simulated soil moisture from open loop (OL) and data assimilation with joint state and parameter updating (DA) for the year 2018.

Site	BIAS (cm <sup>3</sup> /cm <sup>3</sup> )		MAE (cm <sup>3</sup> /cm <sup>3</sup> )		R		RMSE (cm <sup>3</sup> /cm <sup>3</sup> )		ubRMSD (cm <sup>3</sup> /cm <sup>3</sup> )	
	OL	DA	OL	DA	OL	DA	OL	DA	OL	DA
Merzenhausen	0.025	0.000	0.042	0.023	0.786	0.937	0.056	0.030	0.050	0.030
Rollesbroich	0.045	0.019	0.060	0.039	0.794	0.874	0.072	0.048	0.055	0.044

Gevenich	0.003	0.004	0.043	0.026	0.785	0.929	0.053	0.034	0.053	0.034
Ruraue	-0.015	-0.015	0.056	0.024	0.711	0.957	0.066	0.030	0.064	0.026
Wildenrath	0.021	0.001	0.028	0.014	0.892	0.943	0.035	0.021	0.028	0.021
Wüstebach	-0.017	-0.014	0.049	0.024	0.702	0.886	0.062	0.033	0.060	0.030
Heinsberg	0.007	0.005	0.037	0.017	0.833	0.959	0.044	0.024	0.044	0.024
Kall	0.072	0.014	0.073	0.024	0.812	0.926	0.086	0.034	0.047	0.031
Selhausen	-0.031	0.006	0.051	0.025	0.780	0.939	0.064	0.031	0.056	0.031
Schönseiffen	0.015	0.007	0.036	0.023	0.908	0.950	0.045	0.028	0.043	0.028
Kleinau	0.061	0.005	0.062	0.025	0.875	0.953	0.076	0.032	0.046	0.032
Aachen	-0.113	-0.018	0.115	0.028	0.810	0.917	0.123	0.038	0.049	0.033

Table A.3 Comparison of CRNS soil moisture measurements and simulated soil moisture from open loop (OL) and jackknife simulations (DA) for the year 2016.

Site	BIAS (cm <sup>3</sup> /cm <sup>3</sup> )		MAE (cm <sup>3</sup> /cm <sup>3</sup> )		R		RMSE (cm <sup>3</sup> /cm <sup>3</sup> )		ubRMSD (cm <sup>3</sup> /cm <sup>3</sup> )	
	OL	DA	OL	DA	OL	DA	OL	DA	OL	DA
Merzenhausen	-0.036	-0.022	0.041	0.031	0.870	0.846	0.048	0.039	0.032	0.032
Rollebroich	-0.010	-0.017	0.036	0.030	0.854	0.912	0.044	0.038	0.043	0.034
Gevenich	-0.050	-0.022	0.055	0.037	0.879	0.889	0.063	0.046	0.038	0.040
Ruraue	-0.074	-0.039	0.075	0.042	0.779	0.936	0.084	0.047	0.041	0.026
Wildenrath	-0.012	-0.015	0.024	0.023	0.856	0.891	0.030	0.029	0.028	0.025
Wüstebach	-0.083	-0.076	0.084	0.076	0.731	0.859	0.097	0.085	0.051	0.039
Heinsberg	-0.064	0.033	0.065	0.038	0.854	0.839	0.072	0.047	0.033	0.033
Kall	0.014	0.004	0.031	0.028	0.886	0.874	0.037	0.036	0.034	0.035
Selhausen	-0.120	-0.015	0.121	0.039	0.852	0.832	0.127	0.050	0.040	0.047
Schönseiffen	-0.050	0.033	0.057	0.040	0.843	0.901	0.067	0.047	0.045	0.033
Kleinau	0.002	-0.005	0.027	0.026	0.915	0.909	0.035	0.034	0.034	0.033
Aachen	-0.126	-0.002	0.126	0.030	0.883	0.848	0.130	0.039	0.034	0.039

Table A.4 Comparison of CRNS soil moisture measurements and simulated soil moisture from open loop (OL) and jackknife simulations (DA) for the year 2018.

Site	BIAS (cm <sup>3</sup> /cm <sup>3</sup> )		MAE (cm <sup>3</sup> /cm <sup>3</sup> )		R		RMSE (cm <sup>3</sup> /cm <sup>3</sup> )		ubRMSD (cm <sup>3</sup> /cm <sup>3</sup> )	
	OL	DA	OL	DA	OL	DA	OL	DA	OL	DA
Merzenhausen	0.025	0.041	0.042	0.043	0.786	0.904	0.056	0.054	0.050	0.035
Rollebroich	0.045	0.019	0.060	0.040	0.794	0.873	0.072	0.048	0.055	0.044
Gevenich	0.003	0.019	0.043	0.052	0.785	0.773	0.053	0.063	0.053	0.060
Ruraue	-0.015	-0.041	0.056	0.045	0.711	0.919	0.066	0.055	0.064	0.036
Wildenrath	0.021	0.007	0.028	0.027	0.892	0.835	0.035	0.034	0.028	0.034
Wüstebach	-0.017	-0.044	0.049	0.048	0.702	0.853	0.062	0.058	0.060	0.038

Heinsberg	0.007	0.029	0.037	0.045	0.833	0.818	0.044	0.057	0.044	0.049
Kall	0.072	0.048	0.073	0.056	0.812	0.812	0.086	0.068	0.047	0.047
Selhausen	-0.031	0.038	0.051	0.043	0.780	0.888	0.064	0.056	0.056	0.041
Schönseiffen	0.015	0.017	0.036	0.038	0.908	0.868	0.045	0.048	0.043	0.045
Kleinau	0.061	0.027	0.062	0.051	0.875	0.770	0.076	0.068	0.046	0.062
Aachen	-0.113	-0.061	0.115	0.063	0.810	0.831	0.123	0.079	0.049	0.051

Table A.5 Root mean square error (RMSE) for open loop (OL), data assimilation with state updates (State), joint state-parameter updates (Joint), and jackknife simulations with joint state-parameter updates (Jackknife) for the assimilation periods of 2016 and 2018. The seasonal indicator was averaged over all sites with CRNS soil moisture observations.

Year	Season	Mean observed soil moisture (cm <sup>3</sup> /cm <sup>3</sup> )	RMSE (cm <sup>3</sup> /cm <sup>3</sup> )			
			OL	State	Joint	Jackknife
2016	Spring	0.32	0.078	0.043	0.032	0.045
	Summer	0.31	0.075	0.043	0.030	0.047
	Autumn	0.25	0.070	0.039	0.030	0.048
	Winter	0.36	0.087	0.052	0.035	0.045
2018	Spring	0.31	0.076	0.046	0.037	0.062
	Summer	0.18	0.067	0.041	0.033	0.073
	Autumn	0.22	0.066	0.047	0.029	0.046
	Winter	0.33	0.065	0.042	0.031	0.043

Table A.6 Comparison of daily measured evapotranspiration and simulated evapotranspiration from open loop (OL), data assimilation with state updates (State), and joint state-parameter updates (Joint) for two assimilation periods (2016 and 2018).

Site	Year	Simulation	BIAS (mm/day)	MAE (mm/day)	R	RMSE (mm/day)	ubRMSD (mm/day)
Rollesbroich	2016	OL	-0.334	0.464	0.896	0.612	0.513
		State	-0.318	0.444	0.902	0.582	0.487
		Joint	-0.264	0.428	0.897	0.559	0.493
	2018	OL	-0.420	0.547	0.900	0.747	0.618
		State	-0.426	0.553	0.900	0.753	0.621
		Joint	-0.469	0.593	0.880	0.816	0.668
Wüstebach	2016	OL	-0.590	0.707	0.821	0.914	0.698
		State	-0.553	0.675	0.826	0.876	0.679
		Joint	-0.528	0.663	0.824	0.860	0.679
	2018	OL	-0.377	0.619	0.773	0.849	0.760

		State	-0.377	0.610	0.769	0.844	0.755
		Joint	-0.345	0.601	0.775	0.825	0.750
		OL	-0.464	0.643	0.724	0.843	0.704
	2016	State	-0.278	0.585	0.719	0.768	0.716
		Joint	-0.239	0.568	0.723	0.753	0.714
	<b>Selhausen</b>	OL	-0.268	0.629	0.710	0.871	0.829
	2018	State	-0.183	0.612	0.713	0.842	0.822
		Joint	-0.127	0.585	0.731	0.811	0.801

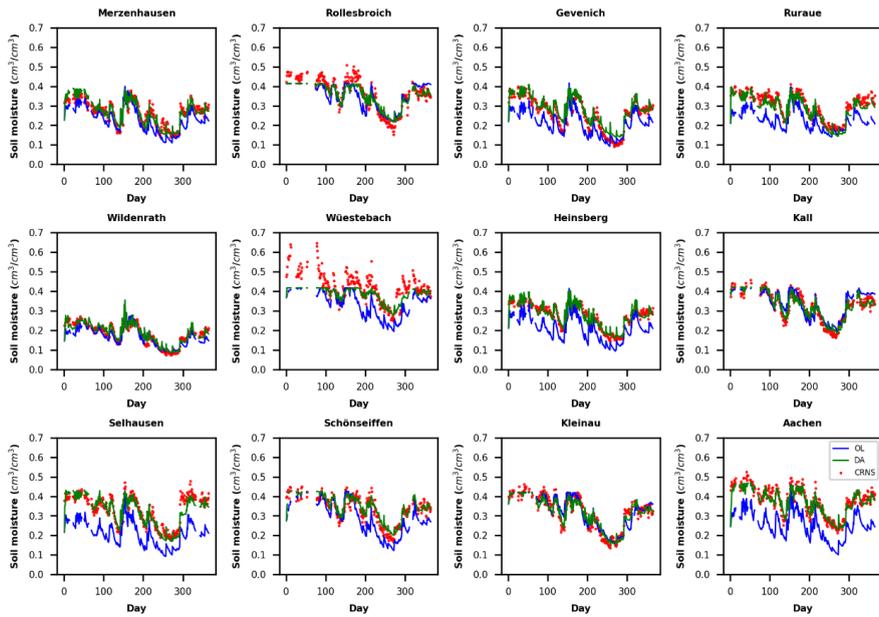


Figure A.1 Temporal evolution of mean simulated soil moisture from the open loop (OL, blue), joint state-parameter estimation (DA, green), together with observed soil moisture from CRNS (red), for the year 2016 at the CRNS sites. Simulated soil moisture was vertically weighted using the revised method.

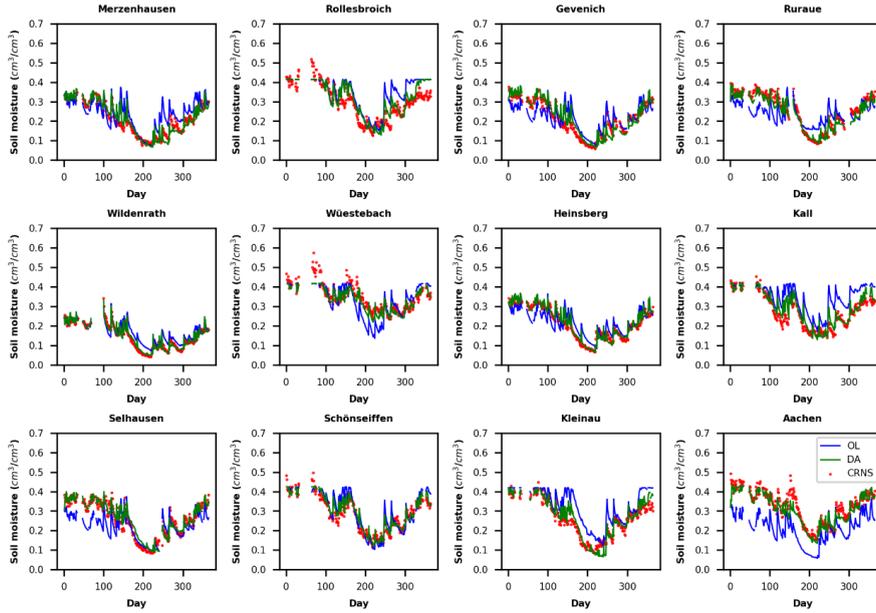


Figure A.2 Temporal evolution of mean simulated soil moisture from the open loop (OL, blue), joint state-parameter estimation (DA, green), together with observed soil moisture from CRNS (red), for the year 2018 at the CRNS sites. Simulated soil moisture was vertically weighted using the revised method.

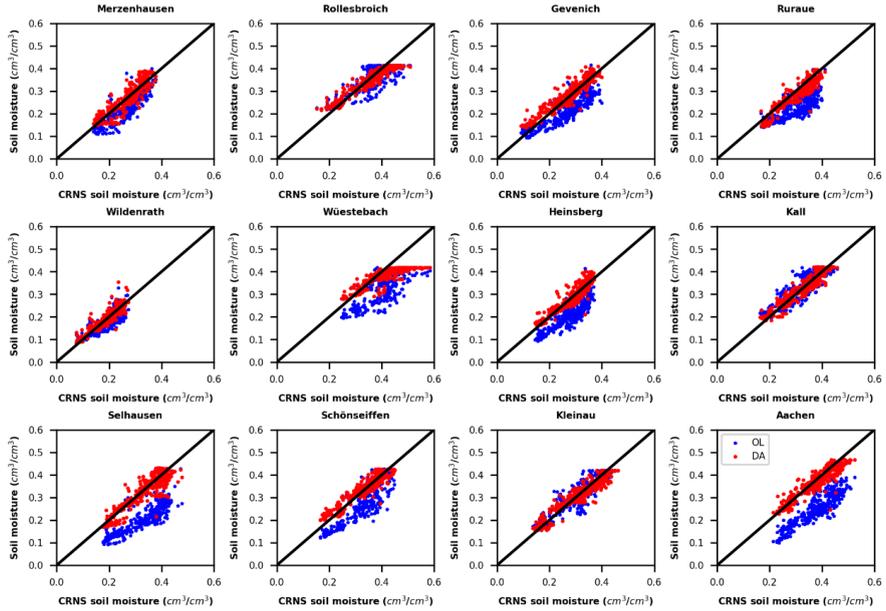


Figure A.3 Soil moisture scatter plots for CRNS observations versus ensemble mean soil moisture from the open loop run (OL, blue) and ensemble mean soil moisture from joint state-parameter estimation (DA, red) for 2016.

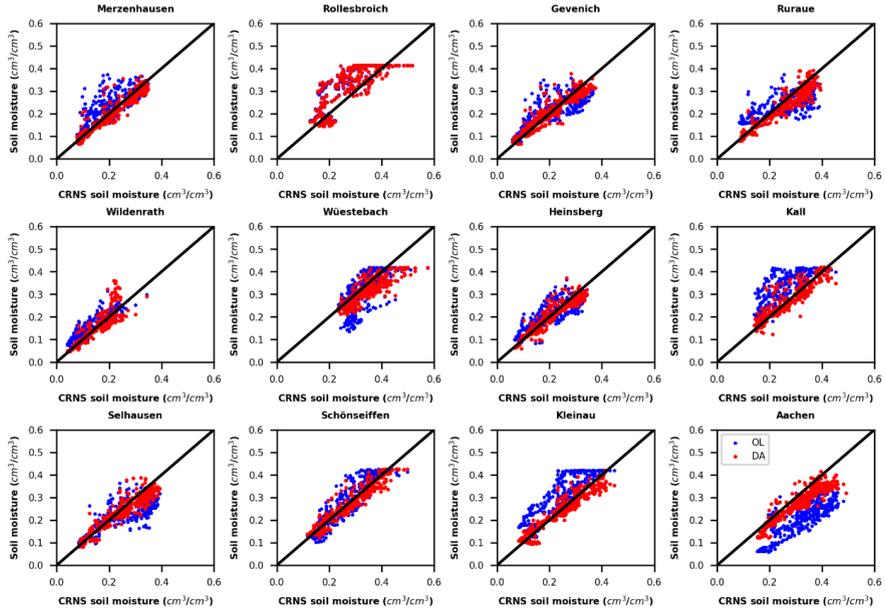


Figure A.4 Soil moisture scatter plots for CRNS observations versus ensemble mean soil moisture from the open loop run (OL, blue) and ensemble mean soil moisture from joint state-parameter estimation (DA, red) for 2018.

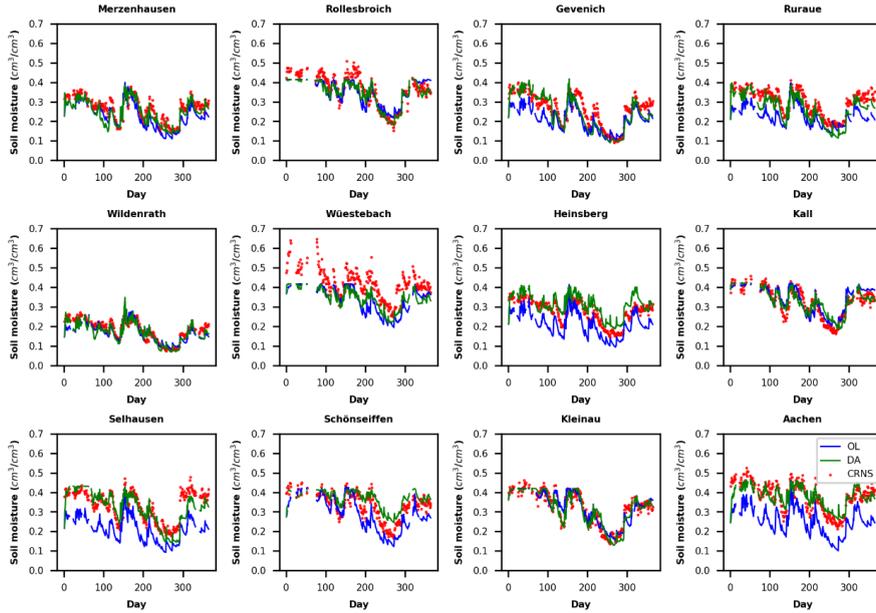


Figure A.5 Temporal evolution of mean simulated soil moisture from the open loop run (OL, blue), jackknife simulations (DA, green), together with the observed soil moisture from CRNS (red) for the year 2016 at the CRNS sites. Simulated soil moisture was vertically weighted using the revised method.

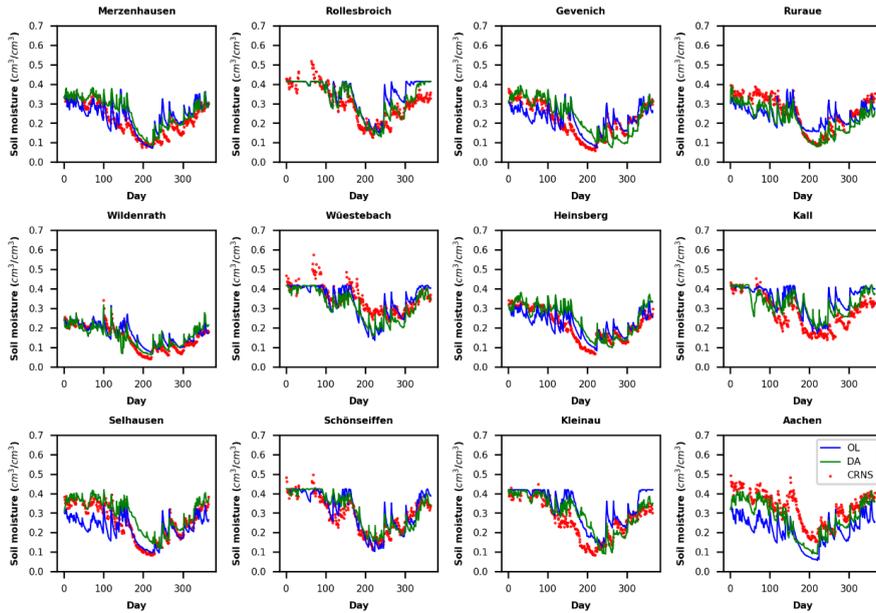


Figure A.6 Temporal evolution of mean simulated soil moisture from the open loop run (OL, blue), jackknife simulations (DA, green), together with the observed soil moisture from CRNS (red) for the year 2018 at the CRNS sites. Simulated soil moisture was vertically weighted using the revised method.

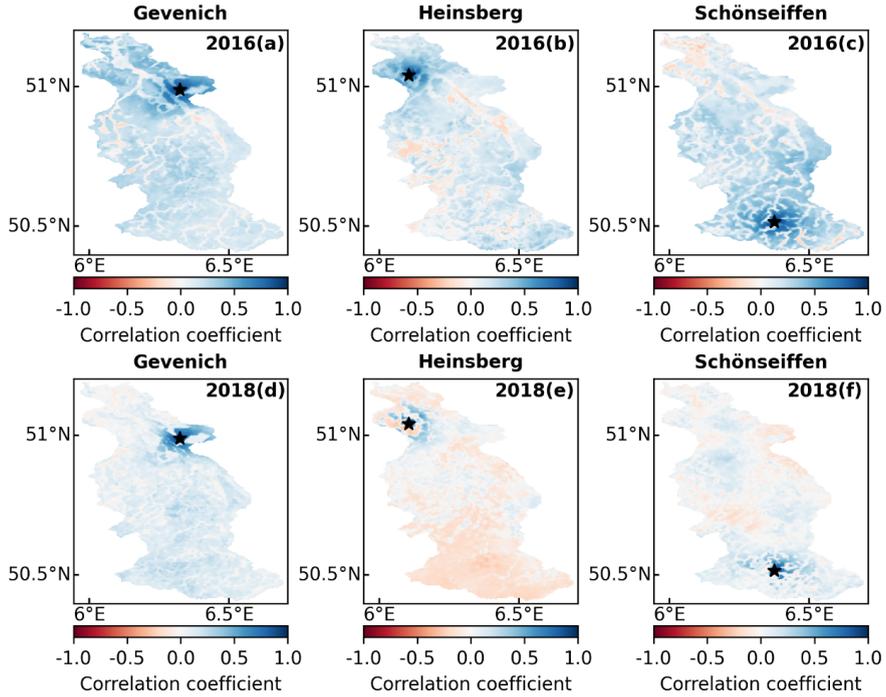


Figure A.7 Examples of the spatial correlations of soil moisture between CRNS sites and other grid cells over the Rur catchment, for the open loop run. Subplots a) and d) are from Gevenich on the 29<sup>th</sup> of June in 2016 and 2018, b) and e) are from Heinsberg on the 2<sup>nd</sup> of August in 2016 and 2018, and c) and f) are from Schöneiffen on the 28<sup>th</sup> of July in 2016 and 2018. The black asterisk is the location of the CRNS sites.

## **Chapter 5: A new approach for joint assimilation of cosmic-ray neutron soil moisture and groundwater level data into an integrated terrestrial model**

\*adapted from: Li, F., Bogena, H. R., Keller, J., Bayat, B., Raj R., and Hendricks Franssen, H.-J.: A new approach for joint assimilation of cosmic-ray neutron soil moisture and groundwater level data into an integrated terrestrial model, submitted to Journal of Hydrology.

### **5.1 Introduction**

Groundwater level (GWL) and root zone soil moisture (RZSM) are two crucial variables in hydrological and land surface modeling (Zhang et al., 2016). Shallow groundwater can have a significant impact on SM through water exchange between the aquifer and the unsaturated zone and also impact ET (Chen and Hu, 2004). Therefore, accurate information on the spatio-temporal variability of GWL and RZSM is important for a detailed understanding of hydrological processes in terrestrial systems and a better estimate of the water and energy fluxes from the subsurface to the land surface and the atmosphere (Vereecken et al., 2022). Nevertheless, many land surface models do not explicitly consider the impact of groundwater dynamics on land surface processes (Kollet and Maxwell, 2008). Integrated models like the TSMP (Shrestha et al., 2014), which simulate the groundwater-soil-vegetation-atmosphere system, are particularly suitable for exploring the impacts of spatio-temporal variations of GWL on terrestrial ecosystems. The capabilities of this model for simulating atmosphere-land-subsurface interactions and surface water-groundwater dynamics have been demonstrated in many studies (Shrestha et al., 2015; Keune et al., 2016; Shrestha et al., 2018; Furusho-Percot et al., 2019; Zhao et al., 2021; Naz et al., 2023).

However, such highly complex coupled models often involve a large number of parameters, which often contribute to model uncertainties and affect prediction accuracy. Particularly in groundwater systems, the simplifications and assumptions that have to be used in the model parameterization process introduce significant uncertainties in the model due to the strong spatial heterogeneity of the hydraulic parameters and the limited availability of direct measurements (de Marsily, 1986). Furthermore, modeling uncertainties may emerge from various sources, including model forcings, model structural components, and initial conditions (Freeze, 1975; D. Baatz et al., 2017). To mitigate uncertainties in model states and parameters, DA can be employed, leveraging measurements to refine and correct model simulations (De

Lannoy et al., 2014). The EnKF (Evensen, 1994; Evensen, 2003), a sequential filtering algorithm, is one of the most widely used DA techniques. It has been shown to successfully handle high-dimensional nonlinear hydrological and land surface simulation problems (Camporese et al., 2009a; Schöniger et al., 2012). Previous research has proven the effectiveness of EnKF in improving SM forecasts for stand-alone land surface models (De Lannoy et al., 2007; Yin et al., 2015; De Lannoy and Reichle, 2016; Naz et al., 2019) or GWL estimation for groundwater models (Chen and Zhang, 2006; Hendricks Franssen and Kinzelbach, 2008).

SM can be assessed or measured across different scales using a variety of techniques, and various SM data types have been used for assimilation into land surface models, including in-situ measurements (Fu et al., 2023; Strebel et al., 2024) and remote sensing (RS) data products, such as Soil Moisture Active Passive (SMAP) (Reichle et al., 2017; He et al., 2021) and Soil Moisture Ocean Salinity (SMOS) (De Lannoy and Reichle, 2016; Lievens et al., 2016). Despite important advances in these previous studies, there is still a lack of high-quality SM data that can be assimilated into models as in-situ SM observations are often very small-scale and therefore not representative of larger areas (Nicolai-Shaw et al., 2015). In addition, in-situ data often have temporal gaps, making it difficult to accurately capture the spatiotemporal SM characteristics, which affects DA and the quality of model predictions. Although RS provides global SM information, either the spatial or temporal resolution is often coarse, and the spatial mismatch between coarse-resolution RS data and high-resolution models can lead to significant data assimilation challenges (Sahoo et al., 2013; Naz et al., 2019). In addition, only information for the upper few cm is provided, and the measurement is affected by larger bias and uncertainty than in situ measurements. Therefore, this study explores an alternative source of SM information that can more accurately measure SM. CRNS (Zreda et al., 2008), a new technology for monitoring SM that bridges the gap between in-situ measurements and RS products, has gained broader acceptance and usage (Bogena et al., 2022). CRNS can provide non-invasive and real-time SM at the field scale, with a footprint of up to 18 hectares and a penetration depth of up to 80 cm (Zreda et al., 2008; Bogena et al., 2015; Köhli et al., 2015). It provides SM data at deeper soil depths with less error than RS products while covering a larger footprint than in-situ measurements. Therefore, it is well suited for DA in land surface models to improve SM estimation (Shuttleworth et al., 2013; Rosolem et al., 2014; Han et al., 2015; Han et al., 2016; Roland Baatz et al., 2017; Patil et al., 2021).

Groundwater table depth is commonly measured in-situ in observation wells. However, studies on the assimilation of groundwater table depth (or hydraulic head) were mainly

conducted using synthetic data. Chen and Zhang (2006) assimilated the synthetic pressure head data into transient flow models and demonstrated the ability of EnKF to estimate hydraulic conductivity fields. Hendricks Franssen and Kinzelbach (2008) conducted synthetic studies using a 2D transient groundwater flow model to explore effective ways to reduce the filter inbreeding problem when using EnKF for parameter estimation. Tong et al. (2011) employed LEnKF in a 2D synthetic transient groundwater flow model and found that EnKF with localization can solve the problem of filter divergence and capture the heterogeneous hydraulic conductivity field reliably with higher efficiency and a smaller ensemble size than EnKF. Panzeri et al. (2013, 2014) proposed EnKF variants (direct solution of nonlocal (integrodifferential) stochastic ensemble moment equations or direct computation of stochastic ensemble moment equations governing the space-time evolution of ensemble means and covariances of hydraulic heads and fluxes) for 2D transient groundwater flow models to address issues arising in groundwater DA, including heavy computation and filter inbreeding, and the validity of these approaches was validated in a field application (Panzeri et al., 2015). The above groundwater assimilation studies show that EnKF can handle high-dimensional nonlinear relationships related to groundwater systems and is suitable for groundwater assimilation. However, even when using synthetic groundwater data, EnKF requires tuning including for example localization to address filter divergence and inbreeding issues. For more complex real-world cases, the assimilation of groundwater observations may be more challenging, and the EnKF algorithm needs to be adapted to address specific challenges to improve the accuracy of the data assimilation and to achieve optimal simulation results.

