



KRISTINA KORONACI

**DATA ASSIMILATION FOR IMPROVED DISCHARGE
PREDICTIONS WITH THE WFLOW_SBM MODEL: A
CASE STUDY OF THE OVERIJSELSE VECHT
RIVER (THE NETHERLANDS)**

MASTER'S THESIS

MASTER STUDY PROGRAMME FLOOD RISK MANAGEMENT



IHE
DELFT



TECHNISCHE
UNIVERSITÄT
DRESDEN



UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Univerza v Ljubljani



Ljubljana, 2022



Candidate

KRISTINA KORONACI

**DATA ASSIMILATION FOR IMPROVED DISCHARGE
PREDICTIONS WITH THE WFLOW_SBM MODEL: A
CASE STUDY OF THE OVERIJSSELSE VECHT
RIVER (THE NETHERLANDS)**

Master's thesis no. : 556.166+556.536(492)(043.3)

Supervisor:

Matthijs den Toom, PhD

Co-supervisor:

Prof. Albrecht Weerts, PhD

Academic supervisor:

Nejc Bezak, PhD

Commission member:

Schalk Jan van Andel, PhD

Chairman of the Commission:

Asst. Prof. Simon Rusjan, PhD

Ljubljana, 08.09.2022



UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Univerza v Ljubljani



I, Kristina Koronaci, registration number (26125053), author of the written thesis of the study entitled:

I DECLARE

1. Circle one of variants a) or b)

- a) that the written final work of the study is the result of my independent work;
- b) that the written final work of the study is the result of the own work of several candidates and meets the conditions set out in the UL Statute for joint final work of the study and is in the required proportion the result of my independent work;

2. that the printed form of the written final part of the study is identical to the electronic form of the written final part of the study;

3. that I have obtained all the necessary permissions for the use of data and copyrighted works in the written final part of the study and clearly marked them in the written final part of the study;

4. that in preparing the written final part of the study I acted in accordance with ethical principles and, where necessary, obtained the consent of the ethics commission for the research;

5. I agree that the electronic form of the written final part of the study is used to check the similarity of the content with other parts with the software for checking the similarity of the content related to the study information system of the member;

6. to transfer to UL free of charge, non-exclusively, spatially and indefinitely the right to store the author's work in electronic form, the right to reproduce and the right to make the written final part of the study available to the public on the World Wide Web through the UL Repository;

7. dissertations. UL consents enable free, non-exclusive, space and time unlimited storage of the author's work in electronic form and reproduction and making the dissertation available to the public on the World Wide Web through the UL Repository;

8. that I allow the publication of my personal data, which are stated in the written final part of the study and this statement, together with the publication of the written final part of the study.

V / Na:

Date: 01-09-2022

Student's signature / -to:



"This page is intentionally blank"

ERRATA

Error page	Error bar	Instead	Let it be
------------	-----------	---------	-----------

ACKNOWLEDGMENTS

My two-year Flood Risk Management Erasmus Mundus Master's adventure, which would not have been possible without the invaluable support and contributions of loads of people, has come to an end with the submission of this work.

Firstly, I would like to express my sincerest gratitude to Prof. Albrecht Weerts and Dr. Matthijs den Toom for their continuous guidance and support over the past six months. I am particularly thankful for the patience and the effort they put forth to deliver complex concepts and tools in the best way possible. I am grateful to Deltares for providing me with the opportunity to connect with a very knowledgeable community, where I gained valuable knowledge from each interaction I had during my internship. A special thanks to Dr. Joost Buitink, who made scripting seem less intimidating and was always available to clarify my queries. My sincere appreciation goes to Dr. Schalk Jan van Andel and Dr. Nejc Bezak for their insightful feedback and comments.

I also wish to thank everyone who contributed to making this journey possible, particularly Prof. Biswa Bhattacharya for his support and direction during this challenging but rewarding journey, and Prof. Simon Rusjan for putting in all the work to make the thesis defense at the University of Ljubljana run smoothly. To all my professors and coordinators from the Technische Universität Dresden, IHE Delft Institute for Water Education, UPC Universitat Politècnica de Catalunya, and the University of Ljubljana, I am honored and forever grateful for every lesson learned over the past two years.

I like to express my appreciation to my FRM group—my new family away from home—for the experiences, lessons, and adventures we shared. Thanks to my parents, grandpa, and my sisters, Laura and Dea, for always being there to lift me up when I need it the most.

Lastly, I would like to dedicate this work to my beloved Grandma, my first and favorite teacher, who is no longer here to cheer me on, but whose lessons will live on.

Kristina Koronaci
Delft, the Netherlands

BIBLIOGRAPHIC-DOCUMENTALISTIC INFORMATION AND ABSTRACT

UDC: 556.166+556.536(492)(043.3)
Author: Kristina Koronaci
Supervisor: Matthijs den Toom, Ph.D.
Co-supervisor: Prof. Albrecht Weerts, Ph.D.
Title: Data assimilation for improved discharge estimates with the wflow_sbm model: A case study of the Overijsselse Vecht river (The Netherlands).
Document type: Master thesis
Notes: XI, 90 p., 19 tab., 34 fig., 13eq., 3 ann., 38 ref.
Keywords: Data assimilation, Ensemble Kalman Filter, wflow_sbm model, discharge prediction, uncertainty, Vecht River, OpenDA

Abstract

Extreme hydrological events have become more frequent, as evidenced by the European floods of July 2021, which affected the southern provinces of the Netherlands. The need for improved discharge predictions to be used in operational water management to avoid potential adverse effects of flooding has encouraged researchers to employ several ways to improve hydrological model estimates, including data assimilation. This thesis explores the data assimilation effects in the discharge predictions of the wflow_sbm distributed hydrological model of the Vecht river basin. Additionally, effects on other hydrological states and fluxes like subsurface flow, saturated water depth, and soil moisture were explored spatially.

This work presents a methodology for applying data assimilation in a model where water is routed from the surface and subsurface. In contrast, previous studies used a model in which water is routed only via surface water. Ensemble Kalman Filter is used to update the model's discharge predictions by assimilating external discharge observations. This methodology also explores how the data assimilation effect is influenced by the uncertainty characterization considered in the assimilation framework and other factors like the length of the assimilation window and the number of assimilation locations. A preliminary study of the rainfall data is performed to determine the uncertainties of the chosen rainfall product. A benchmark simulation scenario is then selected after the review of deterministic and ensemble model predictions. Finally, data assimilation experiments are developed after discussing the characterization of the uncertainty model.

The results of the model output analysis indicate that streamflow assimilation typically has a positive effect on improving model discharge estimations. Additionally, the Ensemble Kalman Filter update effectively captures the system's spatial state dynamics for subsurface states and fluxes, such as saturated water depth, soil moisture, etc. Two alternative experimental setups with different assimilation intervals and numbers of assimilated observations are examined concerning how this effect varies over other flow gauge locations. As demonstrated by both experiments, longer assimilation times give better results, with the assimilation effect significantly improving in the final timesteps of the assimilation frame. Furthermore, it is concluded that assimilation of observations near the outlet and interior gauges will improve discharge predictions, whereas assimilation of observations only near the outlet will only improve discharge predictions at a number of stations, typically those that are closer to the assimilation location and those where the wflow_sbm model exhibits the same trend as the assimilation station. An uncertainty factor of 2.5 for the precipitation error and 0.1 for the observation error yielded the best results for both experiments.

However, this study has several limitations, including assumptions of a perfect model and initial conditions; the way the precipitation and observations error model was derived. As a result, the model gives unrealistic discharge predictions when compensating for the neglected errors. Additionally, a limited number of experiments due to the extensive computational times, attributed to the combination of the OpenDA tool with the distributed model, and the algorithm choice, does not allow the DA impact on the discharge predictions to be judged accurately.

Therefore, the final section of this study provides recommendations for future research, suggesting additional experiments with longer assimilation windows; analysis of the spatial correlation structure of precipitation, the use of more statistically reliable techniques to assess the precipitation uncertainties; consideration of the model parameter and initial conditions uncertainty; etc.

BIBLIOGRAFSKO – DOKUMENTACIJSKA STRAN IN IZVLEČEK

UDK: 556.166+556.536(492)(043.3)
Avtor: Kristina Koronaci
Mentor: dr. Matthijs den Toom
Somentor: prof. dr. Albrecht Weerts
Naslov: Asimilacija podatkov za izboljšane ocene razrešnice z wflow_sbm modelom: Študija primera reke Overijsselse Vecht (Nizozemska).
Tip dokumenta: magistrsko delo
Obseg in oprema: 90 str., 19 pregl., 34 sl., 3 pril.
Ključne besede: Asimilacija podatkov, Ensemble Kalman Filter, model wflow_sbm, napoved pretoka, negotovost, reka Vecht, OpenDA

Izvleček:

Ekstremni hidrološki dogodki so postali vse pogostejši, kar dokazujejo evropske poplave julija 2021, ki so prizadele južne nizozemske province. Potreba po izboljšanih napovedih izpustov, ki se uporabljajo pri operativnem upravljanju voda, da bi se izognili morebitnim škodljivim učinkom poplav, je spodbudila raziskovalce k uporabi več načinov za izboljšanje ocen hidroloških modelov, vključno z asimilacijo podatkov. Diplomsko delo raziskuje učinke asimilacije podatkov pri napovedih pretoka porazdeljenega hidrološkega modela wflow_sbm porečja reke Vecht. Poleg tega so bili prostorsko raziskani učinki na druga hidrološka stanja in tokove, kot so podzemni tok, globina nasičene vode in vlažnost tal.

To delo predstavlja metodologijo za uporabo asimilacije podatkov v modelu, kjer je voda usmerjena s površine in pod površino. V nasprotju s tem so prejšnje študije uporabile model, v katerem je voda speljana le po površinski vodi. Ensemble Kalmanov filter se uporablja za posodobitev napovedi izpustov modela z asimilacijo zunanjih opazovanj izpustov. Ta metodologija raziskuje tudi, kako na učinek asimilacije podatkov vpliva karakterizacija negotovosti, upoštevana v asimilacijskem okviru, in drugi dejavniki, kot sta dolžina asimilacijskega okna in število asimilacijskih lokacij. Izvede se predhodna študija podatkov o padavinah, da se določijo negotovosti izbranega produkta padavin. Po pregledu napovedi determinističnih in ansambelskih modelov se nato izbere primerjalni simulacijski scenarij. Po razpravi o karakterizaciji modela negotovosti so razviti poskusi asimilacije podatkov.

Rezultati analize rezultatov modela kažejo, da ima asimilacija toka običajno pozitiven učinek na izboljšanje ocen pretoka modela. Poleg tega posodobitev filtra Ensemble Kalman učinkovito zajame dinamiko prostorskega stanja sistema za stanja in tokove pod površino, kot so globina nasičene vode, vlažnost tal itd. Dve alternativni eksperimentalni nastavitvi z različnimi intervali asimilacije in številom asimiliranih opazovanj sta preučeni glede tega, kako ta učinek razlikuje glede na druge lokacije merilnika pretoka. Kot sta dokazala oba poskusa, dajejo daljši časi asimilacije boljše rezultate, pri čemer se učinek asimilacije znatno izboljša v končnih časovnih korakih okvira asimilacije. Poleg tega je ugotovljeno, da bo asimilacija opazovanj v bližini izpusta in notranjih merilnikov izboljšala napovedi pretoka, medtem ko bo asimilacija opazovanj samo v bližini iztoka izboljšala le napovedi pretoka na številnih postajah, običajno tistih, ki so bližje lokaciji asimilacije, in tistih kjer model wflow_sbm kaže enak trend kot asimilacijska postaja. Faktor negotovosti 2,5 za napako padavin in 0,1 za napako opazovanja je dal najboljše rezultate za oba poskusa.

Vendar ima ta študija več omejitev, vključno s predpostavkami o popolnem modelu in začetnih pogojih; način, kako je bil izpeljan model napak padavin in opazovanj. Posledično daje model nerealne napovedi praznjenja pri kompenzaciji zanemarjenih napak. Poleg tega omejeno število poskusov zaradi obsežnih računskih časov, pripisanih kombinaciji orodja OpenDA s porazdeljenim modelom, in izbire algoritma ne omogoča natančne ocene vpliva DA na napovedi praznjenja.

Zato zadnji del te študije podaja priporočila za prihodnje raziskave in predlaga dodatne poskuse z daljšimi asimilacijskimi okni; analiza prostorske korelacijske strukture padavin, uporaba statistično zanesljivejših tehnik za ocenjevanje padavinske negotovosti; upoštevanje negotovosti parametrov modela in začetnih pogojev; itd.

TABLE OF CONTENTS

ERRATA.....	I
ACKNOWLEDGMENTS	II
BIBLIOGRAPHIC-DOCUMENTALISTIC INFORMATION AND ABSTRACT	III
TABLE OF CONTENTS.....	V
LIST OF FIGURES	VII
LIST OF TABLES	X
ABBREVIATIONS AND SYMBOLS	XI
1 INTRODUCTION	1
1.1 Motivation.....	1
1.2 Objectives.....	2
1.3 Research questions.....	2
1.4 Practical value and innovation.....	2
1.5 Operational flood forecasting in the Netherlands	3
2 LITERATURE REVIEW	4
2.1 Data assimilation concepts and methods.....	4
2.2 Uncertainty definition	6
2.3 Verification methods	8
2.3.1 Performance verification of deterministic forecasts	8
2.3.2 Performance verification of ensemble forecasts	9
2.4 OpenDA.....	10
3 CATCHMENT DESCRIPTION	12
3.1 Geul river basin	12
3.2 Vecht river basin	13
4 DATA AND MODEL SETUP	15
4.1 Data availability	15
4.1.1 Precipitation	15
4.1.2 Additional forcing data	17
4.1.3 Discharge	18
4.2 Wflow_sbm hydrologic model.....	19
5 RESEARCH METHODOLOGY	21
6 RESULTS AND DISCUSSION.....	23
6.1 Precipitation data analysis	23
6.2 Deterministic simulations	28
6.2.1 Model States	34
6.3 Experiments with DA.....	35
6.3.1 Experimental setup	35
6.3.2 Selection of a reference simulation.....	37
6.3.3 Model performance regarding discharge prediction	38
6.3.3.1 Experiment 1	39

6.3.3.2	Experiment 2	43
6.3.4	Model performance regarding state updating.....	47
7	CONCLUSIONS AND RECOMMENDATIONS.....	51
7.1	Summary	51
7.2	Limitations	52
7.3	Conclusions	52
7.4	Recommendations.....	54
8	REFERENCES	55
9	APPENDICES.....	58
Appendix A.	Streamflow gauge station locations	58
A.1	Vecht	58
A.2	Geul	58
Appendix B.	Precipitation data analysis.....	59
B.1	DWD rain gauge observations VS E-OBS gridded dataset.....	59
B.2	Radolan vs EOBS	61
	Station ID 357.....	61
	Station ID 1223.....	62
	Station ID 1230.....	62
	Station ID 3640.....	63
	Station ID 5131.....	63
	Station ID 4667.....	64
	Station ID 15927.....	64
B.3	Annual max and mean precipitation	65
Appendix C.	Simulations with DA	67
C.1	Sensitivity analysis of the simulations with DA, for a run with 8 ensemble members, with different values of uncertainties for forcing and observations	67
C.2	Discharge prediction results with DA in different station locations – experiment 1	70
	River station – Bilk.....	70
	River station – Gronau.....	70
	River station – Lagel.....	70
	River station – Ohne	71
	River station – Osterwald	71
	River Station – Wehr Neuenhaus.....	71
C.3	Discharge prediction results with DA in different station locations – experiment 2.....	72
	River station – Bilk.....	72
	River station – Lagel	72
	River station – Ohne	72

LIST OF FIGURES

Figure 1: Classic Kalman Filter Algorithm.....	5
Figure 2: Diagram of the model update differences between the EnKF and the AEnKF (Rakovec et al., 2015)	6
Figure 3: Modular components of OpenDA: method (the DA method), observations (the stochastic observer for processing the observations), and the model (van Velzen et al., 2016)	11
Figure 4: Geul river basin location.....	13
Figure 5: Vecht river basin location.....	14
Figure 6: Snapshot of the interactive map made available by the rdwd R package, displaying a rain gauge station located near the border (catchment area), for which no public data are accessible	16
Figure 7: a) Schematic illustration of the wflow_sbm processes and states (Schellekens et al., 2020), b) model states and fluxes in terms of internal variable names.....	20
Figure 8: Schematization of the research methodology workflow.....	22
Figure 9: Comparison of the cumulative and daily rainfall time series between the DWD rain gauge observations and the EOBS observational gridded rainfall dataset for Station ID 357 in the Vecht river catchment for a) year 2021; b) the complete time series from 2011 to 2021	23
Figure 10: Comparison of the cumulative and daily rainfall time series between the DWD rain gauge observations and the EOBS observational gridded rainfall dataset for Station ID 15000 in the Geul catchment for a) year 2021; b) the complete time series from 2011 to 2021	24
Figure 11: EOBS and Radolan precipitation resampled into annual maximum precipitation values for the station ID 15000 in the Geul catchment	24
Figure 12: EOBS and Radolan precipitation resampled into annual mean precipitation values for the station ID 15000 in the Geul catchment.....	25
Figure 13: a) Daily precipitation at station point location with ID 5131 in the Vecht river catchment derived from the Radolan gridded dataset and the EOBS gridded observational datasets. Uncertainty of the Radolan is presented as a 95% uncertainty band; b) Annual values of selected performance metrics (NSE, RMSE, KGE); c) Correlation scatter plots of the Radolan and EOBS daily time series for the 10-year period.	25
Figure 14: a) Daily precipitation at station point location with ID 15000 in the Geul catchment derived from the Radolan gridded dataset and the EOBS gridded observational datasets. Uncertainty of the Radolan is presented as a 95% uncertainty band; b) Annual values of selected performance metrics (NSE, RMSE, KGE); c) Correlation scatterplots of the Radolan and EOBS daily time series for the 10-year period.	26
Figure 15: E-OBS (on the left) vs. RADOLAN (on the right) spatial comparison for the grid covering both catchments a) annual mean rainfall b) annual max rainfall c) annual max rainfall resampled (11km Radolan grid to match the EOBS grid).....	27

Figure 16: Hourly Observations of the Emlichheim discharge and wflow_sbm model difference before and after correction with ERA5. In red, gaps in the uncorrected model output are highlighted.	29
Figure 17: Simulated VS Observed Mean daily runoff, for selected stations in the Vecht catchment, for 2016	29
Figure 18: Correlation of simulated and observed discharge scatter plots at selected station location in the Vecht catchment. Note that the correlation is computed for a different period in different stations, depending on the period of record available at each station.....	31
Figure 19: Simulated VS Observed Mean daily runoff, for selected stations in the Geul catchment, for the first half of the year 2021.....	32
Figure 20: Correlation of simulated and observed discharges scatter plots at selected station location in the Geul river catchment. Correlation is computed for a different period in different stations, depending on the period of record available at each station.....	33
Figure 21: Hydrological model states variation at Emlichheim for a selected year (2016)	34
Figure 22: Model state correlation at Emlichheim	35
Figure 23: Observed discharge in the outlet (Emlichheim) and other interior flow gauges in the Vecht river basin during 2016. Two characteristic winter and summer flood events are selected for the analysis	36
Figure 24: Open loop simulation results for three different combinations of uncertainties: a) forcing standard deviation 2.5, spatial correlation 30 km; b) forcing standard deviation 2, spatial correlation 30 km; c) forcing standard deviation 2, spatial correlation 10 km	38
Figure 25: DA simulation results, experiment 1, Emlichheim. The observed discharges are given in the dotted line, the open loop reference simulation ensemble mean in the red line, the ensemble mean of the DA simulation (updated with the EnKF) in the blue line, and the ensemble spread, representing a 95% uncertainty interval, in the sky blue.	39
Figure 26: DA simulation results, experiment 1, Gronau. The observed discharges are given in the dotted line, the open loop reference simulation ensemble mean in the red line, the ensemble mean of the DA simulation (updated with the EnKF) in the blue line, and the ensemble spread, representing a 95% uncertainty interval, in the sky blue.	40
Figure 27: Boxplot of RMSE of the mean ensemble for each day between the updated with EnKF run and the daily observed discharges in selected stations. The red lines in the boxes represent the RMSE median, box borders are the 25 and 75 percentiles, whiskers are the 9 and 95 percentiles, and the circles show the extreme values of RMSE.	41
Figure 28: Boxplot of RMSE of the mean ensemble for each day between the open loop reference run and the daily observed discharges in selected stations. The red lines in the boxes represent the RMSE median, box borders are the 25 and 75 percentiles, whiskers are the 9 and 95 percentiles, and the circles show the extreme values of RMSE.	42

Figure 29: Boxplot of RMSE of the mean ensemble for each day between the updated with EnKF run and the open loop reference run in selected stations. The red lines in the boxes represent the RMSE median, box borders are the 25 and 75 percentiles, whiskers are the 9 and 95 percentiles, and the circles show the extreme values of RMSE.....	43
Figure 30: DA simulation results, experiment 2, Emlichheim. The observed discharges are given in the dotted line, the open loop reference simulation ensemble mean in the red line, the ensemble mean of the DA simulation (updated with the EnKF) in the blue line, and the ensemble spread, representing a 95% uncertainty interval, in the sky blue.....	44
Figure 31: DA simulation results, experiment 2, Gronau. The observed discharges are given in the dotted line, the open loop reference simulation ensemble mean in the red line, the ensemble mean of the DA simulation (updated with the EnKF) in the blue line, and the ensemble spread, representing a 95% uncertainty interval, in the sky blue.....	45
Figure 32: Boxplot of RMSE of the mean ensemble for each day between the updated with EnKF run and the daily observed discharges in selected stations. The green lines in the boxes represent the RMSE median, box borders are the 25 and 75 percentiles, whiskers are the 9 and 95 percentiles, and the circles show the extreme values of RMSE.....	46
Figure 33: Boxplot of RMSE of the mean ensemble for each day between the open loop reference run and the daily observed discharges in selected stations. The green lines in the boxes represent the RMSE median, box borders are the 25 and 75 percentiles, whiskers are the 9 and 95 percentiles, and the circles show the extreme values of RMSE.....	46
Figure 34: Difference between the predicted (instates) and updated (outstates) model states on three selected days (23, 25, 29 June 2016). Only the most sensitive model states are shown, discharge (Q), land flow (Qland), subsurface flow (ssf), saturated water depth (swd), and soil moisture in three different layers	49

LIST OF TABLES

Table 1: General characteristics of the Geul and Vecht river catchments.....	12
Table 2: Summary of the available gridded precipitation datasets that provide coverage of the study areas.....	16
Table 3: Rain gauge stations nearby Geul (Source: DWD).....	17
Table 4: Rain gauge stations nearby Vecht (Source: DWD).....	17
Table 5: River gauge stations in the Vecht river	18
Table 6: River gauge stations in the Geul.....	19
Table 7: Performance metrics of the Radolan, EOBS rainfall dataset comparison for the complete daily time-series (2011-2021).....	28
Table 8: Performance statistics to evaluate the model's predictive ability in selected stream gauge stations in the Vecht river catchment	31
Table 9: Performance statistics to evaluate the model's predictive ability in selected stream gauge stations in the Geul river catchment.	34
Table 10: An overview of the experiment periods selected for the DA analysis	37
Table 11: Error model setup for the DA experiments	39
Table 12: An overview of the mean, maximum and minimum flows corresponding to observed, open loop, and filtered discharges at Emlichheim	40
Table 13: An overview of the mean, maximum and minimum flows corresponding to observed,.....	40
Table 14: The coefficient of determination R^2 [%] for seven different stations after DA application..	42
Table 15: An overview of the mean, maximum and minimum flows corresponding to observed,.....	44
Table 16: An overview of the mean, maximum and minimum flows corresponding to observed,.....	45
Table 17: The coefficient of determination R^2 [%] for five different stations after DA application.....	47
Table 18: Overview of the streamflow gauge station in the Vecht river basin	58
Table 19: Overview of the streamflow gauge station in the Geul river basin	58

ABBREVIATIONS AND SYMBOLS

AEnKF	Asynchronous Ensemble Kalman Filter
A.s.l.	Above sea level
CDC	Climate Data Center
CDS	Climate Data Store
CRPS	Continuous Rank Probability Score
DA	Data Assimilation
DWD	German national weather service
ECMWF	European Centre for Medium-Range Weather Forecast
EKF	Extended Kalman Filter
ELWAS	Water Information System of the North Rhine-Westphalia
EnKF	Ensemble Kalman Filter
E-OBS	European daily gridded observational dataset
FEWS	Flood forecasting and Early Warning System
KF	Kalman Filter
KGE	Kling-Gupta Efficiency
KNMI	Royal Netherlands Meteorological Institute
NLWKN	Lower Saxony State Office for Water Management, Coastal Protection and Nature Conservation
NSE	Nash-Sutcliffe Efficiency
MAE	Mean Absolute Error
MSE	Mean Squared Error
OI	Optimal Interpolation
pBIAS	Percentage bias (%)
PET	Potential evapotranspiration
QPE	Quantitative Precipitation Estimate
RMSE	Root Mean Square Error
RADOLAN	Radar-Online-Adjustment
R ²	Coefficient of determination
SWP	Agriculture, Natural Resources and Environment entity of the Public Service of Wallonia
Var	Variational
WL	Waterschap Limburg
3D-Var	3-D variational
4D-Var	4-D variational

1 INTRODUCTION

Accurate and reliable discharge predictions have proven challenging in operational hydrology. The recent floods of July 2021 in Europe, especially in the province of Limburg, highlighted the necessity for improved flood forecasting quality and accuracy.

The primary purpose of data assimilation is to predict a system's correct state. For example, a state of interest in operational hydrology can vary from river discharges, which can then be translated into water levels, to other fluxes and states such as subsurface water flow, soil moisture, saturated water depth, and so on, based on the objective of the study. Given that a hydrological model aims to predict such states based on physical and mathematical relationships assigned to each model, one can question why data assimilation is used. In fact, models are unable to make realistic state estimates, as they are prone to a wide range of errors, which could stem from various sources, including structural errors caused by assumptions and approximations made for the physical and numerical relations between system components, uncertainties of the model input (e.g., imperfect forcing data), model parameters and so on.

While it is possible to monitor states through ground-based observations or remote sensing, this may not be enough to capture all states of our interest. States with a high degree of spatial and temporal variability cannot be scaled spatially accurately from point observations, as in the case of precipitation upscaling from rain gauge point observations. On the other hand, satellite data can provide global coverage. Still, in practice, it is not always the case when the observations provide direct information for our state of interest, as with river discharge, where a hydrological model is applied to convert precipitation data into flows.

Data assimilation (from now on, referred to as DA), by observation merging with model predictions, can produce state estimates with a good prediction accuracy and uncertainty quantification (Liu and Gupta 2007a). The correct initial conditions for a future forecast are obtained by updating the model states with external measurements (Rakovec, Weerts, et al., 2012). The uncertainty inherent to the new estimates is intended to be lower than that associated with the model forecasts or observations alone.

Different DA techniques have been developed and are widely used in hydrological forecasting, including approaches like interpolation (e.g., optimal interpolation), filtering (e.g., the Classic, Extended, and Ensemble Kalman Filter), and smoothing (4DVar, Ensemble Kalman Smoother) algorithms. The majority of operational hydrological forecasting systems now use lumped models (with manual or deterministic state updates). However, a substantial shift toward distributed models with hydrological ensemble predictions is noticed, as they present more flexibility in terms of how uncertainty is accounted for (Liu et al., 2012).

1.1 Motivation

The main motivation for this research is to see the potential impact and uses of a better forecasting model could have. This study will focus on data assimilation (DA) effects on distributed hydrological model states and fluxes (e.g., wflow_sbm). The primary advantage of applying data assimilation techniques with distributed models lies in the potential of such models to be forced using spatially observed data, which are now widely accessible due to remote sensing and radar applications. Moreover, distributed model states provide a better representation of the model states, resembling actual observations within the catchment (e.g., subsurface water flow, soil moisture, and discharges) more than the lumped states over the entire catchment. (Rakovec, Weerts, et al., 2012). A data assimilation application technique using spatially distributed models, as recommended by Liu et al. (2012), is tested in two selected small-scale experiments, using the Ensemble Kalman Filter for assimilating the observations.

1.2 Objectives

This study's main objective is to use assimilation of the discharge observations with a distributed hydrological model (wflow_sbm) to derive better quality forecasting and increased accuracy for a flood-prone area (River Vecht). It is necessary to complete the following specific objectives to attain the overall goal:

- Assess the effect of streamflow assimilation on discharge prediction quality for the Vecht river.
- Estimate the effect of applying DA on hydrological model states and fluxes using a model where water is routed via surface and subsurface.
- Examine the influence of uncertainty selection on the discharge prediction quality for the Vecht river.

1.3 Research questions

The primary research goal of this study is to examine the effects of data assimilation in enhancing model flow predictions when used with a distributed model. For this objective to be accomplished, the following research questions are identified and addressed in this work:

- What is the effect of assimilating streamflow observations with the wflow_sbm distributed hydrological model on the streamflow predictions for the Vecht river using the EnKF? What is the influence of the assimilation window length and the number of assimilation locations?
- What effect does discharge data assimilation have on other hydrological model states and fluxes, given that we use a model in which water is routed laterally via subsurface and surface water, whereas previous studies used a model in which water was routed only via surface water?
- What is the influence of the error model specification of the input forcing and observations on the discharge predictions with the wflow_sbm model?

1.4 Practical value and innovation

Extensive research on the assimilation of discharge measurements into lumped hydrological models has been reported, e.g., Weerts and El Serafy (2006). Additionally, several studies have examined DA applications within spatially distributed hydrological models, considering conceptual and operational aspects (e.g., Rakovec et al. 2012). This work assesses DA techniques applied to the wflow_sbm distributed hydrological model using computationally effective algorithms. However, only a few studies utilizing DA with distributed hydrological models have been reported, like the wflow_sbm, where the water is laterally routed through subsurface and surface water. In contrast, previous studies on DA with distributed models only used models that only routed water through surface water.

In light of the floods in Europe in July 2021, improved discharge predictions may facilitate better flood forecasting and decision-making. Predictions of discharge may alert water officials to impending high water levels, allowing them time to take quick action, monitor the situation, and potentially prevent severe consequences. In addition, if the research findings demonstrate that the employment of DA with the wflow_sbm model considerably improves the quality and accuracy of discharge prediction, the method employed in this study may become a suitable alternative. The model used in this work (wflow_sbm), together with the assimilation tool OpenDA is open source, so if no restrictions on the data availability are encountered, it can be applied to other river basins. This work is relevant to hydrologists, water experts, decision-makers, and other fields related to operational hydrology.