Most DA studies focused only on one compartment of the terrestrial system, e.g., the land surface, and assimilated only one variable into the models. However, in natural systems, the components of the water cycle, such as groundwater and SM, are closely linked. Therefore, efforts should be made to improve the simulation of all interrelated variables in the model by means of assimilating multiple variables. Consequently, a number of recent studies have initiated the implementation of multivariate DA in coupled models, such as the joint assimilation of groundwater and SM. The use of multivariate DA techniques in integrated models can potentially better exploit the value of different measurements, enhance terrestrial model predictions, and improve the estimation of model parameters (Zhang et al., 2016). Camporese et al. (2009b) assimilated synthetic pressure head and SM data into the coupled surface water-groundwater model CATHY (CATchment HYdrology) for a small catchment (4.64 km<sup>2</sup>), and the results showed that assimilation of either pressure head or SM can improve the characterization of subsurface states in the vicinity of the measurement locations. Botto et al. (2018) employed the model CATHY to investigate the assimilation of multi-source data

(pressure head, SM, and subsurface outflow) from an artificial hillslope. Their results demonstrated the efficiency of EnKF to correct states and parameters under conditions of strong nonlinearity but also showed that multisource DA may lead to the degradation of model predictions for other variables. Shi et al. (2014, 2015) employed EnKF to assimilate multivariate hydrological observations, either from synthetic experiments or a real-world case (a small catchment of 0.08 km<sup>2</sup>), into a coupled physically-based land surface hydrologic model (Flux-PIHM) to identify the observations critical for parameter estimation. The integrated hydrological model MIKE-SHE was also applied to assimilate groundwater head and SM, and the importance of localization (distance and variable localization) was demonstrated (Zhang et al., 2016). However, the DA experiments in Zhang et al. (2016) were conducted either using synthetic data or the real observations were processed (i.e., the average difference between observations and model simulations was subtracted from the original data). Furthermore, in MIKE-SHE, the unsaturated flow is still only calculated in one dimension.

The coupled model TSMP, in combination with the Parallel Data Assimilation Framework (PDAF) (Nerger et al., 2005), has also been used to assimilate synthetic or real SM or GWL data at different scales (e.g., hillslope and catchment scale). Kurtz et al. (2016) showed the promise of this DA framework for simulating and estimating uncertainties in predicted states and fluxes of the terrestrial system by using synthetic SM observations to jointly update soil water content and saturated hydraulic conductivity in TSMP-PDAF using EnKF. Since then, with TSMP-PDAF, Gebler et al. (2019) assimilated in-situ SM at the hillslope scale and Li et al. (2023a) assimilated CRNS soil moisture at the larger catchment scale, both of which significantly improved SM simulations. Brandhorst and Neuweiler (2023) investigated the influence of updating different soil hydraulic parameters on the accuracy of SM estimation in a three-dimensional heterogeneous hillslope model using TSMP-PDAF, and the synthetic studies showed that the best estimates were obtained when porosity and van Genuchten parameters, as well as saturated conductivity coefficients (optionally), were jointly updated. In terms of GWL assimilation, Li et al. (2023b) attempted to assimilate real GWL data at the catchment scale with LEnKF in TSMP-PDAF and showed a considerable improvement in GWL estimation. However, the established works on the assimilation of both SM and GWL into TSMP were only conducted in synthetic experiments (Zhang et al., 2018; Hung et al., 2022). Zhang et al. (2018) concluded in a synthetic experiment with only four soil columns that the assimilation of both pressure and SM gave the best estimate of RZSM. Hung et al. (2022) assimilated GWL and SM observations in a much more complex synthetic study with a fully 3D heterogeneous subsurface mimicking southwestern Germany. However, they found that the weakly coupled DA strategy, in which only saturated subsurface states were updated, performed better than the fully coupled

DA suggested by Zhang et al. (2018). These examples show that there is still a need to investigate the benefits of jointly assimilating real GWL and SM observations into integrated land surface-subsurface models at the catchment scale.

Considering the additional challenges that may be encountered in performing the joint assimilation of GWL and SM into a coupled land surface-subsurface model (i.e., TSMP) in a real-world case (Rur catchment, Germany), a new multivariate DA approach is proposed. The main objective of the current study is to investigate: (i) the potential of joint assimilation of GWL and SM data in a coupled land surface-subsurface model with the newly proposed multivariate DA approach for the Rur catchment; (ii) how the DA performance may differ across various multivariate DA approaches; (iii) and the advantages of the new multivariate DA approach on the prediction of GWL, SM, and ET compared to univariate DA strategies. To our knowledge, this is the first time that both real GWL and CRNS soil moisture observations from in-situ networks are assimilated into the integrated terrestrial model TSMP at the catchment scale.

## **5.2 Materials and methods**

### **5.2.1 Study area**

The Rur catchment, with an area of 2354 km<sup>2</sup>, is situated mostly in western Germany, with smaller parts in Belgium and the Netherlands. The Rur river flows from south to north as the altitude gradually decreases from 690 m in the south to 15 m a.s.l. in the north (Fig 5.1). The main land uses in the Rur catchment are arable agriculture in the northern lowlands (mainly maize and wheat) and grasslands, and coniferous and deciduous forests in the southern mountainous regions (Waldhoff and Lussem, 2015; Roland Baatz et al., 2017). Other important land uses include open-cast lignite quarries and urban areas (Shukla et al., 2023). Corresponding to the topography from north to south, the mean annual air temperature decreases from 10 to 7 °C, the annual precipitation increases from 650 to 1300 mm, and the annual potential ET decreases from 850 to 450 mm (Montzka et al., 2008; Bogena et al., 2018). The permeability of the upper unconfined aquifer and groundwater recharge in the southern low-mountain range are relatively low due to the prevailing consolidated bedrock, while they are relatively high in the northern lowland region due to the presence of unconsolidated bedrock (Bogena et al., 2018).

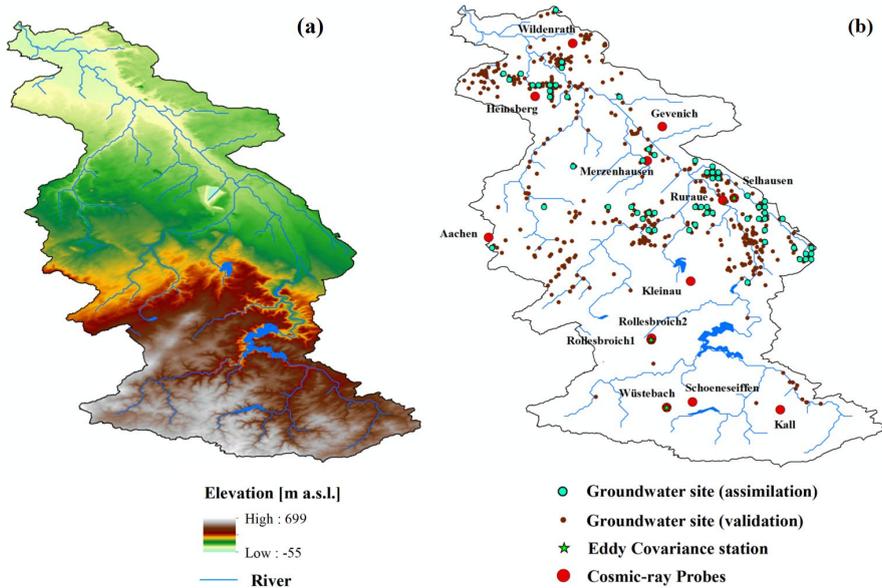


Figure 5.1 Topography of the Rur catchment (a) and locations of the hydrological stations (b), including groundwater wells, cosmic-ray neutron sensors, and eddy covariance stations.

### 5.2.2 Terrestrial System Modeling Platform (TSMP)

The TSMP model is a fully integrated biogeophysical terrestrial system model designed to simulate the transport processes of water and energy among the atmosphere, land surface, and subsurface (Shrestha et al., 2014). TSMP comprises an atmospheric model (Consortium for Small Scale Modelling, COSMO) (Baldauf et al., 2011), a land surface model (Community Land Model (CLM, version 3.5)) (Oleson et al., 2004; Oleson et al., 2008) from the National Center for Atmospheric Research, and the 3D variably saturated groundwater flow model ParFlow (Kollet and Maxwell, 2006) for the subsurface. These three models are two-way coupled by the Ocean Atmosphere Sea Ice Soil - Model Coupling Toolkit (OASIS-MCT, version 3) (Valcke, 2013). In TSMP, the coupler OASIS-MCT is employed to exchange variables and fluxes between the sub-models. TSMP offers the flexibility to operate diverse combinations of component models. It can be set up in a fully coupled configuration (COSMO-CLM-ParFlow), partly coupled configurations (COSMO-CLM or CLM-ParFlow), or stand-alone model components (COSMO, CLM, or ParFlow). In this work, only the land surface model CLM 3.5 and the subsurface model ParFlow were activated.

The land surface model CLM simulates the biophysical processes within the terrestrial system, including energy and water exchange between the land and atmosphere, snow accumulation and melting, water and energy transport in the soil, and stomatal physiology and photosynthesis (Oleson et al., 2004; Oleson et al., 2008). In CLM, the representation of spatial land surface heterogeneity utilizes a nested subgrid hierarchy (Oleson et al., 2008). Each grid cell is subdivided into various land units, including glacier, lake, wetland, urban, and vegetated areas (Oleson et al., 2004; Oleson et al., 2008). Notably, each land unit may contain a different number of snow/soil columns. Within each column, multiple PFTs with different plant physiological parameters can be defined, which could capture the biogeophysical and biogeochemical differences between different vegetation types (Oleson et al., 2004; Oleson et al., 2008). The main purpose of CLM within TSMP is to account for water and heat fluxes to estimate ET from both ground and vegetation, as well as snow accumulation and melting.

In TSMP, ParFlow replaced CLM to simulate the soil hydrological processes and model the surface runoff and groundwater flow (Ashby and Falgout, 1996; Jones and Woodward, 2001; Kollet and Maxwell, 2006; Maxwell, 2013). ParFlow evolved from the parallel 3D variably saturated subsurface flow code ParFlow with a 2D overland flow simulator (Ashby and Falgout, 1996; Kollet and Maxwell, 2006). ParFlow uses the Newton-Krylov nonlinear solver (Jones and Woodward, 2001) to solve the coupled partial differential equations for groundwater flow and surface water flow, including the kinematic wave equation (Lighthill and Whitham, 1955) for overland flow and the 3D Richards equation (Richards, 1931) for groundwater flow in the unsaturated and saturated zones. In addition, ParFlow was developed specifically for parallel computing systems, so it can effectively solve highly heterogeneous large-scale problems at high resolution.

Coupler OASIS-MCT controls the exchange of state variables and fluxes between different model components, utilizing temporal integration/averaging and spatial interpolation operations to ensure that the spatial and temporal scales of fluxes exchanged by different components are consistent (Valcke, 2013; Shrestha et al., 2014). In this work, only flux exchanges between the land surface and the subsurface were considered. The upper boundary condition for ParFlow is supplied by CLM and is represented as net infiltration or exfiltration, taking into account precipitation, interception, total evaporation, and total transpiration (Kurtz et al., 2016; Zhang et al., 2018). In turn, ParFlow provides CLM with calculated pressure and saturation values for the upper ten subsurface layers. More detailed information regarding the implementation of the coupler in CLM-ParFlow can be found in Kollet and Maxwell (2008).

## **5.2.3 Model input data and measurements**

### **5.2.3.1 Atmospheric forcing**

TSMP was forced by the high-resolution atmospheric reanalysis dataset COSMO-REA6 (0.055° [6 km], hourly) (Bollmeyer et al., 2015; Wahl et al., 2017) provided by the German Meteorological Service (Deutscher Wetterdienst; DWD). The numerical weather prediction (NWP) model COSMO (Baldauf et al., 2011) was used to generate this reanalysis dataset. Forcing data for TSMP include precipitation, air temperature, air pressure, wind velocity, specific humidity, incoming shortwave radiation, and incoming longwave radiation.

### **5.2.3.2 Land surface and subsurface data**

The digital elevation model (DEM) for the Rur catchment was obtained from SRTM 90m version 4 (Jarvis et al., 2008), as shown in Fig 5.1. Land cover information was derived from Sentinel-2 satellite data (Drusch et al., 2012) and then aggregated and transferred to the CLM-prescribed PFTs using the approach developed by Montzka et al. (2021). Monthly LAI values for each year (2016-2018) were computed for individual PFTs using LAI maps retrieved from Sentinel-2 satellite observations (Drusch et al., 2012). A retrieval algorithm was employed using the Sentinel-2 Level 2 Prototype Processor (SL2P: Weiss and Baret, 2020) embedded in the freely available Sentinel Application Platform (SNAP). SL2P implements backpropagation artificial neural networks (ANN) trained with a global range of inputs (LAI and other biophysical characteristics) and simulated canopy reflectance from the optical radiative transfer model PROSAIL (Verhoef and Bach, 2007; Jacquemoud et al., 2009; Bayat et al., 2018). To retrieve the LAI at each pixel of Sentinel-2, the trained ANN was provided with the top of canopy reflectance from Sentinel-2, along with the sun and view geometry obtained from the satellite orbit characteristics and swath.

The sand and clay contents of the soil (see Fig 5.2) were taken from the high-resolution regional soil map BK50 at a scale of 1:50,000 (Geologischer Dienst NRW, 2009), and the bulk density was obtained from the European Soil Database (ESDB) (Pano, 2006). The soil textures and bulk densities were used to calculate soil hydraulic parameters using the Rosetta pedotransfer functions (Schaap et al., 2001; Zhang and Schaap, 2017). The digital hydrogeological map HK100 of North Rhine-Westphalia at a scale of 1:100,000 (Geologischer Dienst NRW, 2011) was used to parameterize the hydraulic conductivity ( $K_s$ ) of the bedrock (Fig 5.2).

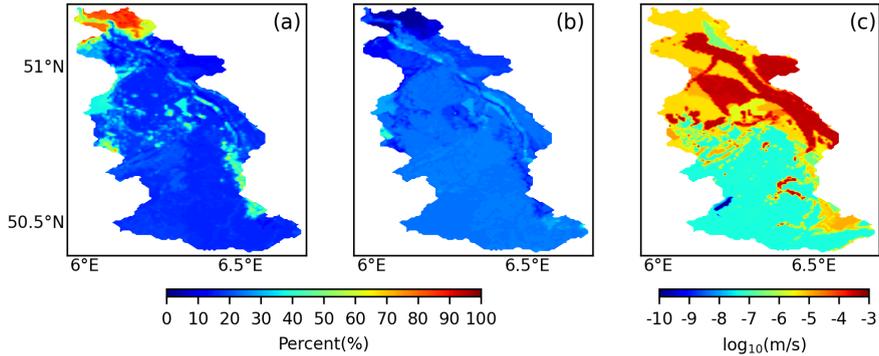


Figure 5.2 Sand (a) and clay (b) content, hydraulic conductivity (c) of the bedrock for the Rur catchment.

### 5.2.3.3 Soil Moisture, Groundwater Level and Evapotranspiration Measurements

The SM data assimilated in this study were obtained from a network of CRNS stations in the Rur Hydrological Observatory of the TERrestrial Environmental Observatories (TERENO) initiative (Bogena et al., 2018). For this study, we used readily processed CRNS soil moisture data from the COSMOS-Europe initiative (Bogena et al., 2022) and selected 13 CRNS stations across the Rur catchment (see Table 5.1). Due to their close proximity, the CRNS stations Rollesbroich1 and Rollesbroich2 were treated as a single site in this work, and the average SM was used. As a result, 12 CRNS stations were used for DA.

In addition, GWL observations from the monitoring network Geoportal NRW ([www.geoportal.nrw](http://www.geoportal.nrw); last accessed on August 26, 2024) were used in this study for DA and independent validation. Due to the intricate nature of the deep confined aquifer, which exhibits minimal interaction with RZSM, our assimilation approach focuses solely on the upper unconfined aquifer. Our screening process specifically targeted groundwater sites in the upper aquifer, with measurement depths ranging from 0 to 20 meters and a minimum monthly data frequency. A total of 616 sites were identified between 2016 and 2018 (Fig 5.1). Due to the 500 m coarse horizontal resolution and the concentration of groundwater monitoring sites along the river, numerous sites are situated within the same model grid cell or within river grid cells. Therefore, it is necessary to screen the data before proceeding with assimilation. In cases where multiple sites are situated within the same grid cell, the groundwater site with the median GWL value was selected for DA. Moreover, grid cells containing sites within or directly adjacent to

the river were excluded from the DA experiments. This decision was made because these grid cells are permanently saturated in these areas of the model domain due to topography-related flow concentration, which led to strong discrepancies with groundwater observations. Ultimately, only 78 sites were retained for DA of GWL, while the other 465 sites were used for GWL independent validation.

The ET simulations were validated using flux data from three EC stations (Rollesbroich, Wüstebach, and Selhausen) obtained from TERENO (<https://www.tereno.net/>; last accessed on August 26, 2023). EC measurements consist of a sonic anemometer (CSAT3, Campbell Scientific Inc., Logan, USA) to measure the 3D wind components, an open-path gas analyzer (Li7500, LI-COR Inc., Lincoln, USA) to determine the H<sub>2</sub>O and CO<sub>2</sub> concentrations in the air, and an air temperature and humidity sensor (HMP45C, Vaisala Inc., Helsinki, Finland). The post-processing of flux data, including EC conversions and uncertainty estimation, followed the methodology in Mauder et al. (2013). The missing data in daily EC measurements were gap-filled by grass reference ET computed via the FAO Penman-Monteith approach (Allan et al., 1998). No corrections were made to address the energy balance non-closure in the EC data. More details on EC measurement and processing can be found in Bogena et al. (2018).

Table 5.1 CRNS sites used in this study, including key site characteristics. The mean air temperature and mean annual precipitation were obtained from the ECMWF climate reanalysis data product ERA5-Land (Muñoz Sabater, 2021), resulting in identical precipitation and air temperature values for some of the sites.

Name	Latitude (degr)	Longitude (degr)	Altitude (m)	Mean annual precipitation (mm y <sup>-1</sup> )	Mean air temperature (°C)	Land use
Merzenhausen	50.93	6.30	91	718	10.3	crop
Rollesbroich1	50.62	6.30	515	1018	7	grassland
Rollesbroich2	50.62	6.31	506	1018	7	grassland
Gevenich	50.99	6.32	107	718	10.3	crop
Ruraue	50.86	6.43	100	718	10.3	grassland
Wildenrath	51.13	6.17	72	722	10.3	needleleaf
Wüstebach	50.51	6.33	605	1401	7	spruce
Heinsberg	51.04	6.10	58	722	10.3	crop
Kall	50.50	6.53	505	857	8	grassland
Selhausen	50.87	6.45	101	718	10.3	crop
Schönesseiffen	50.52	6.38	611	870	7	grassland
Kleinhau	50.72	6.37	374	614	9	grassland
Aachen	50.80	6.03	232	865	10.3	crop

#### 5.2.4 Data assimilation: localized EnKF

DA consists of two steps: the forecast step and the analysis step (McLaughlin, 2002). During the forecast step, the state estimation relies solely on past information. In the subsequent analysis step, the states and/or parameters are estimated using the information from current measurements and the prior forecast, and the probability density of the states is propagated forward (McLaughlin, 2002). In this fashion, the forecast step (model prediction) and analysis step (filter application) are sequentially alternated during DA.

The augmented state vector approach described in Hendricks Franssen et al. (2011) is used for multivariate DA and also for joint state and parameter updating. The model states for the fully coupled DA applied in this work are piezometric heads ( $h$ ) and soil moisture ( $\theta$ ). Also, the parameter uncertainty of hydraulic conductivities ( $K_s$ ) can be considered. The state-parameter vector in the EnKF is formulated as follows:

$$\boldsymbol{\Psi} = \begin{pmatrix} \mathbf{x} \\ \log_{10}(K_s) \end{pmatrix} = \begin{pmatrix} h \\ \theta \\ \log_{10}(K_s) \end{pmatrix} \quad (5.1)$$

To update states and parameters, measurements of piezometric heads (on the basis of groundwater table depths) and SM are used, which are combined in the measurement vector.

The updating equation for  $\boldsymbol{\psi}$  is calculated for each ensemble member  $j$  ( $j=1, \dots, N$ ) (Evensen, 2003). In this work, ensemble members are generated by taking into account uncertainties in the atmospheric forcings and parameters (including  $K_s$  and porosity). The updating equation for an ensemble member is given by:

$$\boldsymbol{\psi}_j^a = \boldsymbol{\psi}_j^f + \alpha \mathbf{K} (\hat{\mathbf{y}}_j - \mathbf{H} \boldsymbol{\psi}_j^f) \quad (5.2)$$

where  $\boldsymbol{\psi}_j^f$  and  $\boldsymbol{\psi}_j^a$  are the forecasted and updated state-parameter vectors of the  $j^{\text{th}}$  realization,  $\hat{\mathbf{y}}_j$  is the perturbed measurement vector of piezometric heads (and SM), and  $\mathbf{K}$  is the Kalman gain.  $\alpha$  is a damping factor for the state (piezometric heads) or parameter update that takes values between 0 and 1. The damping factor ( $\alpha$ ) is used to reduce the modification of the

forecast with the Kalman gain and limit the intensity of the perturbation of the state (pressure head) or parameter ( $\log_{10} K_s$ ) (Gebler et al., 2019; Hung et al., 2022). This is needed because filter inbreeding may occur if the EnKF analysis is applied repeatedly in combination with a limited ensemble size, and this is related to the underestimation of the ensemble covariance (Hendricks Franssen and Kinzelbach, 2008).

The Kalman gain is given by:

$$\mathbf{K} = \mathbf{P}\mathbf{H}^T(\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R})^{-1} \quad (5.3)$$

The  $\mathbf{H}$  is the observation operator that connects the observation and state vectors.  $\mathbf{P}$  is the model covariance matrix of model states and uncertain parameters, and  $\mathbf{R}$  is the measurement error covariance matrix. The accuracy of the filter is determined by the model covariance matrix  $\mathbf{P}$ , calculated from the ensemble (Turner et al., 2008).

The estimation of the covariances with a limited ensemble size is affected by spurious correlations between spatially distant grid cells. Therefore, in this study, the LEnKF proposed by Houtekamer and Mitchell (1998) is adopted to use distance-dependent localization to limit the influence of observation points to a certain radius only (Hamill et al., 2001).  $\mathbf{P}\mathbf{H}^T$  is replaced by  $\boldsymbol{\rho}\mathbf{P}\mathbf{H}^T$  in Eq 5.3, and  $\boldsymbol{\rho}\mathbf{P}\mathbf{H}^T$  is the Schur product of the correlation matrix  $\boldsymbol{\rho}$  and covariance matrix  $\mathbf{P}\mathbf{H}^T$ , where  $\boldsymbol{\rho}$  is a correlation matrix containing correlations between the grid cells and calculated by a fifth-order piecewise function, as given by Gaspari and Cohn (1999).

The correlation  $\omega$  between a grid point and an observation, i.e., an element in  $\boldsymbol{\rho}$ , can be approximated as:

$$\omega(l, e) = \begin{cases} 1 - \frac{1}{4}\left(\frac{e}{l}\right)^5 + \frac{1}{2}\left(\frac{e}{l}\right)^4 + \frac{5}{8}\left(\frac{e}{l}\right)^3 - \frac{5}{3}, & 0 \leq e \leq l \\ \frac{1}{12}\left(\frac{e}{l}\right)^5 - \frac{1}{2}\left(\frac{e}{l}\right)^4 + \frac{5}{8}\left(\frac{e}{l}\right)^3 + \frac{5}{3}\left(\frac{e}{l}\right)^2 - 5\left(\frac{e}{l}\right) + 4 - \frac{2}{3}\left(\frac{e}{l}\right)^{-1}, & l < e \leq 2l \\ 0, & e > 2l \end{cases} \quad (5.4)$$

where  $l$  is the defined localization radius and  $e$  is the Euclidean distance between an analyzed grid point and an observation location. The correlations  $\omega$  are distance-dependent and vary between 1 at observation locations and 0 at distances greater than twice the influence

radius  $l$ .

In this work, the assimilated SM data were from the CRNS. Since the penetration depth of CRNS depends on SM, before assimilation, the measurement depth of CRNS observations was calculated using the method suggested by Schrön et al. (2017). Subsequently, PDAF specifies CRNS soil moisture observations for the soil layers until the measurement depth (see Fig 5.3). The observed data are used to update the simulated SM of those soil layers within the measurement depth. After the final assimilation, the average weighted SM from the simulations is compared with the SM data from CRNS to evaluate model performance. A more comprehensive description of the weighted SM calculations can be found in Schrön et al. (2017). For SM, the localization radii range (100 km) is larger than the domain size. Therefore, even though localization is applied, SM is updated throughout the whole domain. However, the localized correlations still lead to smaller updates further away from the observations, according to the fifth-order polynomial from Eq 5.4. The SM was updated on a daily basis, and the SM measurement error was  $0.03 \text{ cm}^3/\text{cm}^3$ .

For groundwater assimilation in TSMP-PDAF (see Fig 5.3), GWL data need to be converted to pressure heads in the layers of the saturated zone, assuming a hydrostatic pressure distribution (Zhang et al., 2018). The delineation between the saturated and unsaturated zones is determined by the deepest groundwater table depth across all ensemble members, as suggested in Zhang et al. (2018). For groundwater assimilation, a horizontal localization radius of 5 km determined by the spatial correlations of GWL was used. The pressure head was updated weekly, considering the slow dynamics of GWL. The groundwater measurement error was set to 0.05 m.

In the fully coupled DA approach with TSMP, all subsurface states are updated when assimilating GWL and/or SM (Zhang et al., 2018). Follow-up work with a large-scale CLM-ParFlow model including realistic 3D subsurface heterogeneity showed that when GWL only updated the state of the saturated zone, it outperformed fully coupled DA (Hung et al., 2022).

In this study, a new strategy called the weakly coupled approach was introduced in the multivariate DA of GWL and SM data to make the updates more stable. In the new approach, only the states corresponding to the saturated grid cells are updated with piezometric head measurements, whereas the states corresponding to the unsaturated zone (SM) are updated with SM observations. In the new multivariate DA strategy, the assimilation of GWL and SM employed the same assimilation parameters that were used in their respective univariate assimilations, including the localization radius and update frequency. Moreover, the parameter

$K_s$  was updated weekly with a constant damping factor of 0.1. In addition, during assimilation, the saturated river grid cells were masked from the update to avoid numerical instabilities.

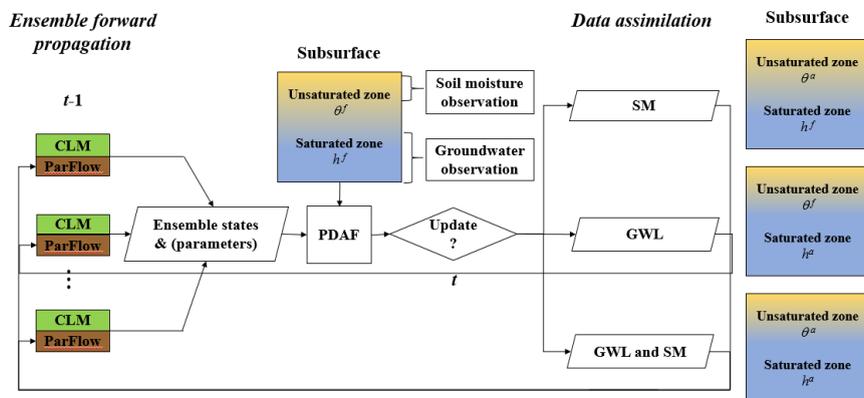


Figure 5.3 Schematic overview of the assimilation of soil moisture from CRNS and groundwater (pressure head) with PDAF into TSMP (CLM-ParFlow).  $\theta^f$  and  $\theta^a$  are the forecasted and analyzed soil moisture in the unsaturated zone;  $h^f$  and  $h^a$  are the forecasted and analyzed pressure heads in the saturated zone. The observed pressure heads are derived from the measured groundwater levels.

## 5.3 Model and Experiment Setup

### 5.3.1 Ensemble generation

To account for input uncertainties, the model CLM-ParFlow was perturbed with respect to model forcings and soil hydraulic parameters (hydraulic conductivity and porosity) to generate an ensemble of 128 members. Table 5.2 summarizes the statistics for atmospheric forcing perturbations. Four atmospheric forcing variables, including precipitation, shortwave radiation, longwave radiation, and air temperature, were perturbed, assuming a multivariate normal distribution and considering the temporal correlations of the atmospheric variables. The temporal correlations were induced by a first-order autoregressive model (Reichle et al., 2010; Han et al., 2015). The temporal correlations and standard deviations of the perturbations were determined based on prior DA experiments conducted at the regional scale (Reichle et al., 2010; Han et al., 2013; Han et al., 2015; Roland Baatz et al., 2017; Li et al., 2023a). Additionally, the

perturbed precipitation and shortwave radiation were multiplied by the corrected lognormally distributed noises to achieve mass and energy balance (Yamamoto, 2007).

Table 5.2 Statistics for the perturbed atmospheric variables, including cross correlations between the atmospheric variables in the last column. The order of the variables in the last column is as indicated in the left column of the table.

Variables	Noise	Standard deviation	Time correlation scale	Cross correlation
Precipitation	Multiplicative	0.3	24 h	[ 1.0, -0.8, 0.5, 0.0,
Shortwave radiation	Multiplicative	0.2	24 h	-0.8, 1.0, -0.5, 0.4,
Longwave radiation	Additive	20 W m <sup>-2</sup>	24 h	0.5, -0.5, 1.0, 0.4,
Air temperature	Additive	1 K	24 h	0.0, 0.4, 0.4, 1.0]

The Rur domain in the model TSMP has a 500 m horizontal spatial resolution and consists of 25 layers with a total depth of 100 m, whereby the vertical resolution of the layers decreases with increasing depth. The topmost 10 layers extend to a depth of 3 meters and align with CLM soil layers to ensure that consistency is maintained. In contrast, the lower layers are considered to be bedrock. Separate perturbations were made to the  $K_s$  and porosity of the soil and bedrock layers.