1.5 Operational flood forecasting in the Netherlands

The river Vecht is a completely rain-fed catchment as it depends on precipitation. As it is a transboundary river flowing from Germany to the Netherlands, many stakeholders engaged in the water management issues of the river, resulting in it lacking an operational forecasting system until 2011, when with the help of the Delft-FEWS system, it was finally built. Delf-FEWS is a data model integration platform that collects real-time fluvial and meteorological observations. HBV-96, a spatially lumped rainfall-runoff model, has been operationally running since 2012 to forecast floods in the Vecht river. The system can forecast five days using a 5-day ECMWF ensemble forecast (Heerinen et al., 2013).

While the forecasting system has proved successful in multiple small flood occasions, in the context of a changing climate and advances in hydrological models and operational hydrology in general, it is sensible to develop and evaluate more effective ways of forecasting in the area. The operational forecasts generated from the HBV-96 model are employed as lateral inputs into the 1D hydrodynamic model (Sobek). The HBV-96 model does not use discharge assimilation for sequential model state updating; instead, an automated auto-regressive error correction approach is used to update the model with discharge measurements. The application of this automatic error correction is analogous to what is described in the work of Broersen and Weerts (2005) for the Rhine catchment. Like the Vecht river forecasting system, the operational forecasting for the Geul is derived using a lumped HBV-96 and a hydrodynamic 1D Sobek model. As it is a small steep catchment fed from precipitation, with quick response and a time skill of just several hours, it is currently aimed by the water authorities to establish a nowcasting system to avert future flooding as severe as July 2021.

The research presented in this thesis explores discharge data assimilation for state updating with wflow_sbm that can be used to derive an updated forecasting system for Vecht.

2 LITERATURE REVIEW

This chapter reports the state-of-the-art on data assimilation and various verification methods for deterministic and ensemble forecast evaluation. First, a general overview of the theoretical backgrounds of data assimilation algorithms, focusing on the Ensemble Kalman Filter (EnKF), is given. Finally, some remarks on the OpenDA toolbox and some details on its applications with hydrological models are presented.

2.1 Data assimilation concepts and methods

Data assimilation is defined as "the process of employing observational data to update a model" (WMO, 1992). Data assimilation is often regarded as the way of maintaining a model "on track" by continuously adjusting it with new observations (University of Reading, 2022).

The main ways to update a model are as follows:

- Input update: forcing data errors are corrected using observation-based forcing, resulting in improved model predictions.
- State update: model states are corrected to get closer to observational data.
- Parameter estimate: model parameters are corrected using observations (e.g., calibration).
- Error correction: model simulations are corrected with a long-term experimental error to diminish systemic model bias (e.g., bias correction).

In hydrology, DA techniques can be sequential or variational. Sequential approaches are typically used for state update of a hydrological model by assimilating external observations as they become available at each time step. Their impacts reflect the observational data and model states' uncertainties. Variational approaches aim to optimize a cost function throughout an experiment's duration. The initial construction of a first-guess model is followed by developing an adjoint model that propagates in time and integrates model-observation misfit (Rakovec, Weerts, et al., 2012).

Optimal interpolation (OI), also known as statistical interpolation, is one of the most used static approaches in DA, as it is a reasonably simple method. OI is a minimum variance estimator that determines the optimal solution based on weighted least squares fit observations and a first-guess simulation (Lorenz 1981). However, the term "optimal" is misleading, as it is challenging to specify error covariances accurately in practice.

The **3-D variational (3D-Var)** method directly solves the recurrent optimization problem, as observations are retrieved at the analysis time, thereby requiring less extensive processing capabilities. The background covariance matrix is typically approximated using the same method as OI.

The **4-D variational (4D-Var)** is an upgrade of the 3-D variational (3D-Var) system wherein observations are assimilated at the exact time they are measured. Within the assimilation window, this assimilation algorithm computes the initial condition in which the prediction most closely reflects the observations. Both methods determine the optimal guess at the time of analysis by optimizing a cost function, which narrows the distance between measurements and a model's forecasted track.

In contrast to OI, variational assimilation approaches (3D-Var, 4D-Var) allow greater flexibility for assimilating multiple variables with a nonlinear relationship to the model, accommodate a large number of observations, and provide a solution for the entire system without dividing it. In addition, they establish the foundation for adopting more complex background-error covariance models (Weaver, 2003). On the downside, the 4D-Var method assumes a perfect model and assigns equal weight to older and more recent observations at the start and end of the assimilation interval.

The **Kalman Filter**, developed on Kalman's (1960) work on optimum control of linear systems, is a popular sequential algorithm for data assimilation. It is a discrete, recursive algorithm for state estimation that optimizes mean squared error. The dynamic update of the predicted (background) error

covariance over time defines the Kalman filters from more static algorithms such as OI. The classic Kalman filter (KF) employs a standard error propagation theory and a linear tangent model. As mentioned in Liu et al. (2012), referring to Liu and Gupta (2007), the generic dynamic state-space update can be summarized as below:

$$x_{k+1} = M_{k+1}(x_k, \theta, u_{k+1}) + \eta_{k+1} \quad (1)$$

$$z_{k+1} = H_{k+1}(x_{k+1}, \theta) + \varepsilon_{k+1} \quad (2)$$

where x_k and x_{k+1} are the system's true state vectors at the time t_k and t_{k+1} respectively; M_{k+1} is a nonlinear operator representing the system propagation from time t_k to t_{k+1} responding to the input vector of the model u_{k+1} ; θ consists of a time-invariant model parameter vector that is derived from physical concepts or calibration of parameters; z_{k+1} is the observation vector referring to model properties and states through the use of an observation operator H_{k+1} ; η_{k+1} is the term used for the model error with mean $\bar{\eta}_{k+1}$ and a covariance Q_{k+1} ; the observation uncertainty is represented as ε_{k+1} with a mean $\bar{\varepsilon}_{k+1}$ and covariance R_{k+1} .

The prediction (forecast) and update steps are the two main components of the KF algorithm. The algorithm's prediction step uses the initial estimates for the state and the error covariance. The results of the prediction step are integrated with measurement data in the update step to create an updated state estimate, which is then used to initialize the subsequent prediction, which will then be utilized in the following update. Figure 1 presents the equations and steps of the KF algorithm are presented.

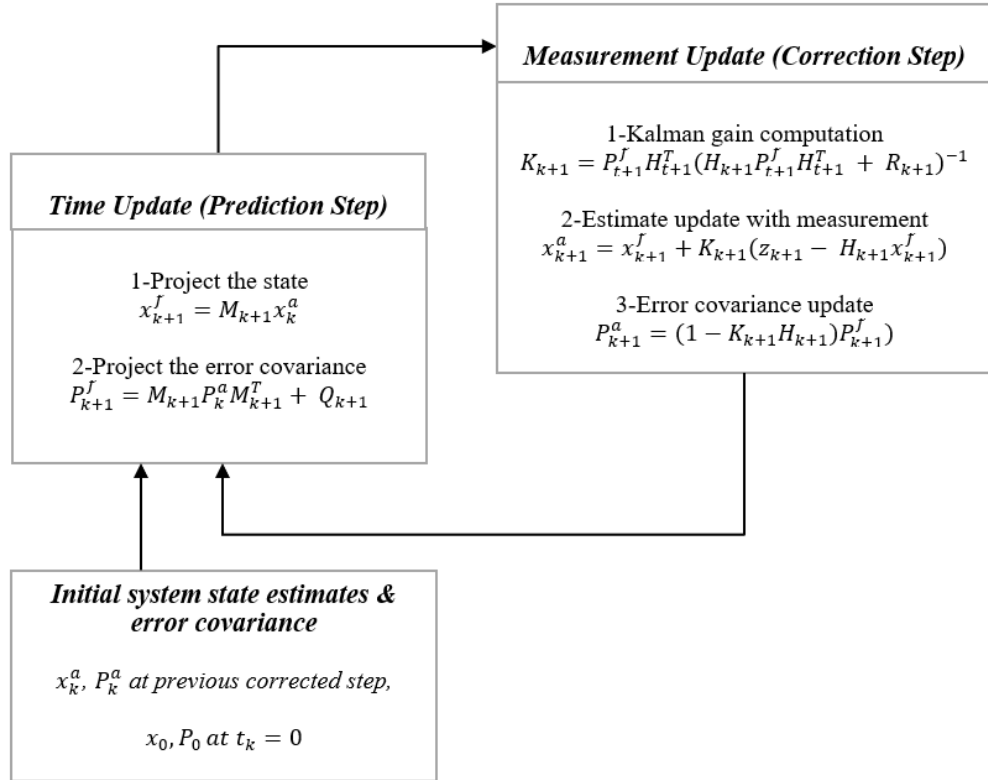


Figure 1: Classic Kalman Filter Algorithm

M and H operators are assumed linear in the Kalman Filter (KF); yet, hydrological systems are typically nonlinear, restricting the application of KF. Variations of the KF method to provide approximated solutions for non-linear situations like the extended Kalman Filter (EKF), Ensemble Kalman Filter (EnKF), etc., have been developed.

A nonlinear variation of the KF is the **Extended Kalman Filter (EKF)**. However, in the case of EKF, local tangent linear formulations of the model and observation operators (M and H) are applied. Despite specific successful uses of the EKF, as cited in Liu and Gupta (2007), Evensen (1994) considered that the EKF can lead to instability or even divergence since it disregards the model's second and higher-order derivatives.

Evensen (1994) proposed the **Ensemble Kalman Filter (EnKF)** as a substitute for the EKF and has since become the most widespread DA technique in hydrology because of its simple conceptual formulation and robustness when used with DA applications (Liu et al., 2012). It is a Bayesian recursive estimation approach that uses observations to assess model states' proper probability density function. Several state forecast ensembles are employed rather than a single discrete estimate of covariances. Adding observation perturbations is critical in EnKF to avoid ensemble members collapsing.

The **Asynchronous Ensemble Kalman Filter (AEnKF)** is developed from the adaptation of the (synchronous) Ensemble Kalman Filter (EnKF) using a data-driven state updating method. The AEnKF incorporates all measurements in a single update and uses a longer assimilation window. The AEnKF is attractive for operational application since additional measurements can be used with little added computation time. Rakovec (2015) examined the benefit of an extended assimilation window in his work.

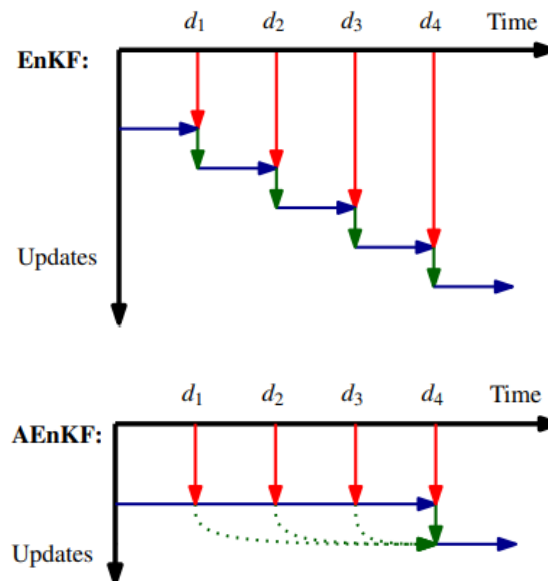


Figure 2: Diagram of the model update differences between the EnKF and the AEnKF (Rakovec et al., 2015)

2.2 Uncertainty definition

While there are recommendations and methodologies for characterizing model prediction uncertainties in the literature, selecting the optimal model error remains challenging in most DA applications. To prevent the filter from failing, arbitrary model uncertainty is typically defined. Inconsistent data coverage, observation errors, improper DA structure, and model initialization are among the sources of DA uncertainties.

The accuracy of error estimates of the model input forcing and flaws in the model structure because of parameterization of physical phenomena or unaddressed scaling issues influence how efficient the DA approach is (Liu et al., 2012). Furthermore, an assumed perfect model structure does not guarantee a perfect prediction, as the model parameter estimates could also be uncertain. As a result, the effectiveness of a DA scheme is highly influenced by the quality of the input and model uncertainty estimates, the accurate representation of the dependencies between model parameters, etc.

In hydrology and operational water management, it is essential to quantify precipitation accurately in terms of its intensity and location. Given the small temporal and spatial correlation scales, precipitation is thought to be the most uncertain input for hydrologic models. Traditionally, spatial precipitation data is retrieved from rain gauge recordings based on estimates at the point scale. Yet, the application of weather radars at increasing scales over recent times has enhanced the knowledge of precipitation patterns. Radar precipitation estimates are prone to error, resulting in the need for operational nowcasting/forecasting systems to employ rain gauge data for minimizing errors in radar QPEs, e.g., as in the Radolan product (after Kreklow et al. 2019). Considering the strong influence of precipitation uncertainties on the model predictions, reliable estimates of these uncertainties are essential for optimal DA applications. Recent advancements in radar performance and rainfall uncertainty estimation are described by Van de Beek et al. (2010) and other authors.

Stochastic perturbations are typically employed to quantify precipitation uncertainty in DA applications (Weerts & el Serafy, 2006). A stochastic perturbation approach typically determines the size of precipitation error noise primarily by considering the precipitation scale. For instance, the standard deviation of precipitation errors is considered equal to fifty percent of the rainfall at each model time step (e.g., Reichle et al. 2002).

As precipitation estimates errors present a high spatial and temporal variation, such order-of-scale-based approaches might lack statistical validity. Therefore, the estimates derived from conditional simulation techniques will be more appropriate as they are statistically more reliable. However, conditional simulation techniques may necessitate additional data parametrization and be computationally inefficient and costly compared to ad hoc stochastic perturbations (Liu et al., 2012).

Model conceptualizations and mathematical structures are accountable for the model's inherent uncertainty. Model error quantification is challenging due to its complex nature, highly dependent on the interactions between the model's various sources of uncertainty. As referred in Rakovec (2014), according to Liu et al. (2012), the hydrological community has adopted the following four methods to measure model error:

(1) stochastic perturbation of the model states analogous to stochastic perturbation of precipitation (e.g., Reichle et al. 2002).

(2) inverse techniques for deriving probability distribution functions of the model parameters (e.g., Vrugt et al., 2003).

In inverse approaches, model parameters and initial conditions are assumed to be perfect. Thus, unrealistic values might be assigned to model variables to compensate for model parameters and initial conditions errors.

(3) synchronous model states and parameters updating and uncertainty quantification in a sequential or recursive DA system (e.g., Vrugt et al., 2005).

The real-time update of model state and parameter values in this method enables the model to correctly simulate the system response at each observation time, depending on the applied updating method (e.g., linear update in EnKF vs. sequential Bayesian update and resampling in particle filtering). Numerous uses of these techniques for river discharge prediction, soil moisture, other state estimation, groundwater flow modeling, etc., have been documented.

(4) multi-model ensembles application (e.g. Georgakakos et al., 2004).

This stochastic perturbation approach accounts for the model uncertainty term η_{k+1} (See equation 1) by

adding perturbations to the model physics, alternative methods account for the model uncertainty by quantifying model parameter errors.

2.3 Verification methods

The hydrological model prediction performance is assessed by different verification methods, most commonly using the differences between the predicted and observed variables. A plethora of verification methods are available for both deterministic and probabilistic forecasts, including scatter plots, error measures, and skill scores. Some popular verification methods for evaluating DA effectiveness are summarized in this section. While deterministic verification methods (e.g., RMSE, ME, etc.) evaluate an ensemble's mean using a single observation, probabilistic methods (e.g., CRPSS) evaluate the entire ensemble using a single observation. In this work, deterministic and probabilistic simulations are used. However, deterministic verification methods are used to evaluate the model predictions in each case (for the probabilistic simulations, the ensemble mean is used).

2.3.1 Performance verification of deterministic forecasts

For performance evaluation of deterministic predictions, metrics like the Root Mean Square Error (RMSE), Percent Bias (pBIAS), Nash-Sutcliffe Efficiency (NSE), Kling-Gupta Efficiency (KGE), etc. are commonly used.

The **Root Mean Square Error (RMSE)** is an accuracy metric used to evaluate a model's performance compared to observed values. It is calculated using equation (3) as below:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i^s - Y_i^o)^2} \quad (3)$$

RMSE is a commonly used indicator that provides an idea about the error distribution and determines whether the model is over or under-predicting. It certainly has several drawbacks, like being scale-dependent; hence it should only be used to evaluate methods within the same dataset and not between scales (Hyndman & Koehler, 2006). As it allocates higher weights to errors of higher values, it is also more sensitive to outliers. RMSE ranges from zero to infinite, with 0 being the ideal value, and it employs the same units as the compared variables.

The **Percent Bias (P bias)** is a mathematical relationship between the mean error and the average observed value. It indicates whether the simulated values are more or less likely to be greater or smaller than the observation values. The Pbias is computed using the following equation:

$$pBIAS = \left[\frac{\sum_{i=1}^n (Y_i^o - Y_i^s) * 100}{\sum_{i=1}^n Y_i^o} \right] \quad (4)$$

P bias has a range of infinite negative to infinite positive values, with P bias = 0 being the optimal value. A positive P bias value suggests the observed values being overestimated compared to the simulated values, whereas a negative P bias value indicates underestimation.

Nash-Sutcliffe Efficiency (NSE) is a metric typically used to assess hydrological models' predictive quality. It indicates how well predictions compare to the mean of the data and is computed using the formula as follows:

$$NSE = 1 - \left[\frac{\sum_{i=1}^n (Y_i^o - Y_i^s)^2}{\sum_{i=1}^n (Y_i^o - Y_i^{mean})^2} \right] \quad (5)$$

This dimensionless indicator ranges from 1 to infinite negative values, with $NSE = 1$ being the ideal value and $NSE = 0$ indicating that the model's performance is the same as if the observed mean values were used as the predictive model.

In equations (3), (4), and (5), Y_i^o are the observed values, Y_i^s the simulated values, while n refers to the number of observations.

Kling Gupta Efficiency (KGE) is an improved variant of the NSE that considers several types of model errors like correlation, bias, and variability. It is typically employed for hydrological model calibration and evaluation. See Gupta et al. (2009) for a thorough description of KGE and its advantages over NSE. There are three main components included in the computation of the KGE index, as shown in the equation below:

$$KGE = 1 - \sqrt{(C_c - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (6)$$

$$C_c = \frac{Cov_{Y^s Y^o}}{\sigma_{Y^s} \sigma_{Y^o}}; \alpha = \frac{\sigma_{Y^s}}{\sigma_{Y^o}}; \beta = \frac{\mu_{Y^s}}{\mu_{Y^o}}$$

where C_c is the Pearson correlation coefficient, with an ideal value of $C_c=1$, computed as the ratio of covariance between simulated and observed values $Cov_{Y^s Y^o}$ with the respective standard deviation of the simulation σ_{Y^s} and observation σ_{Y^o} values; α is the ratio between the standard deviation of the simulated values σ_{Y^s} and the standard deviation of the observed ones σ_{Y^o} , with the ideal value being $\alpha = 1$; and β represents the ratio between the mean of the simulated values μ_{Y^s} and the mean of the observed ones μ_{Y^o} ideally with a value of $\beta = 1$. The optimal KGE indicator will have a value of 1 for the three components, varying from one to negative infinite.

Nash and Sutcliffe's coefficient of determination R^2 (1970) can be used to assess how a DA technique improves a hydrological prediction. While it is identical to the NSE, their key distinction lies in their application, as R^2 is particularly favorable when applied to the ensemble mean. R^2 is a metric of a statistical model's goodness of fit and is computed as follows:

$$R^2 = \left(1 - \frac{MSE_{assimilation}}{MSE_{deterministic}}\right) \times 100 \quad (7)$$

where the Mean Square Error (MSE) is computed as:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i^o - Y_i^s)^2 \quad (8)$$

Where $MSE_{assimilation}$ denotes the mean square error between the ensemble mean and the respective observations after updating with DA, and $MSE_{deterministic}$ is the mean square error between the deterministic simulation (without DA) and the observations. Positive r^2 values imply that the data assimilation is beneficial for forecasting, whereas negative r^2 values suggest a negative effect.

2.3.2 Performance verification of ensemble forecasts

The **Continuous Rank Probability Score (CRPS)** (Brown, 1974; Hersbach, 2000) is a widely used accuracy statistic for ensemble forecasts. Unlike other scores, it compares the cumulative distribution

functions of ensembles and observations. It is defined by the integration of the squared differences between the cumulative density probability functions (cdfs) of the forecasted and observed as follows:

$$CRPS = \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{\infty} (P_i^{fc}(x) - P_i^o(x))^2 dx \quad (9)$$

where $P_i^{fc}(x)$ and $P_i^o(x)$ are the i -th ensemble member probability cdf, and the observation's cdf, respectively. As the observation is a single value, the forecast probability of observations is expressed as a step function ranging from 0 to 1. The unit dimensions of CRPS are the same as the observations. The optimal value of CRPS is zero when the ensemble's cdf equals the observation's cdf.

Skill scores are a common way of evaluating a forecast. A skill score determines the accuracy of the forecast by measuring its relative distance to a reference forecast. The following is the general equation used to calculate forecasting abilities:

$$Skill = \frac{Score_{ref} - Score_f}{Score_{ref}} \quad (10)$$

$$CRPSS = \frac{CRPS_{ref} - CRPS_f}{CRPS_{ref}} \quad (11)$$

$$RMSES = \frac{RMSE_{ref} - RMSE_f}{RMSE_{ref}} \quad (12)$$

where $Score_f$ represents the quality of the ensemble forecasts (calculated using one of the statistics, e.g., RMSE, CRPS) and $Score_{ref}$ is the verification score of the reference forecast. A perfect score is the value of 1 and no skill for the value of 0.

2.4 OpenDA

OpenDA is an open-source generic toolbox used in data assimilation applications, which includes a set of configuration files that define interactions between models, data assimilation algorithms, and observations (OpenDA, 2013). Its modular design supports flexible applications for geophysical process modeling. For example, in hydrology, DA methods can enhance real-time forecasts, calibrate uncertain model parameters, perform uncertainty analysis, and so on.

The three major building blocks of OpenDA are presented in Figure 3, where the method describes the DA (or calibration) algorithm, e.g., 3DVar, EnKF, AEnKF algorithms, etc.; the observations include the stochastic observer for handling the observations, and lastly, the DA algorithm is applied to the model, which in this case is the wflow_sbm model. Still, it can be any external model compatible to the OpenDA (van Velzen et al., 2016).

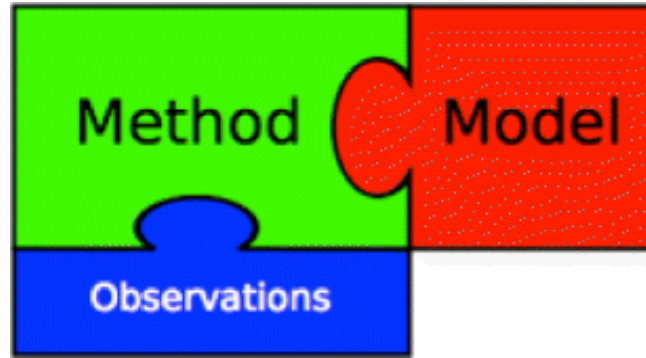


Figure 3: Modular components of OpenDA: method (the DA method), observations (the stochastic observer for processing the observations), and the model (van Velzen et al., 2016)

The model delivers a deterministic run to OpenDA, which can integrate the deterministic model run with the observations in the chosen algorithm to generate the desired output.

OpenDA brings several advantages, including a clear modular framework in which the model, observations, and algorithm are distinct and can be easily configured using an XML schema, from which the user can control the type of algorithm to be used and its settings, like the number of ensemble members; the observations to be assimilated and their relative grid placement inside the model, as well as the error associated with the forcing input, the measurements, and the model itself.

Another benefit is OpenDA's versatility in formats that the observations can be stored (NetCDF, CSV, NOOS, SQL database), and the outputs can be produced (NetCDF, ASCII, Matlab, Python, etc.). A significant downside of OpenDA is its lengthy computing times, the simulation's tendency to fail when longer assimilation windows are employed, and, in my personal experience, the inability to store results when the simulation is run over multiple dates (e.g., simulation starts on 25th of July and ends 26th of July). Using the so-called black box coupling and a model wrapper, OpenDA and the model can be incorporated as a module in the Delft-FEWS.

3 CATCHMENT DESCRIPTION

Experiments will be conducted for the Geul river catchment (Figure 4), which drains an area of approximately 340 km² and is located partially in Belgium, the Netherlands, and Germany (52 percent, 42 percent, and 6 percent, respectively), as well as the Overijsselse Vecht river catchment (Figure 5), which drains a significantly larger area of approximately 1,800 km². Significant elevation changes in the Geul, a rather steep, primarily rain-fed catchment, result in rapidly changing discharges, whereas the Vecht catchment has a slower reaction due to a distinct soil formation and primarily to more minor elevation differences.

Table 1: General characteristics of the Geul and Vecht river catchments

	Geul	Vecht ¹
(Latitude, longitude)	(50 80 N, 05 91 E)	(52 17 N, 07 14 E)
Upstream area [km ²]	344	1980
Annual precipitation [mm]:	800	730
Predominant land use	agricultural area	grassland
Predominant soil type	coarse-silty	loamy sand
Altitude	379–564 m a.s.l.	-3–167 m a.s.l.

3.1 Geul river basin

A tributary of the River Meuse, the Geul river is located in South Limburg, the most southern region of the Netherlands (Figure 4). Originating in the Lichtenbusch region, in northeastern Belgium, near the German border, south of Aachen, it enters the Netherlands at Cotessen in the Vaals municipality. It has a total length of fifty-six kilometers and a catchment area of 380 square kilometers. It is among the few hilly rivers in the region and has a very steep gradient, with the catchment's altitude decreasing from 400 m a.s.l at the source area to 50 m a.s.l where it joins the Meuse River, north of the city of Maastricht.

The gradient range drops from 0.02 m/m close to the source to 0.0015 m/m at the outflow (de Moor et al., 2008). Gulp, Eyserbeek and Selzerbeek are the three tributaries of the Geul. According to the Koppen climate classification system, the climate can be described as Cfb, which indicates a continental climate with warm summers but no dry spells. An average discharge of 4 m³/s characterizes it, varying from 0.8 to 65 m³/s respectively during summer and winter storms, and an annual average rainfall of 800 mm, ranging from 45 mm to 75 mm respectively in March and August (Dautrebande et al., 2000). During the past 50 years, the population more than doubled, urbanization was rapid, the river was straightened, riverbanks were protected to prevent erosion, land use was changed, and heavy agricultural machinery was introduced. All these aspects may have contributed to a change in the hydrological regime of the catchment, resulting in higher peak flows.

To reduce peak discharges, measures like “making room for the river” to meander, construction of water retention basins, repurposing areas that would be flooded during flood events into nature reserves, etc., were implemented after a severe flood event in the Meuse, at the turn of the twentieth century. Today, the river is partially confined (a few short stretches that flow through villages) and partially meandering (van den Munckhof, 2020). The Geul river basin is fed by surface, subsoil, and groundwater flow together. The river's baseflow is controlled by springs formed from groundwater recharge by rainfall traveling laterally through impervious layers to the surface, making Geul a mainly

¹ The statistics refer to the Dutch part of the Vecht river

rain-fed river, where the discharges can change rapidly. Recurring local floods happen with peak discharges as low as forty m³/s; however, the damage is limited to a few submerged meadows along the river.

Typically, rainfall intensity is low, although this is changing due to climate and land use changes. Summer rainfall is becoming more intense and concentrated, and torrential rainfall can cause overland flow on loess-covered slopes and a quick increase in river discharge, as was witnessed in July 2021, when the river flooded, causing extensive damage.

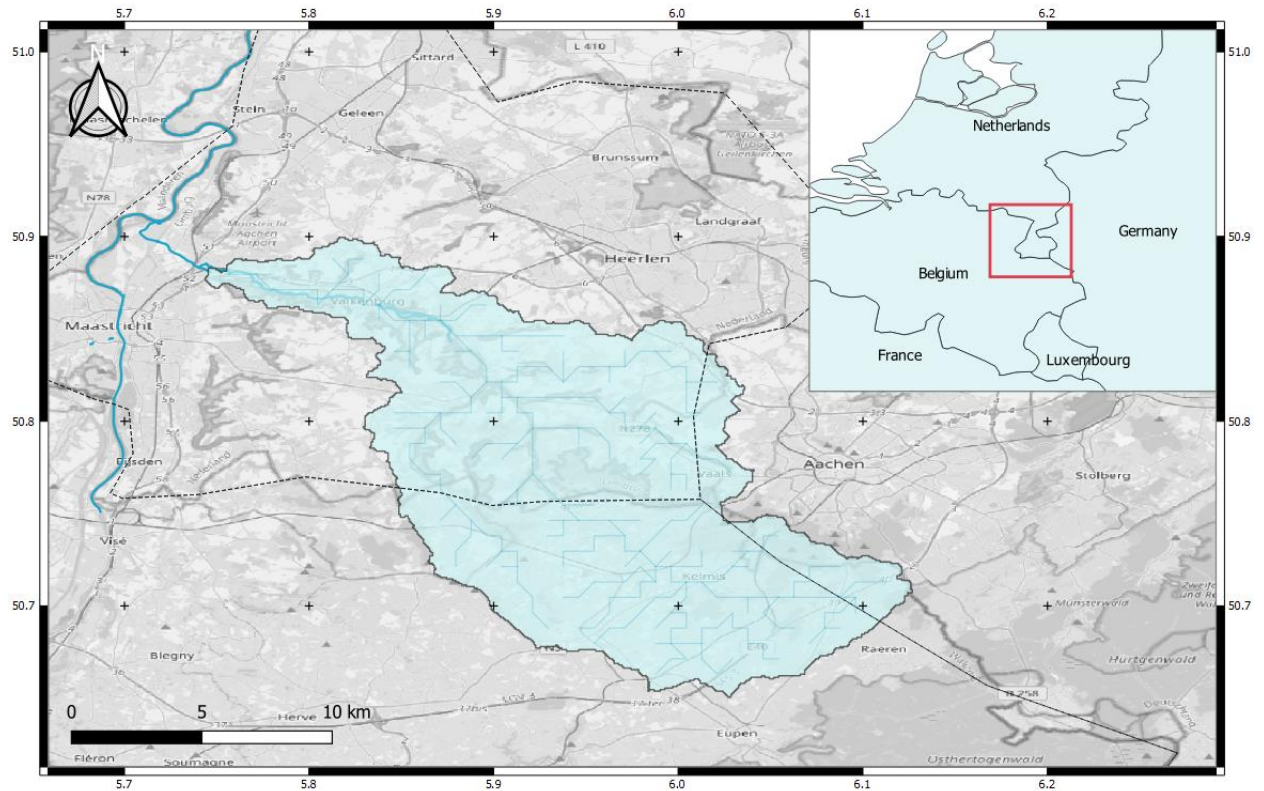


Figure 4: Geul river basin location

3.2 Vecht river basin

The Overijsselse Vecht, a tributary of the Rhine, is a low-energy, sand-bed river that flows partially through the Netherlands and partly through Germany (Figure 5). Originating in the German state of Nordrhein-Westfalen flows through the state of Niedersachsen and, after approximately one hundred kilometers, crosses the Dutch-German border, where it flows through the Dutch provinces of Overijssel and Drenthe before discharging into the Zwarte Water Lake, flowing into lake IJssel, located in the central-north Netherlands. It has a total length of 182 kilometers, of which sixty kilometers are located in the Netherlands, and a total catchment area of 3,780 kilometers squared, of which 1980 squared kilometers are located in the Netherlands (Bijlsma & Blind, 2006).