The soil hydraulic properties of the Mualem-Van Genuchten model were estimated by Rosetta pedotransfer functions (Schaap et al., 2001; Zhang and Schaap, 2017). First, the spatial distributions of sand and clay contents were perturbed separately by geostatistical simulation. Perturbations were simulated by spatially correlated random fields with zero mean, variance 50%<sup>2</sup>, and correlation length of 12.5 km, using a spherical variogram model. To avoid any unrealistic soil texture values, the sand and clay percentages were constrained within a range of 0% to 100%. The spatially distributed silt contents were derived based on the sand and clay contents. Secondly, the spatially distributed  $K_s$  and porosity were then calculated by applying the Rosetta pedotransfer function to the perturbed soil textures for each ensemble member.

For the underlying bedrock layers, the original  $K_s$  values were adopted from a spatially heterogeneous hydrogeological map (Fig 5.2). These  $\log_{10}K_s$  were for each ensemble member perturbed by sampling from a univariate uniform distribution with values between -0.5 and 0.5 and then adding this perturbation to the original  $K_s$  values. The additive perturbation of  $\log_{10}K_s$  was spatially uniform. The porosity of the lower bedrock layers was fixed at a spatially constant value of 0.15.

### 5.3.2 Setup of data assimilation experiments

The spin-up of the model was conducted for each ensemble member to achieve a dynamic hydrologic equilibrium for each realization. The spin-up consists of two steps. Firstly, ParFlow was run for 100 years, initialized with the multi-year average water table depth from Bogen et al. (2005), and repeatedly forced with 30-year average recharge values (derived from gridded German Meteorological Service data on precipitation and actual ET). Secondly, the equilibrium conditions obtained from ParFlow's spin-up were employed to initiate the spin-up for CLM-ParFlow, forced by the atmospheric forcings of 2015 and run repeatedly for 10 years. After the warm-up period, DA experiments were conducted for three years (1<sup>st</sup> of January 2016 until 31<sup>st</sup> of December 2018) to assimilate GWL and SM. In this work, GWL measurements from 78 sites and SM measurements from 12 CRNS stations were assimilated by the localized EnKF.

Eleven different DA experiments (Table 5.3) were carried out to investigate the assimilation performance. The abbreviation PAR indicates that parameters are updated together with model states. In addition to the DA experiments, parameter validation experiments were also conducted, i.e., using the updated  $K_s$  from the DA experiments as input parameters for the other independent years to perform OL and compare the results of the new parameters with the original ones. For example, the updated  $K_s$  from 2016 were validated in 2017 and 2018.

Table 5.3 List of experiments conducted. GWL and SM denote groundwater level and soil moisture observations, respectively.  $h$  is pressure head,  $\theta$  is soil moisture, and  $K_s$  is saturated hydraulic conductivity. The subscripts *sat* and *unsat* denote the saturated and unsaturated zones, respectively. The experiments FC\_DA and FC\_DA\_PAR used the same fully coupled method as Hung et al. (2022).

Experiments (abbrev.)	Observations	State vector	GWL and SM local radius
OL	-	-	-
SM_DA	SM	$\theta$	-
SM_DA_PAR	SM	$\theta, \log K_s$	-
GWL_DA	GWL	$h_{sat}$	-
GWL_DA_PAR	GWL	$h_{sat}, \log K_s$	-
FC_DA	GWL, SM	$\theta, h$	Same
FC_DA_PAR	GWL, SM	$\theta, h, \log K_s$	Same
WC_DA	GWL, SM	$\theta_{unsat}, h_{sat}$	Same
WC_DA_PAR	GWL, SM	$\theta_{unsat}, h_{sat}, \log K_s$	Same
WC_DA_r	GWL, SM	$\theta_{unsat}, h_{sat}$	Different

### 5.3.3 Evaluation of model performance

The performance of the OL and various DA scenarios was assessed by comparing the simulations with daily observations of GWL, SM, and ET. The evaluation utilized root mean square error (RMSE), unbiased root mean square error (ubRMSE), and correlation coefficient (R) as metrics. The RMSE at time step  $t$  is calculated as:

$$RMSE_t = \sqrt{\frac{\sum_{i=1}^{N_{obs}} (y_t^{sim} - y_t^{obs})^2}{N_{obs}}}, \quad (5.5)$$

The ubRMSE at time step  $t$  is calculated as:

$$ubRMSE_t = \sqrt{\frac{\sum_{i=1}^{N_{obs}} [(y_t^{sim} - \bar{y}^{sim}) - (y_t^{obs} - \bar{y}^{obs})]^2}{N_{obs}}}, \quad (5.6)$$

The R is calculated as:

$$R = \frac{\sum_{t=1}^n (y_t^{obs} - \bar{y}^{obs})(y_t^{sim} - \bar{y}^{sim})}{\sqrt{\sum_{t=1}^n (y_t^{obs} - \bar{y}^{obs})^2 \sum_{t=1}^n (y_t^{sim} - \bar{y}^{sim})^2}}, \quad (5.7)$$

where  $y_t^{sim}$  is the ensemble average of the simulated variable (GWL, SM, or ET) at time step  $t$  (from either an OL or a DA run), and  $y_t^{obs}$  is the corresponding observed value at time step  $t$ .  $N_{obs}$  is the total number of observations or validations at time step  $t$ , and  $n$  is the total number of time steps.

## 5.4 Results

### 5.4.1 Univariate data assimilation of soil moisture

Table 5.4 shows the variations in the ubRMSE of SM, ET, and GWL in the OL and univariate assimilation experiments of SM. Table B.1 provides additional details on RMSE and R for SM and ET. Assimilating SM data from CRNS stations alone led to significant improvements in SM forecasts at observation locations for the period from 2016 to 2018 (see Figs B.1-B.3). The reductions in ubRMSE and RMSE for SM in the SM\_DA experiment were substantial, with decreases of over 45% and 50%, respectively. And the joint state parameter estimation performed better than the state update alone. The SM from the DA runs showed improved correlations, with R values ranging from 0.85 to 0.90, compared to R values between 0.61 and 0.63 for the OL (Table B.1). However, the assimilation of SM data only slightly enhanced the ET simulation in RMSE, resulting in less than a 3% reduction compared to the OL. On the other hand, the GWL estimates deteriorate when only SM is assimilated. The ubRMSE of the GWL was increased from 7.23 m (OL) to 7.51 m (SM\_DA). For SM assimilation, the parameter update had little impact on the simulations of the GWL.

Table 5.4 Annual ubRMSE for estimated soil moisture (SM, cm<sup>3</sup>/cm<sup>3</sup>), evapotranspiration (ET, mm/day), and groundwater level (GWL, m) from 2016 to 2018 for OL, SM\_DA, and SM\_DA\_PAR experiments.

Year	Variable	Experiments		
		OL	SM_DA	SM_DA_PAR
2016	SM	0.08	0.05	0.05
2017		0.09	0.04	0.04
2018		0.09	0.05	0.04
2016-2018		0.09	0.05	0.04
2016	ET	0.63	0.65	0.64
2017		0.66	0.66	0.66
2018		0.68	0.70	0.70
2016-2018		0.66	0.67	0.66
2016	GWL	7.30	6.87	6.79
2017		7.24	8.31	7.74
2018		7.16	7.34	7.06
2016-2018		7.23	7.51	7.20

Fig 5.4 shows the differences in SM, ET, and GWL between the OL run and the DA runs (SM\_DA or SM\_DA\_PAR) for the Rur catchment in 2018 (results for 2016 and 2017 are

provided in Figs B.4 and B.5). For the univariate assimilation of SM, significant spatial changes in SM were observed across the catchment after assimilation, becoming wetter in the north and drier in the south. This pattern can be attributed to climatic differences between the regions. The northern catchment, characterized by lower precipitation and higher temperatures, experiences greater moisture stress, leading DA to increase SM to reflect improved water availability. In contrast, the southern catchment, with higher precipitation and lower temperatures, already has relatively abundant SM, resulting in a reduction in SM after assimilation to better align with observed conditions. In 2018, the difference in the spatial variation of annual SM between the SM\_DA and SM\_DA\_PAR experiments was not obvious. Nevertheless, for the years 2016 and 2017, the parameters exhibited a more prominent influence on the characterization of SM. This could be related to the hydrological conditions prevalent in those specific years. In particular, for the wet year 2016, the spatial correlation of SM was stronger when the soil was moist, thereby amplifying the role of parameter updates (Li et al., 2023a).

The spatial pattern of ET differences is similar to that of SM, indicating that SM assimilation has a direct influence on ET. For example, in the southern part of the catchment, both SM\_DA and SM\_DA\_PAR experiments resulted in decreased ET compared to the OL run, which was associated with a decrease in SM. Likewise, in the northern part, ET increased due to the increase in SM after assimilation. However, the effect of SM assimilation on ET in the south was small, i.e., ET changes were mostly less than  $50 \text{ mm yr}^{-1}$ , because ET in the south was generally energy limited. In the northern part of the Rur catchment, where precipitation was generally low, the ET simulation was more strongly influenced by assimilation. For example, an increase in SM due to assimilation resulted in an increase in ET of more than 100 mm per year.

Spatial changes in GWL were also observed in some parts of the catchment. Since the TSMP is a fully integrated model, the univariate assimilation of the SM also has an effect on the GWL. In addition, because the assimilation radius for the SM covers the entire catchment and the TSMP also accounts for lateral groundwater flow, GWL changes do not occur only in the vicinity of the CRNS sites. The variations in GWL were somewhat similar to the spatial variations in SM but not as consistent as for ET.

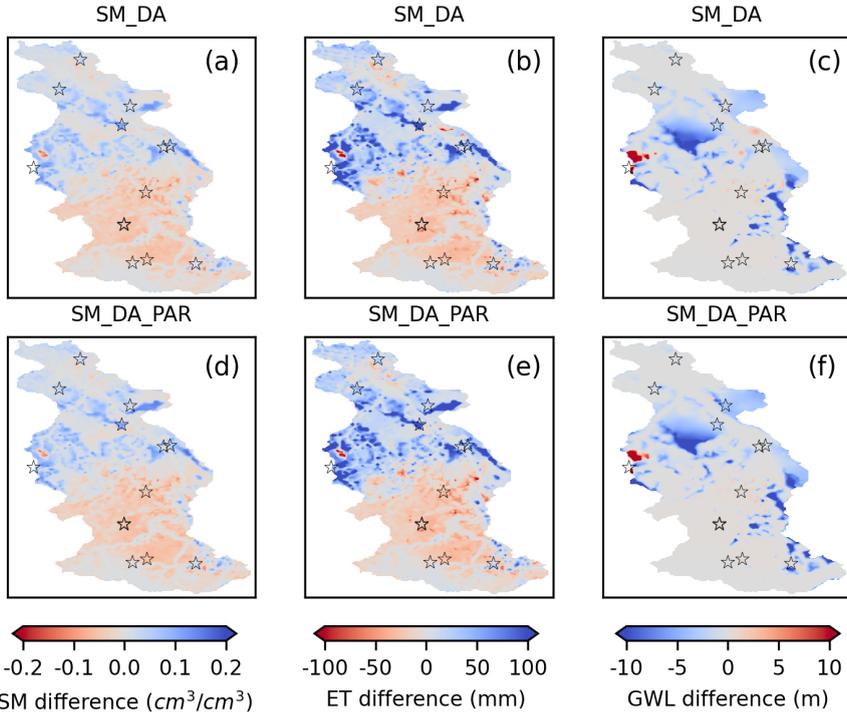


Figure 5.4 Subplots (a) to (c) are the differences (SM\_DA - OL) of annual SM (0-80 cm), ET, and GWL in 2018. Subplots (d) to (f) are the differences (SM\_DA\_PAR - OL) of annual SM (0-80 cm), ET, and GWL in 2018. The locations of the CRNS stations are indicated by the black pentagrams.

#### 5.4.2 Univariate data assimilation of groundwater level

Table 5.5 summarizes the GWL, SM, and ET simulation results over the different years for the GWL\_DA and GWL\_DA\_PAR experiments. The ubRMSE of GWL were calculated at the assimilation locations (distance zero) and also for the validation locations as a function of the distance to the assimilation sites (i.e., 0-0.5 km, 0.5-2.5 km, and 2.5-5 km), showing similar trends. The RMSE for GWL is shown in Table B.2. When GWL was assimilated, this resulted mainly in simulated GWL changes at the groundwater monitoring locations (see Fig B.6, where 12 groundwater monitoring wells are shown). The GWL simulations at the assimilated locations in the GWL\_DA experiment were significantly improved, with the annual ubRMSE reduced from 7.23 m to 2.90 m. Furthermore, the ubRMSE of GWL was reduced to 2.04 m with the

joint state parameter estimation (GWL\_DA\_PAR), and the ubRMSE reduction of GWL was then 70%. Near groundwater assimilation locations, GWL characterization was also improved, with the benefits diminishing further away from the measurement locations. Joint state parameter updating (GWL\_DA\_PAR) always gave better results than state updating alone (GWL\_DA). From 0-0.5 km, the ubRMSE of the GWL for the GWL\_DA\_PAR run was reduced from 6.96 m to 3.78 m, a 46% ubRMSE reduction. For separation distances greater than 0.5 km, the ubRMSE of the GWL from the GWL\_DA\_PAR experiment was still reduced by more than 10% compared to the OL run.

However, when only GWL was assimilated, there was no improvement in the SM estimate. The SM ubRMSEs for the GWL\_DA ( $0.09 \text{ cm}^3/\text{cm}^3$ ) and GWL\_DA\_PAR ( $0.11 \text{ cm}^3/\text{cm}^3$ ) experiments were similar or higher than the ubRMSE in the OL run ( $0.09 \text{ cm}^3/\text{cm}^3$ ). This lack of improvement in SM estimates also extended to ET, as the univariate GWL assimilation did not enhance the SM simulation. Consequently, the ET simulation showed no significant positive effects, with minimal variation in its ubRMSE, RMSE, and R values, as detailed in Table B.1.

Table 5.5 Annual ubRMSE for estimated groundwater level (GWL, m), soil moisture (SM,  $\text{cm}^3/\text{cm}^3$ ), and evapotranspiration (ET, mm/day) from 2016 to 2018 for OL, GWL\_DA, and GWL\_DA\_PAR experiments. *Note:* Distance 0 indicates the groundwater assimilation locations, and 0-0.5 km, 0.5-2.5 km, and 2.5-5 km indicate groundwater validation locations at different distances from the groundwater assimilation sites.

Year	Variable	Distance	Experiments		
			OL	GWL_DA	GWL_DA_PAR
2016	GWL	0	7.30	3.39	2.03
2017			7.24	2.78	2.05
2018			7.16	2.52	2.04
2016-2018			7.23	2.90	2.04
2016		0-0.5 km	7.23	6.54	3.70
2017			6.95	4.46	4.02
2018			6.69	3.89	3.62
2016-2018			6.96	4.97	3.78
2016		0.5-2.5 km	5.32	5.84	4.60
2017			5.26	4.88	4.82
2018			5.09	4.70	4.63
2016-2018			5.22	5.14	4.68
2016	2.5-5 km	6.37	6.36	5.12	
2017		6.31	5.50	4.99	
2018		6.03	5.29	5.16	
2016-2018		6.24	5.72	5.09	

2016			0.08	0.10	0.10
2017	SM	-	0.09	0.09	0.11
2018			0.09	0.10	0.11
2016-2018			0.09	0.09	0.11
2016			0.63	0.63	0.63
2017	ET	-	0.66	0.66	0.66
2018			0.68	0.68	0.68
2016-2018			0.66	0.66	0.66

Fig 5.5 presents the annual differences in GWL, SM, and ET between the GWL assimilation experiments and the OL run for the Rur catchment in 2018. Corresponding results for 2016 and 2017 are available in Figs B.7 and B.8. Since a local radius of 5 km was applied for groundwater assimilation, significant GWL changes were primarily observed in the vicinity of the groundwater assimilation sites. In the mountainous southern part of the Rur catchment, where groundwater monitoring sites are sparse, no notable GWL changes were detected. While the spatial differences in GWL between GWL\_DA and GWL\_DA\_PAR experiments were generally minor, certain areas of the catchment (in the middle) exhibited more pronounced GWL changes after assimilation due to the parameter update.

Assimilation of GWL also affects the estimated SM near the groundwater assimilation sites, with increases and decreases in SM corresponding to changes in GWL. However, since most of the CRNS sites were far from the assimilated groundwater sites, the simulated SM for most of the CRNS sites remained largely unaltered. In addition, the difference between GWL\_DA and GWL\_DA\_PAR experiments was minimal for the estimation of annual SM. The impact on ET simulations was confined to areas near the groundwater assimilation sites, due to the limited localization radius used in the groundwater assimilation. The spatial patterns of SM and ET changes were also similar, reflecting the direct relationship between ET and SM changes.

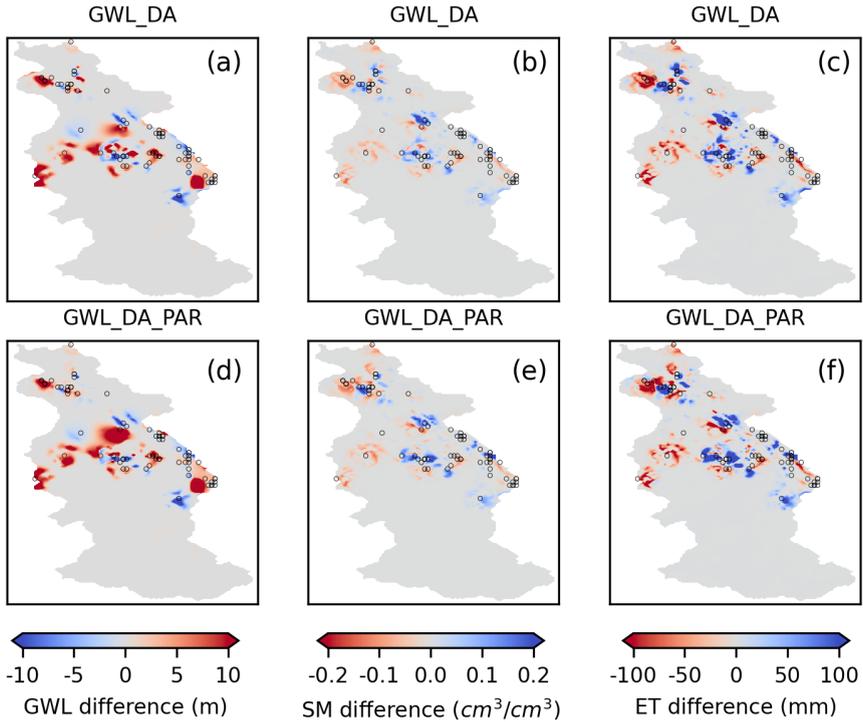


Figure 5.5 Subplots (a) to (c) are the differences (GWL\_DA - OL) of annual GWL, SM (0-80 cm), and ET in 2018. Subplots (d) to (f) are the differences (GWL\_DA\_PAR - OL) of annual GWL, SM (0-80 cm), and ET in 2018. The locations of the assimilated groundwater sites are indicated by the black circles.

#### 5.4.3 Multivariate data assimilation of groundwater level and soil moisture

Table 5.6 presents the ubRMSE for GWL, SM, and ET across various multivariate DA experiments conducted over different years. The RMSE for GWL is shown in Table A.3. Averaged over the period from 2016 to 2018, WC\_DA\_PAR achieved the most accurate GWL estimations at the assimilation locations, with a ubRMSE reduction of 72% (from 7.23 m to 2.05 m). This performance is comparable to that of the univariate GWL assimilation (GWL\_DA\_PAR), which yielded a ubRMSE of 2.04 m. At the validation locations between 0 and 0.5 km, the joint assimilation of GWL and SM was less effective in estimating GWL than the univariate GWL assimilation. However, the difference was not statistically significant. Additionally, the WC\_DA\_r\_PAR experiment showed the lowest ubRMSE (4.56 m) for GWL

at distances 0.5 to 2.5 km from assimilation locations (among all univariate and multivariate DA experiments), while the FC\_DA\_PAR experiment yielded the smallest ubRMSE (4.91 m) for distances between 2.5 and 5 km from assimilation locations.

The characterization of SM in multivariate DA experiments showed notable improvement, with WC\_DA\_PAR achieving the largest reduction in ubRMSE (50%). The corresponding RMSE and R for SM and ET are detailed in Table B.4. However, the statistical metrics of SM in multivariate assimilation experiments (FC\_DA or FC\_DA\_PAR) were not superior to those in the univariate assimilation experiments (SM\_DA or SM\_DA\_PAR). This indicates that incorporating additional groundwater level data does not enhance SM estimation in the fully coupled model. While the multivariate assimilation experiments resulted in a slight improvement in ET simulations, with an approximate reduction in RMSE of 3%, this improvement was not reflected in the ubRMSE. Moreover, the inclusion or exclusion of parameter updates did not significantly impact the ET simulations in the multivariate assimilation experiments.

Table 5.6 Annual ubRMSE for estimated groundwater level (GWL, m), soil moisture (SM, cm<sup>3</sup>/cm<sup>3</sup>), and evapotranspiration (ET, mm/day) from 2016 to 2018 for OL and multivariate assimilation experiments. *Note:* Distance 0 indicates the groundwater assimilation locations, and 0-0.5 km, 0.5-2.5 km, and 2.5-5 km indicate groundwater validation locations at different distances from the groundwater assimilation sites.

Year	Variable	Distance	Experiments						
			OL	FC_DA	FC_DA_PAR	WC_DA	WC_DA_PAR	WC_DA_r	WC_DA_r_PAR
2016	GWL	0	7.30	3.24	2.96	3.14	2.13	3.24	2.37
2017			7.24	4.06	2.88	2.93	1.98	3.01	2.04
2018			7.16	3.44	3.06	3.33	2.03	2.57	2.11
2016-2018			7.23	3.58	2.97	3.13	2.05	2.94	2.17
2016		0-0.5 km	7.23	4.64	4.43	4.16	4.23	4.41	4.52
2017			6.95	3.96	3.69	3.94	4.60	4.27	3.49
2018			6.69	3.25	3.54	3.93	3.62	3.96	3.61
2016-2018			6.96	3.95	3.89	4.01	4.15	4.21	3.87
2016		0.5-2.5 km	5.32	5.57	4.72	4.73	7.56	4.73	4.67
2017			5.26	4.75	4.78	4.65	4.73	4.61	4.48
2018			5.09	4.34	4.46	4.66	4.79	4.61	4.52
2016-2018			5.22	4.89	4.65	4.68	5.70	4.65	4.56
2016	2.5-5 km	6.37	5.65	5.03	5.27	8.24	7.54	5.61	
2017		6.31	5.23	5.02	5.38	5.52	7.01	7.81	
2018		6.03	4.89	4.68	5.30	5.18	5.68	5.10	
2016-2018		6.24	5.26	4.91	5.32	6.31	6.74	6.17	

2016			0.08	0.05	0.05	0.06	0.04	0.06	0.04
2017	SM	-	0.09	0.06	0.06	0.05	0.04	0.05	0.04
2018			0.09	0.08	0.05	0.07	0.04	0.05	0.04
2016-2018			0.09	0.06	0.05	0.06	0.04	0.05	0.04
2016			0.63	0.63	0.64	0.63	0.63	0.64	0.64
2017	ET	-	0.66	0.66	0.66	0.66	0.66	0.66	0.66
2018			0.68	0.70	0.70	0.70	0.70	0.70	0.70
2016-2018			0.66	0.66	0.67	0.66	0.66	0.66	0.66

For a better comparison, the ubRMSE scores for the univariate and multivariate assimilation experiments are illustrated in Fig 5.6. The fully coupled model (FC\_DA and FC\_DA\_PAR) demonstrated improved performance in estimating GWL and SM compared to the OL run. However, these results were less accurate than those achieved through univariate assimilation of GWL or SM for their respective variables. In contrast, the weakly coupled model (WC\_DA and WC\_DA\_PAR) provided more accurate estimates of GWL and SM than the fully coupled model. As shown in Fig 5.6, the WC\_DA\_r and WC\_DA\_r\_PAR experiments yielded overall better results for GWL and SM at the assimilated locations compared to other multivariate assimilation experiments. However, when validating GWL at distances of 2.5 to 5 km from the assimilation sites, the results were slightly less accurate than those from the fully coupled model. The slight decrease in accuracy is likely due to the larger assimilation radius employed for SM, which exerted a more pronounced influence on the GWL estimates. As a comparison, the SM-only assimilation demonstrated that updating SM alone resulted in a less precise estimation of GWL.

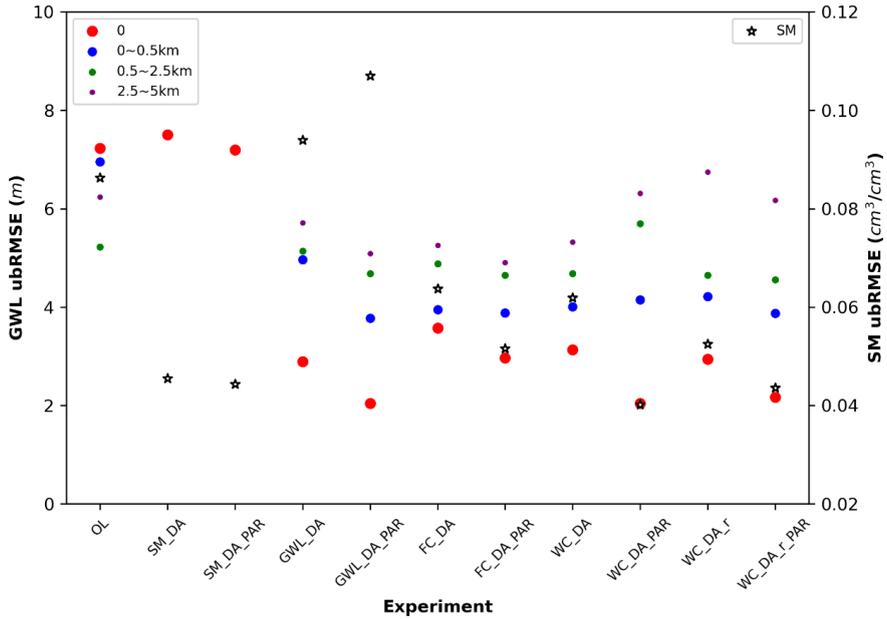


Figure 5.6 Average ubRMSE of GWL (at different distances from the assimilated groundwater sites) and SM for each univariate and multivariate assimilation experiment. The left axis corresponds to GWL and the right axis to SM.

Fig 5.7 presents the variations in annual GWL, SM, and ET across different multivariate assimilation experiments compared to the OL run for the Rur catchment in 2018. Given that the results from jointly updating states and parameters closely mirror those from updating states alone, only the former is displayed. Corresponding graphs for the years 2016 and 2017 are available in Figs B.9 and B.10. In the WC\_DA\_PAR experiment, the alterations observed in the GWL estimates closely resemble those observed in the GWL univariate assimilation experiments. This similarity arises from the identical groundwater updating strategy, which involves updating the pressure head within the saturated zone only. Notably, GWL changes also occurred in regions of the catchment without groundwater assimilation sites. These changes are attributed to the influence of SM updates in the multivariate assimilation experiments, specifically for FC\_DA\_PAR and WC\_DA\_r\_PAR.

The spatial SM pattern observed in the WC\_DA\_r\_PAR experiment closely resembled that of the univariate SM assimilation, as both employed the same assimilation radius for SM updates. However, due to the smaller localization radius (5 km) used in the FC\_DA\_PAR and

WC\_DA\_PAR experiments, SM alterations were mainly concentrated near the CRNS sites. It is noteworthy that differences in annual SM changes were also observed in the vicinity of groundwater sites located in the northern part of the catchment. This indicates that GWL impacted the estimation of SM when GWL and SM were assimilated jointly. With respect to ET, the spatial variation is consistent with SM, as ET is directly affected by SM. Additionally, in the multivariate assimilation experiments, the SM of some locations near the groundwater assimilation sites was also modified by GWL updates. This led to ET changes that may differ from those in the univariate SM assimilation.