The elevation difference ranges from more than 100 m over the entire course to only 10 m for most Dutch parts. It is a rain-fed river system with a mean annual flow of around 22.8 m³/s and a mean annual flood peak of 160 m³/s measured at Mariënberg gauge station from 1995 to 2015. The catchment average annual precipitation in the Overijsselse Vecht varies from 700 to 825 mm, with average evapotranspiration of 525 mm. The mean annual temperature is nine °C, ranging from –22 to 35 degrees Celsius (KNMI). The catchment's outlet near Emlichheim, near the border with Germany, is of particular

interest as it represents the upstream hydrological and flood risk boundary condition for flood risk assessment on the Dutch side. The Overijsselse Vecht valley geomorphology was shaped by fluvioperiglacial sands that were locally topped by aeolian cover sands (Huisink 2000). Aeolian drift sands accumulated near the Overijsselse Vecht from the overexploitation of agriculture during the Late Holocene (Jan Pierik et al., 2018). The Vecht was a meandering river until 1896 when sections of the river were diverted and weirs were constructed. The river was channelized with five weirs managing the water levels after 1914. As part of river restoration efforts aimed at restoring the river's original physical and ecological qualities, meandering side channels that bypass the weirs were recently constructed.

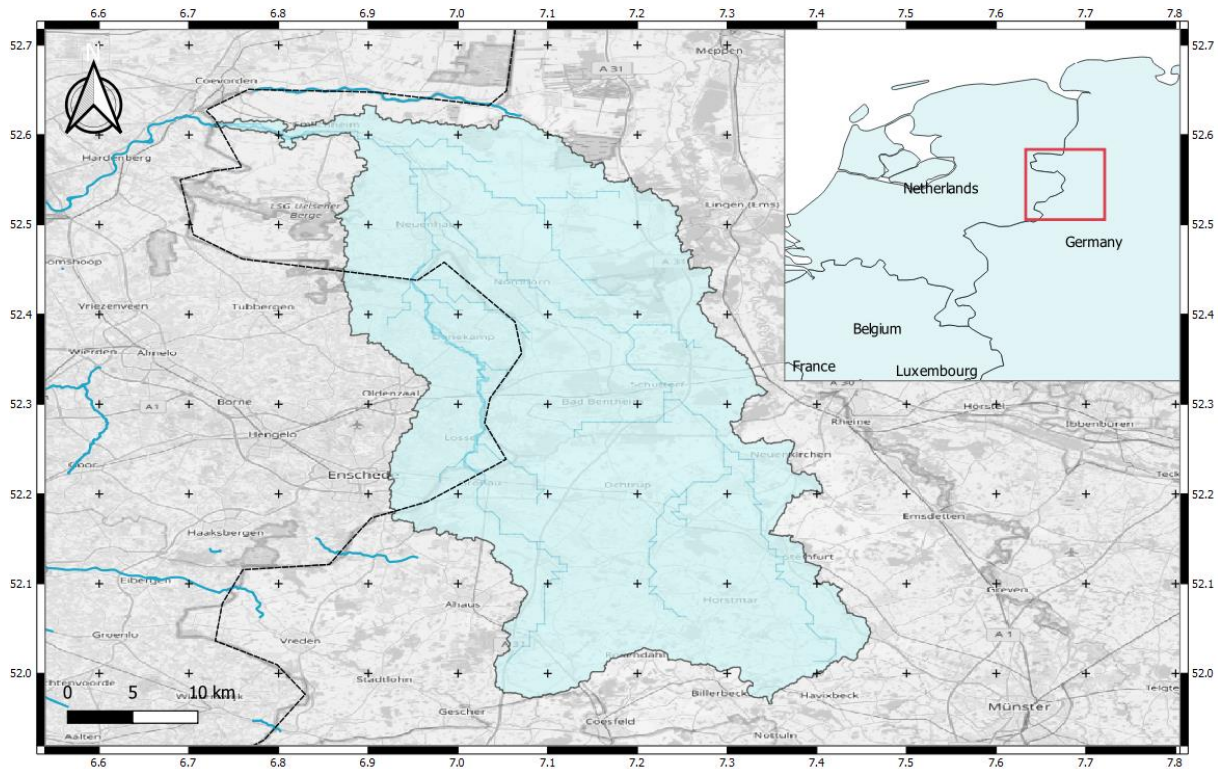


Figure 5: Vecht river basin location

4 DATA AND MODEL SETUP

4.1 Data availability

This section summarizes the available datasets for several meteorological variables, focusing on precipitation, and discharge data. Depending on the dataset type and provider, the spatial and temporal resolution of available rainfall, temperature, and potential evapotranspiration data varies. Before the data can be used to force the models, it is resampled to match its temporal and spatial resolution.

4.1.1 Precipitation

The precipitation product can be obtained from various sources, including the Royal Netherlands Meteorological Institute (KNMI), the Dutch national weather service; the Climate Data Center (CDC) of the German national weather service (DWD); the Climate Data Store (CDS) of the Copernicus Climate Change Service and other global datasets.

The KNMI dataset, also known as Rd1 version 5, is a collection of files containing gridded daily precipitation sums measured at approximately three hundred stations in the Netherlands from 8 a.m. to 8 a.m. the following day. The spatial resolution is 1 km, with a daily temporal resolution of 1 week from 1950-01-01 to the present, with an average delay of 4 weeks.

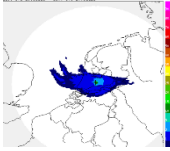

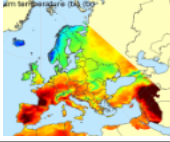
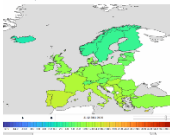
EOBS is an observation-based European daily gridded dataset for precipitation provided by the Copernicus Climate Change Services. The EOBS dataset is available in grids with spatial resolutions of 0.1° and 0.25° and daily resolution. It covers only land precipitation in Europe (including Iceland, as well as parts of Northern Africa, the Middle East, and Russia). ERA5 is the latest reanalysis dataset available at a spatial resolution of 0.25°, made available by the European Centre for Medium-Range Weather Forecasts (ECMWF).

There are uncertainties associated with ERA5 and the EOBS datasets stemming from measurements, spatial interpolation of the observations, and the model itself (reanalysis). REGNIE is a gridded dataset with a 1-kilometer spatial scale and daily temporal resolution, provided by DWD Climate Data Center CDC. The grids are derived from station measurements with the method REGNIE (DWD, 2017).

In addition, RADKLIM high-resolution radar-derived precipitation data can be employed. The data are from the DWD radar-based precipitation climatology based on the RADOLAN technique reprocessing version 2017.002. The precipitation sums (RW) are calculated using radar-based precipitation estimation and corrected using gauge precipitation data. The YW products are quasi-gauge-adjusted five-minute precipitation rates using RW. The data is accessible in a 1 km x 1 km grid and spans the years 2001 to 2017, with extensions until 2021 (DWD, 2022).

The main properties of the available and accessible datasets are listed in the table below. Whether the dataset derived from the actual observations (E-OBS) or the radar estimate is more appropriate to use will be evaluated by a short precipitation data analysis in a subsequent stage of this study.

Table 2: Summary of the available gridded precipitation datasets that provide coverage of the study areas

Source	Name	Domain	Period of record	Available timestep(s)	Available resolution	Available format(s)	Snapshot
KNMI	nl_rdr_data_rtc_or_5m	Netherlands	01.01.1951 to 14.02.2022	5 min	1 km x 1 km	HDF5	
Deutscher Wetterdienst DWD	REGNIE	Germany	01.01.1931 to present	Daily	1 km x 1 km	ASCII	
	RADKLIM		01.01.2001 to 01/01/2021	Hourly/5 min		Binary/ASCII	
CDS, Copernicus Climate Change Service, ECMWF	E-OBS	Europe	01.01.1950 to 30.06.2021	Daily	0.1° x 0.1°	NetCDF	
	ERA5		01.01.1979 to present	Hourly down to yearly	0.25° x 0.25°	NetCDF / csv	

As both river catchments are transboundary and a substantial area lies beyond the Netherlands' country boundaries, it is reasonable to utilize measurements from rainfall gauge stations within the basin area provided by the weather services of the bordering countries (Germany for both the Vecht and the Geul rivers and Belgium for the Geul).

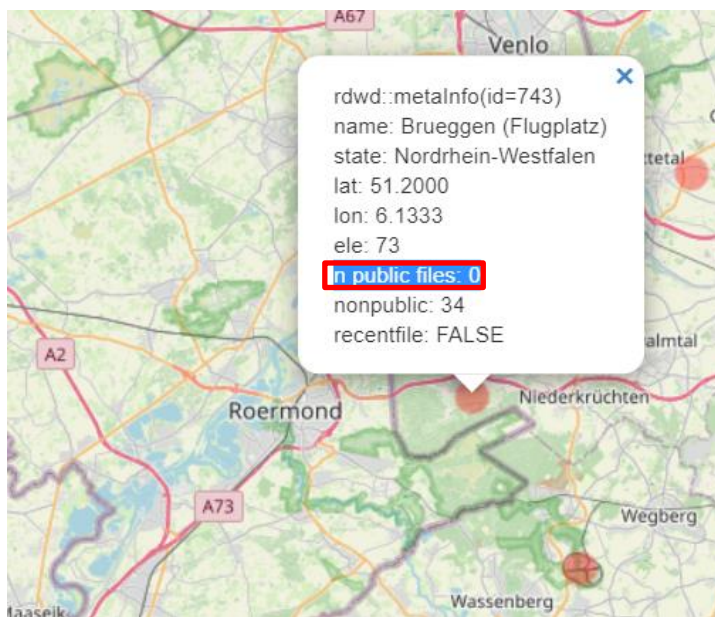


Figure 6: Snapshot of the interactive map made available by the rdwd R package, displaying a rain gauge station located near the border (catchment area), for which no public data are accessible

Below is a summary of the rain gauge station data collected from DWD for each catchment. The nearest stations for each catchment were selected using the rdwd R package's interactive map (Boessenkool, 2022). Unfortunately, it was observed that most border stations lacked accessible records (Figure 6). Nonetheless, a handful of observational stations with recent data were identified (Table 3 & Table 4) and utilized in this study.

Table 3: Rain gauge stations nearby Geul (Source: DWD)

STATION ID	NAME	Scale	LON °	LAT °	HEIGHT (m)	Period of record
4667	Selfkant-Havert	daily	5.91	51.04	46	01.01.1961 to 04.04.2022
15000	Aachen-Orsbach	hourly	6.02	50.80	231	01.04.2011 to 25.04.2022
15927	Übach-Palenberg	daily	6.09	50.93	82	01.10.2017 to 19.01.2022

Table 4: Rain gauge stations nearby Vecht (Source: DWD)

STATION ID	NAME	Scale	LON °	LAT °	HEIGHT (m)	Period of record
357	Bentheim, Bad	hourly	7.13	52.30	50	01.09.2005 to 31.12.2021
1223	Ringe-Großbringe	hourly	6.94	52.59	13	01.09.2005 to 31.12.2021
4849	Steinfurt-Burgsteinfurt	hourly	7.34	52.12	64	01.06.2005 to 05.04.2022
3640	Nordhorn-Blanke	daily	7.06	52.41	24	01.01.1931 to 24.04.2022
1230	Emsbüren-Ahlde	daily	7.29	52.37	48	01.01.1961 to 24.04.2022
5131	Twist	daily	7.11	52.63	21	01.01.1941 to 24.04.2022

4.1.2 Additional forcing data

The wflow_sbm hydrological model utilized in this work employs temperature and potential evapotranspiration as forcing data. The ERA5 dataset provides hourly and daily estimates for a variety of climate variables, such as temperature (T) and solar radiation (K^\downarrow), from which the potential evapotranspiration (E) can be computed. In practice, potential evapotranspiration is estimated from only remotely sensed daily global radiation and air temperature as inputs, using the modified Makkink equation, mentioned by De Bruin and Lablans (1998), as follows:

$$L_v E = 0.65 \frac{s}{s+Y} * K^\downarrow \quad (13)$$

where E denotes evapotranspiration, L_v the latent heat of vaporization, K^\downarrow is the global radiation, s the saturated vapor pressure gradient relative to water, Y the psychrometric constant.

4.1.3 Discharge

Discharge data for the Vecht river (Table 5) are available from 12 gauge stations (outlet and interior gauges), of which eight are obtained from the Lower Saxony State Office for Water Management, Coastal Protection and Nature Conservation NLWKN and four are obtained from the Water Information System of the North Rhine-Westphalia state ELWAS with a daily temporal resolution.

Table 5: River gauge stations in the Vecht river

Station ID	Name	Lon	Lat	Temporal resolution	Period of record	Source
9286106	Ohne	7.287	52.271	daily	01.02.1968 to 31.12.2017	NLWKN
9286127	Wehr Neuenhaus	6.971	52.512	daily	01.01.1950 to 30.12.2017	NLWKN
9286136	Lage I	6.971	52.463	daily	01.03.1963 to 31.12.2017	NLWKN
9286137	Lage II	6.971	52.465	daily	01.05.1963 to 31.12.2017	NLWKN
9286138	Lage III	6.963	52.467	daily	01.04.1972 to 31.12.2017	NLWKN
9286155	Osterwald	7.029	52.545	daily	01.11.1963 to 31.12.2017	NLWKN
9286162	Emlichheim	6.857	52.604	hourly	23.05.2004 to 05.05.2022	NLWKN
9316285	Tinholt	6.950	52.554	daily	01.11.1968 to 31.12.1997	NLWKN
9286110	Darfeld	7.254	52.063	daily	11.01.1957 to 07.03.2010	ELWAS
9286190	Bilk	7.296	52.238	daily	04.07.1957 to 01.01.2020	ELWAS
9286270	Temmingsmühle	7.354	52.112	daily	01.11.1957 to 01.01.2020	ELWAS
9286455	Gronau	7.021	52.221	daily	01.11.1968 to 19.12.2019	ELWAS

For the Geul river, discharge is measured at ten different points at different temporal resolutions, as summarized in Table 6. The data were obtained from the Water Board of Limburg (Waterschap Limburg) and the Agriculture, Natural Resources, and Environment entity of the Public Service of Wallonia (SPW).

Table 6: River gauge stations in the Geul

Station ID	Name	Lon	Lat	Temporal resolution	Period of record	Source
10.Q.29	Cottessen	5.934	50.758	15 mins	1993.01.01 to 12.11.2021	WL
10.Q.30	Hommerich	5.914	50.807	15 mins	03.09.1969 to 12.11.2021	WL
10.Q.36	Meerssen	5.726	50.891	15 mins	01.01.2010 to 12.11.2021	WL
10.Q.63	Schin op Geul	5.869	50.855	15 mins	01.01.2016 to 12.11.2021	WL
11.Q.32	Eyserbeek, Eys	5.929	50.825	15 mins	01.01.1993 to 12.11.2021	WL
12.Q.31	Selzerbeek, Partij	5.922	50.811	15 mins	03.09.1969 to 12.11.2021	WL
12.Q.46	Selzerbeek, molentak	5.912	50.814	15 mins	01.01.1993 to 12.11.2021	WL
13.Q.34	Gulp, Azijnfabriek	5.891	50.814	15 mins	03.09.1969 to 12.11.2021	WL
L6660	Sippenaeken	5.941	50.750	hourly	13.06.1996 to 31.12.2021	SPW
52911002	Kelmis (La calamine)	6.010	50.717	hourly	01.01.2009 to 31.12.2021	SPW

4.2 Wflow_sbm hydrologic model

The wflow_sbm distributed hydrological model described by Schellekens et al. (2020) was used in this study to apply and evaluate the data assimilation method in rainfall-runoff process modeling. Wflow_sbm computes the water balance at each point in each time step using gridded topography, forcing data, soil, and land use, and it is used to model the hydrological processes in the Geul and Vecht river catchments. The kinematic wave equation is used to compute the hydrological routing for the river flow, surface flow, and horizontal subsurface flow.

The structure of the model is illustrated in Figure 7. Snow, soil moisture, upper storage zone, and lower storage zone are the model states considered in each grid cell. Rainfall, snowfall, snowmelt, actual evapotranspiration, infiltration, capillary rise, percolation, direct runoff, quick flow, and baseflow are the fluxes that govern the behavior of the model states.

Built by Deltares, the wflow sbm models are provided for both the Geul and the Vecht river catchments. The wflow_sbm model for the Vecht is calibrated for horizontal hydraulic conductivity fraction in the work of Villareal (2021). As the models are assumed to be perfect (no model parameter uncertainties are considered further in this work), the parameters used in the model, their units and range of values, and the variation of each parameter throughout the catchments are as presented in Villareal (2021) and not detailed again here.

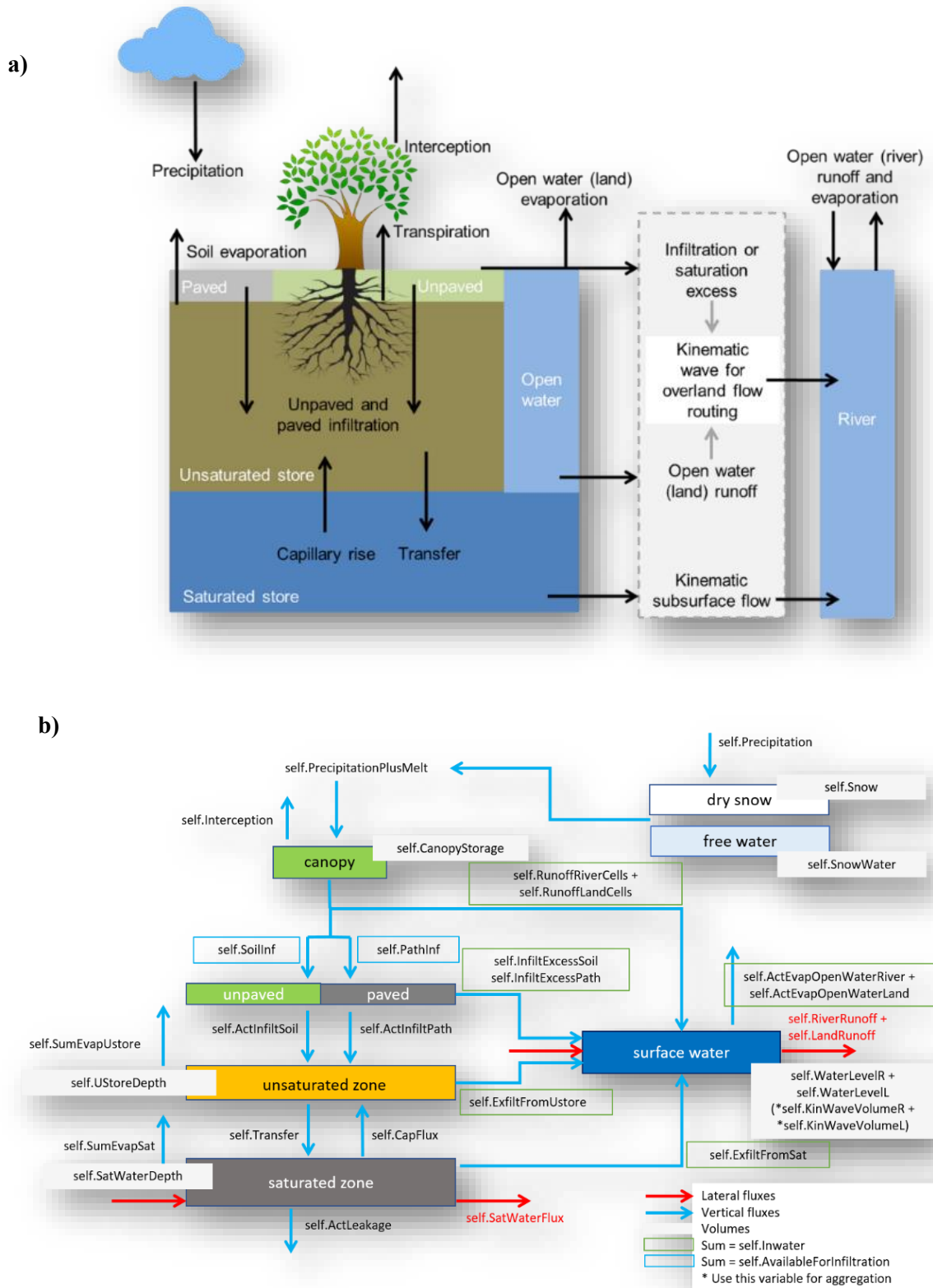


Figure 7: a) Schematic illustration of the wflow_sbm processes and states (Schellekens et al., 2020),
b) model states and fluxes in terms of internal variable names.

5 RESEARCH METHODOLOGY

This study's approach will follow the steps outlined in Figure 8.

The first phase involves input collection, selection, and analysis. Precipitation, discharge, and other relevant data will be selected and evaluated from the available datasets in section 4.1. As the input errors propagate throughout the study, an exploratory precipitation data analysis is carried out for uncertainty estimation of the available gridded datasets and to determine whether it is appropriate to force the wflow sbm model.

In a second step, deterministic discharge predictions are simulated in Delft-FEWS for each model over a 10-year historical simulation period. The results are evaluated at each station where observations are available. Finally, the overall model performance is assessed using appropriate measures, and whether they are suitable for continuing the DA analysis is determined.

Before employing data assimilation, a reference model must be developed. An open-loop simulation will be used as a benchmark, and alternative possibilities (e.g., autoregressive correction) will be investigated. The reference simulation and uncertainty framework selection criteria are based on an ad hoc process that combines literature recommendations and trial-and-error methods. A framework that automates the process is created in advance to be used when running the simulations with DA and processing the results in the consequent stage. For this purpose, and for all the analysis carried out in this study, Python coding is used.

Following the configuration of the DA scheme, repeated runs of the wflow sbm model with OpenDA are completed, and the updated simulation output from the EnKF algorithm is compared to the selected reference simulation discharges, as well as the observations. To analyze the update of the model states with DA, a similar analysis to Rakovec (2015) is conducted. Experiments include the most essential model states in the study, with a special focus on the Upper Zone (subsurface water flow ssf; saturated water depth swd) and Soil Moisture (ustore) states. The performance of DA in spatially capturing model states when used with distributed hydrological models will be examined by observing the changes in model states, thereby accomplishing the second objective of this study. Several probabilistic verification techniques for verifying ensemble predictions of hydro-meteorological and hydrological states will be utilized to assess the outcomes of the DA application in terms of discharge prediction. Metrics such as the KGE, NSE, RMSE, pBIAS and r^2 as described in section 2.3, are utilized in different parts of the results for various purposes (to assess the precipitation data error, compare the deterministic simulation with the observations, and assess the DA effect in the discharge prediction improvement).

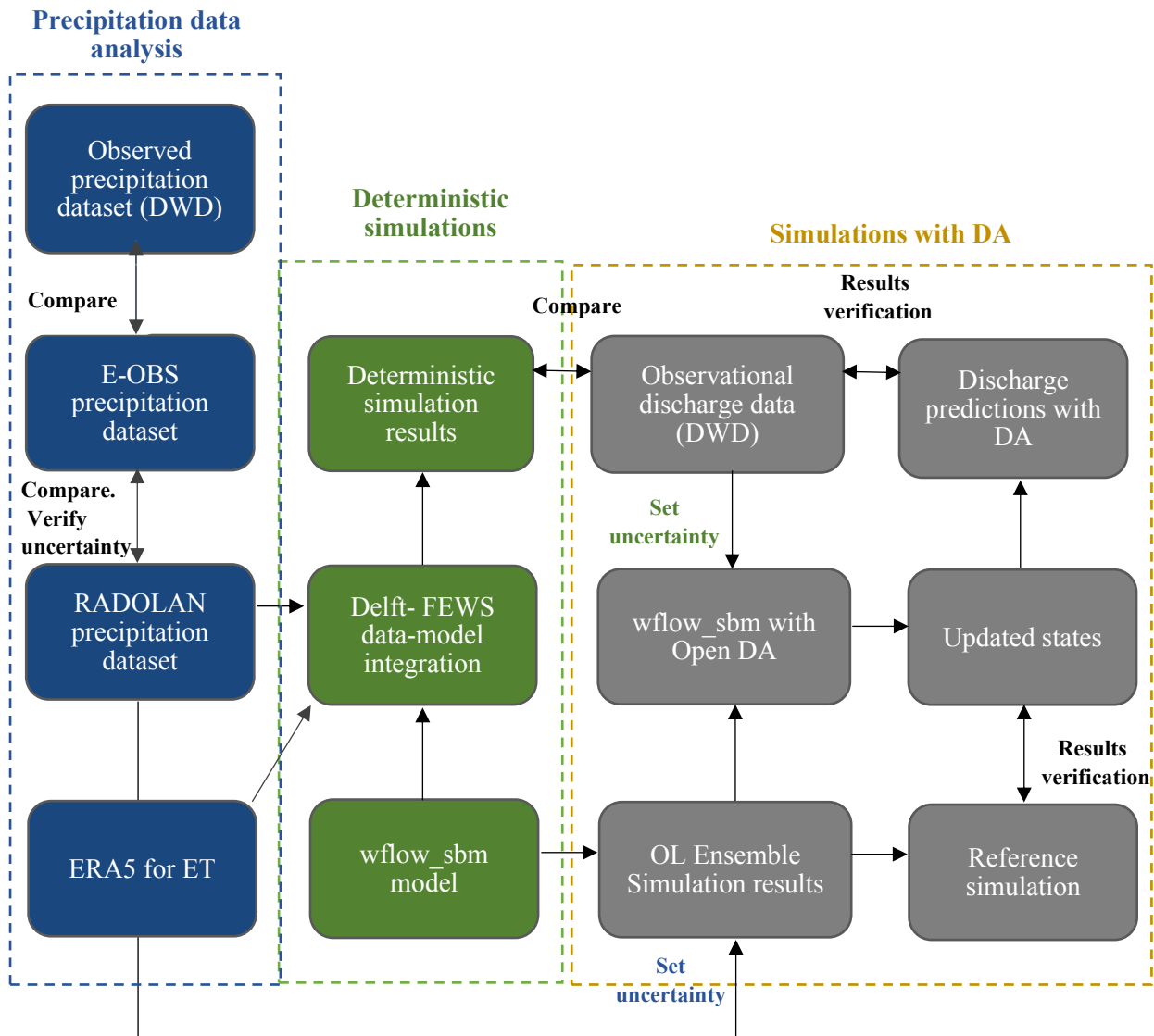


Figure 8: Schematization of the research methodology workflow

6 RESULTS AND DISCUSSION

In this section, results and discussion of the findings of this study are reported. Following the methodological framework in Table 8, the results are presented in three parts. The first section refers to the precipitation data analysis findings, followed by the deterministic simulation results; an analysis of the model performance regarding discharge forecast when applying data assimilation is then discussed, and finally, the DA effect on correctly updating the hydrological model states is assessed.

6.1 Precipitation data analysis

As mentioned in Section 2.2, uncertainties associated with the model forcing, among others, reflect in the DA performance. A precipitation data analysis was conducted to assess the performance and the uncertainties associated with the available gridded rainfall datasets, which would be employed to force the hydrological models. Because of its high temporal and spatial resolution, the hourly dataset of RADKLIM is of greater interest to use among the available datasets. The Radolan gridded dataset is set to be compared with the available observational datasets, both gridded (EOBS) and in rain-gauge stations (DWD), for each catchment.

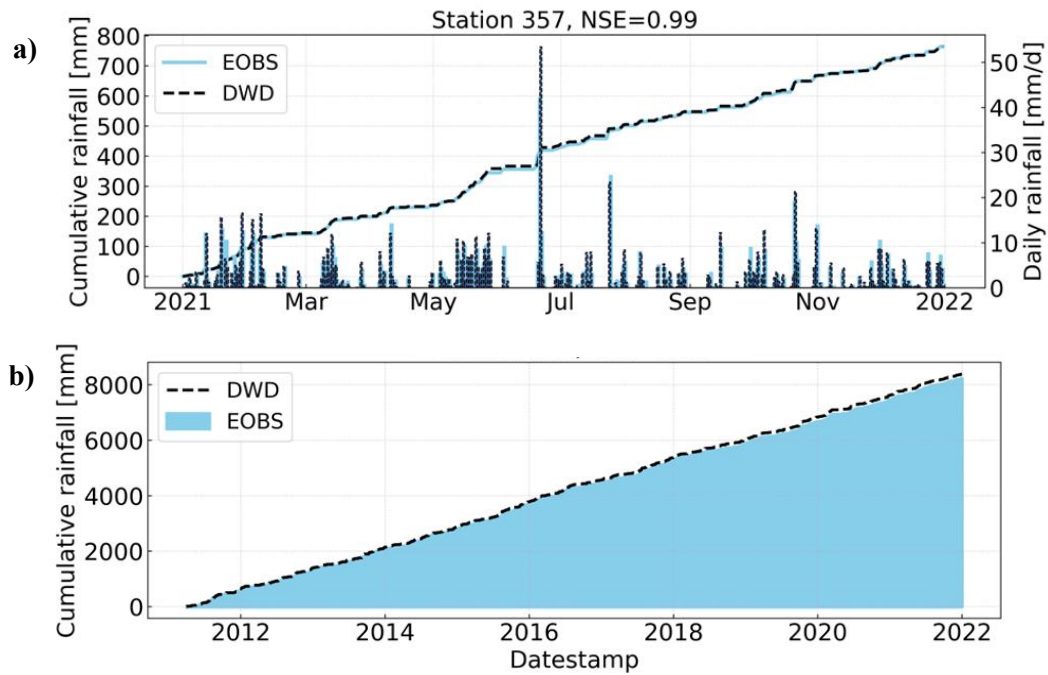


Figure 9: Comparison of the cumulative and daily rainfall time series between the DWD rain gauge observations and the EOBS observational gridded rainfall dataset for Station ID 357 in the Vecht river catchment for a) year 2021; b) the complete time series from 2011 to 2021

Initially, the accumulated rainfall values in the gauge stations were compared to the respective approximated values of the observational gridded dataset EOBS for different years and gauge locations. It was noticed that the point observations obtained from the German Weather Service (DWD) for nearby stations in both catchments present a very high similarity (metrics like Nash Sutcliffe Efficiency $NSE \sim 0.99$) with the respective values from the EOBS dataset, as seen in figures Figure 9 and Table 10 for a selected rain gauge location in the Vecht and Geul river catchment respectively. Furthermore, equivalent results are obtained for all other stations presented in Appendix B.1.