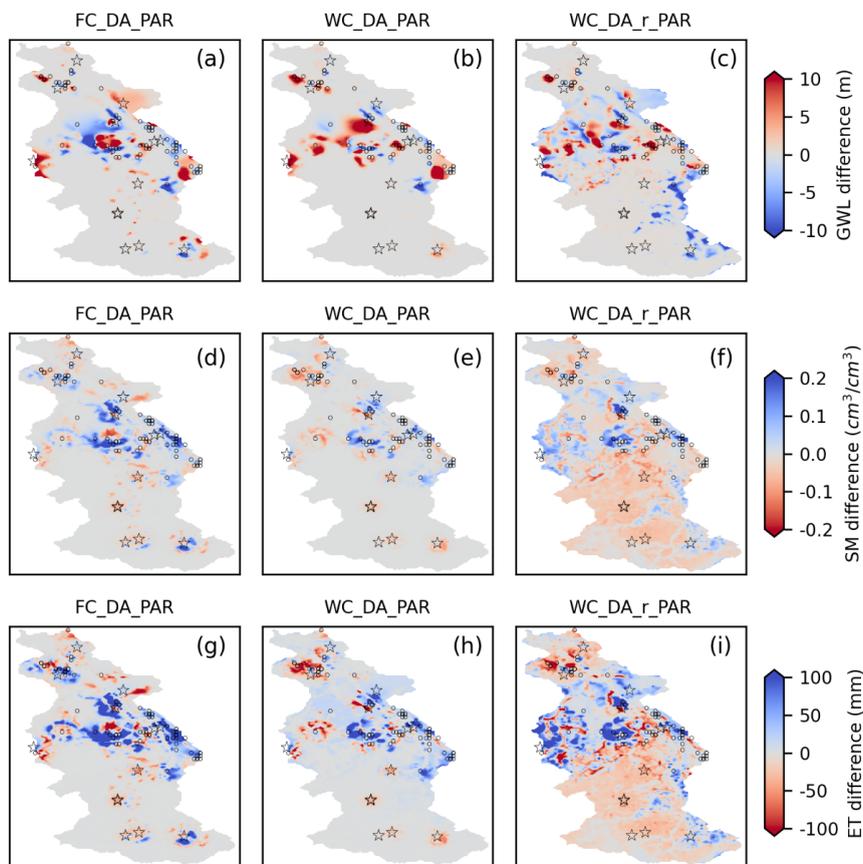


Figure 5.7 Differences in annual GWL, SM, and ET between various multivariate DA experiments and OL run in 2018: (a, d, g) FC\_DA\_PAR; (b, e, h) WC\_DA\_PAR; (c, f, i) WC\_DA\_r\_PAR. The red pentagrams and black circles indicate the locations of the CRNS

stations and the assimilated groundwater sites, respectively.

To provide a more detailed comparison between univariate and multivariate assimilation, Fig 5.8 depicts the temporal evolution of SM and GWL at a CRNS station and a groundwater site across all DA experiments. The time series of predicted SM in the GWL\_DA and GWL\_DA\_PAR experiments are very similar to OL, indicating minimal impact of GWL information on SM prediction in these cases. Similarly, univariate SM assimilation alone does not change the GWL much. When GWL is assimilated, the simulations of GWL become progressively closer to the measurements over time. Among the multivariate assimilation experiments, the fully coupled models (FC\_DA and FC\_DA\_PAR) exhibit the largest discrepancies between the simulated and observed GWL and SM values. Overall, simulated GWL and SM differ not much between the DA experiments involving joint state and parameter updates and those with state updates alone.

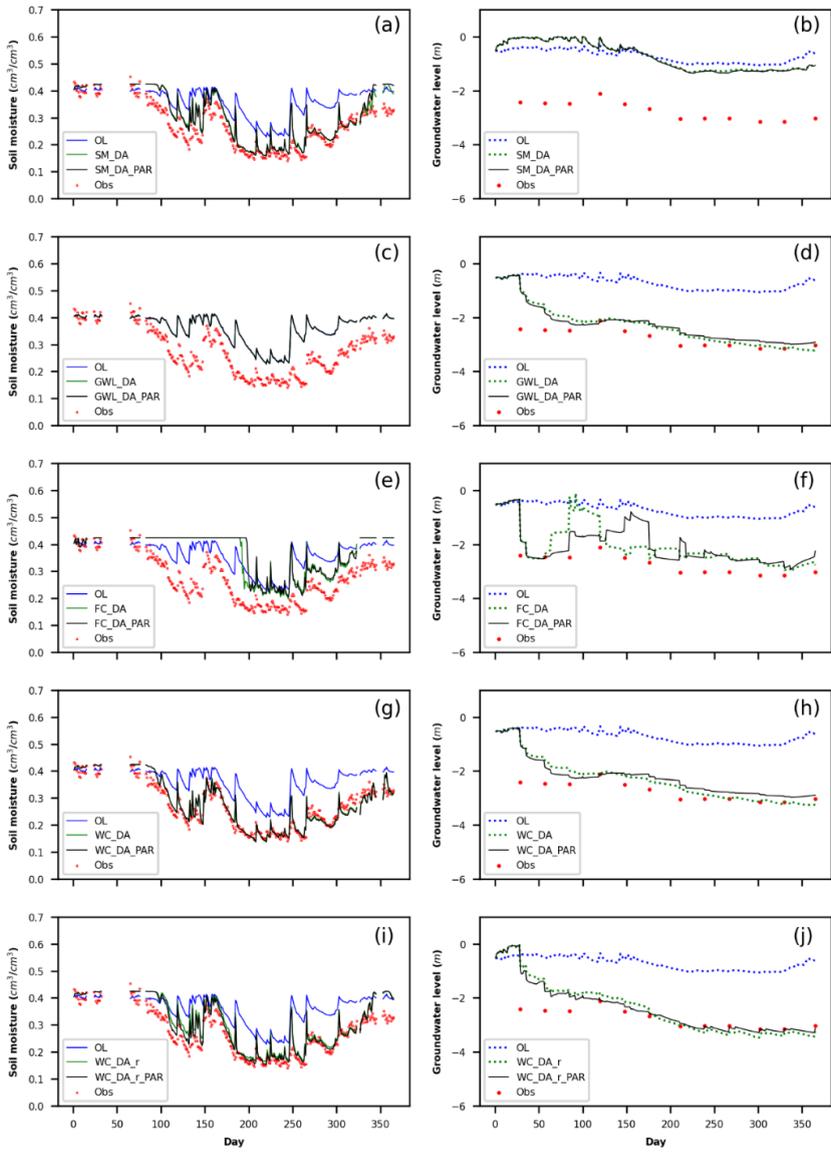


Figure 5.8 Temporal evolution of SM at the CRNS station Kall and GWL at a groundwater site for OL and different DA experiments in 2018: (a, b) SM\_DA and SM\_DA\_PAR; (c, d) GWL\_DA and GWL\_DA\_PAR; (e, f) FC\_DA and FC\_DA\_PAR; (g, h) WC\_DA and WC\_DA\_PAR; (i, j) WC\_DA\_r and WC\_DA\_r\_PAR.

#### 5.4.4 Influence of $K_s$ updates on simulations

In general, the performance of including parameter updates is superior to that of including only state updates in all DA experiments. To further investigate the effectiveness of the estimated parameters, these parameters were applied to OL experiments for independent years and compared with the simulations using the prior (not updated) parameter values. Table 5.7 summarizes the performance statistics, including RMSE, ubRMSE, and R for different variables (GWL, SM, and ET) during the parameter evaluation period. The updated  $K_s$  from the SM\_DA\_PAR experiment improved the estimation of SM, as indicated by enhancements across all evaluation metrics. For instance, the ubRMSE for SM in the OL validation runs decreased from  $0.09 \text{ cm}^3/\text{cm}^3$  (prior  $K_s$ ) to  $0.08 \text{ cm}^3/\text{cm}^3$  (updated  $K_s$  from SM\_DA\_PAR). However, the updated  $K_s$  did not lead to improvement in GWL estimates, nor was there a notable enhancement in ET characteristics.

When the OL runs for the other independent years were performed using the updated  $K_s$  from the GWL\_DA\_PAR experiments, the overall RMSE and ubRMSE of GWL were slightly reduced (less than 2%) compared to using the prior  $K_s$ . Furthermore, the simulation of GWL at unassimilated locations improved with the updated  $K_s$ . Specifically, at distances of 2.5 to 5 km from the assimilated sites, the simulated GWL showed nearly 4% improvement, with the ubRMSE decreasing from 6.24 m to 6.01 m. However, there was no obvious positive effect on the SM and ET simulations with the updated  $K_s$  from the GWL\_DA\_PAR experiment.

The updated  $K_s$  from the WC\_DA\_r\_PAR experiment did not enhance GWL simulations at the assimilation locations during the validation periods in the OL run. However, compared to GWL\_DA\_PAR, the updated  $K_s$  from the WC\_DA\_r\_PAR experiment demonstrated superior performance in GWL estimation at non-assimilated locations. Specifically, the ubRMSE of GWL was reduced by more than 4% at distances of 0~0.5 km and 2.5~5 km from the assimilated sites. Additionally, the updated  $K_s$  contributed to improved SM simulations, with the ubRMSE of SM decreasing from  $0.09 \text{ cm}^3/\text{cm}^3$  (using prior  $K_s$ ) to  $0.08 \text{ cm}^3/\text{cm}^3$  in the WC\_DA\_r\_PAR experiment. Despite these improvements, the updated  $K_s$  from the WC\_DA\_r\_PAR experiment did not significantly influence ET characterization.

Table 5.7 Averaged statistical metrics for the estimated groundwater level (GWL), soil moisture (SM), and evapotranspiration (ET) for all the validation experiments from 2016 to 2018. *Note:* Distance 0 indicates the groundwater assimilation locations, and 0-0.5 km, 0.5-2.5 km, and 2.5-5 km indicate groundwater validation locations at different distances from the groundwater

assimilation sites.

Variable	Distance	Metrics	$K_s$ from SM_DA_PAR	$K_s$ from GWL_DA_PAR	$K_s$ from WC_DA_r_PAR
GWL	0	RMSE (m)	7.90	7.16	7.32
	0-0.5 km			6.92	6.84
	0.5-2.5 km			6.54	6.50
	2.5-5 km			6.98	6.95
	0	ubRMSE (m)		7.09	7.23
	0-0.5 km			6.71	6.63
	0.5-2.5 km			5.21	5.17
	2.5-5 km		6.01	5.97	
SM	-	RMSE (cm <sup>3</sup> /cm <sup>3</sup> )	0.09	0.10	0.09
		ubRMSE (cm <sup>3</sup> /cm <sup>3</sup> )	0.08	0.09	0.08
		R	0.67	0.60	0.68
ET	-	RMSE (mm/day)	0.75	0.76	0.75
		ubRMSE (mm/day)	0.66	0.66	0.66
		R	0.83	0.84	0.83

Fig 5.9 depicts the changes in the spatial ensemble means of  $\log K_s$  at 2 cm and 10 m depth for different experiments (SM\_DA\_PAR, GWL\_DA\_PAR, and WC\_DA\_r\_PAR) compared to the OL run. Results are shown for the year 2018 (results for 2016 and 2017 are provided in Figs B.11 and B.12). The spatial patterns of updates in  $K_s$  for these three years were similar. In the SM\_DA\_PAR experiment,  $K_s$  changed not only in the root zone but also indirectly in the saturated zone at 10 m. These alterations in  $K_s$  could potentially influence the GWL estimate. For the GWL\_DA\_PAR experiment, state updates were confined to the saturated zone, resulting in significant changes in  $K_s$  within the saturated zone at assimilated locations. The unsaturated zone showed no noticeable effect from these updates.

In the WC\_DA\_r\_PAR experiment, SM and GWL variables were assimilated independently. Consequently, the changes in  $K_s$  in the unsaturated zone were expected to resemble those observed in the SM\_DA\_PAR run, while changes in the saturated zone would be following those in the GWL\_DA\_PAR run. The results demonstrate that the spatial patterns of  $K_s$  updates in these zones are indeed similar to those seen in the respective univariate assimilation experiments. However, since the updates for GWL and SM in the multivariate assimilation are interrelated, the updated  $K_s$  values reflect more complex interactions than a simple superposition of the updates from the two univariate assimilation experiments. This led to more pronounced variations in  $K_s$  across different subsurface depths in certain areas of the study region.

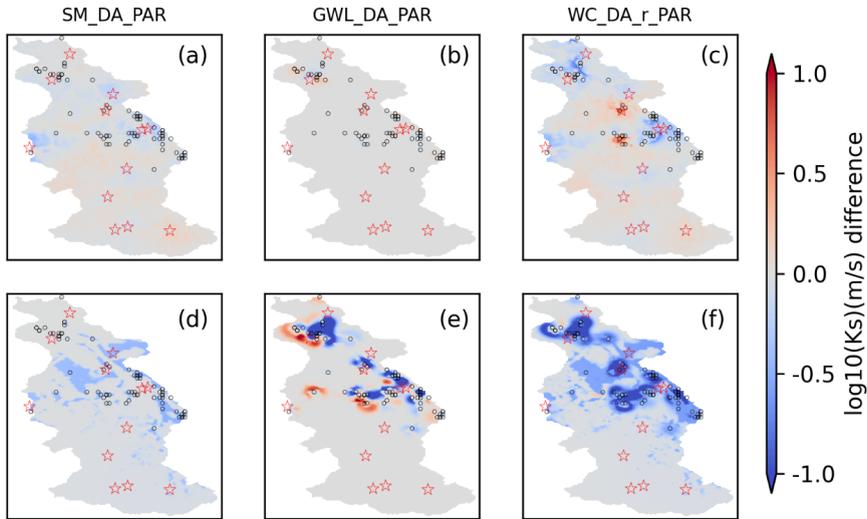


Figure 5.9 Differences in ensemble averaged  $\log K_s$  between OL and different DA experiments in 2018: (a) and (d) SM\_DA\_PAR; (b) and (e) GWL\_DA\_PAR; (c) and (f) WC\_DA\_r\_PAR. The first row is from 2 cm depth, and the second row is from 10 m depth. The red pentagrams and black circles indicate the locations of CRNS stations and the assimilated groundwater sites, respectively.

## 5.5 Discussions

### 5.5.1 Strengths and limitations of new multivariate data assimilation approach

This study demonstrates a new multivariate assimilation approach that improves the estimates of both GWL and SM. In contrast, univariate assimilation, which focuses solely on either SM or GWL, tends to improve the estimation of the targeted variable but often degrades the accuracy of the non-assimilated variable. This degradation may be attributed to unrealistic inter- and cross-variable correlations introduced during the assimilation process. It may also be related to the use of real in-situ measurements. CRNS and groundwater sites are scattered throughout the catchment. Thus, the assimilation of a single variable (e.g., GWL) makes it difficult to improve the estimation of variables (e.g., SM) from other distant grid points.

Hung et al. (2022) already concluded for a complex synthetic case for a region in southwestern Germany that updating only saturated subsurface states outperforms fully coupled

DA in GWL estimation, in contrast to earlier experiments with highly simplified synthetic setups (Zhang et al., 2018). However, the GWL and SM synthetic data for the assimilated sites in Hung et al. (2022) lie within the same grid. This study was conducted for a real catchment where most of the groundwater and CRNS soil moisture monitoring sites are not located on the same grid, providing a more realistic representation of the distribution of soil moisture and groundwater level observation locations. Our study shows that the newly proposed multivariate assimilation approach performs better in state prediction than the fully coupled DA method used by Hung et al. (2022). The advantage of the new approach is the use of independent updates during the multivariate assimilation. This method enables the saturated zone pressure to be updated using GWL measurements, while the unsaturated zone SM is updated using CRNS SM measurements. Moreover, using different localization radii for GWL and SM assimilation can better capture their respective spatial characteristics, reduce the negative impact of distant errors on assimilation, and ensure more accurate state updates, thereby improving the performance of joint assimilation. In addition, asynchronous assimilation allows the use of different update frequencies for various variables. For example, soil moisture can be updated daily to capture its more rapid variations, while groundwater, which typically changes more slowly, can be updated weekly. By assimilating data asynchronously, the timescales of both fast-changing and slow-changing process can be more appropriately considered in the coupled models. As a result, the predictions of multiple variables can be corrected in a coupled model using different data sources, minimizing the risk of confounding factors and reducing the impact of spurious correlations. Therefore, our study highlights applying DA across terrestrial compartments in integrated models when dealing with real-world cases.

The advantage of the new approach is the use of independent updates during the multivariate assimilation. This method enables the saturated zone pressure to be updated using GWL measurements, while the unsaturated zone SM is updated using CRNS SM measurements. As a result, the predictions of multiple variables can be corrected in a coupled model using different data sources, minimizing the risk of confounding factors and reducing the impact of spurious correlations.

The multivariate DA experiments were conducted over three years (2016-2018) under varying hydrological conditions. While some year-to-year variations were observed, the results remained robust and consistent across the study period. The new multivariate assimilation method demonstrated improved predictions of multiple variables within the integrated terrestrial system. However, it is important to note that the ubRMSE of GWL at certain distances (2.5~5 km) was larger in the multivariate assimilation experiments compared to the univariate

GWL assimilation runs (6.17 m vs. 5.09 m). Thus, while multivariate assimilation incorporates additional types of measurements compared to univariate assimilation, it does not always provide additional benefits.

These results are also consistent with the study on an artificial hillslope by Botto et al. (2018), in which the CATHY model was used. This study showed that the assimilation of additional variables may lead to a degradation of model predictions for other variables. Their results indicated that the poor quality of pressure head measurements limited the performance of the filter. However, in the study by Zhang et al. (2016), the poor model predictions of the joint GWL and SM assimilation were likely caused by unrealistic cross-variable correlations due to limited ensemble sizes. The reasons for the limited benefits of multivariate assimilation (compared to univariate assimilation) may therefore differ from model to model.

In addition to the assimilation variables, the related variable ET in the integrated model was also evaluated. However, the results indicated that the GWL assimilation did not improve the estimation of ET. The difficulty is related to the limited improvement in SM characterization. In areas where the GWL is relatively deep, the near-surface SM and ET are less influenced by GWL assimilation. In this study, only the assimilation of SM observations into integrated models yielded some improvement in ET estimates, but to a limited extent. The multivariate DA experiments did not further improve ET simulations but retained the same degree of positive effect on ET as the univariate assimilation of SM.

### **5.5.2 Uncertainties and potential improvement**

This study introduces innovative approaches for the implementation of multivariate assimilation within integrated hydrological models. However, there are still uncertainties that need to be addressed in future studies. The coarse spatial discretization of the model is considered as one of the sources of uncertainty. Coarser model resolution tends to flatten the topography, thus reducing the gradients of surface and groundwater flow, and is expected to introduce a systematic bias in the simulated GWL. In the future, higher model resolutions (e.g., 100 m) could be tested to better represent groundwater bodies associated with narrow valleys, thus eliminating the bias due to coarser model resolution. Moreover, potential biases in the observed data are not considered in this study. In practical applications, various methods can be used to assimilate observations with consistent biases, such as rescaling observations or subtracting long-term averages from observations (Hain et al., 2012; Pauwels et al., 2013; Zhang et al., 2016; Pauwels et al., 2020).

Additionally, structural errors within the model can pose additional challenges for DA when using real data. In this study, GWL is assimilated under the assumption of hydrostatic conditions, but the reality is much more complex. Multiple aquifers may be present, stacked on top of each other, and separated by aquitards. In addition, aquifers can be separated horizontally by fault lines, and these disruptions can affect the continuity of aquifers, potentially influencing water flow and distribution. Furthermore, aquifers can be strongly affected by artificial water extraction. This is particularly the case in the Rur catchment, where water management to keep open-cast lignite mines free of water has a major impact on hydrogeology (Bogena et al., 2018). These problems can be mitigated by GWL assimilation, which can help calibrate the model, adjust parameters to match observed data, and improve the predictive capabilities of the model to better represent the complexities of multiple aquifers, water withdrawals on aquifer systems, and mining activities.

The SM data assimilated in this study were derived from the CRNS observations. Therefore, the quality of DA depends on the accuracy of the CRNS calibration and also on the depth weighting function applied (i.e., Schrön et al., 2017). A more direct way is to assimilate the CRNS neutron intensity with the COSMIC operator (Shuttleworth et al., 2013), which is currently being implemented in TSMP-PDAF and will be available for future DA studies.

The EnKF was specifically developed to deal with nonlinear models and is therefore attractive for predictions with integrated terrestrial system models, but it has already encountered limitations in applications to single terrestrial compartments. In this study, univariate assimilation of GWL or SM resulted in a deterioration in the simulation accuracy of other variables. The nonlinear relationship between GWL (pressure head) and SM increases the complexity of the model-data fusion problem, and therefore multivariate assimilation may lead to significant trade-offs that further limit filter performance (Brandhorst et al., 2017; Botto et al., 2018). In this work, separate DA analyses were applied to the interrelated hydrologic processes to improve GWL and SM predictions together. As more RS and terrestrial field data become available at lower cost, these data sources can be used for DA. However, determining the most appropriate data types for assimilation and examining the trade-offs that may arise when assimilating different variables in a multivariate assimilation framework remain critical challenges for future research. Multivariate assimilation may potentially be improved in several ways, including the use of variants of EnKF, hybrid approaches of EnKF and other filters, or bias-aware filters.

The ultimate goal of multivariate assimilation is to correct the estimates of not only the states but also the relevant parameters. In this work, only the most sensitive parameter of the

subsurface groundwater model,  $K_s$ , was updated, and independent validation results with the updated  $K_s$  also show improved estimates of GWL or SM. Although it is theoretically possible to estimate more parameters, Brandhorst and Neuweiler et al. (2023) found numerical convergence problems in synthetic studies when assimilating SM to update multiple soil hydraulic parameters. Therefore, it is more difficult to update all van Genuchten parameters for real-world cases. The synthetic study by Shi et al. (2015) also showed that as the number of soil hydraulic parameters increases, it becomes increasingly difficult to estimate different parameters simultaneously with EnKF. However, they found that it is helpful to assimilate more types of data to improve the estimates of soil hydraulic parameters. In the future, multivariate assimilation using multi-source data should be considered to update relevant parameters in coupled land surface-subsurface models to improve the accuracy of model predictions.

This study was conducted in the Rur catchment, which benefits from a dense and precise in-situ measurement network for CRNS and groundwater level stations. The availability of such comprehensive datasets provides a distinctive opportunity to validate the novel multivariate assimilation method at the catchment scale. To the best of our knowledge, no other catchment provides a similarly comprehensive combination of well-established monitoring infrastructure. To broaden the applicability of our method, future studies should explore the use of more accessible datasets, e.g., remote sensing data such as groundwater storage variations from GRACE/GRACE-FO (Li et al., 2024) or soil moisture from SMAP.

## 5.6 Conclusions

In this study, different methods were tested to assimilate GWL and SM observations from a dense observation network into the integrated land surface-subsurface model TSMP for the Rur catchment in Germany. The advantages and disadvantages of using these observations individually or in joint DA approaches were analyzed. Furthermore, a new multivariate data assimilation approach is proposed, in which the weakly coupled assimilation of GWL and SM with the LEnKF is performed separately so that the update becomes more stable. Groundwater observations are assimilated to update the position of the unsaturated-saturated zone interface and the saturated zone states (and possibly parameters). SM observations are assimilated to update the unsaturated zone states (and possibly parameters). For this study, 128 ensemble members were generated by perturbing atmospheric forcings and subsurface hydraulic parameters, and DA experiments were conducted for the years 2016-2018. In addition to GWL and SM observations, ET data from eddy covariance stations were used to evaluate the influence of univariate or multivariate DA on the estimates of GWL, SM, and ET.

The various DA experiments resulted in different levels of improvement in model predictions, with results varying across different years. In general, the univariate assimilation experiments were more successful in characterizing the variable for which measurements were assimilated. For example, SM assimilation resulted in 50% ubRMSE reductions in SM at the observation sites. In the case of GWL assimilation, the GWL ubRMSE was reduced by 70% at the observation sites, by almost 50% at a distance of 500 m, and by about 20% at 5 km. However, the univariate assimilation of GWL degraded the prediction accuracy of SM, and the univariate assimilation of SM also diminished the estimation of GWL.

Joint assimilation of GWL and SM using the original method of the fully coupled model does not provide additional benefits over univariate assimilation but rather is much less effective. Nevertheless, the newly proposed multivariate assimilation approach effectively combines the advantages of the different univariate assimilation schemes. As a result, the accuracy of the assimilated variables in the multivariate assimilation is comparable to that achieved through univariate assimilation. Overall, it can be concluded that the joint assimilation of GWL and SM using the new approach has an advantage over univariate assimilation. In addition, the characterization of ET only could be improved when SM is assimilated for univariate or multivariate assimilation.

This study demonstrates the potential of joint assimilation of SM and GWL observations from CRNS and groundwater monitoring networks to improve hydrological modeling of terrestrial systems with process-based integrated model systems. Future work should concentrate on multivariate DA with more data types from diverse sources (e.g., RS data and terrestrial observations) to improve the estimation of multiple terrestrial compartments at higher spatial resolutions. This involves exploring the interactions between different variables in integrated models during multivariate assimilation and developing more effective DA strategies to avoid degradation of non-assimilated variables.

## Appendix B

Table B.1 RMSE and R for estimated SM and ET from 2016 to 2018 for OL and univariate data assimilation experiments.

Year	Variable	OL	SM_DA	SM_DA_PAR	GWL_DA	GWL_DA_PAR
2016	SM (RMSE, cm <sup>3</sup> /cm <sup>3</sup> )	0.11	0.06	0.05	0.12	0.12
2017		0.10	0.05	0.04	0.10	0.11
2018		0.09	0.05	0.04	0.10	0.11
2016-2018		0.10	0.05	0.05	0.10	0.11

2016		0.62	0.85	0.85	0.50	0.47
2017	SM (R)	0.63	0.88	0.88	0.58	0.43
2018		0.61	0.89	0.90	0.56	0.46
2016-2018		0.62	0.87	0.88	0.55	0.46
2016	ET (RMSE, mm/day)	0.76	0.72	0.71	0.76	0.76
2017		0.79	0.77	0.77	0.79	0.79
2018		0.73	0.75	0.75	0.73	0.73
2016-2018		0.76	0.74	0.74	0.76	0.76
2016	ET (R)	0.82	0.80	0.81	0.82	0.82
2017		0.86	0.86	0.86	0.86	0.86
2018		0.83	0.82	0.82	0.83	0.83
2016-2018		0.84	0.83	0.83	0.84	0.84

Table B.2 Annual RMSE for estimated GWL (m) from 2016 to 2018 for OL and univariate data assimilation experiments.

Year	Distance	OL	GWL_DA	GWL_DA_PAR	SM_DA	SM_DA_PAR
2016	0	7.30	3.40	2.04		
2017		7.24	2.77	2.05	7.39	7.92
2018		7.24	2.50	2.05		
2016-2018		7.26	2.89	2.05		
2016	0-0.5 km	7.27	6.51	3.66		
2017		6.98	4.43	3.87	8.43	7.96
2018		6.77	3.83	3.56		
2016-2018		7.01	4.92	3.70		
2016	0.5-2.5 km	6.49	6.49	5.55		
2017		6.45	5.84	5.69	7.84	7.69
2018		6.31	5.52	5.52		
2016-2018		6.42	5.95	5.59		
2016	2.5-5 km	7.12	6.89	5.92		
2017		7.10	6.29	5.82	7.89	7.86
2018		6.92	6.05	5.89		
2016-2018		7.05	6.41	5.88		

Table B.3 Annual RMSE for estimated GWL (m) from 2016 to 2018 for OL and multivariate data assimilation experiments. *Note:* Distance 0 indicates the groundwater assimilation locations, and 0-0.5 km, 0.5-2.5 km, and 2.5-5 km indicate groundwater validation locations at different distances from the groundwater assimilation sites.

Year	Distance	OL	FC_DA	FC_DA_PAR	WC_DA	WC_DA_PAR	WC_DA_r	WC_DA_r_PAR
2016		7.30	3.36	3.27	3.17	2.17	3.25	2.38
2017	0	7.24	4.29	3.19	2.93	1.98	3.02	2.05
2018		7.24	3.64	3.32	3.38	2.03	2.62	2.15

2016-2018		7.26	3.76	3.26	3.16	2.06	2.96	2.19
2016		7.27	4.73	4.55	4.11	4.24	4.40	4.52
2017	0-0.5 km	6.98	4.00	3.71	3.92	4.60	4.25	3.46
2018		6.77	3.15	3.70	3.85	3.56	3.93	3.58
2016-2018		7.01	3.96	3.98	3.96	4.13	4.19	3.85
2016		6.49	6.53	6.01	5.66	8.00	5.82	5.65
2017	0.5-2.5 km	6.45	5.84	5.74	5.55	5.68	5.72	5.58
2018		6.31	5.44	5.75	5.53	5.60	5.72	5.55
2016-2018		6.42	5.93	5.83	5.58	6.43	5.75	5.59
2016		7.12	6.55	6.18	6.09	8.54	7.92	6.20
2017	2.5-5 km	7.10	6.03	5.91	5.99	6.18	7.46	8.17
2018		6.92	5.77	5.92	6.01	5.92	6.37	5.91
2016-2018		7.05	6.12	6.00	6.03	6.88	7.25	6.76

Table B.4 RMSE and R for estimated SM and ET from 2016 to 2018 for OL and multivariate data assimilation experiments.

Year	Variable	OL	FC_DA	FC_DA_PAR	WC_DA	WC_DA_PAR	WC_DA_r	WC_DA_r_PAR
2016		0.11	0.05	0.05	0.07	0.05	0.06	0.05
2017	SM (RMSE, cm <sup>3</sup> /cm <sup>3</sup> )	0.10	0.07	0.06	0.05	0.04	0.05	0.04
2018		0.09	0.08	0.05	0.07	0.04	0.05	0.04
2016-2018		0.10	0.07	0.05	0.06	0.04	0.05	0.05
2016		0.62	0.85	0.87	0.75	0.90	0.74	0.88
2017	SM (R)	0.63	0.74	0.81	0.85	0.89	0.86	0.88
2018		0.61	0.69	0.85	0.71	0.92	0.88	0.89
2016-2018		0.62	0.76	0.84	0.77	0.90	0.83	0.88
2016		0.76	0.71	0.71	0.71	0.71	0.71	0.71
2017	ET (RMSE, mm/day)	0.79	0.77	0.77	0.77	0.77	0.77	0.77
2018		0.73	0.75	0.75	0.75	0.75	0.75	0.75
2016-2018		0.76	0.74	0.75	0.74	0.74	0.74	0.74
2016		0.82	0.81	0.81	0.81	0.81	0.81	0.81
2017	ET (R)	0.86	0.86	0.86	0.86	0.86	0.86	0.86
2018		0.83	0.82	0.82	0.82	0.82	0.82	0.82
2016-2018		0.84	0.83	0.83	0.83	0.83	0.83	0.83

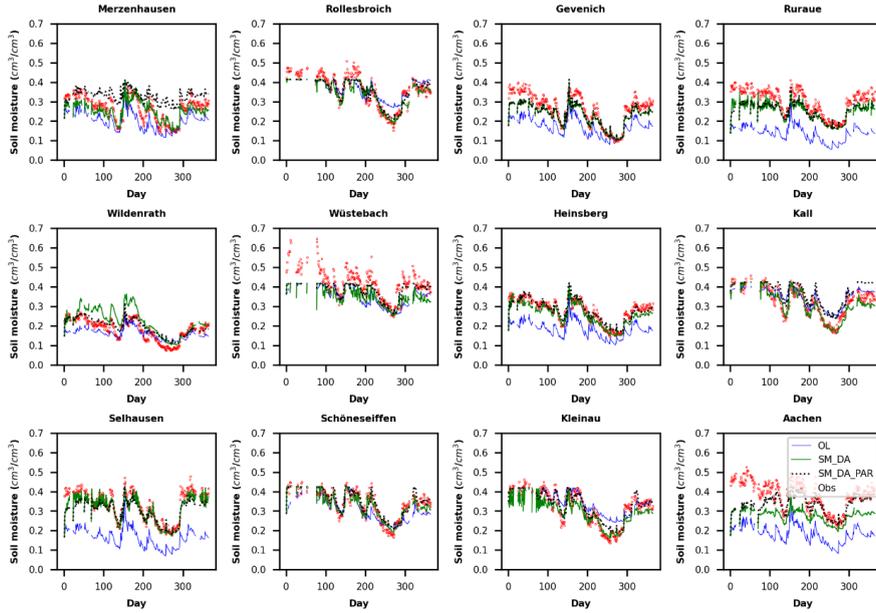


Figure B.1 Temporal evolution of simulated soil moisture from the OL (blue), SM\_DA (green), and SM\_DA\_PAR (black) experiments, together with observed soil moisture from CRNS (Obs, red) for the year 2016 at the CRNS sites. Simulated soil moisture was vertically weighted.

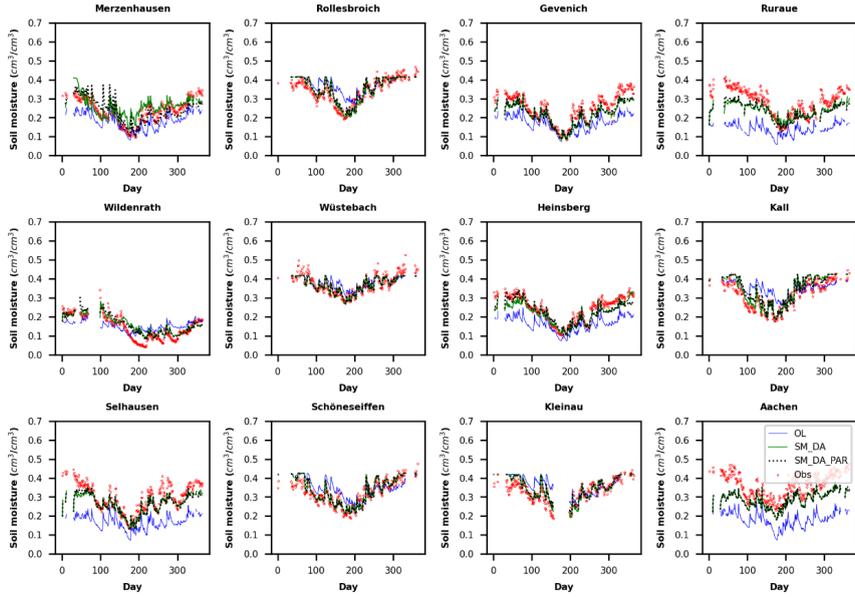


Figure B.2 Temporal evolution of simulated soil moisture from the OL (blue), SM\_DA (green), and SM\_DA\_PAR (black) experiments, together with observed soil moisture from CRNS (Obs, red) for the year 2017 at the CRNS sites. Simulated soil moisture was vertically weighted.

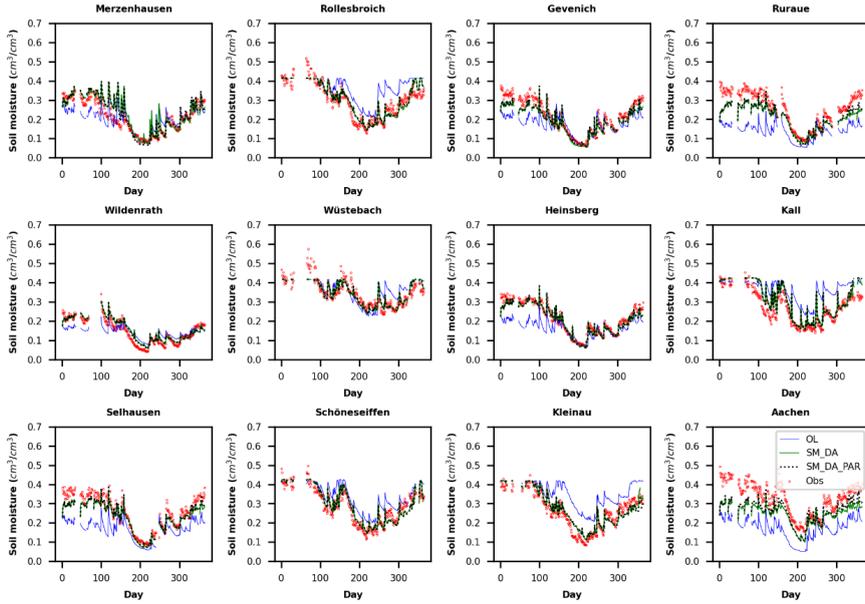


Figure B.3 Temporal evolution of simulated soil moisture from the OL (blue), SM\_DA (green), and SM\_DA\_PAR (black) experiments, together with observed soil moisture from CRNS (Obs, red) for the year 2018 at the CRNS sites. Simulated soil moisture was vertically weighted.

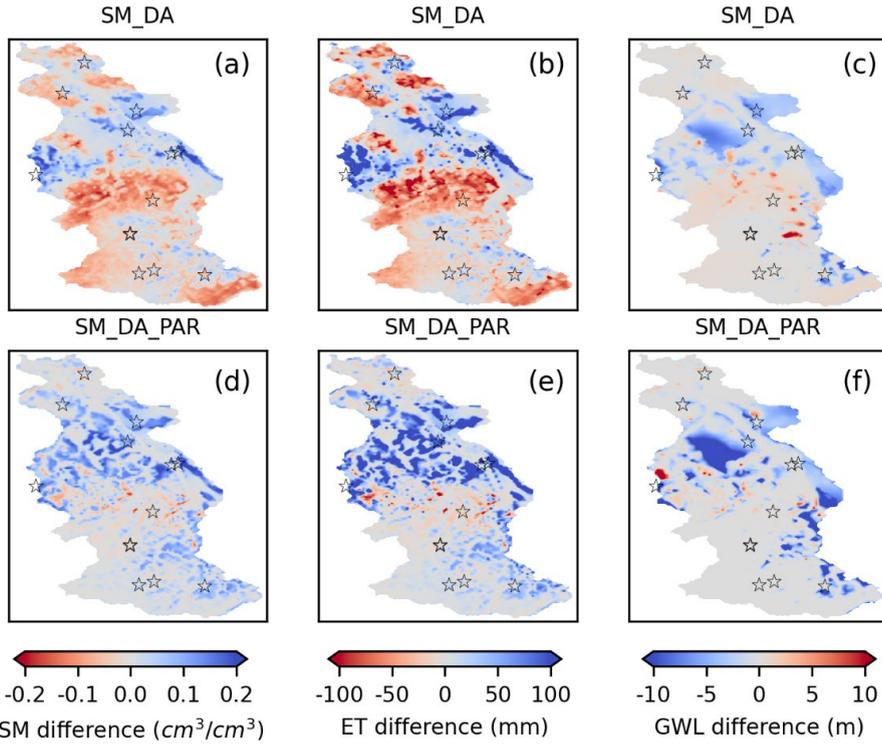


Figure B.4 Subplots (a) to (c) are the differences (SM\_DA - OL) of annual SM (0-80 cm), ET, and GWL in 2016. Subplots (d) to (f) are the differences (SM\_DA\_PAR - OL) of annual SM (0-80 cm), ET, and GWL in 2016. The locations of the CRNS stations are indicated by the black pentagrams.

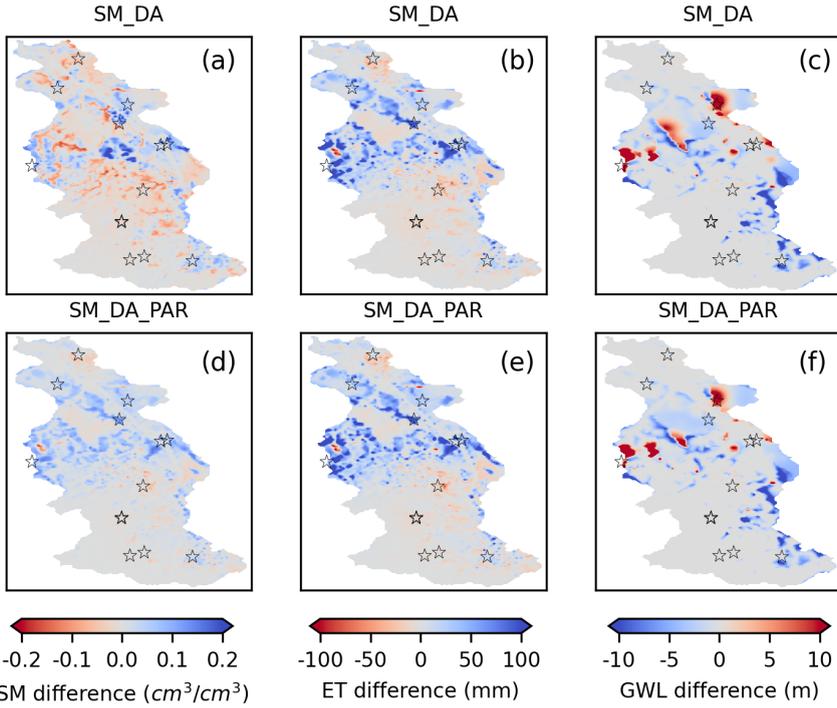


Figure B.5 Subplots (a) to (c) are the differences (SM\_DA - OL) of annual SM (0-80 cm), ET, and GWL in 2017. Subplots (d) to (f) are the differences (SM\_DA\_PAR - OL) of annual SM (0-80 cm), ET, and GWL in 2017. The locations of the CRNS stations are indicated by the black pentagrams.

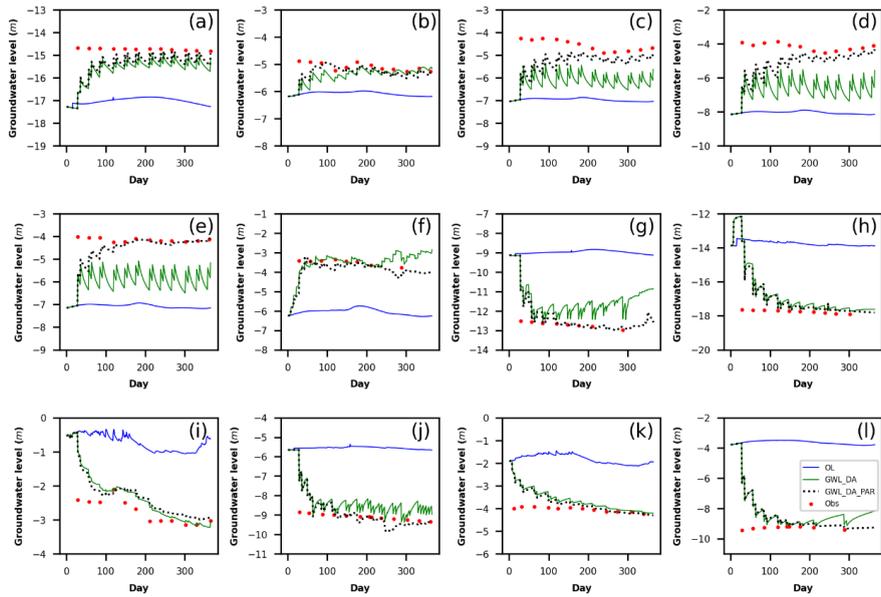


Figure B.6 Temporal evolution of simulated groundwater level from the OL (blue), GWL\_DA (green), and GWL\_DA\_PAR (black) experiments, together with observed groundwater level (Obs, red) at the 12 selected assimilated groundwater sites in 2018.

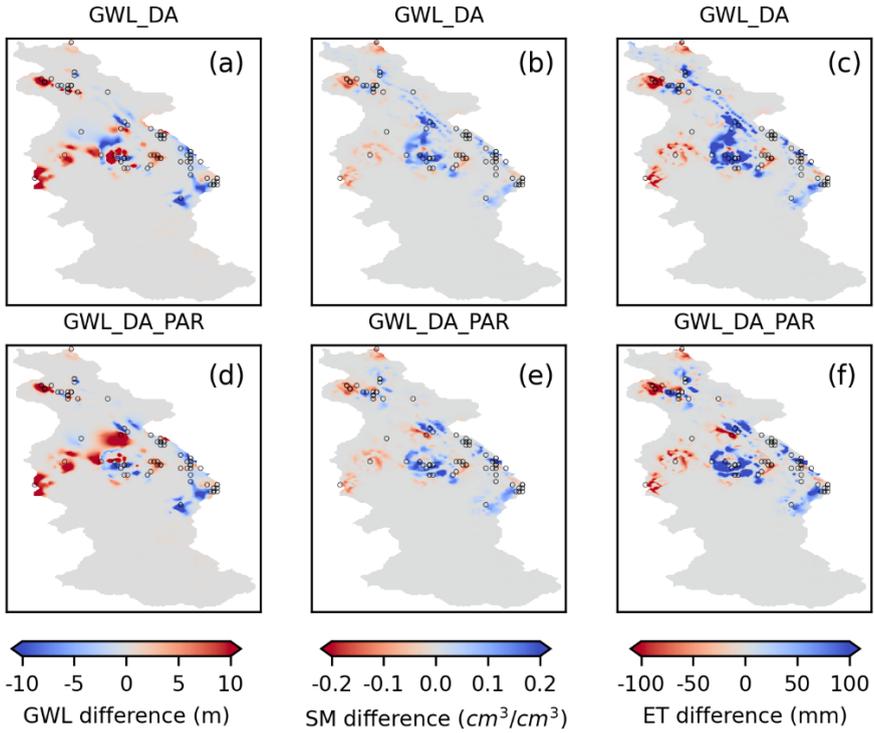


Figure B.7 Subplots (a) to (c) are the differences (GWL\_DA - OL) of annual GWL, SM (0-80 cm), and ET in 2016. Subplots (d) to (f) are the differences (GWL\_DA\_PAR - OL) of annual GWL, SM (0-80 cm), and ET in 2016. The locations of the assimilated groundwater sites are indicated by the black circles.

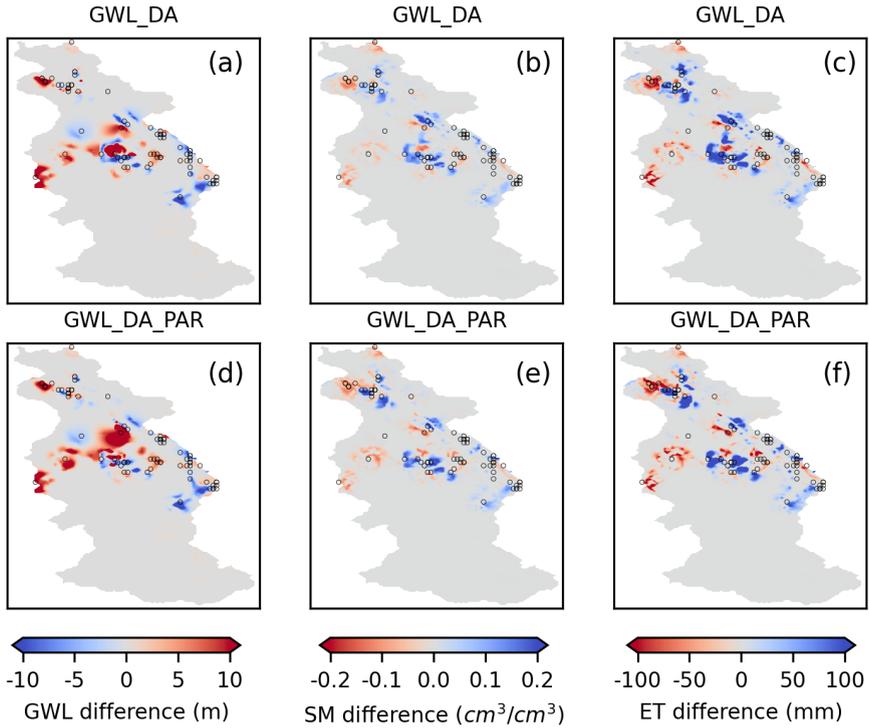


Figure B.8 Subplots (a) to (c) are the differences (GWL\_DA - OL) of annual GWL, SM (0-80 cm), and ET in 2017. Subplots (d) to (f) are the differences (GWL\_DA\_PAR - OL) of annual GWL, SM (0-80 cm), and ET in 2017. The locations of the assimilated groundwater sites are indicated by the black circles.

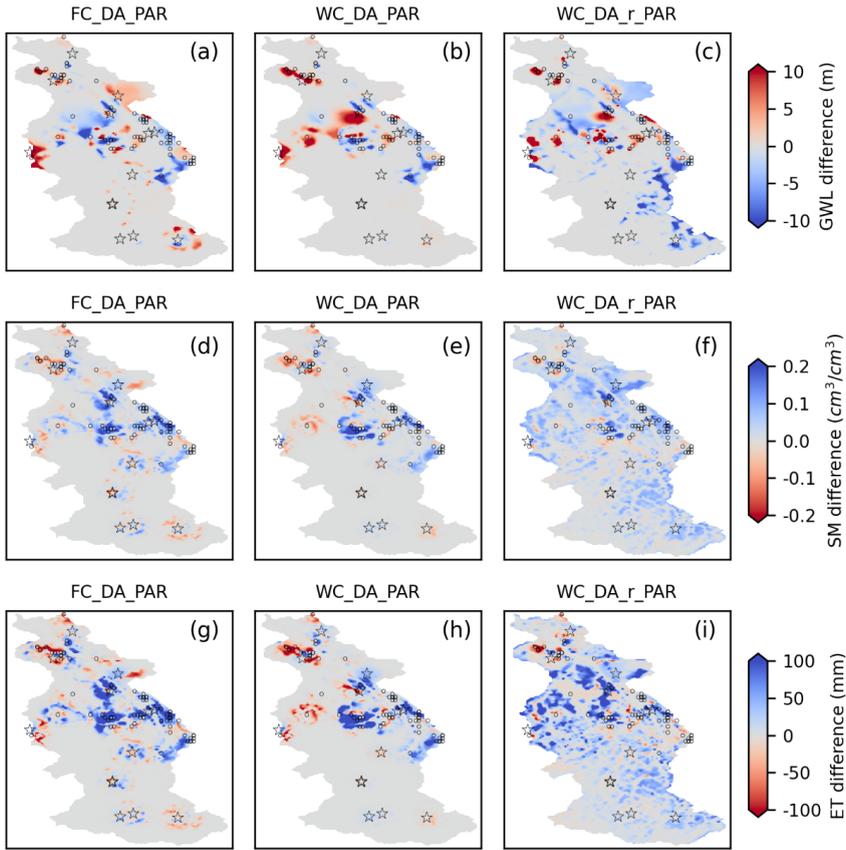


Figure B.9 The differences of annual GWL, SM and ET for OL and different DA experiments in 2016: FC\_DA\_PAR (a, d, g); WC\_DA\_PAR (b, e, h); WC\_DA\_r\_PAR (c, f, i). The red pentagrams and black circles indicate the locations of the CRNS stations and the assimilated groundwater sites, respectively.

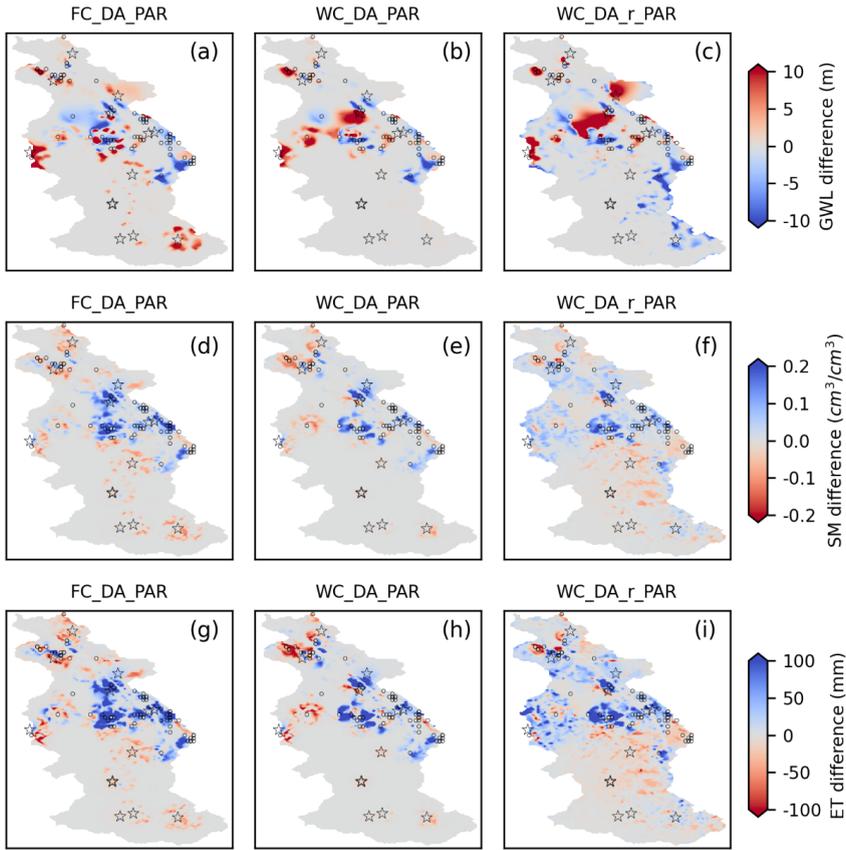


Figure B.10 The differences of annual GWL, SM and ET for OL and different DA experiments in 2017: FC\_DA\_PAR (a, d, g); WC\_DA\_PAR (b, e, h); WC\_DA\_r\_PAR (c, f, i). The red pentagrams and black circles indicate the locations of the CRNS stations and the assimilated groundwater sites, respectively.

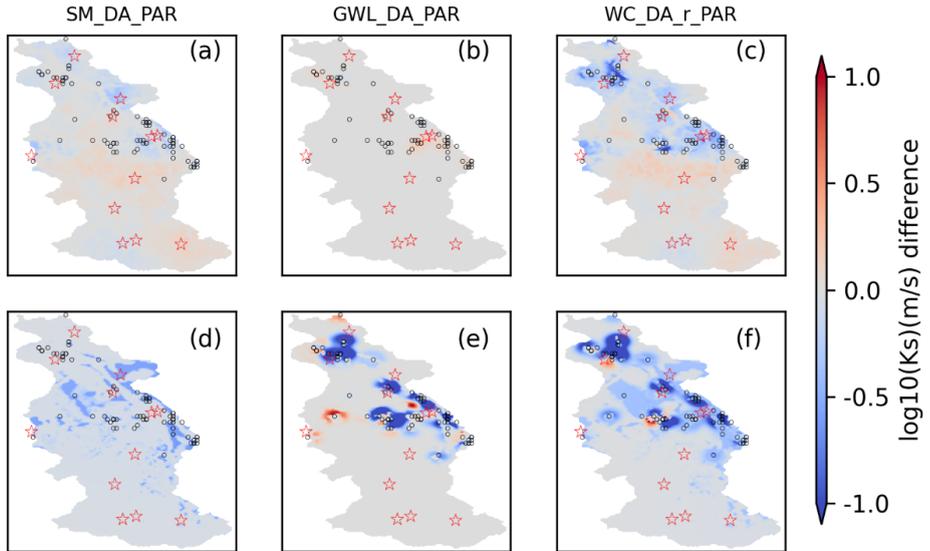


Figure B.11 Differences in ensemble averaged  $\log K_s$  between OL and different DA experiments in 2016: (a) and (d) SM\_DA\_PAR; (b) and (e) GWL\_DA\_PAR; (c) and (f) WC\_DA\_r\_PAR. The first row is from 2 cm depth, and the second row is from 10 m depth. The red pentagrams and black circles indicate the locations of CRNS stations and the assimilated groundwater sites, respectively.

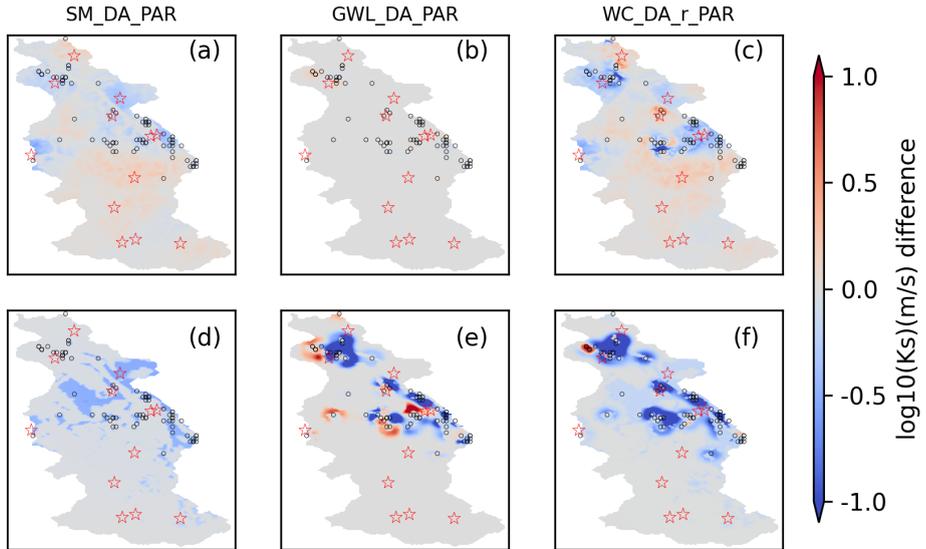


Figure B.12 Differences in ensemble averaged  $\log K_s$  between OL and different DA experiments in 2017: (a) and (d) SM\_DA\_PAR; (b) and (e) GWL\_DA\_PAR; (c) and (f) WC\_DA\_r\_PAR. The first row is from 2 cm depth, and the second row is from 10 m depth. The red pentagrams and black circles indicate the locations of CRNS stations and the assimilated groundwater sites, respectively.

## Chapter 6: Summary and outlook

GWL and RZSM are fundamental components of hydrologic modeling and play key roles in both terrestrial water cycling and broader hydrologic processes. Accurate information on the spatial and temporal variability of GWL and SM in the root zone is essential for a comprehensive understanding of water movement and distribution. This PhD work builds on previous simplified synthetic experiments using the TSMP model for assimilation and continues to use EnKF and its variants to improve the characterization of GWL and SM in a real-world case. The research is carried out for the Rur catchment, situated mainly in western Germany, which covers an area of more than 2000 km<sup>2</sup>. This region shows strong spatial variations in climatic conditions, soil types, and land use. In addition, the catchment has a well-established monitoring infrastructure, including a network of groundwater wells, CRNS stations, eddy covariance stations, and discharge stations. The rich data set provides valuable information for DA with integrated models. The measurement data can be used as assimilation data to correct model predictions and as validation data to evaluate the effectiveness of the assimilation. A central goal of this PhD work is to assimilate GWL and CRNS SM data into the coupled land surface and subsurface model CLM-ParFlow at the large catchment scale to investigate the potential of limited measurements from observation networks to improve the characterization of hydrological variables like SM, GWL, and ET over the entire catchment.

In the first study, real groundwater data were assimilated into the TSMP model via LEnKF for the Rur catchment. The LEnKF is used with a localization radius to avoid unphysical updates associated with spurious correlations. The localization radius was determined by the spatial autocorrelations of the GWL derived from the measurements and the OL run. The DA experiments show that the GWL simulation in the TSMP model can be improved by localized EnKF at the catchment scale. However, the positive impact of assimilation on GWL is limited to the vicinity of the groundwater-assimilated locations. In addition to the fact that a localized EnKF approach was used, the unevenly distributed groundwater observations in the real-world case may be the main reason for the limited improvement. The simulated SM was validated by the CRNS measurements and showed that the simulated SM at the CRNS locations was not much influenced by the GWL assimilation. This relatively low sensitivity is most likely attributed to the quite restrictive localization of groundwater data, i.e., CRNS sites were outside the localization radius of the GWL measurements. In addition, for individual CRNS sites close to the assimilated groundwater sites, the effects of GWL assimilation on the SM simulations were also not significant, which is related to the fact that the SM in the unsaturated zone is only indirectly updated during the assimilation process. This type of weakly coupled DA approach

was used to ensure stability and to avoid the occurrence of anomalous pressure values in the unsaturated zone related to pressure updating during the analysis step.