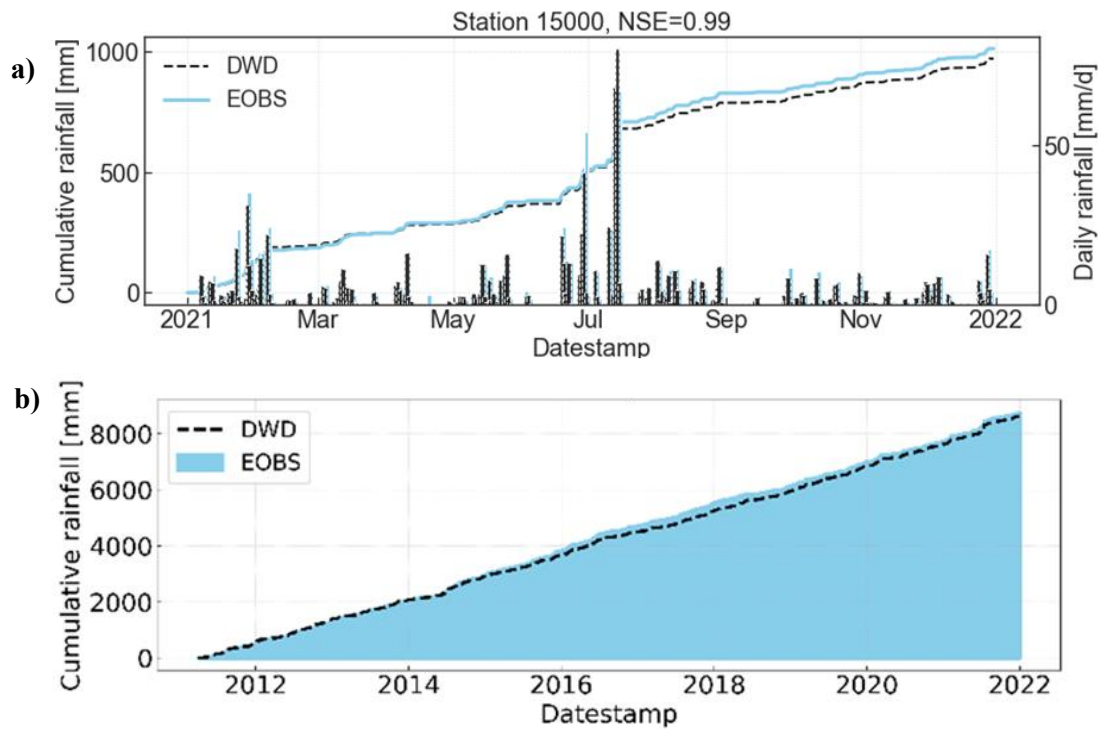


Figure 10: Comparison of the cumulative and daily rainfall time series between the DWD rain gauge observations and the EOBS observational gridded rainfall dataset for Station ID 15000 in the Geul catchment for a) year 2021; b) the complete time series from 2011 to 2021

The high similarity of the datasets is attributed to the methodology used to derive the EOBS dataset, which uses the observed precipitation in the rain stations to interpolate the data and obtain the gridded data. Assuming a perfect agreement between the E-OBS gridded dataset and the corresponding point observations in the catchments, it is more sensible for the Radolan dataset to be further evaluated in comparison with the gridded EOBS dataset, to construct a fuller picture of it, not only in compared in points but also spatially.

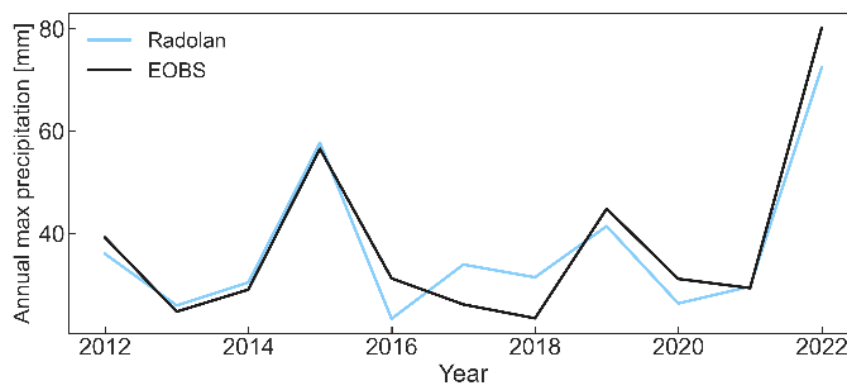


Figure 11: EOBS and Radolan precipitation resampled into annual maximum precipitation values for the station ID 15000 in the Geul catchment

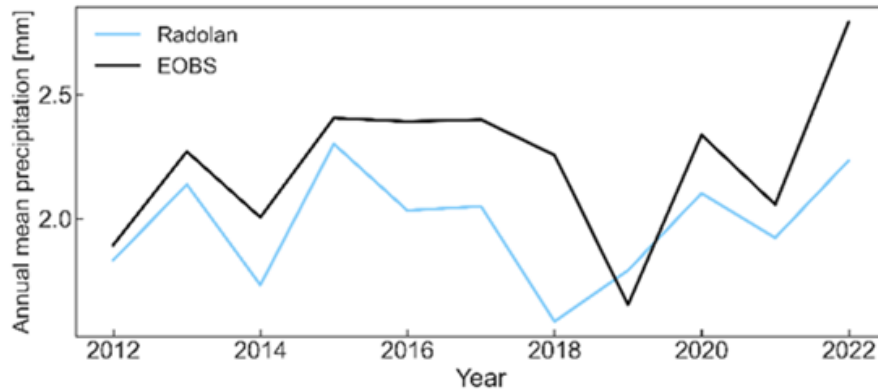


Figure 12: EOBS and Radolan precipitation resampled into annual mean precipitation values for the station ID 15000 in the Geul catchment

The annual maximum and mean precipitation trends of both datasets for station 15000 in the Geul river catchment are observed in Figure 11 and Figure 12, respectively. For completeness, the results for all stations are presented in Appendix B.

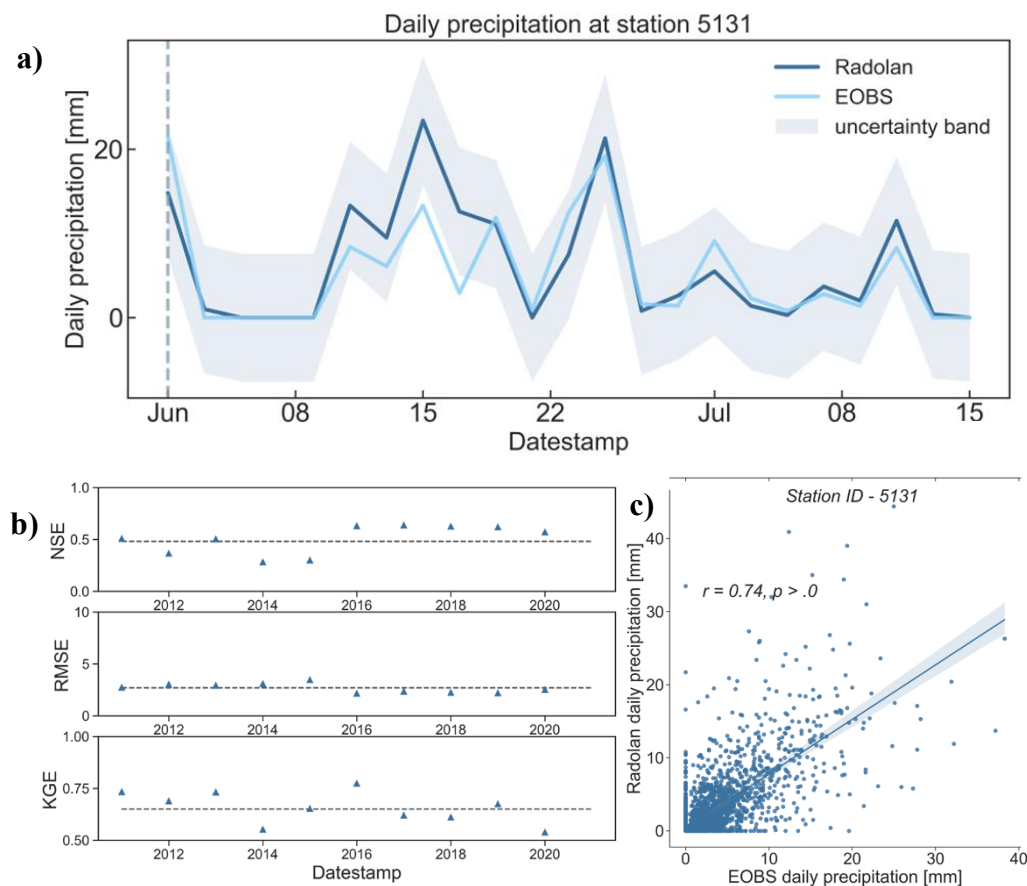


Figure 13: a) Daily precipitation at station point location with ID 5131 in the Vecht river catchment derived from the Radolan gridded dataset and the EOBS gridded observational datasets. Uncertainty of the Radolan is presented as a 95% uncertainty ban; b) Annual values of selected performance metrics (NSE, RMSE, KGE); c) Correlation scatter plots of the Radolan and EOBS daily time series for the 10-year period.

Daily precipitation time series are presented below for a selected point location in each catchment for the period of June-July 2021 at Geul; and June, July 2016 for the Vecht. Both gridded datasets are interpolated in the point locations with coordinates corresponding to nearby rain gauge stations. In addition, a 95% uncertainty band is presented for the Radolan dataset. While in Figure 13a and Figure 14a, the daily precipitation is presented only for a selected period, including the July 2021 event in the Geul catchment and the summer months of 2016 in the Vecht, the correlation between the datasets and the other metrics are computed and plotted in Figure 13b,c and Figure 14b,c for the complete time-series (2011-2021).

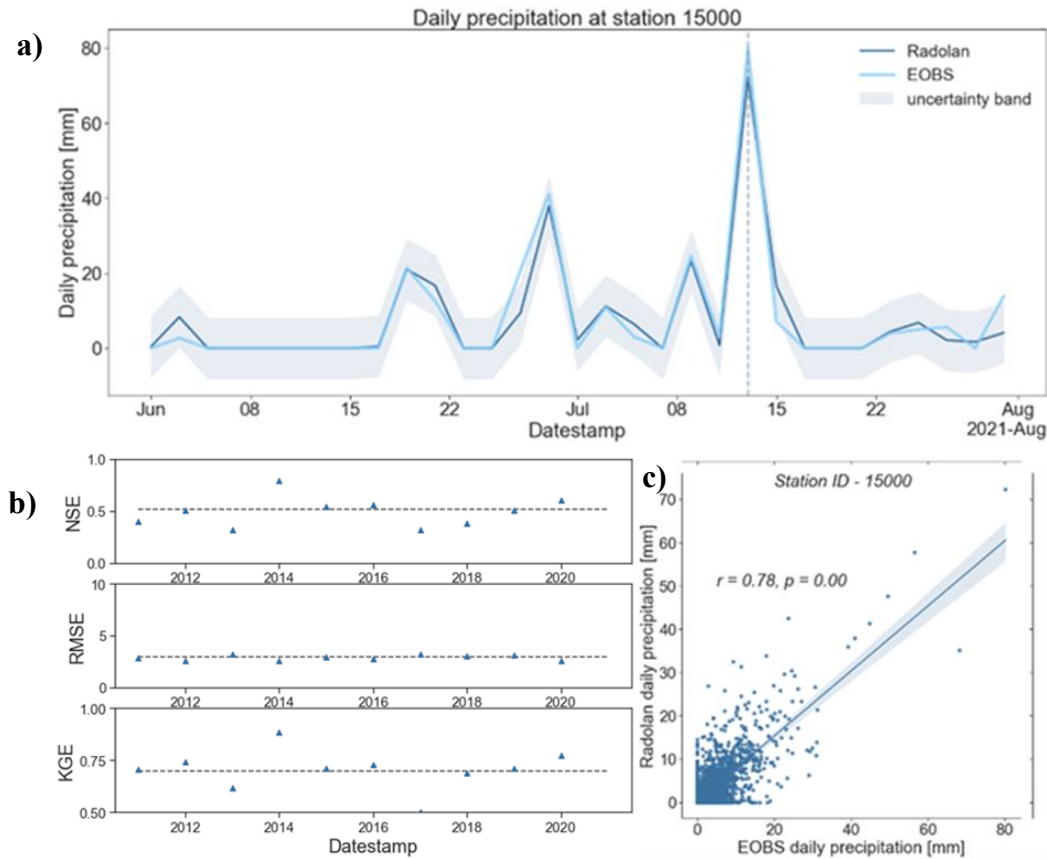


Figure 14: a) Daily precipitation at station point location with ID 15000 in the Geul catchment derived from the Radolan gridded dataset and the EOBS gridded observational datasets. Uncertainty of the Radolan is presented as a 95% uncertainty band; b) Annual values of selected performance metrics (NSE, RMSE, KGE); c) Correlation scatterplots of the Radolan and EOBS daily time series for the 10-year period.

In Figure 15, a spatial comparison of the datasets is presented, where the maximum and mean annual values for 2021 are computed for the same grid covering both basins. The EOBS grid is given on the left, as it can be distinguished from the grid resolution (11 km), which is coarser than the Radolan (1km) on the right. The mean annual precipitation values in Figure 15a and the maximum annual values in Figure 15b present a considerable similarity in the spatial spread of the precipitation of the two datasets. For a better understanding of the differences in the maximum values, Radolan is down-sampled in the 11 km grid as seen in Figure 15c, where the Radolan seems to be underestimating the observational dataset estimates.

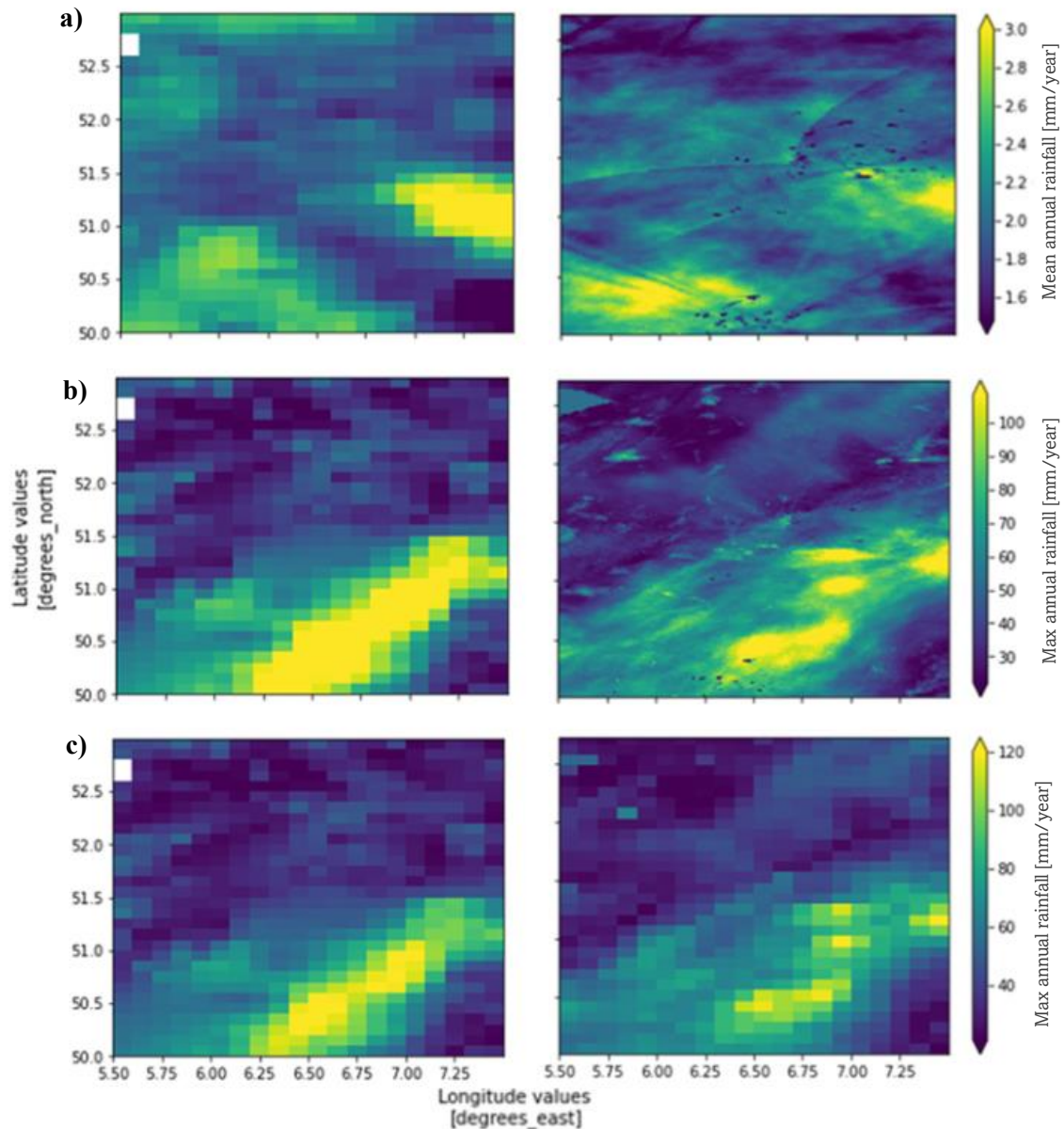


Figure 15: E-OBS (on the left) vs. RADOLAN (on the right) spatial comparison for the grid covering both catchments a) annual mean rainfall b) annual max rainfall c) annual max rainfall resampled (11km Radolan grid to match the EOBS grid)

After the Radolan and the EOBS observational precipitation dataset were analyzed for the entire available period, both in point locations and spatially, we can now qualitatively derive judgment based on their agreement. As noticed throughout the analysis in this section, the Radolan gridded precipitation fails to capture the extreme precipitation events, presenting either overestimation or underestimation of peak flows. At the same time, it shows more consistency in representing smaller events. Despite the radar data being calibrated with gauge observations, there might be a need for bias adjustment correction of the radar dataset.

Table 7: Performance metrics of the Radolan, EOBS rainfall dataset comparison for the complete daily time-series (2011-2021)

	KGE [-]	NSE [-]	RMSE [mm/day]
Catchment averaged (Geul)	0.68	0.52	2.82
Station 15000	0.74	0.54	2.94
Station 4667	0.53	0.47	3.07
Station 15927	0.69	0.52	2.75
Catchment averaged (Vecht)	0.73	0.51	2.70
Station 5131	0.75	0.48	2.72
Station 357	0.72	0.58	2.75
Station 1230	0.71	0.49	2.81
Station 1223	0.73	0.56	2.64
Station 3640	0.68	0.45	2.96

Overall, as indicated by the metrics presented in the plots in this section (See also Appendix B.2) and summarized in Table 7 for all stations, the Radolan presents a reliable performance with KGE values varying from 0.7 to 0.58 and NSE values ranging from 0.45 to 0.58 and RMSE values ranging from 2.64 to 3.07 mm/day. Given its high spatial and temporal resolutions of 0.1° (~ 10 km) and 1 hour, it is thought to be adequate for use in a precipitation-fed, fast-responding catchment like the Geul. Additionally, it indicates good performance in the case of a larger catchment, such as the Vecht; hence rainfall input for the wflow_sbm model simulation in this work is based on Radolan forcing. The Radolan rainfall estimate error metrics will be used as a guide when selecting the input uncertainty in the DA experiments conducted later in this work.

6.2 Deterministic simulations

To demonstrate that the implemented models can simulate reasonable discharge estimates using the Radolan rainfall product as forcing, this section presents the observed discharges and simulated ones for selected gauge locations in both catchments. In the Delft-FEWS platform, where the hydrological models of both catchments are integrated with the respective input data, the first open loop simulations (without data assimilation) were run.

The wflow_sbm models were initially forced with Radolan precipitation data and ERA5 data for potential evapotranspiration. The historical simulations are run for ten years, starting January 2011 to December 2021, with a 10-day timestep, assuming perfect models and perfect forcing data. From the first results of the open loop simulations (without DA and without the Radolan data gap filling with ERA5), it is noticed that for the case of Emlichheim, the station presented in Figure 16 for illustration, as is also noticed in all other stations for which observational data are available, the simulated flows differ significantly from the observed ones. Due to gaps in the hourly rainfall dataset, this discontinuity is reflected in the simulated discharges, as for every gap in the input forcing, the model restarts with a cold state, often overestimating the next time step, resulting in the simulated discharge not being reliable, and the model failing to capture both flow peaks and the baseflow over the entire simulation period. This was resolved by setting up an alternative configuration so that ERA5 precipitation data are used to fill in any gaps in the Radolan dataset that are longer than two hours. Following the application of this

adjustment, the simulated discharges exhibit a distinct consistency for the whole simulation period. The rest of the analysis uses the “corrected” Radolan forcing.

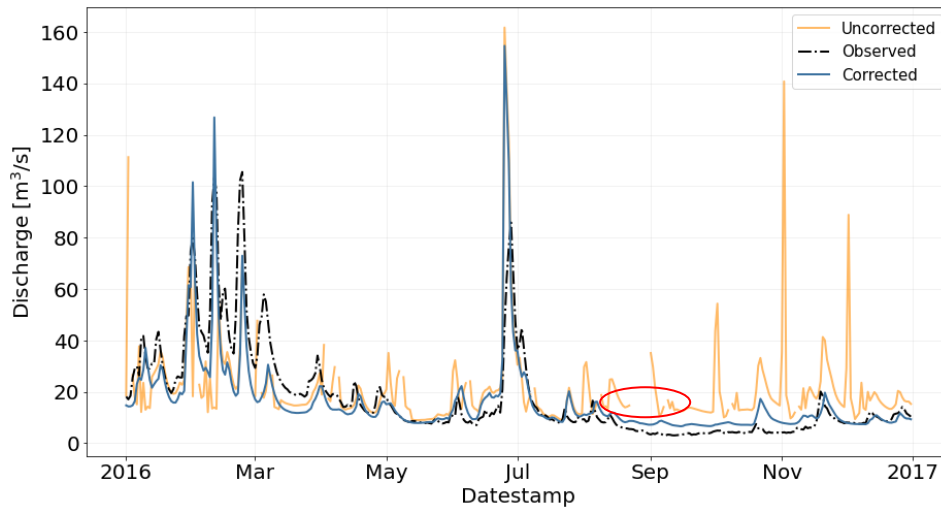


Figure 16: Hourly Observations of the Emlichheim discharge and wflow_sbm model difference before and after correction with ERA5. In red, gaps in the uncorrected model output are highlighted.

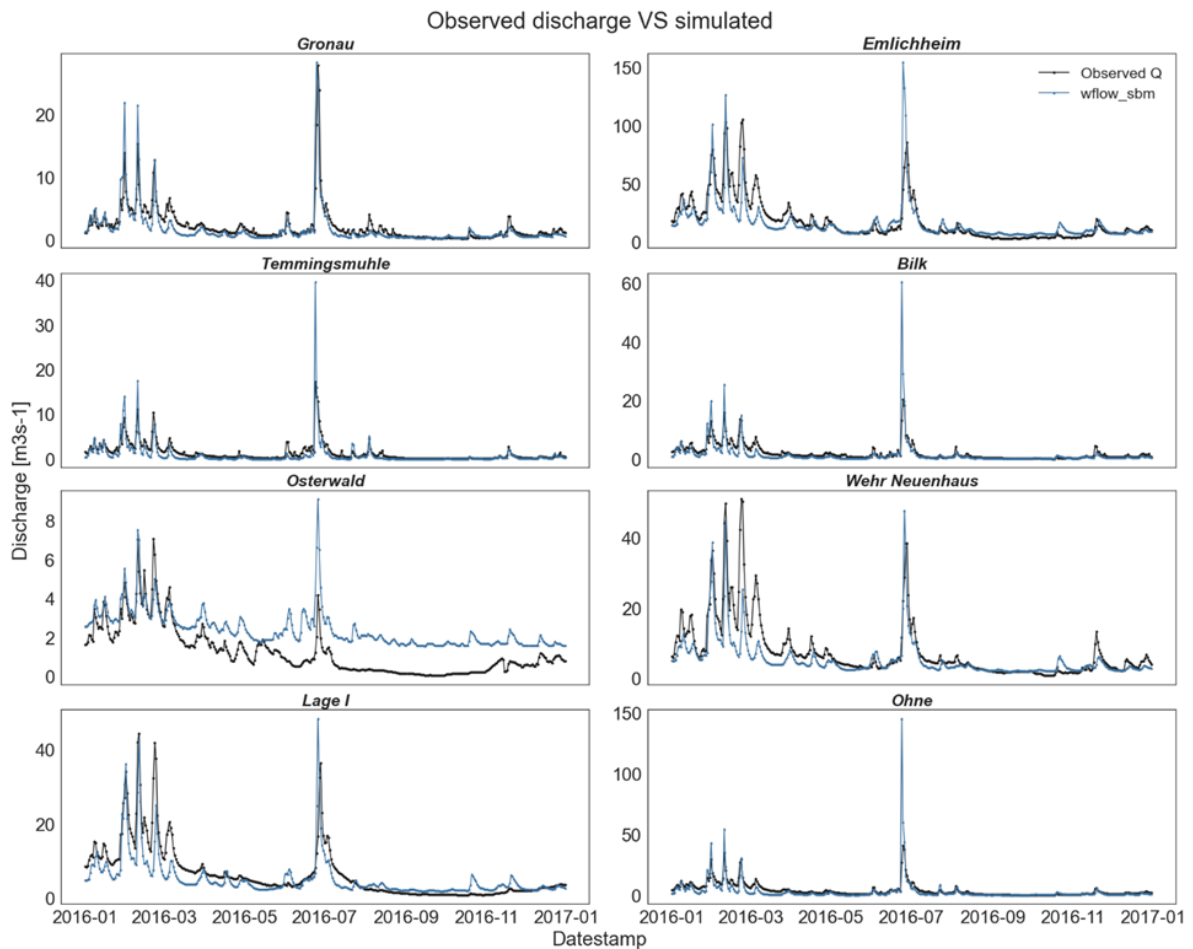
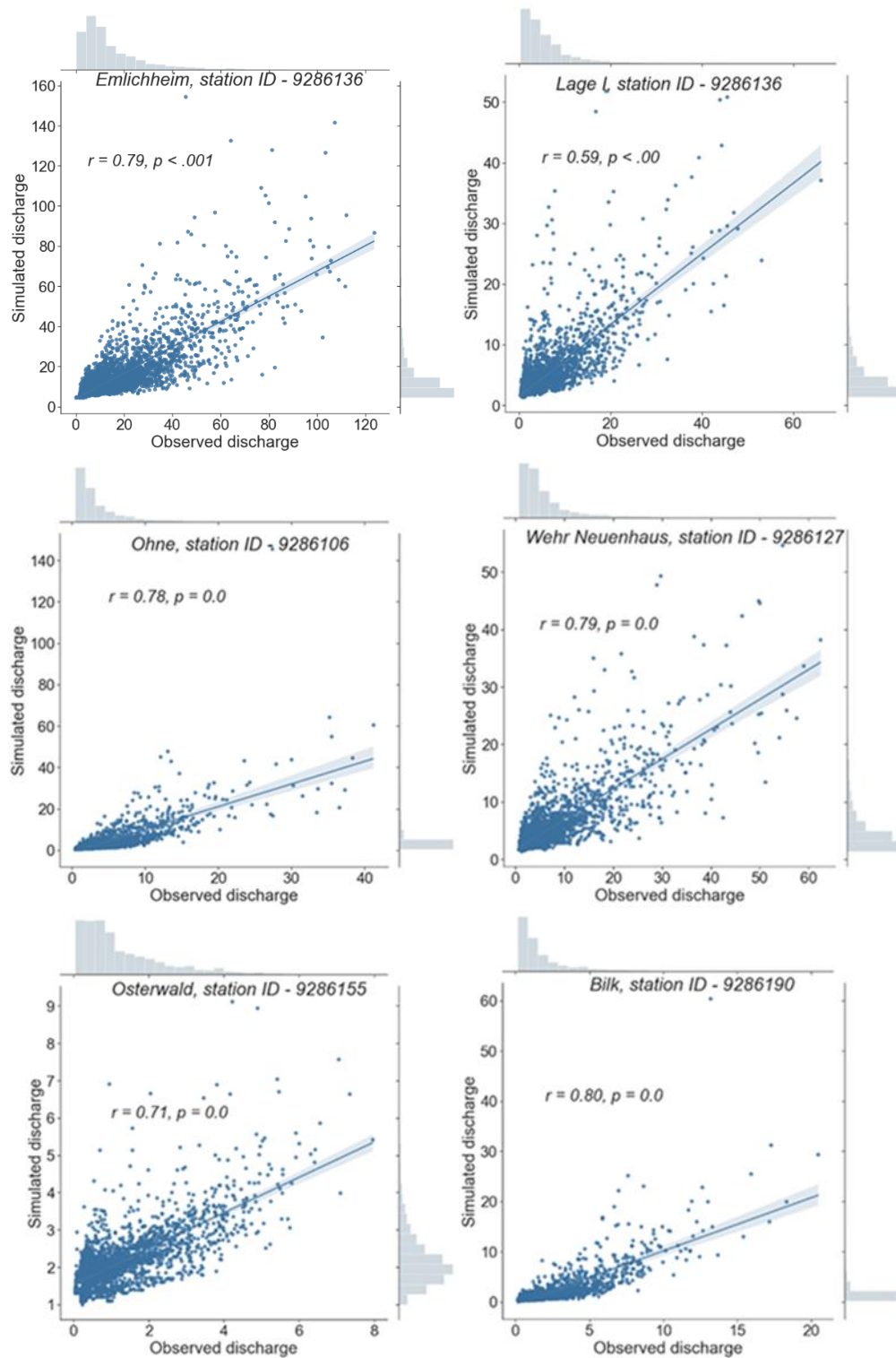


Figure 17: Simulated VS Observed Mean daily runoff, for selected stations in the Vecht catchment, for 2016

Simulated flows with the wflow_sbm model, along with the corresponding observed flows for a few gauge locations in the Vecht, are presented in Figure 17 for a particular year (2016). The historical simulations were run for a 10-year period (2011-2021), and depending on the availability of measurements at each gauge, the correlation statistics are derived as shown in Figure 18 for each station.