In the second study, SM data from a distributed network of 12 CRNS in the Rur catchment were assimilated into the TSMP model via EnKF. The DA results showed that a significant improvement in the SM characteristics at the measurement locations can be achieved by assimilating the SM data from CRNS. In addition, the jackknife simulations demonstrated the potential of the CRNS network to improve modeled SM at unassimilated locations throughout the model domain. However, its performance varied between wet and dry years. The improvement during the dry year was less pronounced compared to the wet year. It was related to the fluctuating spatial correlations of SM under different hydrological conditions and was shown to be weakened under drought conditions. This suggests that the CRNS monitoring network ( $\sim 1$  site per  $200 \text{ km}^2$ ) is not dense enough to effectively cover the Rur catchment. In addition, evaluations of ET and discharge from observational data showed that CRNS assimilation has the potential to improve flux estimates in the integrated terrestrial model, but to a lesser extent, indicating limited sensitivity of ET and discharge to SM. Joint state parameter estimation improved the modeling of hydrologic variables more effectively than state estimation alone, demonstrating the necessity and importance of parameter estimation. This study demonstrates the potential of a CRNS observational network to improve the estimation of hydrological variables and parameters at a larger catchment scale and suggests promising prospects for the application of CRNS. Given the findings that the CRNS monitoring network was not dense enough to effectively cover the whole catchment, it is suggested that the density of CRNS stations be increased to reduce the spacing between monitoring sites, thereby improving spatial coverage and the accuracy of hydrological simulations. In addition, a higher model resolution should be used in the future, as the 500 m resolution is still coarse compared to the footprint of the CRNS measurement. Spatial weighting should also be considered during the assimilation process to compare simulated SM and SM measured by the CRNS probe to account for vertical (and possible also horizontal) spatial heterogeneity in the SM distribution. This can maximize the influence of the data on the model state estimation to improve the accuracy and representativeness of the assimilated data, leading to more robust model predictions.

In the third study, GWL and SM were assimilated into the TSMP model in the Rur catchment using the newly proposed multivariate assimilation method. Univariate assimilation was also performed separately for GWL and SM for comparison. The results showed that the univariate assimilation of GWL or SM only improves the predictions of the assimilated variable

but might even deteriorate the simulation of the other variables (SM or GWL), which may be related to unrealistic cross-variable correlations, small ensemble sizes, time-lagged responses in the system, and model structural errors. Therefore, there is a need to explore the interactions between different variables in the integrated model of the assimilation process and to propose better DA strategies to avoid the degradation of other variables. By employing cross-correlation analysis, time-series regression, wavelet analysis, or other techniques, the time-lagged correlations and dependencies between variables can be identified, which is helpful to handle the time-dependent relationships between variables during assimilation. Inverse modeling (e.g., Iterative Ensemble Smoother) provides a viable way to mitigate the risk of degrading simulations of other variables as it also considers explicit temporal correlations, potentially improving overall model performance. However, it is important to recognize that there are challenges to this approach, such as the need to thoroughly consider various factors, ensuring an adequate ensemble size to accurately capture the complexity of the system, and managing the expensive computational costs associated with iterative processes, especially for large-scale models. Overall, while the Iterative Ensemble Smoother is a promising approach for improving predictions and accounting for time-lagged responses, it is essential to consider both its benefits and challenges when designing an effective assimilation framework. The negative effect may also be related to the use of real in situ measurements. CRNS and groundwater sites are scattered throughout the catchment. Thus, the assimilation of a single variable (e.g., GWL) makes it difficult to improve the estimation of variables (e.g., SM) which are evaluated at other distant grid points. However, the estimates of GWL and SM can be significantly and simultaneously improved using the new multivariate DA approach. The advantage of the new approach is the use of separate updates during the multivariate assimilation, allowing the saturated zone pressure update based on GWL measurements and the unsaturated zone SM update based on SM measurements. As a result, the predictions of multiple variables can be corrected sequentially in a coupled model using different data sources, minimizing the risk of confounding factors and reducing the impact of spurious correlations. Furthermore, the characterization of ET can only be improved by assimilating SM, whether through univariate assimilation or multivariate assimilation, reflecting the direct effects of SM on ET. The low sensitivity of the GWL assimilation to ET simulation is primarily due to the fact that groundwater effects on ET are generally indirect and slow. In addition, joint state and parameter estimation consistently yielded better results than updating the state alone, and the updated parameters outperformed the original parameters in validation. Thus, the benefits of parameter estimation were double-confirmed. Overall, the combined assimilation of GWL and SM showed superior performance compared to the separate assimilation of GWL or SM in the

integrated model. Therefore, this study also illustrates the feasibility of integrating a novel CRNS soil moisture observation network with a traditional groundwater monitoring network to improve the modeling of an integrated system.

The overall results of this thesis show that the use of DA to improve the simulation of important variables such as SM and GWL within integrated models is a promising approach. However, it is important to recognize that the effectiveness of DA is limited by the complexity of the real-world catchment. Examples are: (i) Groundwater wells were mostly clustered in the northern flat plains and rarely in the southern mountains. The uneven distribution of monitoring stations resulted in an improvement of the GWL simulation only in certain areas within the catchment rather than over the entire region. In addition, the uneven distribution of the different types of monitoring stations makes it difficult to verify the actual simulation impact on other variables in the model. For example, when the CRNS station is not within the localization radius of the assimilated groundwater site, it is difficult to clarify the interactions related to assimilation. (ii) The horizontal spatial resolution used for the coupled model in the study area is set at 500 m, taking into account several factors, including the resolution of atmospheric forcing (6 km) and soil texture (1:50,000), as well as the computational costs associated with performing integrated model ensemble runs on supercomputers and the significant amount of computer storage required. However, the model resolution of 500 m is still too coarse for the Rur catchment to represent small valleys and hill slopes. The coarse resolution limits the representation of topographic features and cannot accurately capture gravity-driven lateral flow processes. Downscaling the input data (atmospheric forcing and subsurface hydraulic parameters) to a higher resolution can mitigate the problem of the model being too coarse. However, it is important to note that this may introduce some approximation errors. (iii) The existence of multiple and complex aquifers and artificial water extraction has a significant impact on groundwater flow and groundwater levels, which is not accounted for in the model. By incorporating detailed geological information, such as permeability, we can better characterize aquifers and understand their behavior. However, it is essential to recognize that geological models are highly uncertain due to the inherent heterogeneity and complexity of subsurface formations. Furthermore, the performance of the integrated model in combination with DA in practical applications is affected by model structural errors and multiple uncertain parameters, which affect, for example, the simulation of ET. Inaccurate LAI parameters are likely a contributing factor to the systematic underestimation of ET in this study. The omission of irrigation in the model and the use of simplified crop parameters can also affect the accuracy of evapotranspiration simulations.

For further studies of DA in combination with integrated models, the following possible aspects can be considered to improve the reliability and accuracy of the assimilation:

(i) The coupled model could be extended to the entire hydrological terrestrial system by adding the atmospheric model component. Atmospheric models provide higher spatial and temporal resolution compared to reanalysis data, allowing a more detailed representation of atmospheric processes. The finer resolution allows better characterization of local atmospheric dynamics, which can significantly influence hydrological processes. When integrated with atmospheric models, the coupled model can capture complex interactions between the atmosphere and the land surface, including feedback mechanisms such as land-atmosphere coupling and the impact of land surface properties on atmospheric conditions. In contrast, reanalysis data can only provide the boundary conditions for the land surface without interacting with it. In this way, the exchange of energy and water between different components of the Earth system cannot be fully accounted for, leading to less accurate predictions of hydrological variables. Integrating an atmospheric model with hydrologic modeling allows for a more comprehensive representation of atmospheric processes and their interactions with the land surface and subsurface. For example, it is possible to use the atmospheric module already included in the TSMP platform. Extending the parameterization of the land surface model to include different vegetation types and crops allows for more detailed consideration of agroecosystem processes. Version 5.0 of the CLM can be used instead of version 3.5 in the TSMP model. The new version 5.0 provides a more nuanced description of vegetation characteristics and dynamics, which helps to simulate land-atmosphere interactions more accurately. An updated version of the TSMP model, version 2.0, will include CLM 5.0, and this new modeling platform is in the testing phase and soon will be available for broader use. The mismatch of spatio-temporal scales between different processes should be carefully addressed. The new model platform should be applied at different scales (i.e., global, regional, and catchment scales) to test the validity and suitability of the model for various fields and scales.

(ii) Due to the scarce in situ hydrological observations, especially in remote regions, more diverse types of observations should be considered for DA. The data to be assimilated can be collected at different scales and from different sources, like RS products. For SM, the Soil Moisture Active Passive (SMAP, 9 km) mission (Entekhabi et al., 2010) and the European Radar Observatory Sentinel-1 (1 km) (Torres et al., 2012; Bauer-Marschallinger et al., 2019; Balenzano et al., 2021) provide global high-resolution SM datasets that closely match the scale of catchments. These two datasets are well suited for DA in integrated models. When it comes to LAI, most existing products use MODIS and AVHRR satellite data to produce LAI using

different approaches and methodologies (Fang et al., 2019; Bayat et al., 2021), such as GLOBMAP V1 (500 m) (Liu et al., 2012) and MOD15A2H (500 m) (Myneni et al., 2015). With the Sentinel-2 Level 2 Prototype Processor (SL2P: Weiss and Baret, 2020) embedded in the Sentinel Application Platform (SNAP), the LAI product can be derived at 20 m resolution, offering new potential for monitoring plant characteristics at fine spatial resolution and allowing more accurate parameterization of vegetation in terrestrial models. In addition, a new generation of measurements from UAVs and small satellites can overcome the disadvantages of traditional sensing techniques and provide data with higher spatial and temporal resolution at local and catchment scales. Recently, ultralight Light Detection and Ranging (LiDAR) systems have been able to be mounted on UAVs, offering affordability for measuring various variables such as vegetation density, canopy height, and LAI (You et al., 2017; ten Harkel et al., 2019; Bates et al., 2021). UAVs can obtain observations at fine spatial scales and with frequent temporal resolution, allowing for localized measurements within specific regions or catchments. These localized observations can complement existing RS data and ground-based measurements, providing additional constraints for model assimilation and improving the representation of small-scale processes. In addition to directly updating state observations, other related variables or parameters (e.g., hydraulic and vegetation related parameters) can be updated to improve model estimates. It is important to note that the assimilation of new observations requires the development of new suitable observation operators and the assignment of appropriate uncertainties to different types of measurements. In multivariate assimilation, it is important to reconcile spatial and temporal scale discrepancies between different types of observations, as well as scale discrepancies between observations and model simulations. For example, different update frequencies can be used for different variables based on observed data to overcome inconsistent temporal resolution, and interpolation and weighting calculations can be used to compensate for discrepancies in spatial scale. The assimilation of multivariate and multi-source observations is expected to further improve the estimation of model states and parameters and increase the accuracy of integrated model predictions.

(iii) Explore advanced DA algorithms for solving multivariate DA problems. Based on our study, it is still suggested to consider the EnKF with localization. By incorporating localization, we can focus the assimilation process on relevant variables within spatial areas, thus reducing computational costs and improving the accuracy of the estimates. Although the assignment of different localization radii for different variables proposed in this study alleviates the impact of spurious cross-correlations among different variables, correlations might also change in time and could be taken into account via temporally variable localization radii. Furthermore, because multi-source DA involves the integration of information from multiple sources (e.g., satellite

observations and ground-based measurements), there are trade-offs in deciding how to effectively combine these different sources of information to improve the overall quality of the assimilated data. Sensitivity analysis or optimization techniques can be used to quantify the relative importance of different data sources, and the desired balance can be achieved between the various assimilated data sets. Hybrid approaches to DA with EnKF and other filters, such as Ensemble Variational (En-Var) DA, could also be beneficial for multivariate assimilation (Bannister, 2017; Pinnington et al., 2020). En-Var combines the strengths of EnKF and variational approaches by using ensemble-based background error covariance estimation within the variational framework. This approach provides the flexibility to handle different types of observations and model uncertainties and promises to effectively address the challenges of multivariate assimilation in a modeling framework while maintaining computational efficiency. Filter performance could also be improved through ensemble generation, parameter tuning, and inflation. In addition, DA can be integrated with new data analytics, such as ML, to address the existing challenges of large-scale DA. ML algorithms can learn complex patterns and relationships between input variables and model outputs. Therefore, ML techniques can be used to replace certain processes of the forward model, especially those that are computationally expensive or poorly understood. Integrating the hydrologic model with ML can improve the model's ability to describe hydrologic processes and provide better predictions for various hydrologic variables. In addition, the integration of ML with Bayesian techniques can serve as the analysis process in DA. For example, in DA, the observation operator is used to map model forecasts to the observation space, which is usually assumed to be linear but often exhibits nonlinear behavior. ML techniques can provide non-linear relationships for the observation operator, thereby enhancing Bayesian techniques with a more accurate representation of these relationships and leading to better assimilation results. In addition, using Bayesian techniques alone to estimate parameters may require manual adjustments to better calibrate the parameters if the model is very complex or the parameter space is very large. In such cases, determining the optimal parameter settings may require specialized knowledge and experience. On the contrary, ML optimization methods are usually adaptive, allowing automatic parameter adjustments, thus reducing the need for manual tuning. Therefore, ML techniques can replace the parameter- updating part of Bayesian techniques. More flexible calibration of model parameters can lead to better model performance and significantly reduce computational complexity and cost. The integration of ML with DA in integrated models offers a promising approach for advancing our understanding of complex environmental systems and improving the reliability of model predictions. However, it is important to note that ML does not necessarily replace the entire DA process or forward model. Instead, ML techniques

complement the inadequacies of hydrological models and DA methods. Experimentation and validation on real data are crucial to identifying the most effective strategies for a particular modeling system. The ability to simultaneously improve all relevant state and parameter estimates in the system is the ultimate goal.

## Bibliography

- Abdalati, W., Zwally, H.J., Bindschadler, R., Csatho, B., Farrell, S.L., Fricker, H.A., Harding, D., Kwok, R., Lefsky, M., Markus, T., Marshak, A., Neumann, T., Palm, S., Schutz, B., Smith, B., Spinirne, J., Webb, C., 2010. The ICESat-2 Laser Altimetry Mission. *Proceedings of the IEEE*, 98(5): 735-751. 10.1109/JPROC.2009.2034765.
- Adams, K.H., Reager, J.T., Rosen, P., Wiese, D.N., Farr, T.G., Rao, S., Haines, B.J., Argus, D.F., Liu, Z., Smith, R., Famiglietti, J.S., Rodell, M., 2022. Remote Sensing of Groundwater: Current Capabilities and Future Directions. *Water Resources Research*, 58(10). 10.1029/2022wr032219.
- Ajami, H., McCabe, M.F., Evans, J.P., Stisen, S., 2014. Assessing the impact of model spin-up on surface water-groundwater interactions using an integrated hydrologic model. *Water Resources Research*, 50(3): 2636-2656. 10.1002/2013wr014258.
- Allan, R., Pereira, L., Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56.
- Anderson, M.P., Woessner, W.W., Hunt, R.J., 2015. Chapter 2 - Modeling Purpose and Conceptual Model. In: Anderson, M.P., Woessner, W.W., Hunt, R.J. (Eds.), *Applied Groundwater Modeling (Second Edition)*. Academic Press, San Diego, pp. 27-67. <https://doi.org/10.1016/B978-0-08-091638-5.00002-X>.
- Andreasen, M., Jensen, K.H., Desilets, D., Zreda, M., Bogaen, H.R., Looms, M.C., 2017. Cosmic-ray neutron transport at a forest field site: the sensitivity to various environmental conditions with focus on biomass and canopy interception. *Hydrology and Earth System Sciences*, 21(4): 1875-1894. 10.5194/hess-21-1875-2017.
- 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. <https://doi.org/10.13182/nse96-a24230>.
- Baalousha, H., 2008. *Fundamentals of groundwater modelling, Ground-water: modelling, management and contamination*. Nova Science Publishers, New York, pp. 113-130.
- Baatz, D., Kurtz, W., Hendricks Franssen, H.J., Vereecken, H., Kollet, S.J., 2017. Catchment tomography - An approach for spatial parameter estimation. *Advances in Water Resources*, 107: 147-159. <https://doi.org/10.1016/j.advwatres.2017.06.006>.
- Baatz, R., Bogaen, H.R., Hendricks Franssen, H.J., Huisman, J.A., Qu, W., Montzka, C., Vereecken, H., 2014. Calibration of a catchment scale cosmic-ray probe network: A comparison of three parameterization methods. *Journal of Hydrology*, 516: 231-244. 10.1016/j.jhydrol.2014.02.026.
- Baatz, R., Hendricks Franssen, H.-J., Han, X., Hoar, T., Bogaen, H.R., Vereecken, H., 2017. Evaluation of a cosmic-ray neutron sensor network for improved land surface model prediction. *Hydrology and Earth System Sciences*, 21(5): 2509-2530. <https://doi.org/10.1016/j.advwatres.2017.06.006>
- Babacian, E., Sadeghi, M., Jones, S.B., Montzka, C., Vereecken, H., Tuller, M., 2019. Ground, Proximal, and Satellite Remote Sensing of Soil Moisture. *Reviews of Geophysics*, 57(2): 530-616. 10.1029/2018rg000618.
- 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. <https://doi.org/10.1175/mwr-d-10-05013.1>.
- Balenzano, A., Mattia, F., Satalino, G., Lovregine, F.P., Palmisano, D., Peng, J., Marzahn, P., Wegmüller, U., Cartus, O., Dąbrowska-Zielińska, K., Musial, J.P., Davidson, M.W.J., Pauwels, V.R.N., Cosh, M.H., McNairn, H., Johnson, J.T., Walker, J.P., Yueh, S.H., Entekhabi, D., Kerr, Y.H., Jackson, T.J., 2021. Sentinel-1 soil moisture at 1 km resolution: a validation study. *Remote Sensing of Environment*, 263: 112554. 10.1016/j.rse.2021.112554.
- Bannister, R.N., 2017. A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 143(703): 607-633. 10.1002/qj.2982.
- Baroni, G., Zink, M., Kumar, R., Samaniego, L., Attinger, S., 2017. Effects of uncertainty in soil properties on simulated hydrological states and fluxes at different spatio-temporal scales. *Hydrology and Earth System Sciences*, 21(5): 2301-2320. 10.5194/hess-21-2301-2017.
- Bates, J.S., Montzka, C., Schmidt, M., Jonard, F., 2021. Estimating Canopy Density Parameters Time-Series for Winter Wheat Using UAS Mounted LiDAR. *Remote Sensing*, 13(4): 710. 10.3390/rs13040710.
- Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., Modanesi, S., Massari, C., Ciabatta, L., Brocca, L., Wagner, W., 2019. Toward Global Soil Moisture Monitoring With Sentinel-1: Harnessing Assets and Overcoming Obstacles. *IEEE Transactions on Geoscience and Remote Sensing*, 57(1): 520-539. 10.1109/TGRS.2018.2858004.
- Bayat, B., Camacho, F., Nickeson, J., Cosh, M., Bolten, J., Vereecken, H., Montzka, C., 2021. Toward operational validation systems for global satellite-based terrestrial essential climate variables. *International Journal of Applied Earth Observation and Geoinformation*, 95: 102240. 10.1016/j.jag.2020.102240.
- Bayat, B., van der Tol, C., Verhoef, W., 2018. Retrieval of land surface properties from an annual time series of Landsat TOA radiances during a drought episode using coupled radiative transfer models. *Remote Sensing of Environment*, 238: 110917. 10.1016/j.rse.2018.09.030.

- Bear, Jacob, 1975. Dynamics of Fluids in Porous Media. *Soil Science* 120: 162-163.
- Bear, J., Verruijt, A., 1987. Modeling groundwater flow and pollution, United States. <https://doi.org/10.1007/978-1-4020-6682-5>.
- Bisht, G., Riley, W.J., 2019. Development and Verification of a Numerical Library for Solving Global Terrestrial Multiphysics Problems. *Journal of Advances in Modeling Earth Systems*, 11(6): 1516-1542. [10.1029/2018ms001560](https://doi.org/10.1029/2018ms001560).
- Bizhanimanzar, M., Leconte, R., Nuth, M., 2019. Modelling of shallow water table dynamics using conceptual and physically based integrated surface-water-groundwater hydrologic models. *Hydrology and Earth System Sciences*, 23(5): 2245-2260. [10.5194/hess-23-2245-2019](https://doi.org/10.5194/hess-23-2245-2019).
- Blyth, E.M., Arora, V.K., Clark, D.B., Dadson, S.J., De Kauwe, M.G., Lawrence, D.M., Melton, J.R., Pongratz, J., Turton, R.H., Yoshimura, K., Yuan, H., 2021. Advances in Land Surface Modelling. *Current Climate Change Reports*, 7(2): 45-71. [10.1007/s40641-021-00171-5](https://doi.org/10.1007/s40641-021-00171-5).
- Boas, T., Bogena, H., Grünwald, T., Heinesch, B., Ryu, D., Schmidt, M., Vereecken, H., Western, A., Hendricks Franssen, H.-J., 2021. Improving the representation of cropland sites in the Community Land Model (CLM) version 5.0. *Geoscientific Model Development*, 14(1): 573-601. [10.5194/gmd-14-573-2021](https://doi.org/10.5194/gmd-14-573-2021).
- Bogena, H., Kunkel, R., Schöbel, T., Schrey, H.P., Wendland, F., 2005. Distributed modeling of groundwater recharge at the macroscale. *Ecological Modelling*, 187(1): 15-26. [10.1016/j.ecolmodel.2005.01.023](https://doi.org/10.1016/j.ecolmodel.2005.01.023).
- Bogena, H.R., Herbst, M., Hake, J.F., Kunkel, R., Montzka, C., P., T., Vereecken, H., and Wendland, F., 2005. MOSYRUR: Water balance analysis in the Rur basin. *Forschungszentrum Jülich, Jülich*.
- Bogena, H.R., Herrmann, F., Jakobi, J., Brogi, C., Ilias, A., Huisman, J.A., Panagopoulos, A., Pisinaras, V., 2020. Monitoring of Snowpack Dynamics With Cosmic-Ray Neutron Probes: A Comparison of Four Conversion Methods. *Frontiers in Water*, 2. [10.3389/frwa.2020.00019](https://doi.org/10.3389/frwa.2020.00019).
- Bogena, H.R., Huisman, J.A., Baatz, R., Hendricks Franssen, H.J., Vereecken, H., 2013. Accuracy of the cosmic-ray soil water content probe in humid forest ecosystems: The worst case scenario. *Water Resources Research*, 49(9): 5778-5791. [10.1002/wrcr.20463](https://doi.org/10.1002/wrcr.20463).
- Bogena, H.R., Huisman, J.A., Güntner, A., Hübner, C., Kusche, J., Jonard, F., Vey, S., Vereecken, H., 2015. Emerging methods for noninvasive sensing of soil moisture dynamics from field to catchment scale: a review. *Wiley Interdisciplinary Reviews-Water*, 2(6): 635-647. [10.1002/wat2.1097](https://doi.org/10.1002/wat2.1097).
- Bogena, H.R., Montzka, C., Huisman, J.A., Graf, A., Schmidt, M., Stockinger, M., von Hebel, C., Hendricks-Franssen, H.J., van der Kruk, J., Tappe, W., Lücke, A., Baatz, R., Bol, R., Groh, J., Pütz, T., Jakobi, J., Kunkel, R., Sorg, J., Vereecken, H., 2018. The TERENO-Rur Hydrological Observatory: A Multiscale Multi-Compartment Research Platform for the Advancement of Hydrological Science. *Vadose Zone Journal*, 17(1): 180055. [10.2136/vzj2018.03.0055](https://doi.org/10.2136/vzj2018.03.0055).
- Bogena, H.R., Schrön, M., Jakobi, J., Ney, P., Zacharias, S., Andreasen, M., Baatz, R., Boorman, D., Duygu, M.B., Eguibar-Galán, M.A., Fersch, B., Franke, T., Geris, J., González Sanchis, M., Kerr, Y., Korf, T., Mengistu, Z., Mialon, A., Nasta, P., Nityechoruk, J., Pisinaras, V., Rasche, D., Rosolem, R., Said, H., Schattan, P., Zreda, M., Achleitner, S., Albentosa-Hernández, E., Akyürek, Z., Blume, T., del Campo, A., Canone, D., Dimitrova-Petrova, K., Evans, J.G., Ferraris, S., Frances, F., Gisolo, D., Güntner, A., Herrmann, F., Iwema, J., Jensen, K.H., Kunstmann, H., Lidón, A., Looms, M.C., Oswald, S., Panagopoulos, A., Patil, A., Power, D., Rebmann, C., Romano, N., Scheffele, L., Seneviratne, S., Weltin, G., Vereecken, H., 2022. COSMOS-Europe: a European network of cosmic-ray neutron soil moisture sensors. *Earth System Science Data*, 14(3): 1125-1151. [10.5194/essd-14-1125-2022](https://doi.org/10.5194/essd-14-1125-2022).
- Bollmeyer, C., Keller, J.D., Ohlwein, C., Wahl, S., Crewell, S., Friederichs, P., Hense, A., Keune, J., Kneifel, S., Pscheidt, I., Redl, S., Steinke, S., 2015. Towards a high-resolution regional reanalysis for the European CORDEX domain. *Quarterly Journal of the Royal Meteorological Society*, 141(686): 1-15. <https://doi.org/10.1002/qj.2486>.
- Bonan, G.B. 1996. "A Land Surface Model (LSM Version 1.0) for Ecological, Hydrological, and Atmospheric Studies: Technical Description and User's Guide". University Corporation for Atmospheric Research. <https://doi.org/10.5065/D6DF6P5X>.
- Bonan, G.B., Levis, S., Kergoat, L., Oleson, K.W., 2002. Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Global Biogeochemical Cycles*, 16(2): 5-15-23. [10.1029/2000gb001360](https://doi.org/10.1029/2000gb001360).
- Botto, A., Belluco, E., Camporese, M., 2018. Multi-source data assimilation for physically based hydrological modeling of an experimental hillslope. *Hydrology and Earth System Sciences*, 22(8): 4251-4266. [10.5194/hess-22-4251-2018](https://doi.org/10.5194/hess-22-4251-2018).
- Brandhorst, N., Erdal, D., Neuweiler, I., 2017. Soil moisture prediction with the ensemble Kalman filter: Handling uncertainty of soil hydraulic parameters. *Advances in Water Resources*, 110: 360-370. [10.1016/j.advwatres.2017.10.022](https://doi.org/10.1016/j.advwatres.2017.10.022).
- Brandhorst, N., Neuweiler, I., 2023. Impact of parameter updates on soil moisture assimilation in a 3D heterogeneous hillslope model. *Hydrology and Earth System Sciences*, 27(6): 1301-1323. [10.5194/hess-27-1301-2023](https://doi.org/10.5194/hess-27-1301-2023).
- Brocca, L., Melone, F., Moramarco, T., Wagner, W., Nacimi, V., Bartalis, Z., Hasenauer, S., 2010. Improving runoff prediction through the assimilation of the ASCAT soil moisture product. *Hydrology and Earth System Sciences*, 14(10): 1881-1893. [10.5194/hess-14-1881-2010](https://doi.org/10.5194/hess-14-1881-2010).
- Butts, M., Graham, D., 2005. Flexible Integrated Watershed Modeling with MIKE SHE. 245-271. [10.1201/9781420037432.ch10](https://doi.org/10.1201/9781420037432.ch10).