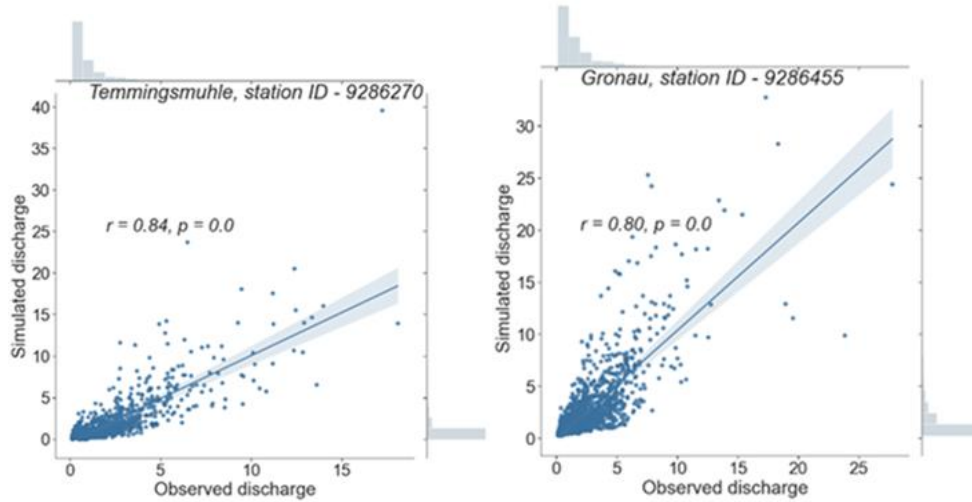


Figure 18: Correlation of simulated and observed discharge scatter plots at selected station location in the Vecht catchment. Note that the correlation is computed for a different period in different stations, depending on the period of record available at each station

The Nash-Sutcliffe model efficiency (NSE) and other metrics are computed at each stream gauge location before the discharges are assimilated to validate the outcomes of the DA experiment. In the case of Emlichheim, the catchment outlet, the NSE model efficiency coefficient between the measured and simulated discharges is approximately 0.8, indicating that the hydrological model performs well. Moreover, it is essential to recognize that NSE is sensitive to extreme values (peak flows) that were not previously filtered. Additionally, for a 10-year simulation interval, the root mean-square error (RMSE) varies from 1.11 m³/s at Temmingsmuhle to 9.5 m³/s at Emlichheim. For completeness, additional performance statistics computed for the long simulation period, which spans from 2011 to 2021 and varies in different stations, are summarized in Table 8.

Table 8: Performance statistics to evaluate the model's predictive ability in selected stream gauge stations in the Vecht river catchment

	r	NSE	RMSE	pBIAS
	[-]	[-]	[m³/s]	[%]
Emlichheim	0.79	0.63	9.5	-8.0
Lage I	0.59	0.54	4.36	-14.47
Ohne	0.78	0.22	3.75	-12.11
Wehr Neuenhaus	0.79	0.57	5.01	-20.7
Osterwald	0.71	-0.02	1.21	65.36
Bilk	0.80	0.36	1.72	-16.85
Temmingsmuhle	0.84	0.54	1.11	-21.74
Gronau	0.80	0.41	1.56	3.65

Similarly to the analysis carried out for the Vecht river catchment gauge locations, results are presented below for the Geul river catchment, where a different period is selected to be presented (part of 2021) because of the interest in the July 2021 flood event.

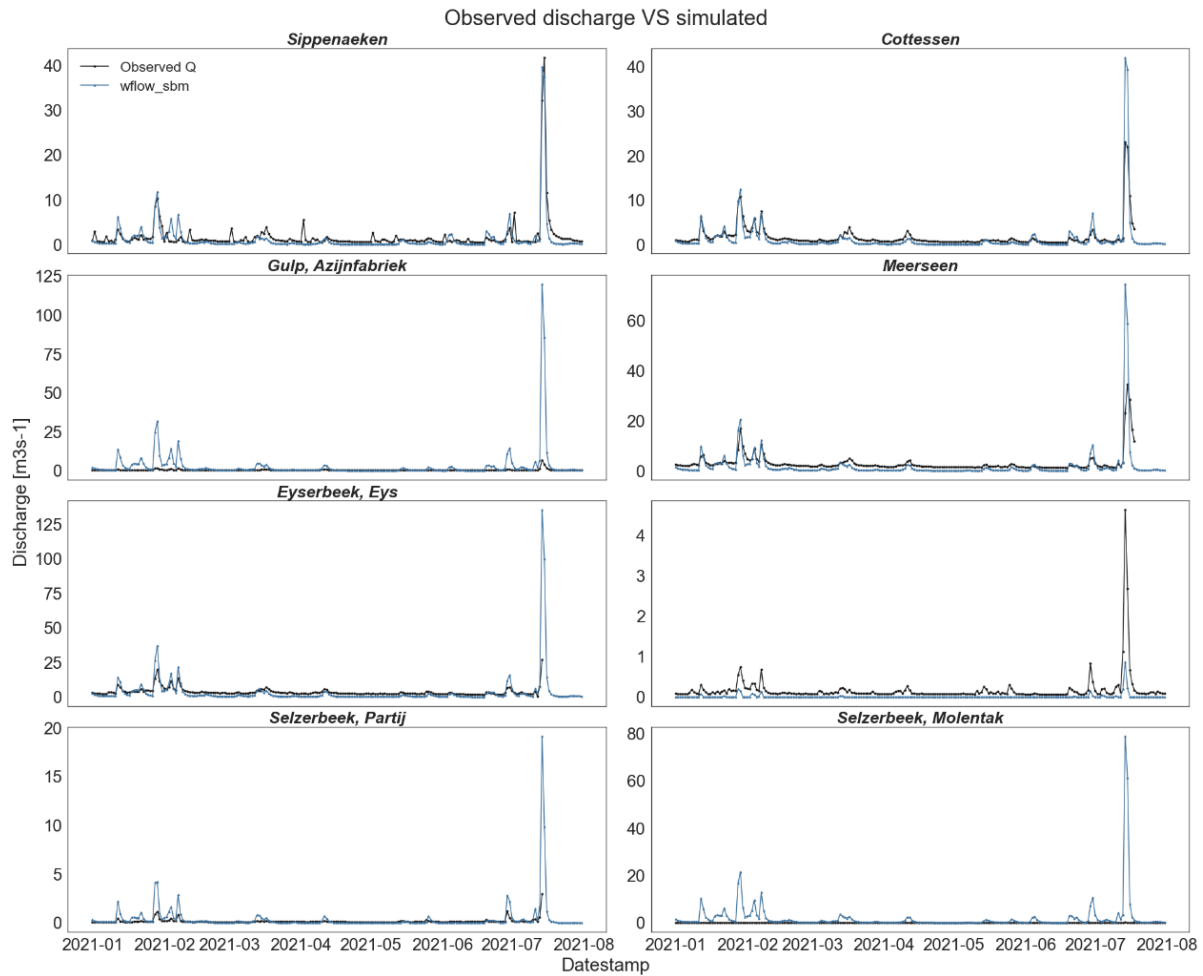
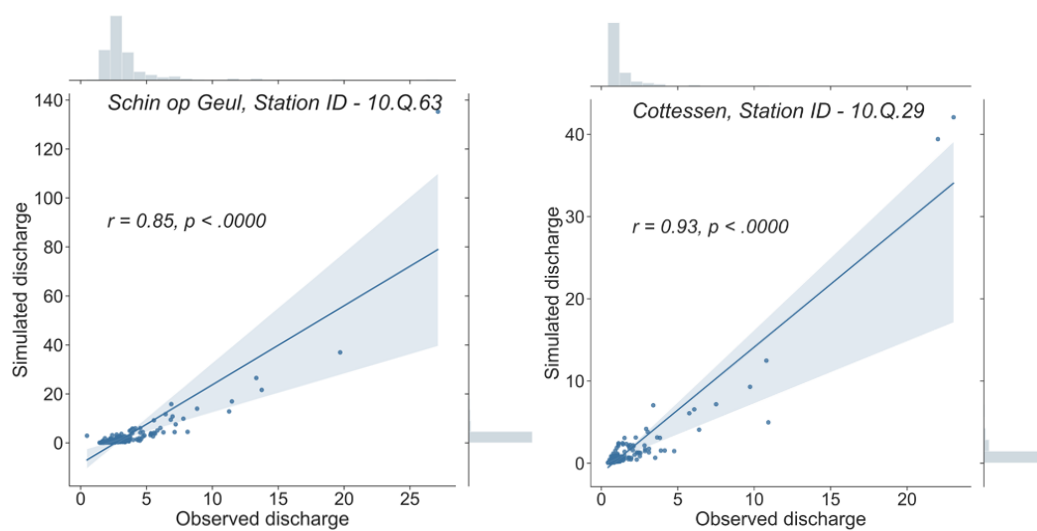


Figure 19: Simulated VS Observed Mean daily runoff, for selected stations in the Geul catchment, for the first half of the year 2021



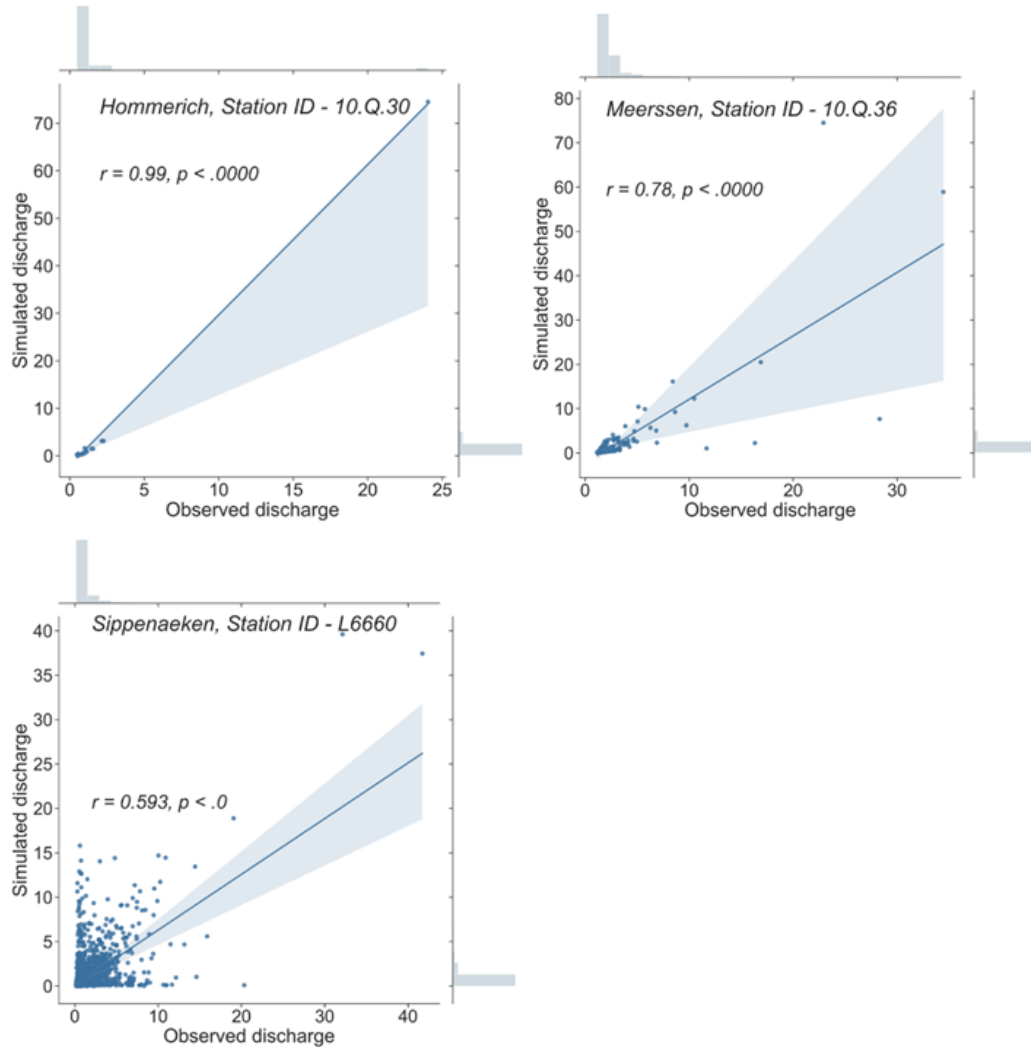


Figure 20: Correlation of simulated and observed discharges scatter plots at selected station location in the Geul river catchment. Correlation is computed for a different period in different stations, depending on the period of record available at each station

The results of the deterministic simulations reported in this section will aid in selecting a benchmark simulation to contrast with the simulated discharges in the following step when observed discharges are assimilated. The grid-based wflow_sbm models produced satisfactory results at the Geul and Vecht catchment scale despite not being recalibrated.

However, the direct comparison of the simulated and measured discharges in various stations presented in Figure 20, together with the metrics summarized in Table 9, indicate that the available model for the Geul River does not correctly represent high peak flows, as in the case of the July 2021 flood event (for other years as well), and, as in this study, uncertainties associated with the model parameters are neglected, in the case of a quick responding catchment with a time skill of just a few hours, proceeding with a data assimilation scheme where we only account the uncertainties associated with the forcing data and observations, results in very unrealistic results.

Moreover, as using DA with wflow_sbm is experimental and incomplete, technical problems were encountered while linking the Geul wflow_sbm to OpenDA. Hence, the data assimilation setup and results given from now onwards are limited to the Vecht River catchment; however, some speculations will be made for the Geul River based on the conclusions drawn from the Vecht River.

Table 9: Performance statistics to evaluate the model's predictive ability in selected stream gauge stations in the Geul river catchment.

	R	NSE	RMSE	pBIAS
	[-]	[-]	[m³/s]	[%]
Sippenaeken	0.59	0.076	1.54	-34
Cottessen	0.93	0.380	1.93	-22
Hommerich	0.99	-3.83	9.07	91.34
Meerssen	0.78	-0.54	4.47	-31
Schin op Geul	0.85	-7.67	8.20	-17.20

6.2.1 Model States

Before updating with DA, we look at the model states of interest. As stated in the study's research objectives, this analysis is limited to the Vecht river catchment. Of particular interest to analyze are the slow-changing components of the models, Subsurface Flow (ssf), Saturated Water Depth (swd), and Soil Moisture (ustorelayerdepth) states, the behavior of which, together with other hydrological states, is depicted in Figure 21 over a year. Since we want to examine the effect of updating with the EnKF in the hydrological states, a preliminary analysis was performed to investigate the states' behavior and whether there is any association between them. Figure 21 hints at a considerable correlation between the river discharge (q_{river}) and the saturated water depth and subsurface flow, becoming more visible during spring, where the flow peaks usually happen. However, there is a time lag between the peaks of those components.

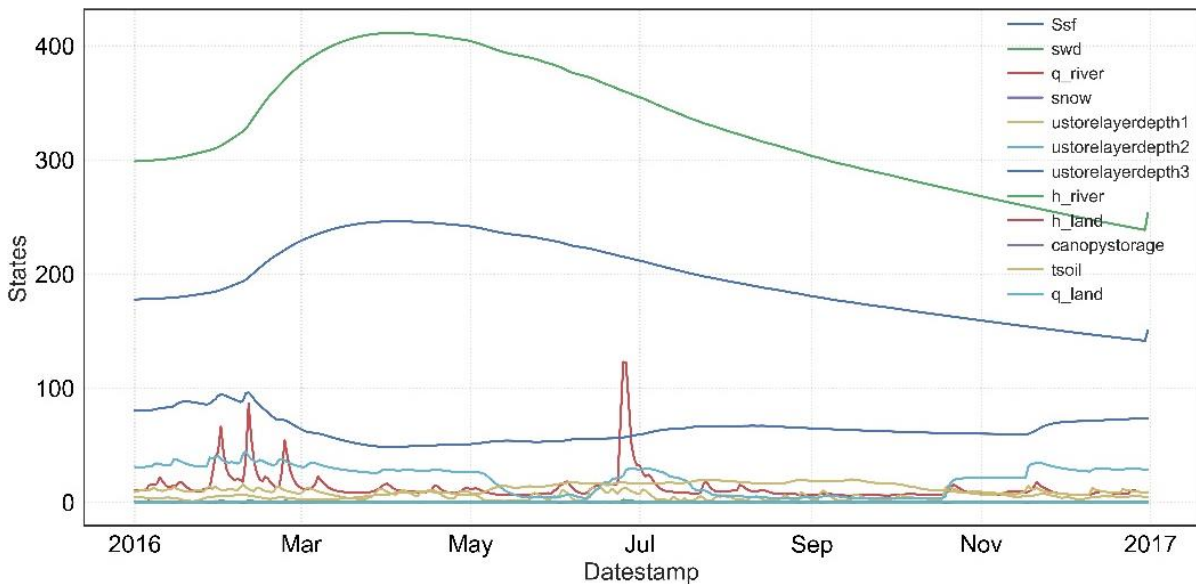


Figure 21: Hydrological model states variation at Emlichheim for a selected year (2016)

As it is challenging to derive judgment from the annual variation of the model states, a correlation heatmap of model states is presented in Figure 22. As expected, the correlation is positive and more substantial between the states of interest (ssf, swd, ustorelayerdepth) and the river discharge (q_{river}) during the whole simulation period. Additionally, there is a considerable correlation between

the surface water flow and depth, with the river flow and river water depth. Therefore, further, in the analysis, only the states that indicated higher correlation values (a correlation threshold of 0.1 was considered) are included in the DA experiment, namely the subsurface water flow (ssf), saturated water depth (swd), overland flow (q_land), overland water depth (h_land), soil moisture in three layers (ustorelayer1, 2 and 3), river flow (q_river) and river water depth (h_river). States with negative correlation (tsoil) or correlation values lower than 0.1 (snow) are excluded from the rest of the analysis. Additionally, the canopy storage state is excluded.

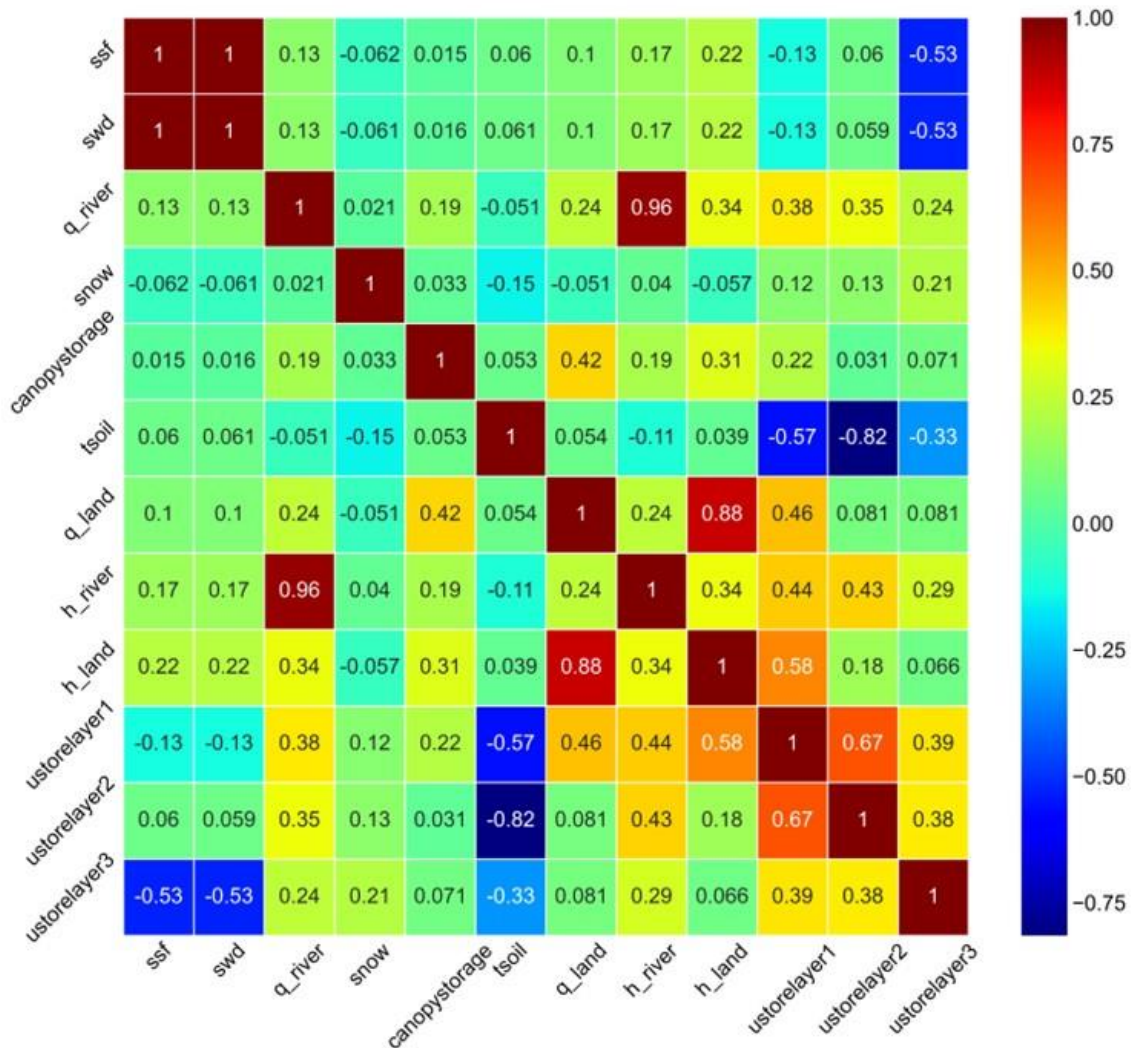


Figure 22: Model state correlation at Emlichheim

6.3 Experiments with DA

6.3.1 Experimental setup

This section presents the filtering methods' configuration setup for the assimilation of observations in the wflow_sbm model to improve the discharge predictions, particularly at the catchment outlet. The hydrological observations used for updating the model states are discharge measurements due to their availability in high temporal resolutions and their good representation of the catchment wetness conditions, which is significant for operational forecasting. Model states are updated with externally measured variables to obtain proper initial conditions in the following timestep. The

observational discharges are assimilated daily for two selected flood events in the Vecht river basin depicted in Figure 23. Additionally, an overview of the selected assimilation periods is given in Table 10. Only discharge data at the catchment outlet (Emlichheim) are assimilated. Another location (Gronau) is presented to validate the results in the first experiment. In contrast, discharge data are assimilated in 5 different locations in the second experiment. Finally, results are presented again for the stations of Emlichheim and Gronau (the other point results are shown in the Appendices).

In this work, only input forcing uncertainty is represented from ensembles, and observation uncertainties are considered, disregarding any uncertainty related to the model parameters or initial conditions. A precipitation ensemble is comprised of a limited number of spatial realizations in time that represent uncertainties related to precipitation's spatial and temporal variation (Rakovec, Hazenberg, et al., 2012). The ensemble of simulations is derived from perturbing the model forcing with stochastic spatially correlated model error. In an open loop simulation, the model is forced with uncertain rainfall input with a simulation memory of forty-eight ensemble members derived using stochastic perturbation for precipitation and deterministic PET to produce an ensemble of simulated discharges out of which a single realization was chosen to be further used as a reference for comparison with the DA ensemble discharges. The selection procedure of this reference simulation is detailed in the next section. The ensemble size was selected to be forty-eight members for the OL simulations and 16 for the filtered experiments for computational reasons. However, tested experiments with 24 and 48 realizations were not necessarily better than those with sixteen ensemble members. The error of the discharge observations is considered to have a normal distribution with a standard deviation of $(kQ_{\text{obs}})^2$, with k ranging between 0.01 and 0.1 (as seen in Weerts & el Serafy, 2006).

A sensitivity analysis was carried out to investigate the influence of the error selection for both the forcing and the observations (Appendix C.1). Based on the findings, the best selection of errors was made to continue the rest of the analysis. The EnKF is used to update the hourly spatially distributed wflow_sbm model's grid-based distributed states. The hydrological cycle is simulated for the Vecht (1800 km²) for two small-scale experiments. First, the EnKF algorithm was applied to a one-month interval from 15th June to 15th of July 2016, which included both a dry and a wet period, with one flood event, and to a second experiment, for the period 15 January - 15 March 2016, with three consecutive flood events.

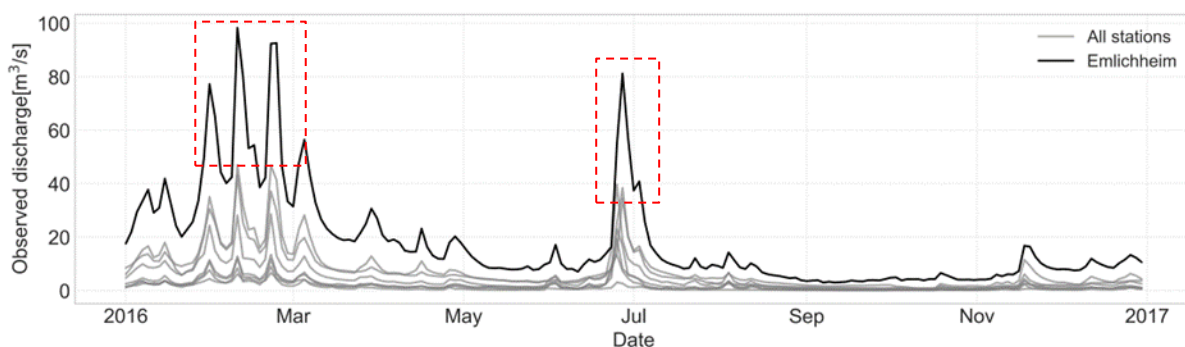


Figure 23: Observed discharge in the outlet (Emlichheim) and other interior flow gauges in the Vecht river basin during 2016. Two characteristic winter and summer flood events are selected for the analysis

Extensive runs are performed in the first experiment to examine the impact of different uncertainty values assigned to the forcing input and observations, select a reference simulation, as well as investigate the effect of DA in the model states. In contrast, in the second experiment, where a longer assimilation window is considered, the perturbation errors and reference simulation are chosen in the same way as in the first experiment. In addition, the effect of DA in the predicted discharges is evaluated.

Table 10: An overview of the experiment periods selected for the DA analysis

	Period	Maximum observed discharge [m³/s]
1	15 Jun 2016 – 15 Jul 2016	81
2	15 Jan 2016 – 15 Mar 2016	98

6.3.2 Selection of a reference simulation.

Open loop ensemble model realizations were generated before DA application to ensure good model results, no numerical instabilities, and select a reference simulation. These simulations were run without introducing any error to the system, and the selection of the benchmark simulation was made considering the ensembles should be as near to the deterministic simulation as possible and have a reasonable ensemble spread. In this work, the selection of the benchmark simulation is made by a visual inspection, while several metrics are available in the literature. Additionally, the standard deviation of ensembles can be evaluated against the measurements; if the spread of the standard deviation covers the observations, it is deemed appropriate (Dharmadasa, 2014). As the deterministic runs produced fair results, as seen in the previous chapter, an open loop probabilistic simulation is considered more suitable as a reference simulation to evaluate the DA effect.

The most challenging part of a DA experiment lies in selecting the optimal input uncertainties. While there are reference values for the precipitation and observation noise to be used in DA experiments, to be able to choose the best reference simulation based on the proximity of the probabilistic simulation with the deterministic one and the most reasonable ensemble spread, a sensitivity analysis was performed in advance to investigate the response of the filtered model when different perturbation noises for the precipitation and observations are given. The results of multiple runs conducted with the first experimental setup to investigate how the filter reacts when increasing or decreasing the noises of both forcing and the observations, summarized in the appendix, were utilized to derive the reference simulations presented in this section. The simulations presented in the appendix are conducted with eight ensemble members for computational efficiency. Their only purpose is to guide us in selecting the correct uncertainty values for the rest of the analysis.

From the extensive work in the appendix, the three best Open Loop scenarios were determined as follows:

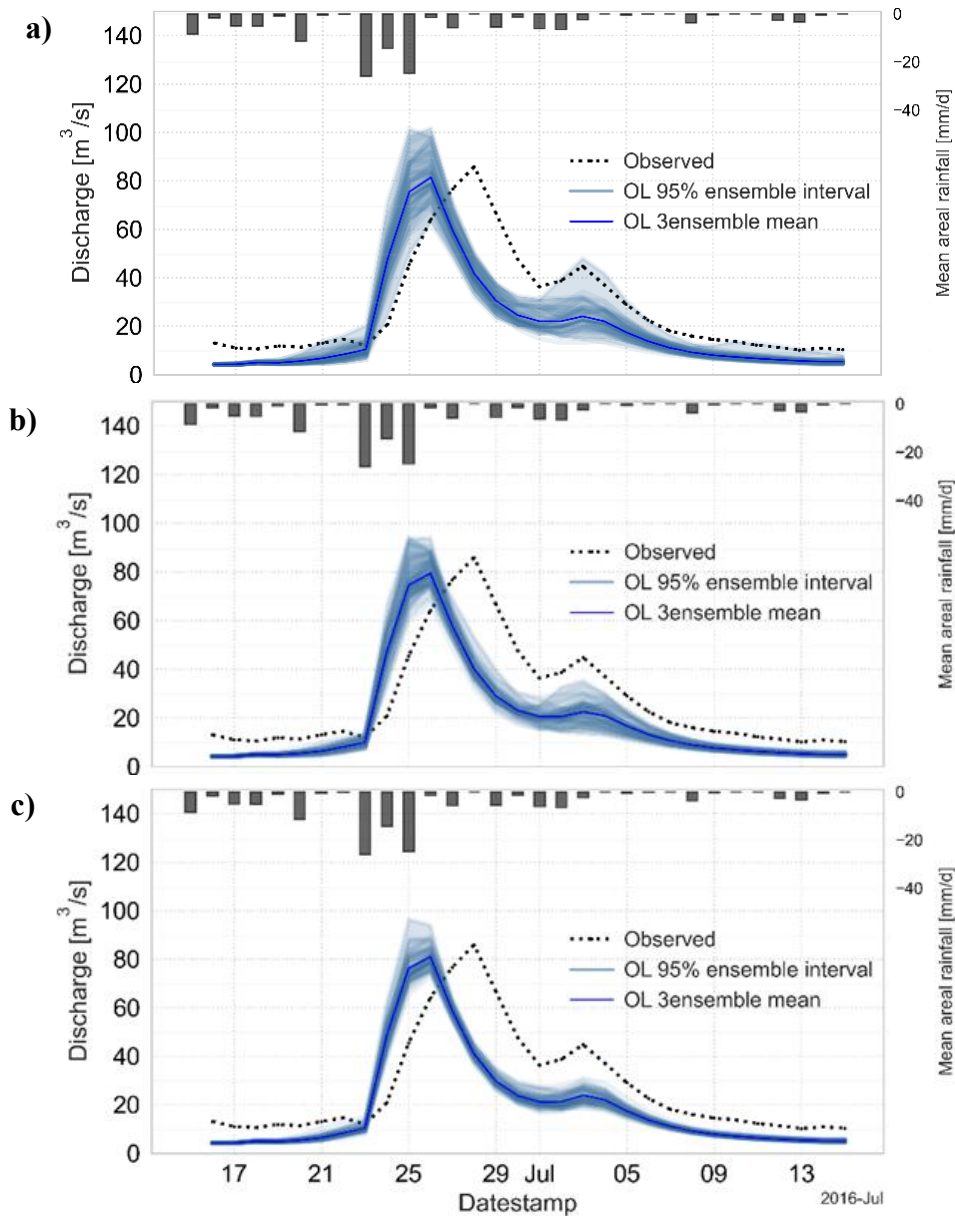


Figure 24: Open loop simulation results for three different combinations of uncertainties: a) forcing standard deviation 2.5, spatial correlation 30 km; b) forcing standard deviation 2, spatial correlation 30 km; c) forcing standard deviation 2, spatial correlation 10 km

While when it comes to metrics, all three scenarios present similar results when compared with the deterministic simulation (RMSE mean value between the ensemble mean of the OL simulations and the deterministic one is 12.49, 12.13, and 12.08 m^3/s , respectively for cases a, b and c in Figure 24), the OL reference simulation was chosen as the one with the most reasonable ensemble spread, which in this case corresponds with the one in Figure 24b. A good ensemble spread has a relatively small spread, where uncertainty is small, but it also should have the ability to span the observations.

6.3.3 Model performance regarding discharge prediction

Different parameter values for the observational noise were investigated (See appendix C.1), and the one that produced the best results (ensemble spread is closer to the observed values) is presented

in this section. The parameters selected for the noise models and the chosen number of realizations are summarized in Table 11. Finally, the DA performance regarding discharge predictions is evaluated using RMSE.

Table 11: Error model setup for the DA experiments

Qobs [m ³ /s]	EnKF [Nr. of ensembles]	Forcing [mm]	Spatial correlation [m]
0.1 multiplicative	8	2.5 multiplicative	30000

Initially, the error model of the forcing was introduced as an additive, considering the uncertainty metrics that resulted from the precipitation data analysis of Radolan in section 6.1. However, this often leads to algorithm failure as negative precipitation values are produced at certain timesteps. To avoid this, the forcing noise was set as multiplicative. At the same time, outcomes from the sensitivity analysis were utilized for determining the optimal value, where trial and error runs were made for a wide range of forcing uncertainty (factors from 0.5 to 3, results from 0.5 to 1.5 are presented in the appendix). A multiplicative error in the range of 2-3 mm produces the best discharge estimates, as seen in the results presented throughout section 6.3. It is difficult to comprehend how such a sizeable multiplicative error produces the best estimates; however, this choice is backed up by the sensitivity analysis results.

Additionally, a spatial precipitation correlation of 30 km was introduced. The range of the observations error model was selected based on literature (Weerts & el Serafy, 2006) and was further defined to be 0.1 based on the sensitivity analysis. Below, the results of the two experiments with the same error model configurations but different assimilation timelines and the number of assimilation locations are presented and discussed.

6.3.3.1 Experiment 1

The first experiment covers 30 days from 15 June 2016 to 15 July 2016, when a flood with a peak discharge of 81 m³s⁻¹ is observed. Discharges are assimilated every 24 hours from the station of Emlichehim, located near the catchment outlet. The error model configuration for both experiments is summarized in Table 11.

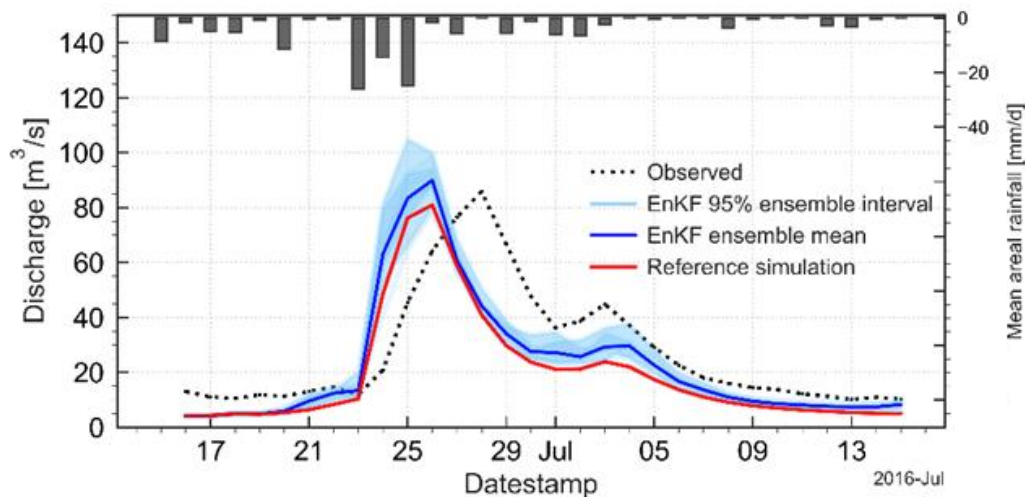


Figure 25: DA simulation results, experiment 1, Emlichehim. The observed discharges are given in the dotted line, the open loop reference simulation ensemble mean in the red line, the ensemble mean of

the DA simulation (updated with the EnKF) in the blue line, and the ensemble spread, representing a 95% uncertainty interval, in the sky blue.

The visible lag may be attributed to uncertainties in the initial conditions, which are not considered in this work, and perhaps the ensemble size. In addition, as the assimilation window progresses, the influence of updating becomes more apparent, and the ensemble spread better captures the observations. This indicates that a longer assimilation window would be more suitable; however, as the time available for this study is restricted, smaller assimilation windows are preferred to avoid extensive computational processing times.

Table 12: An overview of the mean, maximum and minimum flows corresponding to observed, open loop, and filtered discharges at Emlichheim

	Mean	Max	Min
Q_obs [m ³ /s]	27.65	85.85	10.13
Q_ref [m ³ /s]	19.62	96.43	4.08
Q_EnKF [m ³ /s]	23.16	104.89	4.09

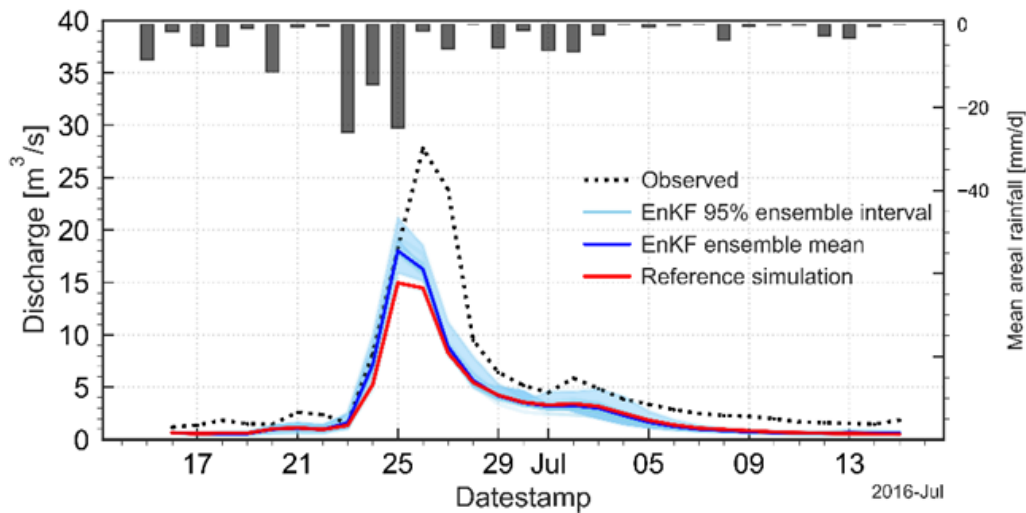


Figure 26: DA simulation results, experiment 1, Gronau. The observed discharges are given in the dotted line, the open loop reference simulation ensemble mean in the red line, the ensemble mean of the DA simulation (updated with the EnKF) in the blue line, and the ensemble spread, representing a 95% uncertainty interval, in the sky blue.

Table 13: An overview of the mean, maximum and minimum flows corresponding to observed, open loop and filtered discharges at Gronau.

	Mean	Max	Min
Q_obs [m ³ /s]	5.02	27.79	1.19
Q_ref [m ³ /s]	2.84	19.34	0.37
Q_EnKF [m ³ /s]	3.05	21.17	0.40

Figure 25 depicts the influence of discharge observations assimilation from Emlichheim at Gronau, an interior gauge upstream of the assimilation location. As observed in the instance of Emlichheim, assimilation brings simulated discharges closer to the observed values. The following

paragraphs will explain whether this effect is less pronounced in the interior gauges than at the assimilation location. Appendix C.2 contains additional results of the DA effect on the expected discharges at different interior gauges of the basin. In the case of Gronau (for location in the catchment, see gauge nr. 5 in Appendix A), the filtered discharge is closer to the observation than the reference; however, the ensembles spread cannot capture the observation peak discharge as seen in Figure 26.

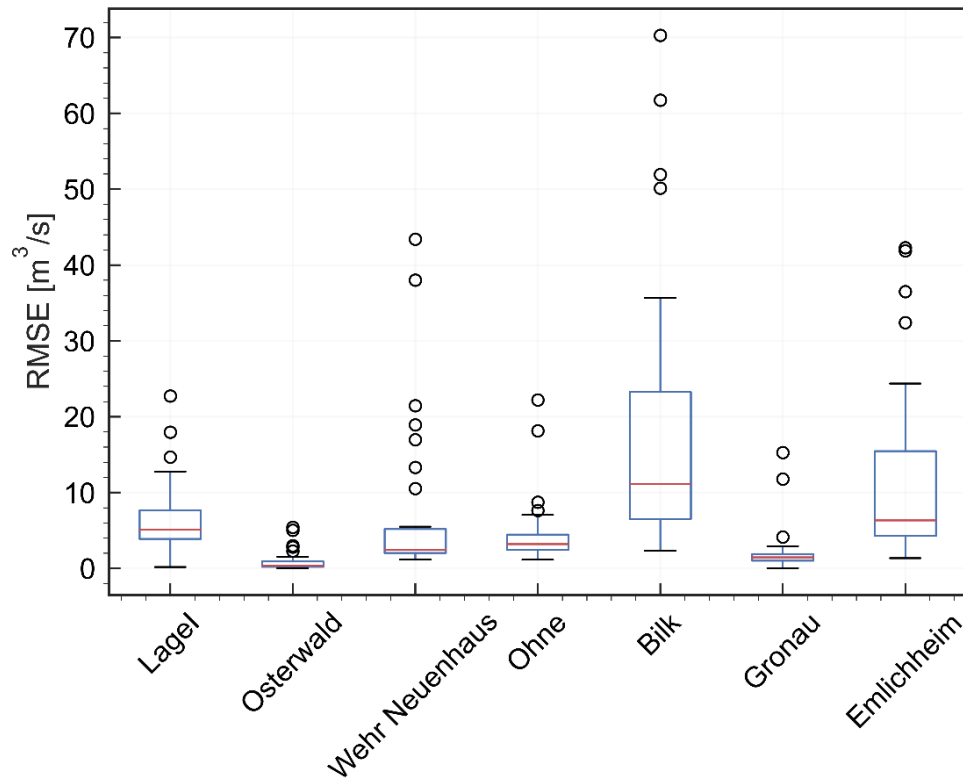


Figure 27: Boxplot of RMSE of the mean ensemble for each day between the updated with EnKF run and the daily observed discharges in selected stations. The red lines in the boxes represent the RMSE median, box borders are the 25 and 75 percentiles, whiskers are the 9 and 95 percentiles, and the circles show the extreme values of RMSE.

A summary of the data assimilation performance for all gauge stations is shown in Figure 27, where boxplots of the RMSE between the ensemble means for each day of the assimilation window between the discharges updated with EnKF and the observations are presented to evaluate the efficacy of DA. Figure 28 also shows the RMSE between the reference simulation ensemble mean and the observations. One approach for evaluating the DA effect is to compare the RMSE (MAE) values at each gauge in the simulation with and without. As seen in Figure 27 and Figure 28, the updated simulation reduces the RMSE values at Emlichheim, whereas at other stations, the effect is less pronounced, or there is no positive effect at all (e.g., Bilk).

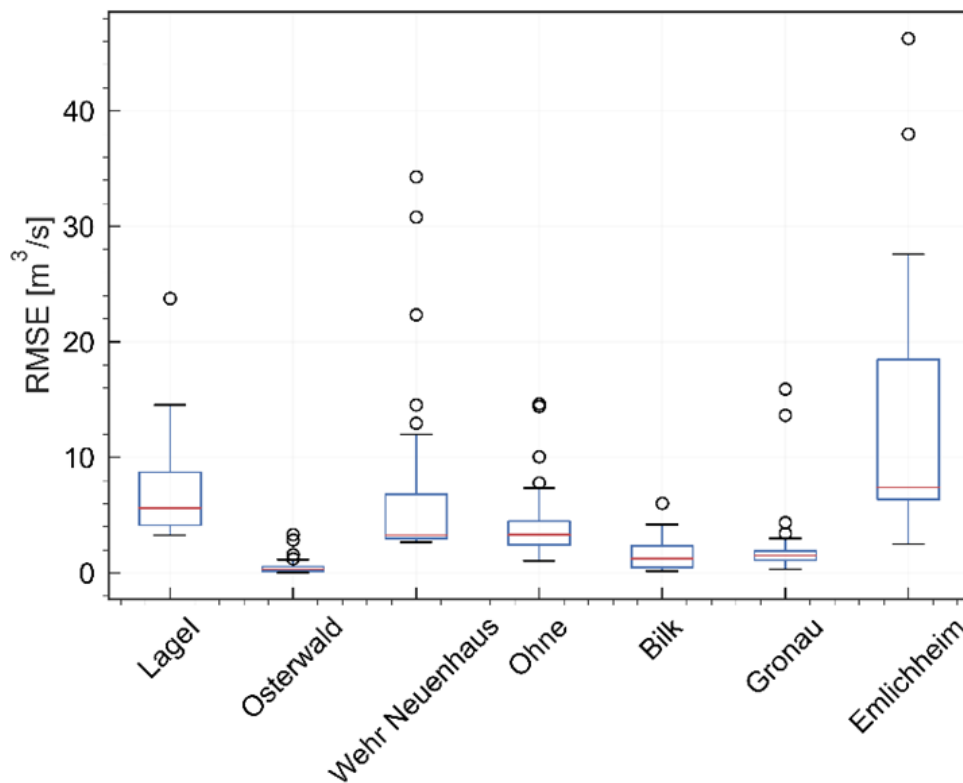


Figure 28: Boxplot of RMSE of the mean ensemble for each day between the open loop reference run and the daily observed discharges in selected stations. The red lines in the boxes represent the RMSE median, box borders are the 25 and 75 percentiles, whiskers are the 9 and 95 percentiles, and the circles show the extreme values of RMSE.

The coefficient of determination R^2 , a goodness of fit measure that represents the percentage of reduction in prediction error, calculated using the RMSE values in Figure 27 and Figure 28, is given in Table 14, where positive values indicate a positive effect of DA, whereas negative values the opposite. As expected, the DA effect when assimilation only in one location downstream has little or no positive impact at all in most upstream gauges. However, as the stations near the outlet present higher interest in flood forecasting in the Netherlands, our focus remains at Emlichheim, where the data assimilation process results in a 21.05 % improvement of the model discharge prediction. The results summarized in the table below are analogous to hydrograph results. Bilk was notably the station where the OL simulation had a significantly better performance than the DA one, as seen in Appendix C.2.

Table 14: The coefficient of determination R^2 [%] for seven different stations after DA application

	Lagel	Osterwald	Wehr Neuenhaus	Ohne	Bilk	Gronau	Emlichheim
R^2 [%]	3.57	16.67	-8.33	-11.76	-511.11	5.26	21.05

Figure 29 depicts a direct comparison of the RMSE between the OL and DA ensemble mean, where it is evident that the RMSE is lowest at the assimilation location. Overall, we may generalize that data assimilation is effective, as there is a considerable reduction in the RMSE for the downstream gauges but less for the upstream ones. That is closely related to the distance from the point of

assimilation, which in this case is the outlet; therefore, as anticipated, the upstream gauges are influenced by DA less or not at all.

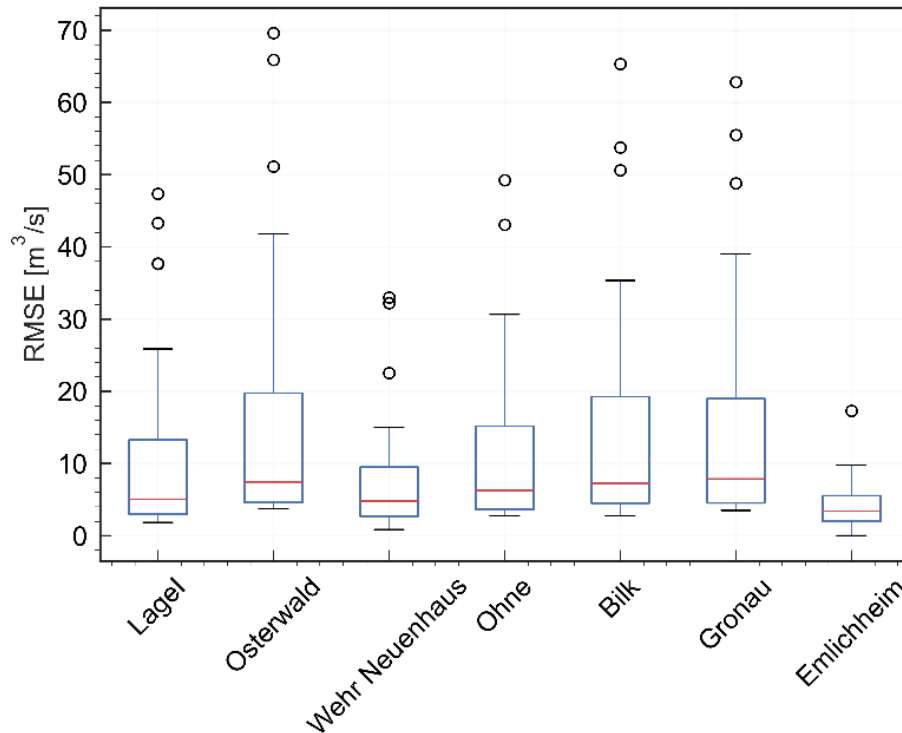


Figure 29: Boxplot of RMSE of the mean ensemble for each day between the updated with EnKF run and the open loop reference run in selected stations. The red lines in the boxes represent the RMSE median, box borders are the 25 and 75 percentiles, whiskers are the 9 and 95 percentiles, and the circles show the extreme values of RMSE.

6.3.3.2 Experiment 2

For the second experiment, a larger assimilation window is considered, covering a 60-day period from 15 January 2016 to 15 March 2016, where three consecutive flood peaks are observed, with a maximum discharge of 98 m³s⁻¹. The discharge observations are assimilated every 24 hours from Emlichheim (near the outlet of the catchment) and four interior gauges (Lage I, Ohne, Bilk, Gronau), and the error model is chosen the same as in the first experiment as summarized in Table 11. The reference simulation selected for the second experiment is developed for a significantly longer simulation time, while the assimilation run is only done for 60 days for computational reasons. The different initial conditions for the reference and DA run contribute to the significant differences between the two in the initial timesteps, as noticed in the results below.

As seen in Figure 30 and the summarized extreme and mean discharge values for the observations, reference simulation, and filtered results in Table 15, it is noticeable that the results with DA are again closer to the actual values (observations) than the ensemble means of the OL simulation. A mismatch is noticed in the first steps, stemming from uncertainties in the initial conditions. The positive effect of a longer assimilation frame is noticed, especially in the last flow peak, where the ensembles capture the observations very well.

Additionally, an apparent positive effect of streamflow assimilation is shown in Figure 31, which illustrates the results for the interior gauge of Gronau, where it is notable that the simulations

demonstrate a very good fit with the observed values. Moreover, the extremes and mean discharge and captured very well at Gronau (see Table 16), with the maximum discharge with DA (Q_{EnKF}) going down to $19.73 \text{ m}^3/\text{s}$ from maximum open loop reference discharge (Q_{ref}) at $21.92 \text{ m}^3/\text{s}$. The maximum discharges refer to the second flood peak with an observed discharge of $15.35 \text{ m}^3/\text{s}$, and this discharge reduction translates to a 10% reduction in the absolute error.

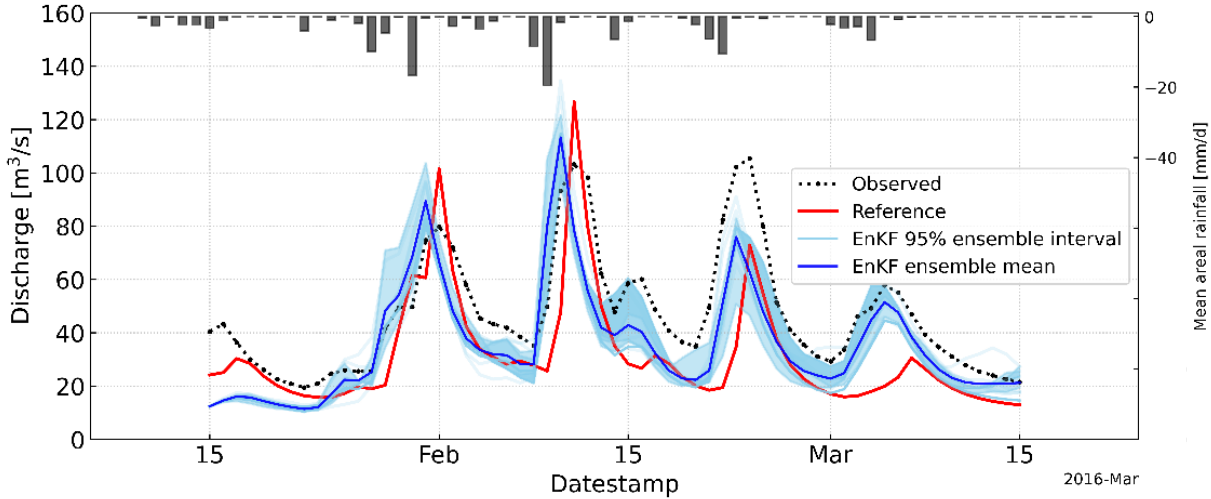


Figure 30: DA simulation results, experiment 2, Emlichheim. The observed discharges are given in the dotted line, the open loop reference simulation ensemble mean in the red line, the ensemble mean of the DA simulation (updated with the EnKF) in the blue line, and the ensemble spread, representing a 95% uncertainty interval, in the sky blue.

Table 15: An overview of the mean, maximum and minimum flows corresponding to observed, open loop and filtered discharges at Emlichheim for the second experiment.

	Mean	Max	Min
Q_obs [m^3/s]	46.74	105.52	19.51
Q_ref [m^3/s]	31.64	126.73	12.95
Q_EnKF [m^3/s]	34.93	114.69	10.36

As was anticipated, the simulation with the discharge measurements assimilation at five selected catchments produces better results, bringing the DA discharge predictions closer to the observed discharges. This positive effect is noticed at all locations where streamflow observations were assimilated (See Appendix C.3). Similarly to the first experiment, the assimilation effect increases as we go further in the assimilation window. Additionally, the lower flows at this station might substantially affect the DA application in this gauge. This is not only attributed to the DA effect but also to the flow magnitude at each assimilation station. It was noticed that the DA effect is more significant in stations with lower flows. This is also related to the model's ability to represent lower flows better, as was noticed in the first experiment and the deterministic run.

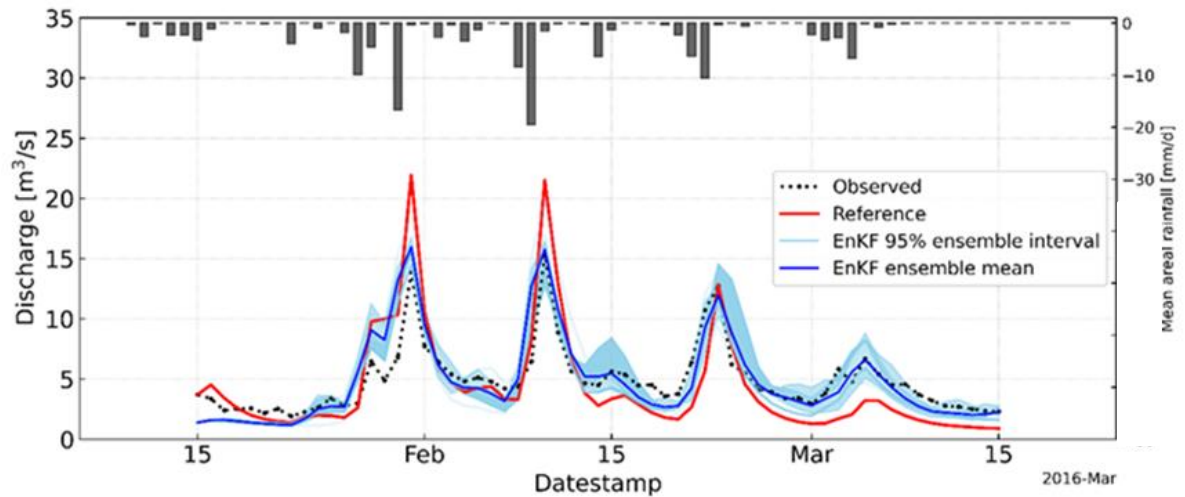


Figure 31: DA simulation results, experiment 2, Gronau. The observed discharges are given in the dotted line, the open loop reference simulation ensemble mean in the red line, the ensemble mean of the DA simulation (updated with the EnKF) in the blue line, and the ensemble spread, representing a 95% uncertainty interval, in the sky blue.

Table 16: An overview of the mean, maximum and minimum flows corresponding to observed, open loop and filtered discharges at Gronau for the second experiment.

	Mean	Max	Min
Q_obs [m³/s]	4.87	15.35	1.91
Q_ref [m³/s]	4.29	21.92	0.9
Q_EnKF [m³/s]	4.89	19.73	1.06

From the additional results presented in Appendix C.3 and summarized in Figure 32, it is apparent that not only the location of the flow gauge influences the DA results but also the magnitude of flows at each station, as the wflow_sbm model has a better predictive skill at lower flows, as also seen in the deterministic simulations in section 6.2. Overall, the mean RMSE between the observations and discharges with DA is slightly lower than the RMSE between the observations and the OL simulation discharges. However, the RMSE spread is distinctively smaller after DA, especially at Emlichheim.

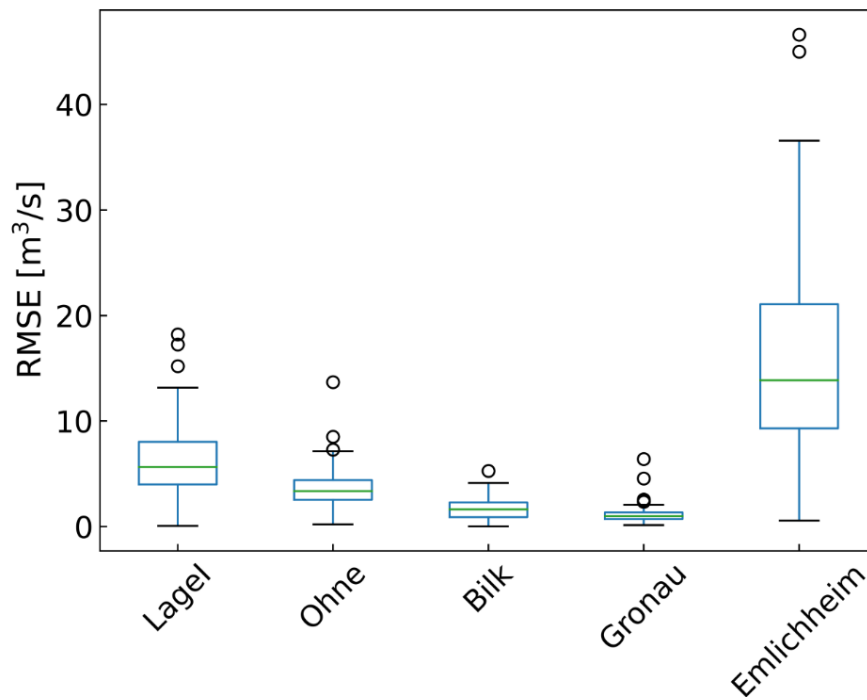


Figure 32: Boxplot of RMSE of the mean ensemble for each day between the updated with EnKF run and the daily observed discharges in selected stations. The green lines in the boxes represent the RMSE median, box borders are the 25 and 75 percentiles, whiskers are the 9 and 95 percentiles, and the circles show the extreme values of RMSE.

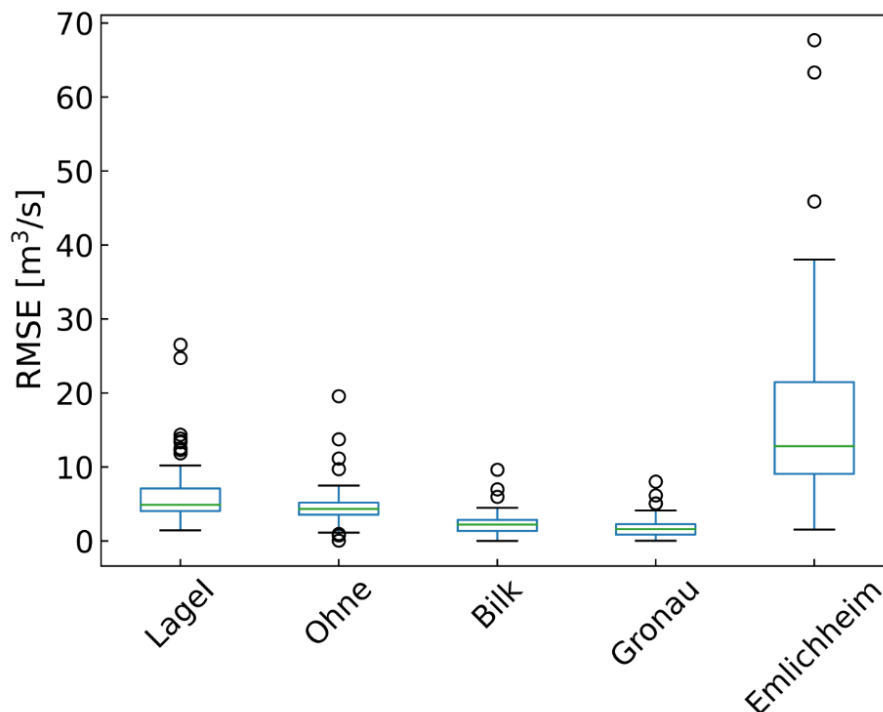


Figure 33: Boxplot of RMSE of the mean ensemble for each day between the open loop reference run and the daily observed discharges in selected stations. The green lines in the boxes represent the RMSE median, box borders are the 25 and 75 percentiles, whiskers are the 9 and 95 percentiles, and the circles show the extreme values of RMSE.