- Camporese, M., Paniconi, C., Putti, M., Salandin, P., 2009a. Comparison of Data Assimilation Techniques for a Coupled Model of Surface and Subsurface Flow. *Vadose Zone Journal*, 8(4): 837-845. <https://doi.org/10.2136/vzj2009.0018>
- Camporese, M., Paniconi, C., Putti, M., Salandin, P., 2009b. Ensemble Kalman filter data assimilation for a process-based catchment scale model of surface and subsurface flow. *Water Resources Research*, 45(10): 1-14. <https://doi.org/10.1029/2008wr007031>.
- Chen, J., Knight, R., Zebker, H.A., 2017. The Temporal and Spatial Variability of the Confined Aquifer Head and Storage Properties in the San Luis Valley, Colorado Inferred From Multiple InSAR Missions. *Water Resources Research*, 53(11): 9708-9720. [10.1002/2017wr020881](https://doi.org/10.1002/2017wr020881).
- Chen, X., Hu, Q., 2004. Groundwater influences on soil moisture and surface evaporation. *Journal of Hydrology*, 297(1-4): 285-300. <https://doi.org/10.1016/j.jhydrol.2004.04.019>.
- 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. [10.1016/j.advwatres.2005.09.007](https://doi.org/10.1016/j.advwatres.2005.09.007).
- Chevalking, L.K.; van Steenbergen Description, F. Ideas for Groundwater Management. Available online: [https://metameta.nl/wp-content/uploads/2010/05/Ideas\\_for\\_Groundwater\\_Management\\_Digital.pdf](https://metameta.nl/wp-content/uploads/2010/05/Ideas_for_Groundwater_Management_Digital.pdf)
- Clapp, R.B., Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. *Water Resources Research*: 14(4), 601-604. <https://doi.org/10.1029/wr014i004p00601>.
- Condon, L.E., Kollet, S., Bierkens, M.F.P., Fogg, G.E., Maxwell, R.M., Hill, M.C., Fransen, H.J.H., Verhoef, A., Van Loon, A.F., Sulis, M., Abesser, C., 2021. Global Groundwater Modeling and Monitoring: Opportunities and Challenges. *Water Resources Research*, 57(12). [10.1029/2020wr029500](https://doi.org/10.1029/2020wr029500).
- Cox, P.M., Betts, R.A., Bunton, C.B., Essery, R.L.H., Rowntree, P.R., Smith, J., 1999. The impact of new land surface physics on the GCM simulation of climate and climate sensitivity. *Climate Dynamics*, 15(3): 183-203. [10.1007/s003820050276](https://doi.org/10.1007/s003820050276).
- Crow, W.T., Berg, A.A., Cosh, M.H., Loew, A., Mohanty, B.P., Panciera, R., de Rosnay, P., Ryu, D., Walker, J.P., 2012. Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Reviews of Geophysics*, 50(2). [10.1029/2011rg000372](https://doi.org/10.1029/2011rg000372).
- Darcy, H., 1856. The public fountains of the city of Dijon. Victor Dalmont, Paris, France.
- De Lannoy, G.J.M., Houser, P.R., Pauwels, V.R.N., Verhoest, N.E.C., 2007. State and bias estimation for soil moisture profiles by an ensemble Kalman filter: Effect of assimilation depth and frequency. *Water Resources Research*, 43(6). [10.1029/2006wr005100](https://doi.org/10.1029/2006wr005100).
- De Lannoy, G.J.M., Reichle, R.H., 2016. Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model. *Hydrology and Earth System Sciences*, 20(12): 4895-4911. [10.5194/hess-20-4895-2016](https://doi.org/10.5194/hess-20-4895-2016).
- De Lannoy, G.J.M., Reichle, R.H., Vrugt, J.A., 2014. Uncertainty quantification of GEOS-5 L-band radiative transfer model parameters using Bayesian inference and SMOS observations. *Remote Sensing of Environment*, 148: 146-157. <https://doi.org/10.1016/j.rse.2014.03.030>.
- de Marsily, G., 1986. Quantitative Hydrogeology: Groundwater Hydrology for Engineers. Academic Press, Orlando, FL.
- Diersch, H.-J.G., 2014. FEFLOW: Finite Element Modeling of Flow, Mass and Heat Transport in Porous and Fractured Media. FEFLOW.
- Dimitrova-Petrova, K., Rosolem, R., Soulsby, C., Wilkinson, M.E., Lilly, A., Geris, J., 2021. Combining static and portable Cosmic ray neutron sensor data to assess catchment scale heterogeneity in soil water storage and their integrated role in catchment runoff response. *Journal of Hydrology*, 601: 126659. [10.1016/j.jhydrol.2021.126659](https://doi.org/10.1016/j.jhydrol.2021.126659).
- Dong, J., Steele-Dunne, S.C., Ochsner, T.E., van de Giesen, N., 2016. Determining soil moisture and soil properties in vegetated areas by assimilating soil temperatures. *Water Resources Research*, 52(6): 4280-4300. [10.1002/2015wr018425](https://doi.org/10.1002/2015wr018425).
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meyret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120: 25-36. [10.1016/j.rse.2011.11.026](https://doi.org/10.1016/j.rse.2011.11.026).
- El-Rawy, M., Zijl, W., Salem, A., Awad, A., Eltarabily, M.G., Negm, A.M., 2022. Fundamentals of Groundwater Modeling Methods and a Focused Review on the Groundwater Models of the Nile Valley Aquifer. In: Negm, A.M., El-Rawy, M. (Eds.), *Sustainability of Groundwater in the Nile Valley, Egypt*. Springer International Publishing, Cham, pp. 39-70. [10.1007/978-3-031-12676-5\\_3](https://doi.org/10.1007/978-3-031-12676-5_3).
- Enemark, T., Peeters, L.J.M., Mallants, D., Batelaan, O., 2019. Hydrogeological conceptual model building and testing: A review. *Journal of Hydrology*, 569: 310-329. [10.1016/j.jhydrol.2018.12.007](https://doi.org/10.1016/j.jhydrol.2018.12.007).
- Entekhabi, D., Njoku, E.G., Neill, P.E.O., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Zyl, J.V., 2010. The Soil Moisture Active Passive (SMAP) Mission. *Proceedings of the IEEE*, 98(5): 704-716. [10.1109/JPROC.2010.2043918](https://doi.org/10.1109/JPROC.2010.2043918).

- Evans, J.G., Ward, H.C., Blake, J.R., Hewitt, E.J., Morrison, R., Fry, M., Ball, L.A., Doughty, L.C., Libre, J.W., Hitt, O.E., Rylett, D., Ellis, R.J., Warwick, A.C., Brooks, M., Parkes, M.A., Wright, G.M.H., Singer, A.C., Boorman, D.B., Jenkins, A., 2016. Soil water content in southern England derived from a cosmic-ray soil moisture observing system - COSMOS-UK. *Hydrological Processes*, 30(26): 4987-4999. 10.1002/hyp.10929.
- Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5): 10,143-10,162. <https://doi.org/10.1029/94jc000572>.
- Evensen, G., 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4): 343-367. <https://doi.org/10.1007/s10236-003-0036-9>.
- Fang, H., Baret, F., Plummer, S., Schaepman-Strub, G., 2019. An Overview of Global Leaf Area Index (LAI): Methods, Products, Validation, and Applications. *Reviews of Geophysics*, 57(3): 739-799. 10.1029/2018rg000608.
- Fersch, B., Francke, T., Heistermann, M., Schrön, M., Döpfer, V., Jakobi, J., Baroni, G., Blume, T., Boga, H., Budach, C., Gränzig, T., Förster, M., Güntner, A., Hendricks Franssen, H.J., Kasner, M., Köhli, M., Kleinschmit, B., Kunstmann, H., Patil, A., Rasche, D., Scheffele, L., Schmidt, U., Szulc-Seyfried, S., Weimar, J., Zacharias, S., Zreda, M., Heber, B., Kiese, R., Mares, V., Mollenhauer, H., Völksch, I., Oswald, S., 2020. A dense network of cosmic-ray neutron sensors for soil moisture observation in a highly instrumented pre-Alpine headwater catchment in Germany. *Earth Syst. Sci. Data*, 12(3): 2289-2309. 10.5194/essd-12-2289-2020.
- Finkele, K., Hochstrasser, T., Murphy, P.N.C., Daly, E., Jarman, C., Richards, K., Fenton, O., Cummins, T., Saunders, M., Johnston, P.M., Bruen, M., Byrne, K.A., Delaney, D.T., Fealy, R., Green, S., Higgins, S., Hunter Williams, N., Lanigan, G., McCarthy, T., McCormack, T., Mellander, P.-E., Nicholson, O., Nugent, C., O'Loughlin, F., Tobin, B., Tuohy, P., and Whetton, R., 2022. The new Irish Soil Moisture Observation Network – ISMON: an Umbrella for Integrating Several Recent Soil Moisture Measurements Initiatives, EMS Annual Meeting 2022, Bonn, Germany. <https://doi.org/10.5194/ems2022-273>.
- Fisher, R.A., Koven, C.D., 2020. Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems. *Journal of Advances in Modeling Earth Systems*, 12(4). 10.1029/2018ms001453.
- Franz, T.E., Zreda, M., Ferre, T.P.A., Rosolem, R., Zweck, C., Stillman, S., Zeng, X., Shuttleworth, W.J., 2012. Measurement depth of the cosmic ray soil moisture probe affected by hydrogen from various sources. *Water Resources Research*, 48(8). 10.1029/2012wr011871.
- Freeze, R.A., 1975. A stochastic-conceptual analysis of the one-dimensional groundwater flow in nonuniform homogeneous media. *Water Resources Research*: 11 (5), 725-742. <https://doi.org/10.1029/WR011i005p00725>.
- Freeze, R.A.a.C., J. A., 1979. *Groundwater*. Hemel Hempstead: Prentice-Hall International. *Progress in Physical Geography: Earth and Environment*, 5(4), 616-617. <https://doi.org/10.1177/030913338100500412>.
- Fu, X., Zhang, Y., Zhong, Q., Lü, H., Ding, Y., Li, Z., Yu, Z., Jiang, X., 2023. Soil Moisture Estimation by Assimilating In-Situ and SMAP Surface Soil Moisture Using Unscented Weighted Ensemble Kalman Filter. *Water Resources Research*, 59(9). 10.1029/2023wr034506.
- Furusho-Percot, C., Goergen, K., Hartick, C., Kulkarni, K., Keune, J., Kollet, S., 2019. Pan-European groundwater to atmosphere terrestrial systems climatology from a physically consistent simulation. *Sci Data*, 6(1): 1-9. <https://doi.org/10.1038/s41597-019-0328-7>.
- Garcia-Menendez, O., Ballesteros, B.J., Renau-Prunonosa, A., Morell, I., Mochales, T., Ibarra, P.I., Rubio, F.M., 2018. Using electrical resistivity tomography to assess the effectiveness of managed aquifer recharge in a salinized coastal aquifer. *Environ Monit Assess*, 190(2): 100. 10.1007/s10661-017-6446-9.
- Gaspari, G., Cohn, S.E., 1999. Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society* (125(554)): 723–757. <https://doi.org/10.1002/qj.49712555417>.
- Gaspari, F., Goergen, K., Shrestha, P., Sulis, M., Rihani, J., Geimer, M., Kollet, S., 2014. Implementation and scaling of the fully coupled Terrestrial Systems Modeling Platform (TerrSysMP v1.0) in a massively parallel supercomputing environment – a case study on JUQUEEN (IBM Blue Gene/Q). *Geoscientific Model Development*, 7(5): 2531-2543. 10.5194/gmd-7-2531-2014.
- Gebler, S., Kurtz, W., Pauwels, V.R.N., Kollet, S.J., Vereecken, H., Hendricks Franssen, H.J., 2019. Assimilation of High-Resolution Soil Moisture Data Into an Integrated Terrestrial Model for a Small-Scale Head-Water Catchment. *Water Resources Research*, 55(12): 10358-10385. <https://doi.org/10.1029/2018wr024658>.
- Geologischer-Dienst-NRW, 2009. Informationssystem Bodenkarte von Nordrhein-Westfalen 1:50.000 (IS BK 50). <https://gdk.gdi-de.org/geonetwork/srv/api/records/D6775754-266F-4396-A22B-14E1E4096F59>.
- Geologischer-Dienst-NRW, 2011. Informationssystem Hydrogeologische Karte von Nordrhein-Westfalen 1:100.000 (IS HK 100 DS). <https://gdk.gdi-de.org/geonetwork/srv/api/records/9FBF9F0A-5D86-4691-8873-D3FC0099BDC6>.
- Hain, C.R., Crow, W.T., Anderson, M.C., Mecikalski, J.R., 2012. An ensemble Kalman filter dual assimilation of thermal infrared and microwave satellite observations of soil moisture into the Noah land surface model. *Water Resources Research*, 48(11). 10.1029/2011wr011268.

- Hamill, T.M., Whitaker, J.S., Snyder, C., 2001. Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter. *Monthly Weather Review*(129(11)): 2776–2790. 10.1175/1520-0493(2001)129<2776>.
- Han, X., Franssen, H.-J.H., Montzka, C., Vereecken, H., 2014. Soil moisture and soil properties estimation in the Community Land Model with synthetic brightness temperature observations. *Water Resources Research*, 50(7): 6081-6105. 10.1002/2013wr014586.
- Han, X., Franssen, H.J.H., Rosolem, R., Jin, R., Li, X., Vereecken, H., 2015. Correction of systematic model forcing bias of CLM using assimilation of cosmic-ray Neutrons and land surface temperature: a study in the Heihe Catchment, China. *Hydrology and Earth System Sciences*, 19(1): 615-629. <https://doi.org/10.5194/hess-19-615-2015>.
- Han, X., Hendricks Franssen, H.-J., Jiménez Bello, M.Á., Rosolem, R., Bogena, H., Alzamora, F.M., Chanzy, A., Vereecken, H., 2016. Simultaneous soil moisture and properties estimation for a drip irrigated field by assimilating cosmic-ray neutron intensity. *Journal of Hydrology*, 539: 611-624. 10.1016/j.jhydrol.2016.05.050.
- 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. <https://doi.org/10.2136/vzj2012.0072>.
- Hawdon, A., McJannet, D., Wallace, J., 2014. Calibration and correction procedures for cosmic-ray neutron soil moisture probes located across Australia. *Water Resources Research*, 50(6): 5029-5043. 10.1002/2013wr015138.
- He, H., Aogu, K., Li, M., Xu, J., Sheng, W., Jones, S.B., González-Teruél, J.D., Robinson, D.A., Horton, R., Bristow, K., Dyck, M., Filipović, V., Noborio, K., Wu, Q., Jin, H., Feng, H., Si, B., Lv, J., 2021. Chapter Three - A review of time domain reflectometry (TDR) applications in porous media. In: Sparks, D.L. (Ed.), *Advances in Agronomy*. Academic Press, pp. 83-155. <https://doi.org/10.1016/bs.agron.2021.02.003>.
- He, L., Chen, J.M., Mostovoy, G., Gonsamo, A., 2021. Soil Moisture Active Passive Improves Global Soil Moisture Simulation in a Land Surface Scheme and Reveals Strong Irrigation Signals Over Farmlands. *Geophysical Research Letters*, 48(8). 10.1029/2021gl092658.
- Heidbüchel, I., Güntner, A., Blume, T., 2016. Use of cosmic-ray neutron sensors for soil moisture monitoring in forests. *Hydrology and Earth System Sciences*, 20(3): 1269-1288. 10.5194/hess-20-1269-2016.
- Heistermann, M., Bogena, H., Francke, T., Güntner, A., Jakobi, J., Rasche, D., Schrön, M., Döpfer, V., Fersch, B., Groh, J., Patil, A., Pütz, T., Reich, M., Zacharias, S., Zengerle, C., Oswald, S., 2022. Soil moisture observation in a forested headwater catchment: combining a dense cosmic-ray neutron sensor network with roving and hydrogravimetry at the TEREÑO site Wüstebach. *Earth Syst. Sci. Data*, 14(5): 2501-2519. 10.5194/essd-14-2501-2022.
- 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). 10.1029/2010wr009480.
- 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): 1-22. <https://doi.org/10.1029/2007wr006505>.
- Houtekamer, P.L., Mitchell, H.L., 1998. Data Assimilation Using an Ensemble Kalman Filter Technique. *Monthly Weather Review*(126(3)): 796–811. 10.1175/1520-0493(1998)126<0796:dauae>.
- Houtekamer, P.L., Mitchell, H.L., 2001. A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*(129(1)): 123–137. [https://doi.org/10.1175/1520-0493\(2001\)129<0123:asekff>2.0.co;2](https://doi.org/10.1175/1520-0493(2001)129<0123:asekff>2.0.co;2).
- Hu, J., Fennel, K., Mattern, J.P., Wilkin, J., 2012. Data assimilation with a local Ensemble Kalman Filter applied to a three-dimensional biological model of the Middle Atlantic Bight. *Journal of Marine Systems*, 94: 145-156. <https://doi.org/10.1016/j.jmarsys.2011.11.016>.
- Hubbert, M King, Rubey, W.W., 1959. Role of fluid pressure in mechanics of overthrust faulting: i. Mechanics of fluid-filled porous solids and its application to overthrust faulting. *GSA Bulletin*, 70(2): 115-166. 10.1130/0016-7606(1959)70[115:rofpim]2.0.co;2.
- Hung, C.P., Schalge, B., Baroni, G., Vereecken, H., Hendricks Franssen, H.J., 2022. Assimilation of Groundwater Level and Soil Moisture Data in an Integrated Land Surface-Subsurface Model for Southwestern Germany. *Water Resources Research*, 58(6). 10.1029/2021wr031549.
- Hunt, B.R., Kostelich, E.J., Szunyogh, I., 2007. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, 230(1-2): 112-126. <https://doi.org/10.1016/j.physd.2006.11.008>.
- Husain, M.M., Cherry, J.A., Fidler, S., Frappe, S.K., 1998. On the long-term hydraulic gradient in the thick clayey aquitard in the Sarnia region, Ontario. *Canadian Geotechnical Journal*: 35(6), 986–1003. 10.1139/t98-057.
- Hwang, C., Yang, Y., Kao, R., Han, J., Shum, C.K., Galloway, D.L., Sneed, M., Hung, W.C., Cheng, Y.S., Li, F., 2016. Time-varying land subsidence detected by radar altimetry: California, Taiwan and north China. *Sci Rep*, 6: 28160. 10.1038/srep28160.

- Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C., Ustin, S.L., 2009. PROSPECT+SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, 113: S56-S66. 10.1016/j.rse.2008.01.026.
- Jakobi, J., Huisman, J.A., Vereecken, H., Dieckkrüger, B., Bogena, H.R., 2018. Cosmic Ray Neutron Sensing for Simultaneous Soil Water Content and Biomass Quantification in Drought Conditions. *Water Resources Research*, 54(10): 7383-7402. 10.1029/2018wr022692.
- Jarvis, A., H.I. Reuter, A. Nelson, Guevara, E., 2008. Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database.
- Jiang, L., Shi, C., Sun, S., Liang, X., 2021. Fusion of In situ Soil Moisture and Land Surface Model Estimates Using Localized Ensemble Optimum Interpolation over China. *Journal of Meteorological Research*, 34(6): 1335-1346. 10.1007/s13351-020-0033-7.
- 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. 10.1016/S0309-1708(00)00075-0.
- Kalbus, E., Reinstorf, F., Schirmer, M., 2006. Measuring methods for groundwater – surface water interactions: a review. *Hydrol. Earth Syst. Sci.* 10, 873–887. <https://doi.org/10.5194/hess-10-873-2006>.
- Köhli, M., Schrön, M., Zreda, M., Schmidt, U., Dietrich, P., Zacharias, S., 2015. Footprint characteristics revised for field-scale soilmoisture monitoring with cosmic-ray neutrons. *Water Resources Research* 10.1002/.
- Keune, J., Gasper, F., Goergen, K., Hense, A., Shrestha, P., Sulis, M., Kollet, S., 2016. Studying the influence of groundwater representations on land surface-atmosphere feedbacks during the European heat wave in 2003. *Journal of Geophysical Research: Atmospheres*, 121(22): 13,301-13,325. <https://doi.org/10.1002/2016jd025426>.
- Kim, H., Wigneron, J.-P., Kumar, S., Dong, J., Wagner, W., Cosh, M.H., Bosch, D.D., Collins, C.H., Starks, P.J., Seyfried, M., Lakshmi, V., 2020. Global scale error assessments of soil moisture estimates from microwave-based active and passive satellites and land surface models over forest and mixed irrigated/dryland agriculture regions. *Remote Sensing of Environment*, 251: 112052. 10.1016/j.rse.2020.112052.
- 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. <https://doi.org/10.1016/j.advwatres.2005.08.006>.
- Kollet, S.J., Maxwell, R.M., 2008. Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model. *Water Resources Research*, 44(2): 1-18. <https://doi.org/10.1029/2007wr006004>.
- Koster, R.D., Dirmeyer, P.A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C.T., Kanac, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C.H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Sud, Y.C., Taylor, C.M., Verseghy, D., Vasic, R., Xue, Y., Yamada, T., Team, G., 2004. Regions of strong coupling between soil moisture and precipitation. *Science*, 305(5687): 1138-40. <https://doi.org/10.1126/science.1100217>.
- Kumar, C.P., 2004. Groundwater Flow Models.
- 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. *Geoscientific Model Development*, 9(4): 1341-1360. <https://doi.org/10.5194/gmd-9-1341-2016>.
- 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. <https://doi.org/10.1002/2013wr014823>.
- 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). 10.1029/2011wr011743.
- Landerer, F.W., Flechtner, F.M., Save, H., Webb, F.H., Bandikova, T., Bertiger, W.I., Bettadpur, S.V., Byun, S.H., Dahle, C., Döbbslaw, H., Fahnestock, E., Harvey, N., Kang, Z., Kruizinga, G.L.H., Loomis, B.D., McCullough, C., Murböck, M., Nagel, P., Paik, M., Pie, N., Poole, S., Strelakov, D., Tamisiea, M.E., Wang, F., Watkins, M.M., Wen, H.Y., Wiese, D.N., Yuan, D.N., 2020. Extending the Global Mass Change Data Record: GRACE Follow-On Instrument and Science Data Performance. *Geophysical Research Letters*, 47(12). 10.1029/2020gl088306.
- Li, D., Schrön, M., Köhli, M., Bogena, H., Weimar, J., Jiménez Bello, M.A., Han, X., Martínez Gimeno, M.A., Zacharias, S., Vereecken, H., Hendricks Franssen, H.J., 2019. Can Drip Irrigation be Scheduled with Cosmic-Ray Neutron Sensing? *Vadose Zone Journal*, 18(1): 190053. 10.2136/vzj2019.05.0053.
- Li, F., Bogena, H.R., Bayat, B., Kurtz, W., Franssen, a.H.-J.H., 2023a. Can a sparse network of cosmic ray neutron sensors improve soil moisture and evapotranspiration estimation at the larger catchment scale? *Water Resources Research*. 10.1029/2023WR035056.
- Li, F., Kurtz, W., Hung, C.P., Vereecken, H., Hendricks Franssen, H.-J., 2023b. Water table depth assimilation in integrated terrestrial system models at the larger catchment scale. *Frontiers in Water*, 5. 10.3389/frwa.2023.1150999.

- Li, S., Zhang, L., Ma, R., Yan, M., Tian, X., 2020. Improved ET assimilation through incorporating SMAP soil moisture observations using a coupled process model: A study of U.S. arid and semiarid regions. *Journal of Hydrology*, 590: 125402. [10.1016/j.jhydrol.2020.125402](https://doi.org/10.1016/j.jhydrol.2020.125402).
- Li, W., Bao, L., Yao, G., Wang, F., Guo, Q., Zhu, J., Zhu, J., Wang, Z., Bi, J., Zhu, C., Zhong, Y., Lu, S., 2024. The analysis on groundwater storage variations from GRACE/GRACE-FO in recent 20 years driven by influencing factors and prediction in Shandong Province, China. *Scientific Reports*, 14(1): 5819. [10.1038/s41598-024-55588-3](https://doi.org/10.1038/s41598-024-55588-3).
- 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. *Remote Sensing of Environment*, 180: 292-304. [10.1016/j.rse.2015.10.033](https://doi.org/10.1016/j.rse.2015.10.033).
- Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves I. Flood movement in long rivers. *Proceedings of the Royal Society of London. Series A. Mathematical, Physical and Engineering Sciences*, 229(1178): 281-316. <https://doi.org/10.1098/rspa.1955.0088>.
- Liu, Y., Liu, R., Chen, J.M., 2012. Retrospective retrieval of long-term consistent global leaf area index (1981–2011) from combined AVHRR and MODIS data. *Journal of Geophysical Research: Biogeosciences*, 117(G4). [10.1029/2012jg002084](https://doi.org/10.1029/2012jg002084).
- Manabe, S., 1969. The atmospheric circulation and the hydrology of the earth's surface, *Mon. Weather Rev.*, 97, 739–774. [10.1175/1520-0493\(1969\)09760:0739:catoc62;2.3.co;2](https://doi.org/10.1175/1520-0493(1969)09760:0739:catoc62;2.3.co;2).
- Masood, A., Tariq, M.A.U.R., Hashmi, M.Z.U.R., Waseem, M., Sarwar, M.K., Ali, W., Farooq, R., Almazroui, M., Ng, A.W.M., 2022. An Overview of Groundwater Monitoring through Point-to Satellite-Based Techniques. *Water*, 14(4): 565. [10.3390/w14040565](https://doi.org/10.3390/w14040565).
- Mauder, M., Cuntz, M., Drüe, 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. <https://doi.org/10.1016/j.agrformet.2012.09.006>.
- 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. <https://doi.org/10.1016/j.advwatres.2012.10.001>.
- Maxwell, R.M., Miller, N.L., 2005. Development of a Coupled Land Surface and Groundwater Model. *Journal of Hydrometeorology*, 6(3): 233-247. [10.1175/jhm422.1](https://doi.org/10.1175/jhm422.1).
- McDonald, M.G., Harbaugh, A.W., 1988. A modular three-dimensional finite-difference ground-water flow model. 06-A1. [10.3133/twri06A1](https://doi.org/10.3133/twri06A1).
- McLaughlin, D., 2002. An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering. *Advances in Water Resources*, 25( 8–12): 1275-1286. [https://doi.org/10.1016/S0309-1708\(02\)00055-6](https://doi.org/10.1016/S0309-1708(02)00055-6).
- Montzka, C., Bayat, B., Tewes, A., Mengen, D., Vereecken, H., 2021. Sentinel-2 Analysis of Spruce Crown Transparency Levels and Their Environmental Drivers After Summer Drought in the Northern Eifel (Germany). *Frontiers in Forests and Global Change*, 4. [10.3389/ffgc.2021.667151](https://doi.org/10.3389/ffgc.2021.667151).
- Montzka, C., Bogena, H., Zreda, M., Monerris, A., Morrison, R., Muddu, S., Vereecken, H., 2017. Validation of Spaceborne and Modelled Surface Soil Moisture Products with Cosmic-Ray Neutron Probes. *Remote Sensing*, 9(2): 103. [10.3390/rs9020103](https://doi.org/10.3390/rs9020103).
- Montzka, C., Canty, M., Kunkel, R., Menz, G., Vereecken, H., Wendland, F., 2008. Modelling the water balance of a mesoscale catchment basin using remotely sensed land cover data. *Journal of Hydrology*, 353(3-4): 322-334. [10.1016/j.jhydrol.2008.02.018](https://doi.org/10.1016/j.jhydrol.2008.02.018).
- 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. [10.1016/j.jhydrol.2011.01.020](https://doi.org/10.1016/j.jhydrol.2011.01.020).
- Muñoz Sabater, J., 2021. ERA5-Land hourly data from 1950 to 1980 Copernicus Climate Change Service (C3S) Climate Data Store (CDS) <https://doi.org/10.24381/cds.e2161bac>.
- Mukhlisin, M., Astuti, H.W., Wardhani, E.D., Matlan, S.J., 2021. Techniques for ground-based soil moisture measurement: a detailed overview. *Arabian Journal of Geosciences*, 14(19). [10.1007/s12517-021-08263-0](https://doi.org/10.1007/s12517-021-08263-0).
- Myneni, R., Knyazikhin, Y., Park, T., 2015. MOD15A2H MODIS/Terra leaf area Index/FPAR 8-Day L4 global 500m SIN grid V006. NASA EOSDIS Land Processes DAAC. [http://doi.org/10.5067/MODIS/MYD15A2H.006](https://doi.org/10.5067/MODIS/MYD15A2H.006).
- Naz, B.S., Kollet, S., Franssen, H.H., Montzka, C., Kurtz, W., 2020. A 3 km spatially and temporally consistent European daily soil moisture reanalysis from 2000 to 2015. *Sci Data*, 7(1): 111. [10.1038/s41597-020-0450-6](https://doi.org/10.1038/s41597-020-0450-6).
- Naz, B.S., Kurtz, W., Montzka, C., Sharples, W., Goergen, K., Keune, J., Gao, H., Springer, A., Hendricks Franssen, H.-J., Kollet, S., 2019. Improving soil moisture and runoff simulations at 3&thinsp;km over Europe using land surface data assimilation. *Hydrology and Earth System Sciences*, 23(1): 277-301. <https://doi.org/10.5194/hess-23-277-2019>.