Utilizing the metrics presented in Figure 32 and Figure 33, the coefficient of determination R^2 is computed as seen in Table 17, where positive values of R^2 are given for all assimilation stations considered in the second experiment. This indicates a positive effect of the DA application. However, the percentage of reduction error is related to the discharge magnitude in each station, so a higher rate of error reduction (36.41% at Gronau) does not necessarily translate into a higher absolute error reduction, as R^2 is not always straightforward. Noticeably, the DA effect is positive at all stations when multiple streamflows are assimilated, unlike the first experiment where measurements only near the outlet were assimilated. However, for the rest of the analysis, the first experiment with a shorter assimilation window and one assimilation location is considered when exploring the effects of DA in hydrological states and fluxes spatially.

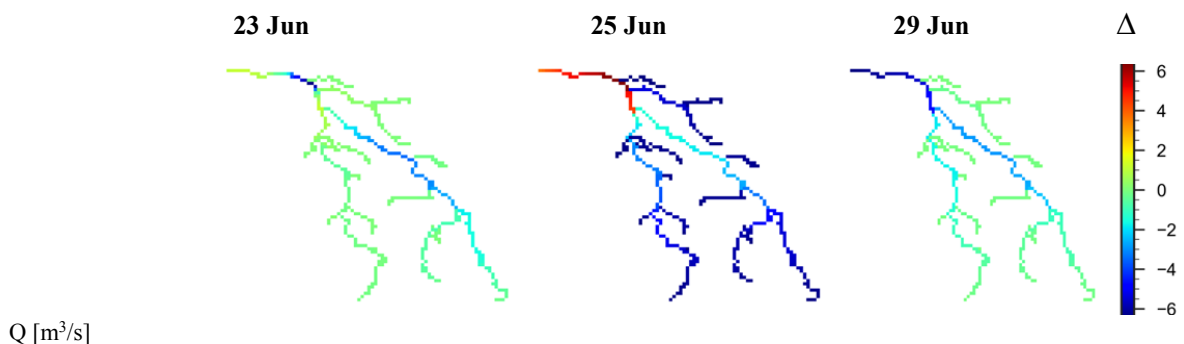
Table 17: The coefficient of determination R^2 [%] for five different stations after DA application

	Lagel	Ohne	Bilk	Gronau	Emlichheim
R^2 [%]	3.98	23.46	29.94	36.41	3.82

6.3.4 Model performance regarding state updating

The following case is presented to explore the effect of the DA on the update of spatially distributed model states. The mean difference between the predicted and updated model states (refers to the EnKF algorithm steps as described in Figure 1) is shown in Figure 34 for three different time steps, for every state included in the analysis, selected after examining the model performance regarding discharge predictions for the first experiment (See 6.3.3.1). These examples aim to aid our understanding of the discharge measurement assimilation impact in the gridded hydrological model. As the model simulations (at the outlet and most of the interior gauges) are mainly underestimating the observations, water is added to the system during the update step; however, this remains location specific as further from the assimilation location, the model overestimates the discharges, so water is removed from the system.

Below we see how the EnKF captures the hydrological system spatially, varying from a positive to a negative difference depending on the location on the grid.



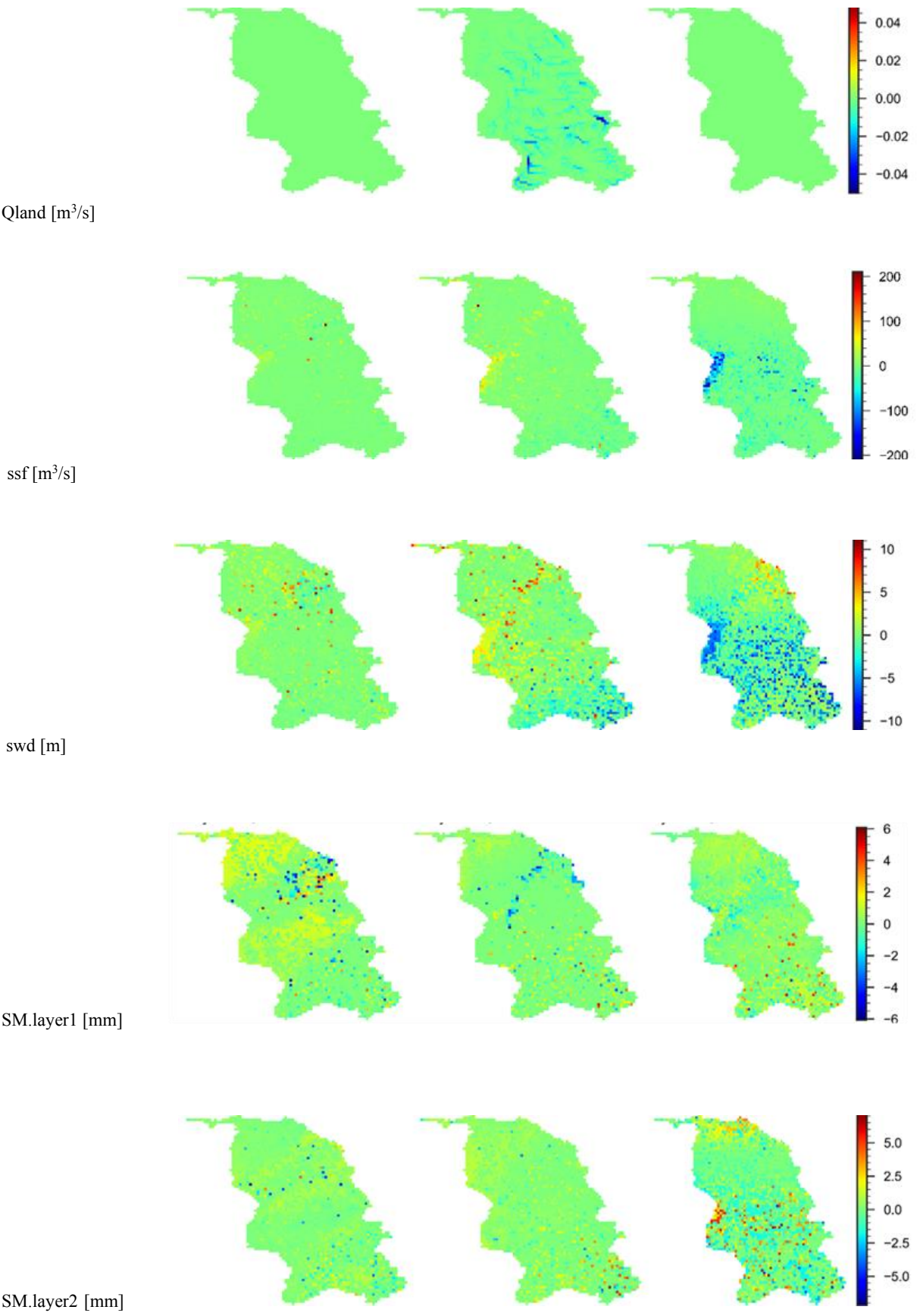




Figure 34: Difference between the predicted (instates) and updated (outstates) model states on three selected days (23, 25, 29 June 2016). Only the most sensitive model states are shown, discharge (Q), land flow (Qland), subsurface flow (ssf), saturated water depth (swd), and soil moisture in three different layers

As we are assimilating near the outlet, the effect of assimilation in nearby gauges will mimic the effect at the assimilation point, where the EnKF adds water to the system as seen in the hydrographs of Emlichheim and other downstream gauges LageI, Wehr Neuenhausen, Osterwald; While the effect of DA is rather positive at downstream stations (during the flood peak, around 25th of June) that present a similar situation with Emlichheim (the model underpredicts the flood peak), in other stations where the model overpredicts (case of Ohne for example) or produces simulations very close to observations (Lage I), the effect of outlet discharge DA is negative as the EnKF updated discharge values are further than the simulated ones without DA (case of Wehr Neuenhaus).

For the more upstream gauges like Bilk, Ohne, and Gronau, spatially, we notice that water starts to be removed from the system on the 25th of June (flood peak is registered earlier in the upstream stations). On 29th, most of the upstream area presents negative values of the hydrological states, making sense as the water is moving from upstream to downstream. Again, as data is assimilated only near the outlet, the updated model discharge predictions do not follow necessarily follow the observations in the upstream stations, resulting in little or no positive effect at all in improving the model discharge predictions.

From the model states investigation, we notice first that the streamflow assimilation impacts not only the overland state components but also considerable changes are detected in the subsurface components. Typically, by adding water to the system near the outlet, this water is drained from the most upstream locations as seen in the spatial presentation of the discharge Q. A negative difference in the overland flow Qland is noticed around the time of the flood peak upstream, which reflects that the model overestimates the Qland. At the same time, it is noticed that subsurface water flow is added in the downstream part of the catchment, indicating that the model correctly redistributes the water in the system when applying the EnKF. This is also noted in the saturated water depth (swd), which is increased downstream near the flood peak time and shows negative values further upstream where the flood peak has passed. The further we move from the flood peak time, the lower the saturated water depth gets upstream, as seen on the column corresponding to the 29th of June. The effect is also noticed in the uppermost soil moisture layer before and near the flood peak.

In contrast, significant changes are only noticed in the deeper layers later in time. Overall the system state dynamics are captured accurately from the update with the EnKF; however, whether this new distribution represents the “true” model states remains unknown as it is difficult to observe all hydrological model states. Additionally, as mentioned above, EnKF capturing the state dynamics

spatially does not necessarily lead to improved discharge predictions in the catchment. However, as was already mentioned assimilating more observations will lead to better discharge estimates and therefore better state estimates in the whole catchment.

However, as the goal of this study is to investigate the effect of DA in improving the state estimates mainly near the outlet (the point of interest for operational flood forecasting in the Netherlands), the experiment where we assimilated only near the outlet was deemed enough to answer the respective research question, also considering the time and computational constraints.

7 CONCLUSIONS AND RECOMMENDATIONS

This section provides a summary of the work reported in this study and the key findings that address the research objectives. The limitations of the study's methodology and approach are also addressed. Finally, conclusions and future research recommendations are outlined.

7.1 Summary

To explore the above, in the first part of this thesis, a precipitation data analysis was conducted to give insight into the selected precipitation product quality and whether it is appropriate to be used in the assimilation experiments. From this analysis, the precipitation product demonstrated a satisfactory performance, confirmed by several metrics employed. For this reason, it was deemed appropriate to be used as input forcing in the DA model. Additionally, utilizing information on the uncertainties associated with the forcing product, the error model for the DA scheme was further selected. Finally, this work also includes a short analysis of the observed and model discharges obtained from a deterministic historical 10-year simulation run in Delft-FEWS, which suggests that the model has a good overall performance. However, peak flows are not always adequately captured.

Utilizing OpenDA, a generic data assimilation toolbox, two data assimilation experiments were conducted using the same error model but differing in from the assimilation window (one short and one longer experiment) and the number of observation locations assimilated (one location for the first experiment and five locations for the second one). Regarding the uncertainty model, as the model is considered to be perfect, no uncertainties related to the model error are considered when applying DA, but on the contrary, the errors (mismatch between observations and model simulation) we attribute only to the observations, the uncertainty value of which was selected based on literature suggestions, as well as the results from the deterministic run and a sensitivity analysis carried out to enhance our understanding of how the DA scheme works. For the input forcing uncertainty, a trial and error approach was the most effective in selecting the optimal forcing uncertainty. Before the assimilation experiments, reference simulations were selected for each experiment, using open-loop probabilistic simulations with 48 ensemble members. The run that produced the most reasonable ensemble spread closer to the observations, and deterministic simulation was selected. The results confirmed that a longer assimilation window would result in better estimates, especially in the later timesteps of the simulation.

Additionally, assimilating only in one location does not always result in improved predictions in the other stations, as the assimilation effect will depend on the distance of each station from the assimilation location and the flow magnitude at each station. However, when multiple streamflow observations were assimilated, the assimilation effect results were positive at every checked location. Lastly, the hydrological state dynamics when updating with the EnKF were investigated, and results suggest that the discharge assimilation has a positive effect in capturing the system dynamics of the wflow_sbm model; however, this is mainly a qualitative assessment, as it is challenging to derive quantitative judgment while we don't know the 'true' system states. To conclude, streamflow assimilation was considered to positively affect the discharge predictions and capture the system state dynamics. However, this study's several limitations are elaborated in the following section, and future research directions are given.

7.2 Limitations

This study's approach and methodology have several shortcomings. First, there are limitations connected with the study's assumptions, such as the assumption of a faultless hydrological model and initial conditions. However, incorporating the unaccounted-for uncertainties into this study will undoubtedly result in more accurate estimations, allowing the DA impact on the discharge predictions to be judged accurately.

Only the precipitation forcing input in the DA experimental setup is assumed to be subject to error, whereas the potential evapotranspiration and temperature forcing uncertainties are disregarded. Another constraint is the precipitation uncertainty characterization. In this study, error characterizations of precipitation forcing in the DA system are merely multiplicative errors applied to precipitation fields. As rainfall uncertainty is the primary source of spread among ensembles, which is required for incorporating data from observations to correct model predictions, and precipitation fields may contain both magnitude and position errors, other statistically more reliable methods described in Section 2.2 can be used. Despite the radar product being calibrated with observational rainfall data, it presents inaccuracies, and radar bias correction techniques can be applied. Correctly accounting for observation uncertainty in data assimilation is also a shortcoming of this study, as the observation uncertainty is assessed within a literature-suggested range, and an ad-hoc error selection method is employed. Other limitations include the errors associated with the model parameters, which are assumed to be ideal in this work.

Lastly, the high computational time required to conduct the experiments is a significant shortcoming of this study. The high computational times could stem from the OpenDA assimilation tool itself, the fact that we work with a distributed model, and the filtering algorithm choice. While OpenDA is a very flexible and simple-to-configure tool, when used with wflow_sbm, it requires exceptionally long computing durations, which rise dramatically (up to 12 hours) in runs with a longer assimilation time (61 days), a greater number of ensembles (more than eight ensembles), and more discharge observation location's data being assimilated (5 assimilation locations). Experiments with a 31-day assimilation window, a single assimilation location, and eight ensemble members necessitate considerable computational time (two to three hours) and capacity. This, together with the limited available time for this research, has reduced the number and duration of the DA experiments presented in this work. A more time-efficient tool must be developed to be utilized in operational forecasting for catchments that respond rapidly, such as the Geul, for which the DA was not performed in this study due to limitations of the model and data available.

7.3 Conclusions

The primary objective of this study was to investigate the effects of assimilating streamflow data in the model predictions in the Vecht river catchment. Thus, this work aimed to enhance the understanding of assimilation in other hydrological states of the wflow_sbm hydrological model. As several uncertainties are to be considered in the DA process, as elaborated throughout this study, this thesis included some preliminary analysis to account for the related uncertainties. Therefore, a precipitation data analysis was carried out for the selected gridded rainfall product (Radolan), comparing at points with the observational dataset to account for the model forcing uncertainty. Then, the ability of the model to capture the observed discharges by conducting historical simulation runs was investigated. However, the model parameter uncertainties were ignored, and the model was assumed

perfect when carrying out data assimilation experiments. Next, effects in the discharge predictions were reported, and finally, implications on other model states and whether the assimilation filter can capture a distributed hydrological model state dynamics were addressed. In the conclusion of this thesis, the research questions identified in Section 1.3 are revisited.

- What is the effect of assimilating streamflow observations with the wflow_sbm distributed hydrological model on the streamflow predictions for the Vecht river using the EnKF? What is the influence of the assimilation window length and the number of assimilation locations?
Results from this work indicate a positive effect on streamflow model predictions. However, the magnitude of improvement is sensitive to the error model selected for the experiments, and it varies depending on the amount of assimilated streamflow observations and the length of the assimilation window. By comparing the outcomes of experiment 1 and experiment 2, it can be concluded that assimilating observations near the outlet and other interior gauges will result in better discharge predictions in the entire catchment, while when assimilating only near the outlet, the EnKF application improves predictions only in a particular station, particularly in stations closer to the assimilation location, and stations where the wflow_sbm model follows the same trend (overpredicting or underpredicting flow peaks) as the assimilation station. For example, there is a 21.05% improvement in the model predictions at Emlichheim in the first experiment after DA. However, an adverse effect is noticed in other stations (e.g., Bilk, Ohne). The first experiment showed that the EnKF implementation does not solve the issue of a too steep rising limb and a peak flow predicted too early.
In the second experiment, the predictions are improved with 3.82% (or 8 m³/s improvements in absolute error term) for the maximum flood peak. Therefore, it can be concluded that a longer assimilation window and assimilation of multiple observations will considerably improve the discharge predictions. However, this conclusion is derived from a limited number of experiments, and for quantitatively assessing the potential of the EnKF for broader applications, more investigations are needed.
- What effect does discharge data assimilation have on other hydrological model states and fluxes, given that we use a model in which water is routed laterally via subsurface and surface water, whereas previous studies used a model in which water was routed only via surface water?
From the results presented in section 6.3.4, it is concluded that the discharge DA effect, as expected, is reflected in all states that demonstrate a considerable correlation with the model discharges. The model is able to properly redistribute the water in the system when updating with the EnKF. Changes are noticed in the subsurface and overland flows corresponding to the theoretical flood response of the system, so for example, differences are detected during the flood peak time in the uppermost soil moisture layer of the model, while these differences are reflected in the deeper soil moisture layers later in time. Overall, it is concluded that the subsurface and surface water changes can be captured appropriately. However, assessing the effect on individual states at particular points is difficult, as we lack information on the actual values of various states.
- What is the influence of the error model specification of the input forcing and observations on the discharge predictions with the wflow_sbm model?
The outputs of the distributed hydrological model used in this work present a high sensitivity in the error model specifications chosen for the assimilation framework, as was noted in the

preliminary analysis carried out, where even for a 10% change in the uncertainty specification (particularly for precipitation), significant impact is noticed in the predicted flows. For this reason, an appropriate error selection of the precipitation uncertainty is crucial. In this work, the combination of uncertainties that produced the best results for both experiments is a precipitation error factor of 2.5 and a 0.1 factor for the observation error. However, as other uncertainties are neglected, these values need to be revised when including uncertainties related to the model parameters and initial conditions. The model might produce unrealistic flow predictions (and other variables) when compensating for the neglected errors.

7.4 Recommendations

During the production of this study, characteristics of suggestions for future research were identified, relating to the limitations and findings presented; as a result, the following recommendations for further research are presented:

1. Given the lack of information on the "best" rainfall data source and given that the precipitation uncertainty propagates throughout the analysis, it is recommended that statistically more reliable techniques, like conditional simulation, are used to assess the precipitation error estimates. Additionally, before uncertainty assessment of the radar precipitation product, applying bias adjustment of the radar forcing data is considered appropriate.
2. While a sensitivity analysis was done to investigate which combination of spatial correlation and uncertainties will result in the most optimal simulation with DA (the one that brings the ensemble closer to the observations), the values chosen for the spatial correlation of precipitation were based on literature suggestion. Therefore, an analysis of the spatial correlation structure of precipitation at the local scale should be conducted to obtain better results.
3. The effect of consideration of other uncertainties, like the model structure uncertainty, initial condition uncertainty, etc., should be explored.
4. More experiments with a larger assimilation window and additional assimilated gauges should be conducted to investigate DA's effects further. While performing those experiments with the current OpenDA version is computationally ineffective, further research should be conducted on how to produce results more effectively.
5. Additional experiments with EnKF DA can be conducted to derive ensemble forecasts to be evaluated based on lead times, using the verification measures mentioned in Section 2.3.2.
6. In view of the estimated positive results of streamflow DA at the Vecht river catchment, it is recommended to develop a similar assimilation setup for the Geul river basin, following the work that is already carried out for the Geul in this study, as time and computational challenges did not allow for the implementation of additional DA experiments. A comparison of the effectiveness of EnKF DA for the two different catchments can then be made. Furthermore, for the Geul, other filtering algorithms like the Asynchronous Ensemble Kalman Filter (explained in section 2.1) are recommended to be explored, as it is a promising one regarding reduced assimilation time.

8 REFERENCES

- Bijlsma, R., & Blind, M. (2006). HarmoniRiB River Basin Data Documentation Chapter 8-Vecht River Basin.
- Broersen, P. M. T., & Weerts, A. H. (2005). Automatic Error Correction of Rainfall-Runoff models in Flood Forecasting Systems. <https://doi.org/10.1109/IMTC.2005.1604281>
- Brown, T. A. (1974). Admissible Scoring Systems for Continuous Distributions. RAND Corporation. <https://www.rand.org/pubs/papers/P5235.html>
- de Bruin, H. A. R., & Lablans, W. N. (1998). Reference crop evapotranspiration determined with a modified Makkink equation.
- de Moor, J. J. W., Kasse, C., van Balen, R., Vandenberghe, J., & Wallinga, J. (2008). Human and climate impact on catchment development during the Holocene - Geul River, the Netherlands. *Geomorphology*, 98(3–4), 316–339. <https://doi.org/10.1016/j.geomorph.2006.12.033>
- Dharmadasa, M. D. V. S. (2014). Flooding in Županjska Posavina. ISRBC.
- DWD Climate Data Center (CDC): Recent hourly RADOLAN grids of precipitation depth (binary), version 2.5. (n.d.).
- Eduardo Villarreal Jaime, D. (2021). Ensemble hydrological forecasts to derive extreme return periods: Case Study of the Overijsselse Vecht River using the wflow_sbm model Master Programme in Flood Risk Management.
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. In *Journal of Geophysical Research* (Vol. 99, Issue C5).
- Georgakakos, K. P., Seo, D. J., Gupta, H., Schaake, J., & Butts, M. B. (2004). Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. *Journal of Hydrology*, 298(1–4), 222–241. <https://doi.org/10.1016/J.JHYDROL.2004.03.037>
- Gupta, H. v., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91. <https://doi.org/10.1016/J.JHYDROL.2009.08.003>
- Heerinen, K.-J., Filius, P., Tromp, G., & Renner, T. (2013). FEWS Vecht, a crossing boundaries flood forecasting system. 13808.
- Huisink, M. (2000). Changing river styles in response to Weichselian climate changes in the Vecht valley, eastern Netherlands. www.elsevier.nl/locate/sedgeo

- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- Jan Pierik, H., J van Lanen, R., TIJ Gouw-Bouman, M., J Groenewoudt, B., Wallinga, J., & Z Hoek, W. (2018). Controls on late-Holocene drift-sand dynamics: The dominant role of human pressure in the Netherlands. <https://doi.org/10.1177/0959683618777052>
- KNMI. (n.d.). Climatology - Measurements and Observations. Retrieved February 1, 2022, from www.knmi.nl
- Kreklow, J., Tetzlaff, B., Kuhnt, G., & Burkhard, B. (2019). A Rainfall Data Intercomparison Dataset of RADKLIM, RADOLAN, and Rain Gauge Data for Germany. <https://doi.org/10.5281/zenodo.3262172>
- Liu, Y., & Gupta, H. v. (2007a). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. 43, 7401. <https://doi.org/10.1029/2006WR005756>
- Liu, Y., & Gupta, H. v. (2007b). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. In *Water Resources Research* (Vol. 43, Issue 7). <https://doi.org/10.1029/2006WR005756>
- Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., Seo, D.-J., Schwanenberg, D., Smith, P., van Dijk, A. I. J. M., van Velzen, N., He, M., Lee, H., Noh, S. J., Rakovec, O., & Restrepo, P. (2012). Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. *Hydrology and Earth System Sciences*, 16(10), 3863–3887. <https://doi.org/10.5194/hess-16-3863-2012>
- Lorenc A.C. (1981). [15200493 - Monthly Weather Review] A Global Three-Dimensional Multivariate Statistical Interpolation Scheme. 109(4), 701–721. [https://doi.org/https://doi.org/10.1175/1520-0493\(1981\)109%3C0701:AGTDMS%3E2.0.CO;2](https://doi.org/https://doi.org/10.1175/1520-0493(1981)109%3C0701:AGTDMS%3E2.0.CO;2)
- Nash, J. E., & Sutcliffe, J. v. (1970). Not to be reproduced by photoprint or microfilm without written permission from the publisher. River flow forecasting through conceptual models part i-a discussion of principles*. In *Journal of Hydrology* (Vol. 10). © North-Holland Publishing Co.
- OpenDA. (2013). The OpenDA data-assimilation toolbox. [Http://Www.Openda.Org](http://Www.Openda.Org).
- Rakovec, O. (2014). Improving operational flood forecasting using data assimilation.
- Rakovec, O., Hazenberg, P., F. Torfs, P. J. J., Weerts, A. H., & Uijlenhoet, R. (2012). Generating spatial precipitation ensembles: Impact of temporal correlation structure. *Hydrology and Earth System Sciences*, 16(9), 3419–3434. <https://doi.org/10.5194/hess-16-3419-2012>
- Rakovec, O., Weerts, A. H., Hazenberg, P., F. Torfs, P. J. J., & Uijlenhoet, R. (2012). State updating of a distributed hydrological model with ensemble kalman Filtering: Effects of

- updating frequency and observation network density on forecast accuracy. *Hydrology and Earth System Sciences*, 16(9). <https://doi.org/10.5194/hess-16-3435-2012>
- Reichle, R. H., Walker, J. P., Koster, R. D., & Houser, P. R. (n.d.). Extended versus Ensemble Kalman Filtering for Land Data Assimilation.
- Schellekens, J. , V. W. van, V. M. , W. H. , E. T. , B. L. , T. C. , V. S. de, B. H. , E. D. , T. D. , W. A. , B. F. , H. P. , L. A. , V. C. ten, J. M. , & B. I. (2021). Openstreams. wflow.
- van de Beek, C. Z., Leijnse, H., Stricker, J. N. M., Uijlenhoet, R., & Russchenberg, H. W. J. (2010). Performance of high-resolution X-band radar for rainfall measurement in The Netherlands. *Hydrology and Earth System Sciences*, 14(2), 205–221. <https://doi.org/10.5194/hess-14-205-2010>
- van den Munckhof, G. (2020). Forecasting river discharge using machine learning methods with application to the Geul and Rur river. <http://repository.tudelft.nl/>.
- van Velzen, N., Muhammad, ., Altaf, U., & Verlaan, . Martin. (2016). OpenDA-NEMO framework for ocean data assimilation. *Ocean Dynamics*, 66, 691–702. <https://doi.org/10.1007/s10236-016-0945-z>
- Vrugt, J. A., Diks, C. G. H., Gupta, H. v, Bouten, W., & Verstraten, J. M. (2005). Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. 41, 1017. <https://doi.org/10.1029/2004WR003059>
- Vrugt, J. A., Gupta, H. v, Bastidas, L. A., Bouten, W., & Sorooshian, S. (2003). Effective and efficient algorithm for multiobjective optimization of hydrologic models. 39(8), 1214. <https://doi.org/10.1029/2002WR001746>
- Weaver, A. T. (2003). 3D-Var and 4D-Var approaches to ocean data assimilation. ECMWF Workshop on the Role of the Upper Ocean in Medium and Extended Range Forecasting, 13-15 November 2002, 57–66. <https://www.ecmwf.int/node/13003>
- Weerts, A. H., & el Serafy, G. Y. H. (2006). Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models. *Water Resources Research*, 42(9). <https://doi.org/10.1029/2005WR004093>
- World meteorological organization., W. M. Organization. (1992). Simulated real-time intercomparison of hydrological models. Secretariat of the World meteorological organization.