- Naz, B.S., Sharples, W., Ma, Y., Goergen, K., Kollet, S., 2023. Continental-scale evaluation of a fully distributed coupled land surface and groundwater model, ParFlow-CLM (v3.6.0), over Europe. *Geoscientific Model Development*, 16(6): 1617-1639. 10.5194/gmd-16-1617-2023.
- Nerger, L., Hiller, W., Oter, J.S., 2005. PDAF-The Parallel Data Assimilation Framework Experiences with Kalman Filtering. *Experiences with Kalman Filtering*: 63-83. [https://doi.org/10.1142/9789812701831\\_0006](https://doi.org/10.1142/9789812701831_0006).
- Nicolai-Shaw, N., Hirschi, M., Mittelbach, H., Seneviratne, S.I., 2015. Spatial representativeness of soil moisture using in situ, remote sensing, and land reanalysis data. *Journal of Geophysical Research: Atmospheres*, 120(19): 9955-9964. 10.1002/2015jd023305.
- Ochsner, T.E., Cosh, M.H., Cuenca, R.H., Dorigo, W.A., Draper, C.S., Hagimoto, Y., Kerr, Y.H., Larson, K.M., Njoku, E.G., Small, E.E., Zreda, M., 2013. State of the Art in Large-Scale Soil Moisture Monitoring. *Soil Science Society of America Journal*, 77(6): 1888-1919. 10.2136/sssaj2013.03.0093.
- Oleson, K., Dai, Y., Bonan, G.B., Bosilovich, M., Dickinson, R., Dirmeyer, P., Zeng, X., 2004. Technical Description of the Community Land Model (CLM). University Corporation for Atmospheric Research: 1-187. <https://doi.org/10.5065/D6N877R0>.
- 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., 2007. CLM3.5 documentation. In: Research, N.C.f.A. (Ed.). National Center for Atmospheric Research, Boulder, CO, Colorado.
- Oleson, K.W., Niu, G.Y., Yang, Z.L., Lawrence, D.M., Thornton, P.E., Lawrence, P.J., Stöckli, 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. *Journal of Geophysical Research: Biogeosciences*, 113(G1): 1-26. <https://doi.org/10.1029/2007jg000563>.
- Paniconi, C., Mario, P., 1994. A comparison of Picard and Newton iteration in the numerical solution of multidimensional variably saturated flow problems. *Water Resources Research*: 30(12), 3357-3374. 10.1029/94wr02046.
- Paniconi, C., Wood, E.F., 1993. A detailed model for simulation of catchment scale subsurface hydrologic processes. *Water Resources Research*: 29(6), 1601-1620. 10.1029/92wr02333
- Pano, P., 2006. The European soil database In: connexion, G. (Ed.). 5. 32-33.
- Panzeri, M., Riva, M., Guadagnini, A., Neuman, S.P., 2013. Data assimilation and parameter estimation via ensemble Kalman filter coupled with stochastic moment equations of transient groundwater flow. *Water Resources Research*, 49(3): 1334-1344. 10.1002/wrcr.20113.
- Panzeri, M., Riva, M., Guadagnini, A., Neuman, S.P., 2014. Comparison of Ensemble Kalman Filter groundwater-data assimilation methods based on stochastic moment equations and Monte Carlo simulation. *Advances in Water Resources*, 66: 8-18. 10.1016/j.advwatres.2014.01.007.
- Panzeri, M., Riva, M., Guadagnini, A., Neuman, S.P., 2015. EnKF coupled with groundwater flow moment equations applied to Lauswiesen aquifer, Germany. *Journal of Hydrology*, 521: 205-216. 10.1016/j.jhydrol.2014.11.057.
- Patil, A., Fersch, B., Hendricks Franssen, H.-J., Kunstmann, H., 2021. Assimilation of Cosmogenic Neutron Counts for Improved Soil Moisture Prediction in a Distributed Land Surface Model. *Frontiers in Water*, 3. 10.3389/frwa.2021.729592.
- Pauwels, V.R.N., Balenzano, A., Satalino, G., Skriver, H., Verhoest, N.E.C., Mattia, F., 2009. Optimization of Soil Hydraulic Model Parameters Using Synthetic Aperture Radar Data: An Integrated Multidisciplinary Approach. *IEEE Transactions on Geoscience and Remote Sensing*, 47(2): 455-467. 10.1109/TGRS.2008.2007849.
- Pauwels, V.R.N., De Lannoy, G.J.M., Hendricks Franssen, H.J., Vereecken, H., 2013. Simultaneous estimation of model state variables and observation and forecast biases using a two-stage hybrid Kalman filter. *Hydrology and Earth System Sciences*, 17(9): 3499-3521. 10.5194/hess-17-3499-2013.
- Pauwels, V.R.N., Hendricks Franssen, H.-J., De Lannoy, G.J.M., 2020. Evaluation of State and Bias Estimates for Assimilation of SMOS Retrievals Into Conceptual Rainfall-Runoff Models. *Frontiers in Water*, 2. 10.3389/frwa.2020.00004.
- Peng, J., Loew, A., 2017. Recent Advances in Soil Moisture Estimation from Remote Sensing. *Water*, 9(7): 530. 10.3390/w9070530.
- Peralta, R.C.; Mazur, V.; Dutram, P. Monitoring of Groundwater Levels for Real-Time Conjunctive Water Management. *Publ.-Arkansas Water Resour. Res. Cent.* 1983, 92, 1-27. <https://scholarworks.uark.edu/awrcrtr/258>.
- Pinnington, E., Amezcua, J., Cooper, E., Dadson, S., Ellis, R., Peng, J., Robinson, E., Morrison, R., Osborne, S., Quaife, T., 2021. Improving soil moisture prediction of a high-resolution land surface model by parameterising pedotransfer functions through assimilation of SMAP satellite data. *Hydrology and Earth System Sciences*, 25(3): 1617-1641. 10.5194/hess-25-1617-2021.
- Pinnington, E., Quaife, T., Lawless, A., Williams, K., Arkebauer, T., Scoby, D., 2020. The Land Variational Ensemble Data Assimilation Framework: LAVENDAR v1.0.0. *Geoscientific Model Development*, 13(1): 55-69. 10.5194/gmd-13-55-2020.
- R. Baatz, H. R. Bogaen, H.-J. Hendricks Franssen, J. A. Huisman, C. Montzka, Vereecken, H., 2015. An empirical vegetation

- correction for soil water content quantification using cosmic ray probes. *Water Resources Research*. <https://doi.org/10.1002/2014WR016443>.
- Reichle, R.H., Crow, W.T., Keppenne, C.L., 2008. An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water Resources Research*, 44(3): 1-13. <https://doi.org/10.1029/2007wr006357>.
- Reichle, R.H., D. B. McLaughlin, Entekhabi, D., 2002. Hydrologic Data Assimilation With the Ensemble Kalman Filter. *Monthly Weather Review*: 130(1): 103-114. [https://doi.org/10.1175/1520-0493\(2002\)130<0103:hdawte>2](https://doi.org/10.1175/1520-0493(2002)130<0103:hdawte>2).
- Reichle, R.H., De Lannoy, G.J.M., Liu, Q., Koster, R.D., Kimball, J.S., Crow, W.T., Ardizzone, J.V., Chakraborty, P., Collins, D.W., Conaty, A.L., Girotto, M., Jones, L.A., Kolassa, J., Lievens, H., Lucchesi, R.A., Smith, E.B., 2017. Global Assessment of the SMAP Level-4 Surface and Root-Zone Soil Moisture Product Using Assimilation Diagnostics. *Journal of Hydrometeorology*, 18(12): 3217-3237. [10.1175/JHM-D-17-0130.1](https://doi.org/10.1175/JHM-D-17-0130.1).
- 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. [10.1175/2010jhm1262.1](https://doi.org/10.1175/2010jhm1262.1).
- Richards, L., 1931. Capillary Conduction of Liquids Through Porous Mediums. *Journal of Applied Physics*: 318-333. <https://doi.org/10.1063/1.1745010>.
- Richey, A.S., Thomas, B.F., Lo, M.H., Reager, J.T., Famiglietti, J.S., Voss, K., Swenson, S., Rodell, M., 2015. Quantifying renewable groundwater stress with GRACE. *Water Resour Res*, 51(7): 5217-5238. [10.1002/2015WR017349](https://doi.org/10.1002/2015WR017349).
- Ridler, M.-E., Madsen, H., Stisen, S., Bircher, S., Fensholt, R., 2014. Assimilation of SMOS-derived soil moisture in a fully integrated hydrological and soil-vegetation-atmosphere transfer model in Western Denmark. *Water Resources Research*, 50(11): 8962-8981. [10.1002/2014wr015392](https://doi.org/10.1002/2014wr015392).
- 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(12). [10.1029/2010wr009111](https://doi.org/10.1029/2010wr009111).
- Román-Cascón, C., Pellarin, T., Gibon, F., Brocca, L., Cosme, E., Crow, W., Fernández-Prieto, D., Kerr, Y.H., Massari, C., 2017. Correcting satellite-based precipitation products through SMOS soil moisture data assimilation in two land-surface models of different complexity: API and SURFEX. *Remote Sensing of Environment*, 200: 295-310. [10.1016/j.rse.2017.08.022](https://doi.org/10.1016/j.rse.2017.08.022).
- Romano, N., 2014. Soil moisture at local scale: Measurements and simulations. *Journal of Hydrology*, 516: 6-20. [10.1016/j.jhydrol.2014.01.026](https://doi.org/10.1016/j.jhydrol.2014.01.026).
- Rosid, M.S.; Keping, A.W. Hydrogeological mapping using the Seismoelectric Method. *Explor. Geophys.* 2005, 36, 245–249. [10.1071/EG05245](https://doi.org/10.1071/EG05245).
- Rosolem, R., Hoar, T., Arellano, A., Anderson, J.L., Shuttleworth, W.J., Zeng, X., Franz, T.E., 2014. Translating aboveground cosmic-ray neutron intensity to high-frequency soil moisture profiles at sub-kilometer scale. *Hydrology and Earth System Sciences*, 18(11): 4363-4379. [10.5194/hess-18-4363-2014](https://doi.org/10.5194/hess-18-4363-2014).
- Sabri, L.M., Sudarsono, B., Aryadi, Y., 2020. Groundwater Change Detection by Gravity Measurement on Northern Coast of Java: A Case Study in Semarang City of Central Java of Indonesia. *IOP Conference Series: Materials Science and Engineering*, 797(1): 012032. [10.1088/1757-899x/797/1/012032](https://doi.org/10.1088/1757-899x/797/1/012032).
- Sahoo, A.K., De Lannoy, G.J.M., Reichle, R.H., Houser, P.R., 2013. Assimilation and downscaling of satellite observed soil moisture over the Little River Experimental Watershed in Georgia, USA. *Advances in Water Resources*, 52: 19-33. [10.1016/j.advwatres.2012.08.007](https://doi.org/10.1016/j.advwatres.2012.08.007).
- Sattari, M.T., Mirabbasi, R., Sushab, R.S., Abraham, J., 2018. Prediction of Groundwater Level in Ardebil Plain Using Support Vector Regression and M5 Tree Model. *Ground Water*, 56(4): 636-646. [10.1111/gwat.12620](https://doi.org/10.1111/gwat.12620).
- 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). [10.1029/2011wr010462](https://doi.org/10.1029/2011wr010462).
- 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. [10.1016/s0022-1694\(01\)00466-8](https://doi.org/10.1016/s0022-1694(01)00466-8).
- Schattan, P., Baroni, G., Oswald, S.E., Schöber, J., Fey, C., Kormann, C., Huttenlau, M., Achleitner, S., 2017. Continuous monitoring of snowpack dynamics in alpine terrain by aboveground neutron sensing. *Water Resources Research*, 53(5): 3615-3634. [10.1002/2016wr020234](https://doi.org/10.1002/2016wr020234).
- Schattan, P., Schwaizer, G., Schöber, J., Achleitner, S., 2020. The complementary value of cosmic-ray neutron sensing and snow covered area products for snow hydrological modelling. *Remote Sensing of Environment*, 239: 111603. [10.1016/j.rse.2019.111603](https://doi.org/10.1016/j.rse.2019.111603).
- Schrön, M., Köhli, M., Scheiffele, L., Iwema, J., Bogena, H.R., Lv, L., Martini, E., Baroni, G., Rosolem, R., Weimar, J., Mai, J., Cuntz, M., Rebmann, C., Oswald, S.E., Dietrich, P., Schmidt, U., Zacharias, S., 2017. Improving calibration and validation of cosmic-ray neutron sensors in the light of spatial sensitivity. *Hydrology and Earth System Sciences*, 21(10): 5009-5030. [10.5194/hess-21-5009-2017](https://doi.org/10.5194/hess-21-5009-2017).

- Seo, E., Lee, M.-I., Reichle, R.H., 2021. Assimilation of SMAP and ASCAT soil moisture retrievals into the JULES land surface model using the Local Ensemble Transform Kalman Filter. *Remote Sensing of Environment*, 253: 112222. 10.1016/j.rse.2020.112222.
- 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. 10.1002/2013wr014070.
- 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. 10.1016/j.advwatres.2015.06.009.
- Shrestha, P., Kurtz, W., Vogel, G., Schulz, J.P., Sulis, M., Hendricks Franssen, H.J., Kollet, S., Simmer, C., 2018a. Connection Between Root Zone Soil Moisture and Surface Energy Flux Partitioning Using Modeling, Observations, and Data Assimilation for a Temperate Grassland Site in Germany. *Journal of Geophysical Research: Biogeosciences*, 123(9): 2839-2862. 10.1029/2016jg003753.
- 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. <https://doi.org/10.1175/mwr-14-00029.1>.
- 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. *Hydrology and Earth System Sciences*, 19(10): 4317-4326. 10.5194/hess-19-4317-2015.
- Shrestha, P., Sulis, M., Simmer, C., Kollet, S., 2018b. Effects of horizontal grid resolution on evapotranspiration partitioning using TerrSysMP. *Journal of Hydrology*, 557: 910-915. <https://doi.org/10.1016/j.jhydrol.2018.01.024>.
- Shukla, S., Meshesha, T.W., Sen, I.S., Bol, R., Bogaena, H., Wang, J., 2023. Assessing Impacts of Land Use and Land Cover (LULC) Change on Stream Flow and Runoff in Rur Basin, Germany. *Sustainability*, 15(12): 9811. 10.3390/su15129811.
- Shuttleworth, J., Rosolem, R., Zreda, M., Franz, T., 2013. The COsmic-ray Soil Moisture Interaction Code (COSMIC) for use in data assimilation. *Hydrology and Earth System Sciences*, 17(8): 3205-3217. 10.5194/hess-17-3205-2013.
- Sridhar, V., Billah, M.M., Hildreth, J.W., 2018. Coupled Surface and Groundwater Hydrological Modeling in a Changing Climate. *Ground Water*, 56(4): 618-635. 10.1111/gwat.12610.
- Smith, R.G., Majumdar, S., 2020. Groundwater Storage Loss Associated With Land Subsidence in Western United States Mapped Using Machine Learning. *Water Resources Research*, 56(7). 10.1029/2019wr026621.
- Strebel, L., Bogaena, H., Vereecken, H., Andreasen, M., Aranda-Barranco, S., Hendricks Franssen, H.-J., 2024. Evapotranspiration prediction for European forest sites does not improve with assimilation of in situ soil water content data. *Hydrology and Earth System Sciences*, 28(4): 1001-1026. 10.5194/hess-28-1001-2024.
- Strebel, L., Bogaena, H.R., Vereecken, H., Hendricks Franssen, H.-J., 2022. Coupling the Community Land Model version 5.0 to the parallel data assimilation framework PDAF: description and applications. *Geoscientific Model Development*, 15(2): 395-411. 10.5194/gmd-15-395-2022.
- Sulis, M., Keune, J., Shrestha, P., Simmer, C., Kollet, S.J., 2018. Quantifying the Impact of Subsurface-Land Surface Physical Processes on the Predictive Skill of Subseasonal Mesoscale Atmospheric Simulations. *Journal of Geophysical Research: Atmospheres*, 123(17): 9131-9151. 10.1029/2017jd028187.
- Sulis, M., Langensiepen, M., Shrestha, P., Schickling, A., Simmer, C., Kollet, S.J., 2015. Evaluating the Influence of Plant-Specific Physiological Parameterizations on the Partitioning of Land Surface Energy Fluxes. *Journal of Hydrometeorology*, 16(2): 517-533. <https://doi.org/10.1175/jhm-d-14-0153.1>.
- Sulis, M., Meyerhoff, S.B., Paniconi, C., Maxwell, R.M., Putti, M., Kollet, S.J., 2010. A comparison of two physics-based numerical models for simulating surface water-groundwater interactions. *Advances in Water Resources*, 33(4): 456-467. 10.1016/j.advwatres.2010.01.010.
- ten Harkel, J., Bartholomeus, H., Kooistra, L., 2019. Biomass and Crop Height Estimation of Different Crops Using UAV-Based Lidar. *Remote Sensing*, 12(1): 17. 10.3390/rs12010017.
- Tong, J., Hu, B.X., Yang, J., 2011. Assimilating transient groundwater flow data via a localized ensemble Kalman filter to calibrate a heterogeneous conductivity field. *Stochastic Environmental Research and Risk Assessment*, 26(3): 467-478. 10.1007/s00477-011-0534-0.
- Torres, R., Snocij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Flourey, N., Brown, M., Traver, I.N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., Rostan, F., 2012. GMES Sentinel-1 mission. *Remote Sensing of Environment*, 120: 9-24. 10.1016/j.rse.2011.05.028.
- Turner, M., Walker, J., Oke, P., 2008. Ensemble member generation for sequential data assimilation. *Remote Sensing of Environment*, 112(4): 1421-1433. 10.1016/j.rse.2007.02.042.
- UNWWAP, 2015. Facts and Figures from the World Water Development Report. Document code: SC/2015/PI/H/2, SC-2015/WS/5

United Nations World Water Assessment Programme. Available online at: <https://unesdoc.unesco.org/ark:/48223/pf0000231823>.

- V., C., 1965. The relation between the lowering of the Piezometric surface and the rate and duration of discharge of a well using ground-water storage. 10.1111/jawr.1965.1.3.9.
- Valcke, S., 2013. The OASIS3 coupler: a European climate modelling community software. *Geoscientific Model Development*, 6(2): 373-388. <https://doi.org/10.5194/gmd-6-373-2013>.
- van Genuchten, M.T., 1980. A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. *Soil Science Society of America Journal*, 44(5): 892-898. <https://doi.org/10.2136/sssaj1980.03615995004400050002x>.
- Vather, T., Everson, C.S., Franz, T.E., 2020. The Applicability of the Cosmic Ray Neutron Sensor to Simultaneously Monitor Soil Water Content and Biomass in an Acacia mearnsii Forest. *Hydrology*, 7(3): 48. 10.3390/hydrology7030048.
- Vereecken, H., Amelung, W., Bauke, S.L., Bogena, H., Brüggemann, N., Montzka, C., Vanderborght, J., Bechtold, M., Blöschl, G., Carninatti, A., Javaux, M., Konings, A.G., Kusche, J., Neuweiler, I., Or, D., Steele-Dunne, S., Verhoef, A., Young, M., Zhang, Y., 2022. Soil hydrology in the Earth system. *Nature Reviews Earth & Environment*, 3(9): 573-587. 10.1038/s43017-022-00324-6.
- 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). 10.1029/2008wr006829.
- Vereecken, H., Maes, J., Feyen, J., 1990. Estimating Unsaturated Hydraulic Conductivity from Easily Measured Soil Properties. *Soil Science*, 149(1): 1-12. 10.1097/00010694-199001000-00001.
- 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):1-57. <https://doi.org/10.2136/vzj2015.09.0131>.
- Verhoef, W., Bach, H., 2007. Coupled soil-leaf-canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data. *Remote Sensing of Environment*, 109(2): 166-182. 10.1016/j.rse.2006.12.013.
- Wahl, S., Bollmeyer, C., Crewell, S., Figura, C., Friederichs, P., Hense, A., Keller, J.D., Ohlwein, C., 2017. A novel convective-scale regional reanalysis COSMO-REA2: Improving the representation of precipitation. *Meteorologische Zeitschrift*, 26(4): 345-361. <https://doi.org/10.1127/metz/2017/0824>.
- Waldhoff, G., Lusse, U., 2015. Preliminary land use classification of 2015 for the Rur catchment. TR32DB. 10.5880/TR32DB.14.
- Weiss, M., Baret, F., 2020. S2ToolBox Level 2 products: LAI, FAPAR, FCOVER, Institut National de la Recherche Agronomique (INRA), Avignon, France. [https://step.esa.int/docs/extra/ATBD\\_S2ToolBox\\_V2.0.pdf](https://step.esa.int/docs/extra/ATBD_S2ToolBox_V2.0.pdf).
- Wen, X.H., Deutsch, C.V., Cullick, A.S., 2002. Construction of geostatistical aquifer models integrating dynamic flow and tracer data using inverse technique. *Journal of Hydrology*, 255(1): 151-168. [https://doi.org/10.1016/S0022-1694\(01\)00512-1](https://doi.org/10.1016/S0022-1694(01)00512-1).
- White, A.M., Gardner, W.P., Borsa, A.A., Argus, D.F., Martens, H.R., 2022. A Review of GNSS/GPS in Hydrogeodesy: Hydrologic Loading Applications and Their Implications for Water Resource Research. *Water Resour Res*, 58(7): e2022WR032078. 10.1029/2022WR032078.
- Yamamoto, J.K., 2007. On unbiased backtransform of lognormal kriging estimates. *Computational Geosciences*, 11(3): 219-234. <https://doi.org/10.1007/s10596-007-9046-x>.
- Yeh, P.J.-F., Eltahir, E.A.B., 2005. Representation of Water Table Dynamics in a Land Surface Scheme. Part II: Subgrid Variability. *Journal of Climate*, 18(12): 1881-1901. <https://doi.org/10.1175/JCLI3331.1>.
- Yin, J., Zhan, X., Zheng, Y., Hain, C.R., Liu, J., Fang, L., 2015. Optimal ensemble size of ensemble Kalman filter in sequential soil moisture data assimilation. *Geophysical Research Letters*, 42(16): 6710-6715. <https://doi.org/10.1002/2015gl063366>.
- Yin, W., Hu, L., Jiao, J.J., 2017. Evaluation of Groundwater Storage Variations in Northern China Using GRACE Data. *Geofluids*, 2017: 1-13. 10.1155/2017/8254824.
- You, H., Wang, T., Skidmore, A., Xing, Y., 2017. Quantifying the Effects of Normalisation of Airborne LiDAR Intensity on Coniferous Forest Leaf Area Index Estimations. *Remote Sensing*, 9(2): 163. 10.3390/rs9020163.
- Yu, H.-L., Wu, Y.-Z., Cheung, S.Y., 2020. A data assimilation approach for groundwater parameter estimation under Bayesian maximum entropy framework. *Stochastic Environmental Research and Risk Assessment*, 34(5): 709-721. <https://doi.org/10.1007/s00477-020-01795-z>.
- Zhang, D., Madsen, H., Ridler, M.E., Kidmose, J., Jensen, K.H., Refsgaard, J.C., 2016. Multivariate hydrological data assimilation

- of soil moisture and groundwater head. *Hydrology and Earth System Sciences*, 20(10): 4341-4357. <https://doi.org/10.5194/hess-20-4341-2016>.
- Zhang, H., 2018. Improved characterization of root zone soil moisture by assimilation groundwater level and surface soil moisture data in an integrated terrestrial system model, RWTH Aachen University, Forschungszentrum Jülich GmbH. <http://hdl.handle.net/2128/19414>.
- Zhang, H., Hendricks Franssen, H.-J., Han, X., Vrugt, J.A., Vereecken, H., 2017. State and parameter estimation of two land surface models using the ensemble Kalman filter and the particle filter. *Hydrology and Earth System Sciences*, 21(9): 4927-4958. 10.5194/hess-21-4927-2017.
- Zhang, H., Kurtz, W., Kollet, S., Vereecken, H., Franssen, H.-J.H., 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. <https://doi.org/10.1016/j.advwatres.2017.11.003>.
- Zhang, Y., Schaap, M.G., 2017. Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3). *Journal of Hydrology*, 547: 39-53. <https://doi.org/10.1016/j.jhydrol.2017.01.004>.
- Zhao, H., Montzka, C., Baatz, R., Vereecken, H., Franssen, H.-J.H., 2021. The Importance of Subsurface Processes in Land Surface Modeling over a Temperate Region: An Analysis with SMAP, Cosmic Ray Neutron Sensing and Triple Collocation Analysis. *Remote Sensing*, 13(16): 3068. 10.3390/rs13163068.
- Zhou, J., Wu, Z., Crow, W.T., Dong, J., He, H., 2020. Improving Spatial Patterns Prior to Land Surface Data Assimilation via Model Calibration Using SMAP Surface Soil Moisture Data. *Water Resources Research*, 56(10). 10.1029/2020wr027770.
- Zreda, M., Desilets, D., Ferré, T.P.A., Scott, R.L., 2008. Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons. *Geophysical Research Letters*, 35(21). 10.1029/2008gl035655.
- Zreda, M., Shuttleworth, W.J., Zeng, X., Zweck, C., Desilets, D., Franz, T., Rosolem, R., 2012. COSMOS: the COsmic-ray Soil Moisture Observing System. *Hydrology and Earth System Sciences*, 16(11): 4079-4099. 10.5194/hess-16-4079-2012.

## Acknowledgements

First of all, I would like to thank Prof. Dr. Harrie-Jan Hendricks-Franssen for his intensive supervision during my four years and three months of PhD study. During this time, we had frequent meetings and many discussions. Prof. Dr. Harrie-Jan Hendricks-Franssen is knowledgeable and always passionate about science. He can always comment on my simulations and calculations from different perspectives and give me useful suggestions to inspire me to further explore my research. For every report, IKO, conference, poster, and paper during my PhD, he was very patient and made revisions many times, not sparing a single small detail and striving for perfection.

I am also grateful to Dr. Heye Bogena for the collaboration on this PhD work since my second paper. Thanks to him for attending my regular PhD meetings and being able to provide scientific advice and insightful suggestions. Thanks to him for the data and help in contacting other data providers. He was always able to give feedback on my work in a very short period of time, which greatly accelerated my work.

I also gratefully acknowledge Dr. Bagher Bayat, who taught me a lot about how to proceed with my doctoral studies, always encouraged me when I was frustrated with my PhD study, and helped me figure out what to do. Every encouragement from him always helped me get energized again and get back into my PhD studies.

I am also thankful to Dr. Johannes Keller for his help with the modeling. Thanks to him for taking the time to have meetings with me, discussing the model issues with me, and always trying to help me modify my model as soon as possible so that I can run the simulations smoothly.

I would also like to thank Dr. Wolfgang Kurtz for providing the model settings and scripts from Rur catchment, which helped me a lot when I first started learning about model TSMP-PDAF. Thanks to him for doing all this preliminary work, I was able to successfully complete my PhD work based on the Rur catchment as the study area.

Many thanks to Dr. Ching Pui Hung for his support in modeling and groundwater data assimilation during the early stage of my PhD study.

Besides, I would like to thank my other colleagues from the Institute of Bio- and Geosciences at the Juelich Research Center, especially:

Prof. Dr. Harry Vereecken, Director of the Institute, for his insightful suggestions.

Dr. Carsten Montzka for the land use data.

Dr. Rahul Raj for the LAI data.

My officemate Phoung Ta for his help in my life in Germany.

All my group members for their help in my study and life in Germany, especially Yafei Huang, Haojin Zhao and Xiao Lu.

My friend Dr. Haoran Wang (Henry), who came to IBG3 at the same time as me, thanks for the help with my daily life and studies.

In addition, I would like to thank the China Scholarship Council (CSC) for sponsoring my work and life in Germany. Furthermore, I gratefully acknowledge the computing time granted by the supercomputers JUWELS and JURECA at Forschungszentrum Jülich. I would also like to thank the Terrestrial Environmental Observatories (TERENO) network for providing valuable data for the Rur catchment.

Last but not least, special thanks to my families for their constant support, encouragement, and love. Thanks to Bingyao Wang for his understanding, patience, trust, and companionship.

Band / Volume 637

**Investigation of Lower Boundary Conditions of Brominated Very Short-lived Species (VSLs)**

S. Zheng (2024), 2, iii, 160 pp

ISBN: 978-3-95806-770-7

Band / Volume 638

**Modellgestützte Analyse zukünftigen Mobilitätsverhaltens**

J. P. Reul (2024), XVI, 291 pp

ISBN: 978-3-95806-771-4

Band / Volume 639

**Insights into Mechanisms of Secondary Organic Aerosol Formation: Approaching Atmospherically Relevant Conditions in an Atmospheric Reaction Chamber**

Y. Baker (2024), XVII, 122 pp

ISBN: 978-3-95806-776-9

Band / Volume 640

**Advancing the representation of agricultural systems in Land Surface Models: systematic model evaluations and technical model developments**

T. S. Boas (2024), xxi, 145 pp

ISBN: 978-3-95806-777-6

Band / Volume 641

**Imaging spatial and temporal soil water content variations of the soil-plant continuum using ground penetrating radar**

L. Lärm (2024), xii, 303 pp

ISBN: 978-3-95806-778-3

Band / Volume 642

**Development of Iridium-based Nanostructures for Oxygen Evolution Reaction in PEM Water Electrolysis**

S. Park (2024), 135 pp

ISBN: 978-3-95806-779-0

Band / Volume 643

**Multi-dimensional GPR full-waveform inversion for small-scale hydrogeophysical soil characterization**

D. Hoven (2024), IX, 163 pp

ISBN: 978-3-95806-781-3

Band / Volume 644

**Analyse des Gastransports in komplexen Membransystemen durch Modellierung und multiskalige Simulation**

K. Wilkner (2024), VIII, 122 pp

ISBN: 978-3-95806-784-4

Band / Volume 645

**Deployment of Fuel Cell Vehicles in Road Transport and the Expansion of the Hydrogen Refueling Station Network: 2024 Update**

T. Grube; M. Rex (2024), iii, 26 pp

ISBN: 978-3-95806-786-8

Band / Volume 646

**Modellgestützte Analyse treibhausgasneutraler Transformationsstrategien für Deutschland**

T. F. Schöb (2024), XII, 228 pp

ISBN: 978-3-95806-789-9

Band / Volume 647

**Future Distribution Grids Using Geo-Referenced Synthetic Network Topologies**

A. Bandam (2024), ix, 237 pp

ISBN: 978-3-95806-790-5

Band / Volume 648

**Multi-scenario, high-resolution Lagrangian transport modeling for the analysis of the Asian tropopause aerosol layer**

J. H. Clemens (2024), 143 pp

ISBN: 978-3-95806-792-9

Band / Volume 649

**Prospektive Lebenszyklusanalysen von Power-to-Gas-Optionen**

J. C. Koj (2024), XVIII, 197 pp

ISBN: 978-3-95806-794-3

Band / Volume 650

**Assimilation of groundwater level and cosmic-ray neutron sensor soil moisture measurements into integrated terrestrial system models for better predictions**

F. Li (2024), xvii, 172 pp

ISBN: 978-3-95806-796-7



Energie & Umwelt / Energy & Environment  
Band / Volume 650  
ISBN 978-3-95806-796-7

Mitglied der Helmholtz-Gemeinschaft