9 APPENDICES

Appendix A. Streamflow gauge station locations

A.1 Vecht

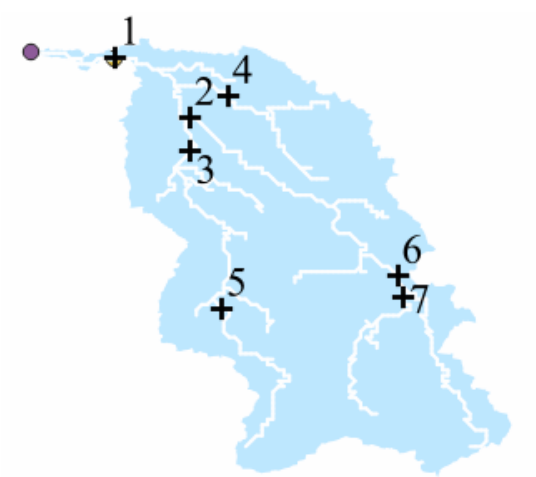


Table 18: Overview of the streamflow gauge station in the Vecht river basin

Map ID	1	2	3	4	5	6	7
Station Name	Emlichheim	Wehr-Neuenhaus	Lage I	Osterwald	Gronau	Ohne	Bilk

A.2 Geul

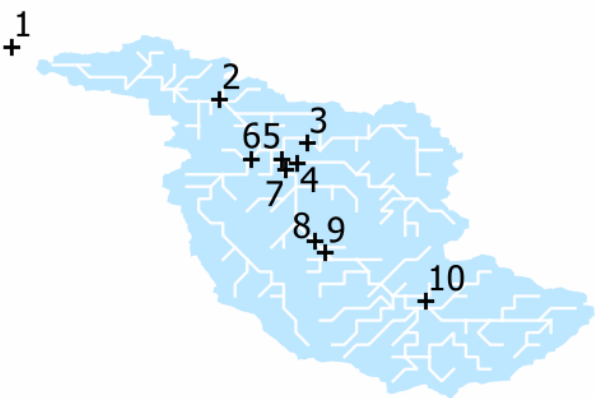


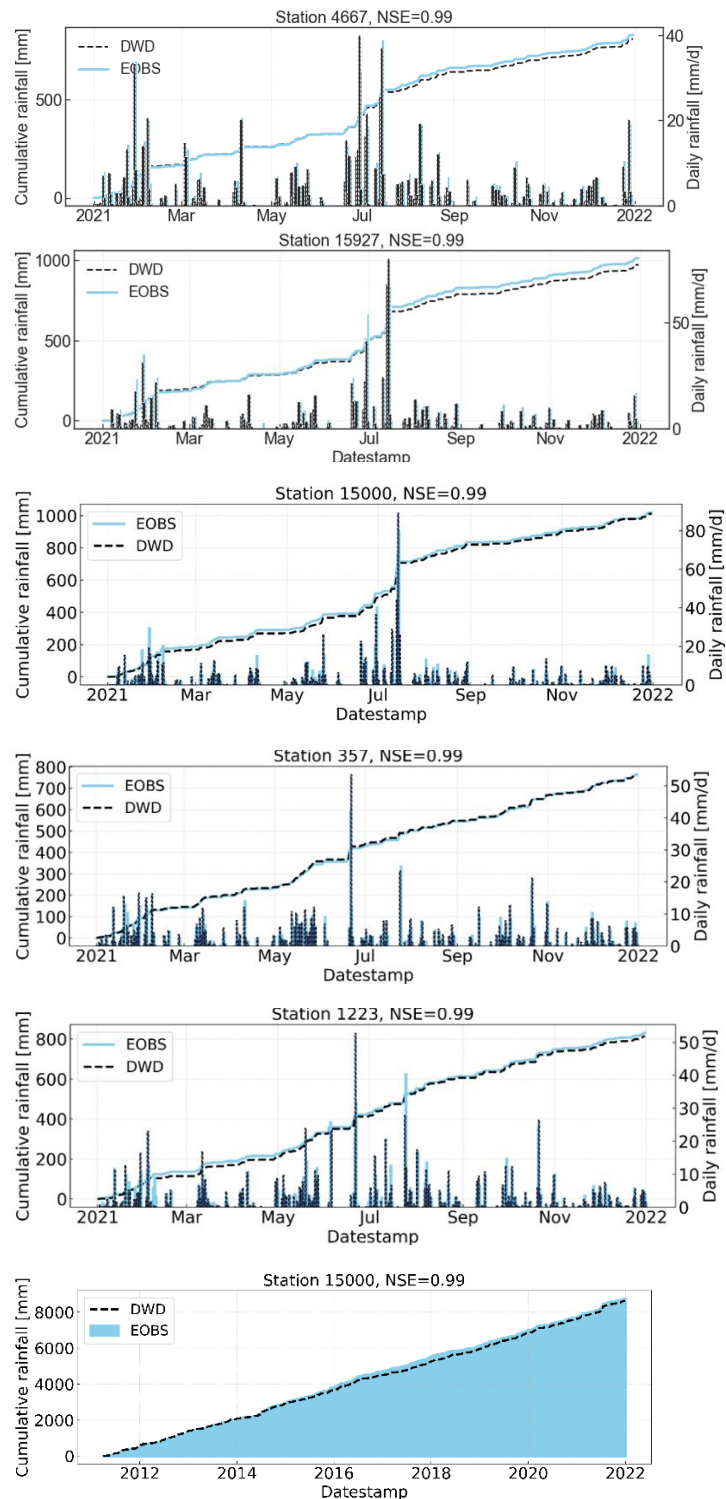
Table 19: Overview of the streamflow gauge station in the Geul river basin

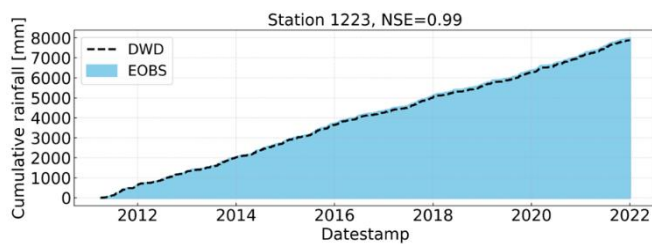
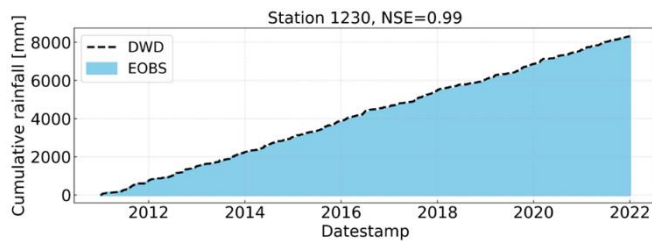
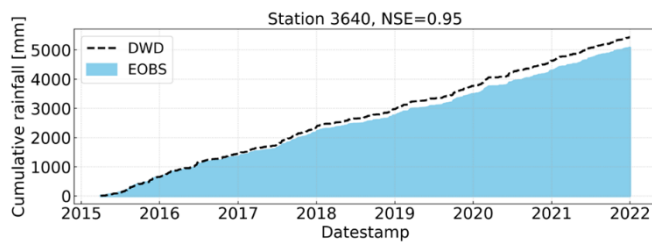
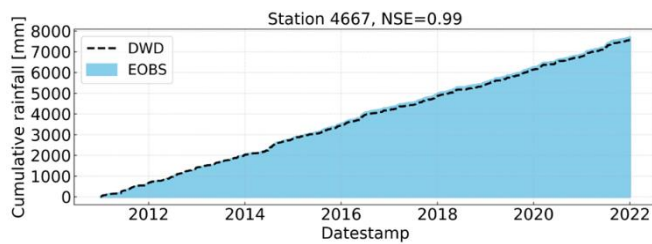
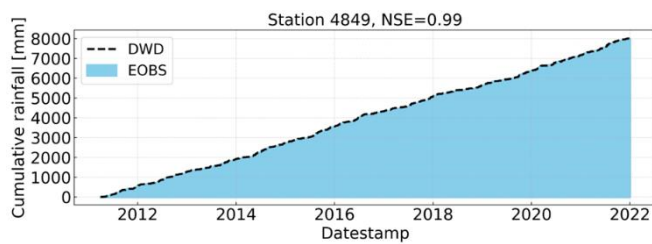
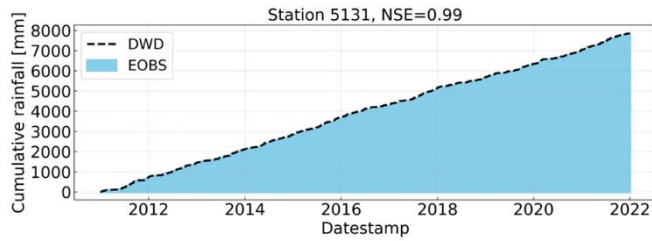
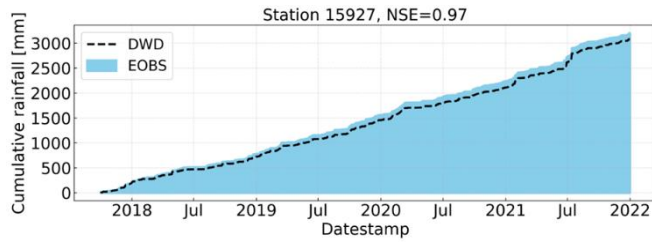
Map ID	1	2	3	4	5
Station Name	Meerssen	Schin op Geul	Eyserbeek Eys	Selzerbeek Partij	Selzerbeek, molentak

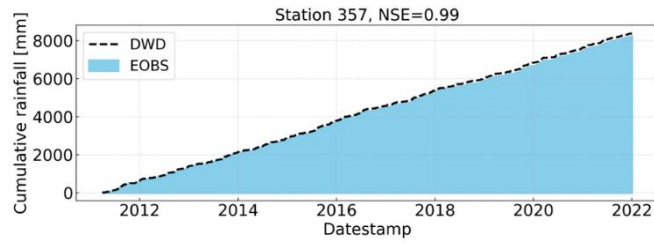
Map ID	6	7	8	9	10
Station Name	Gulp, Azijnfabrik	Hommerich	Cottessen	Sippenaken	Kelmis

Appendix B. Precipitation data analysis

B.1 DWD rain gauge observations VS E-OBS gridded dataset

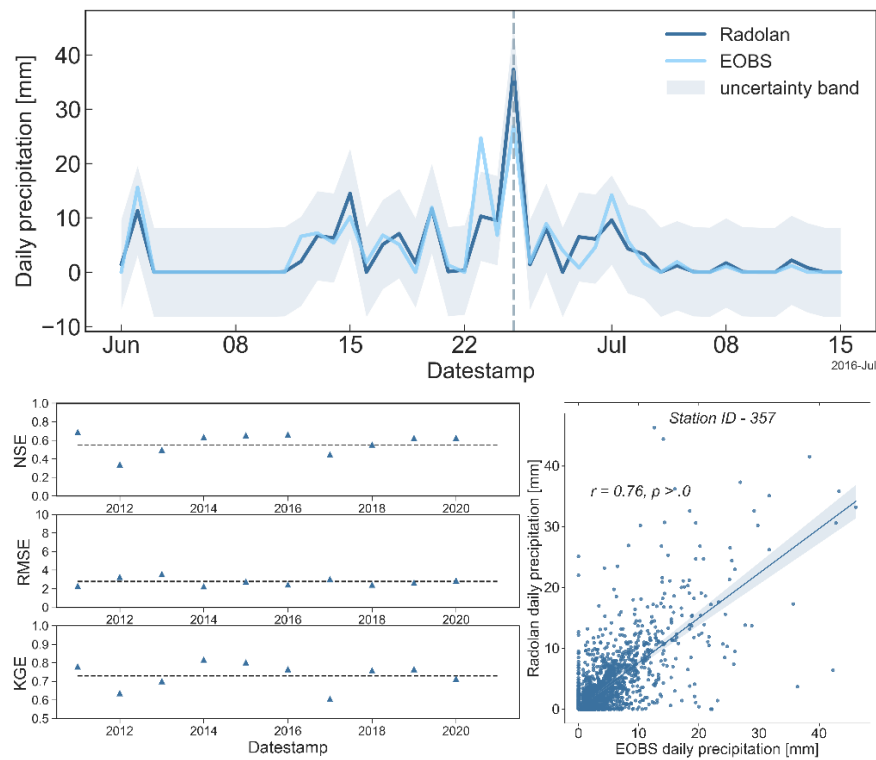


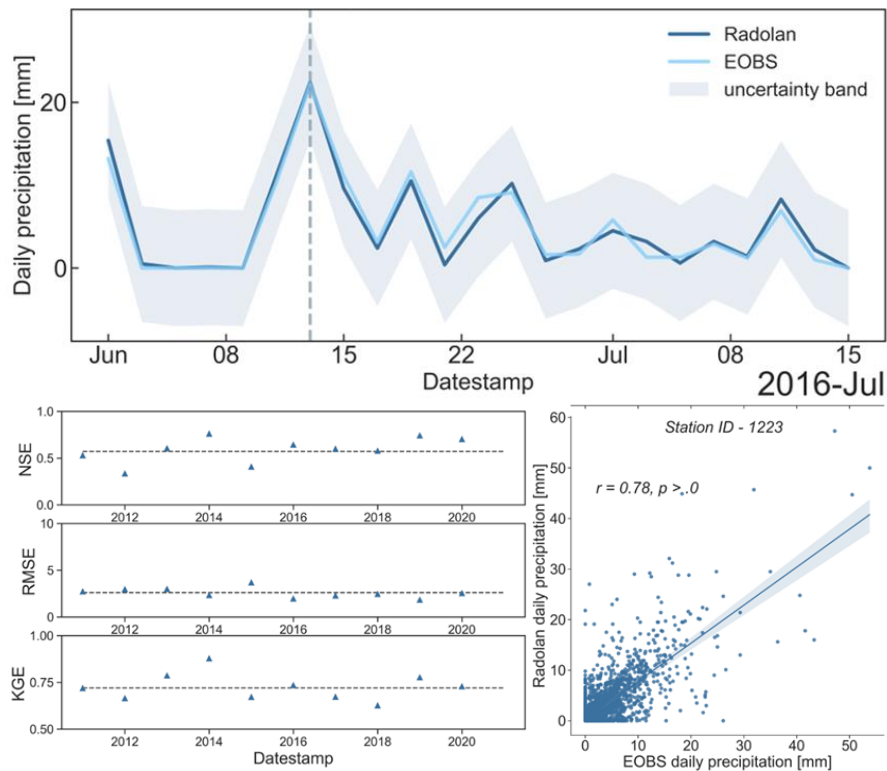
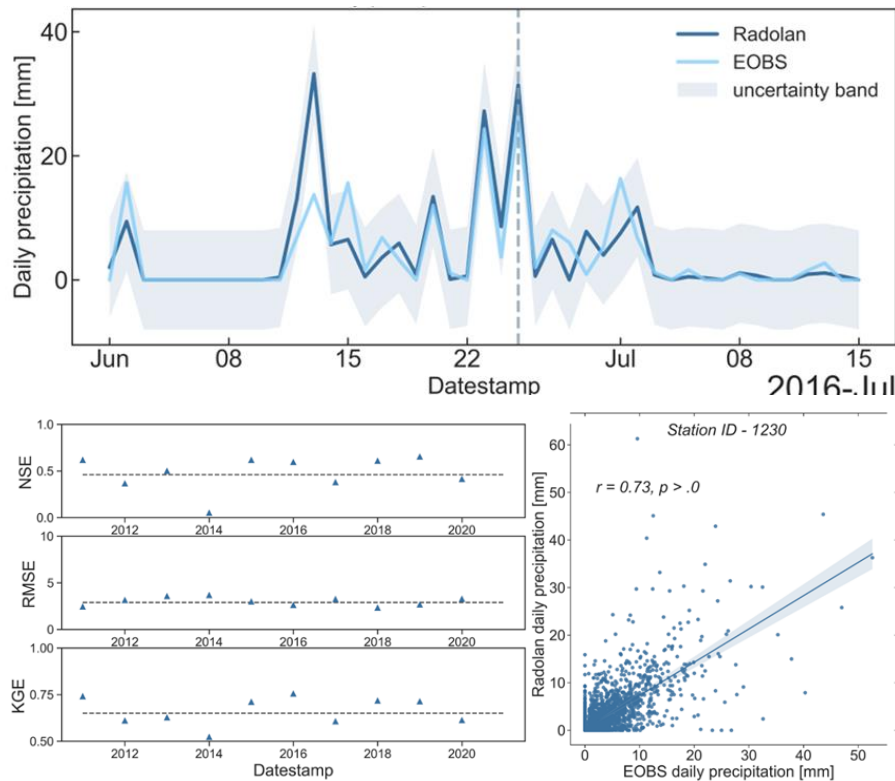




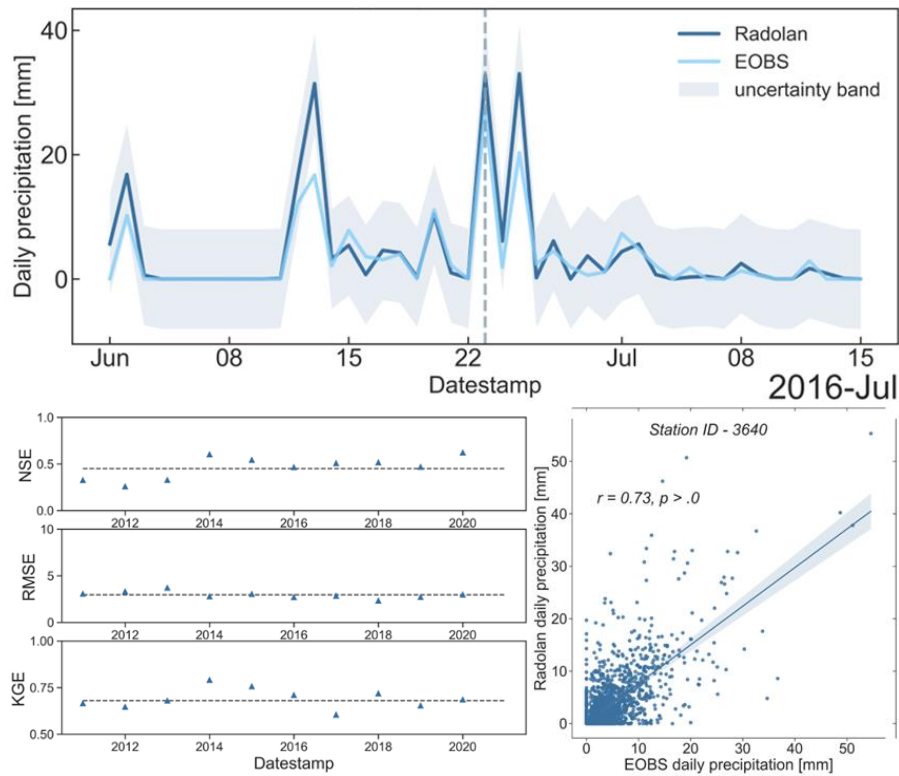
B.2 Radolan vs EOBS

Station ID 357

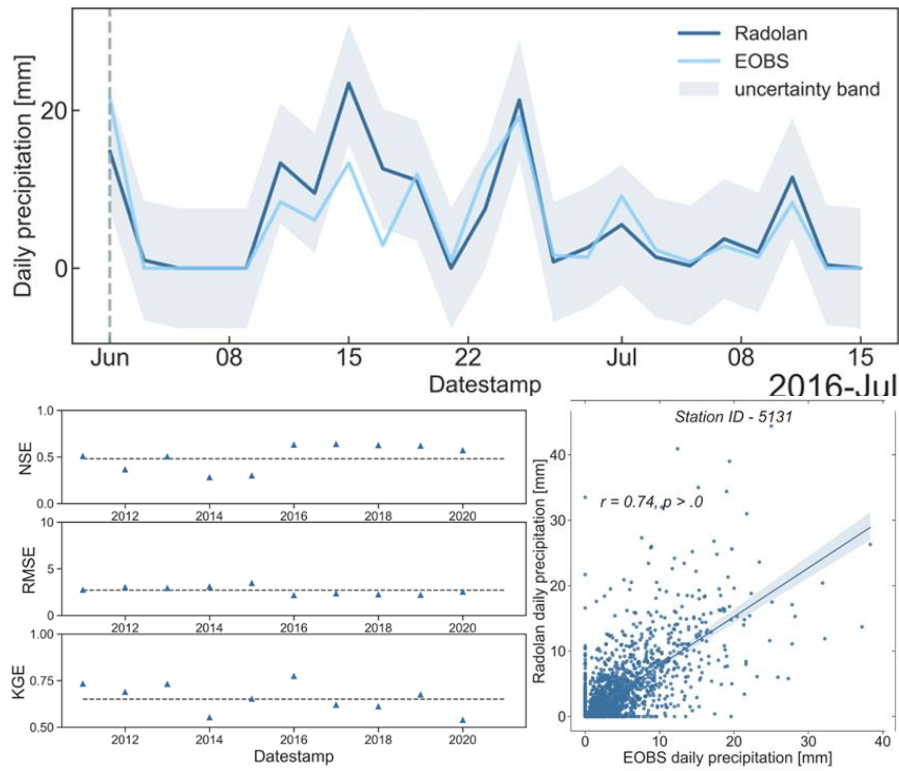


Station ID 1223**Station ID 1230**

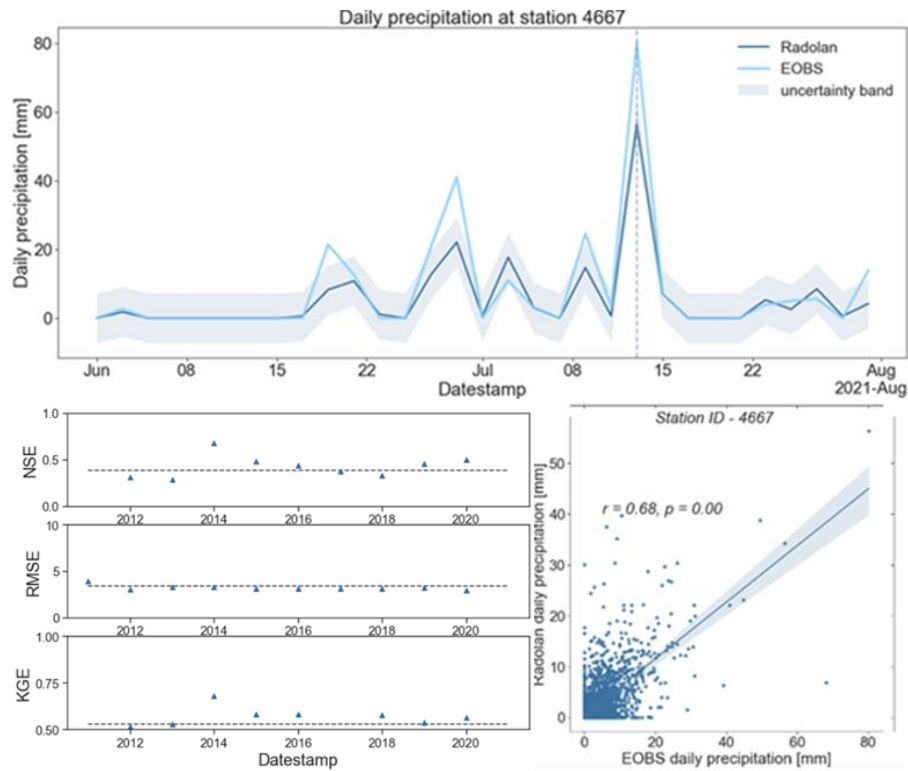
Station ID 3640



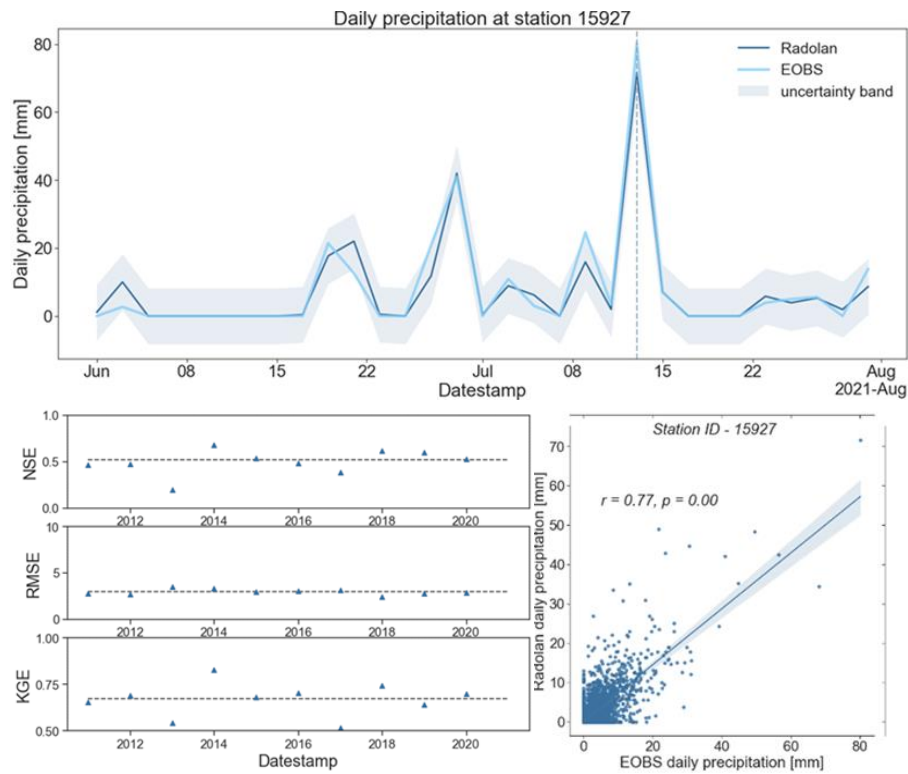
Station ID 5131



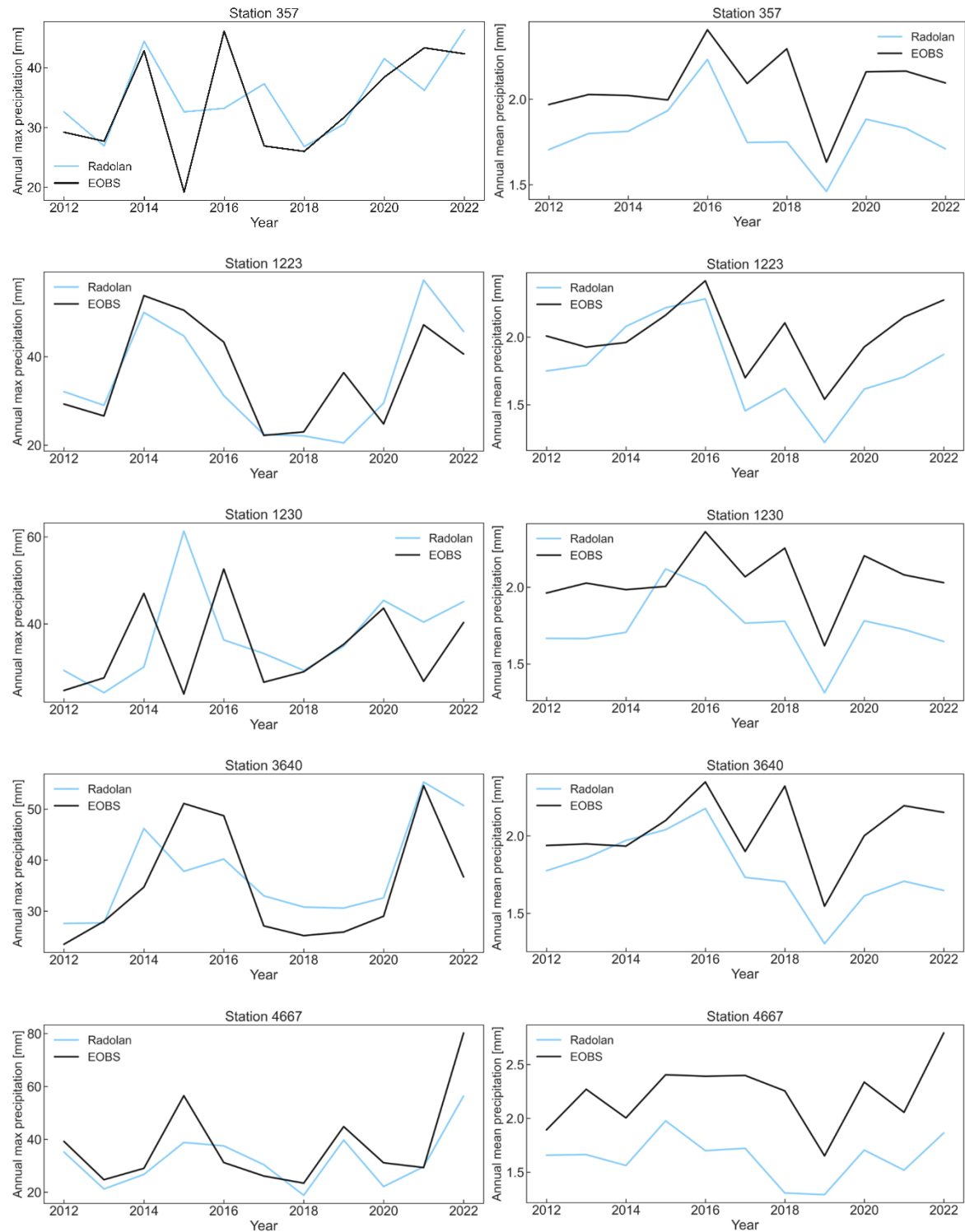
Station ID 4667

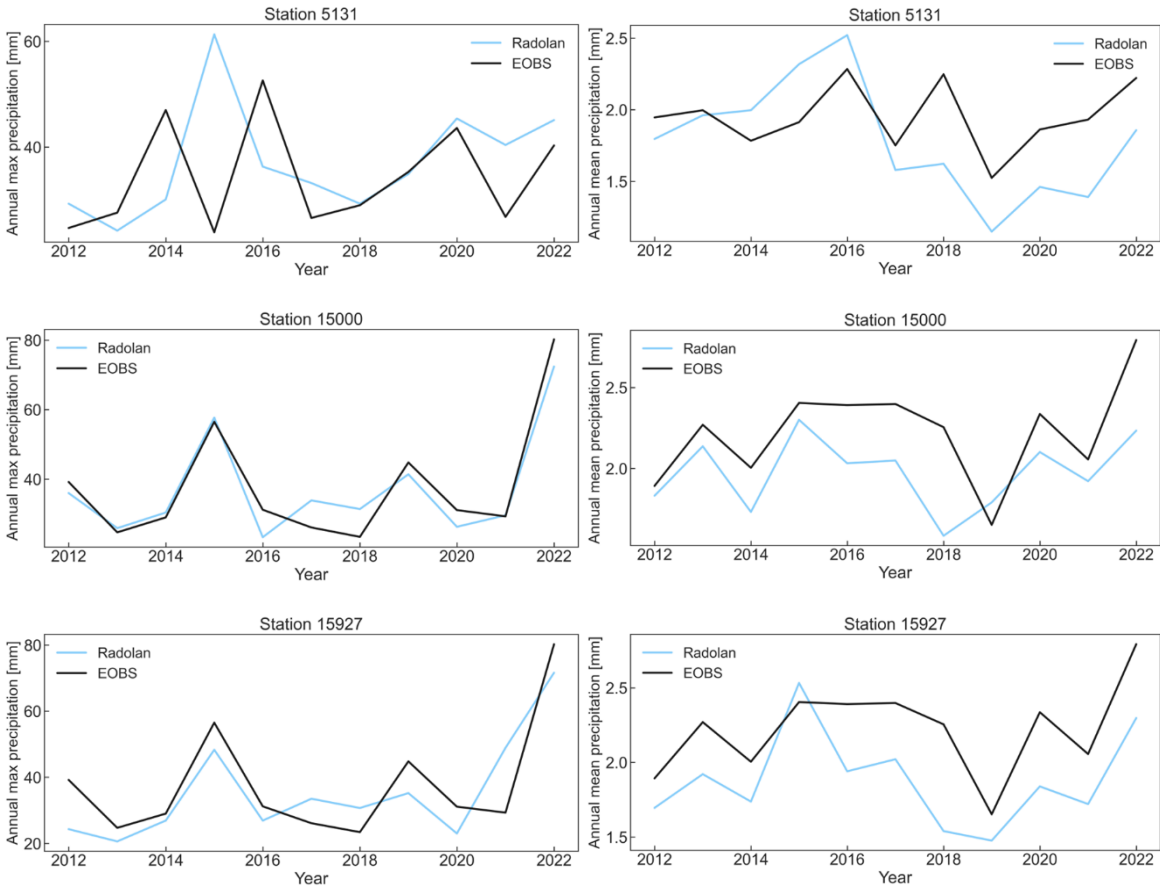


Station ID 15927



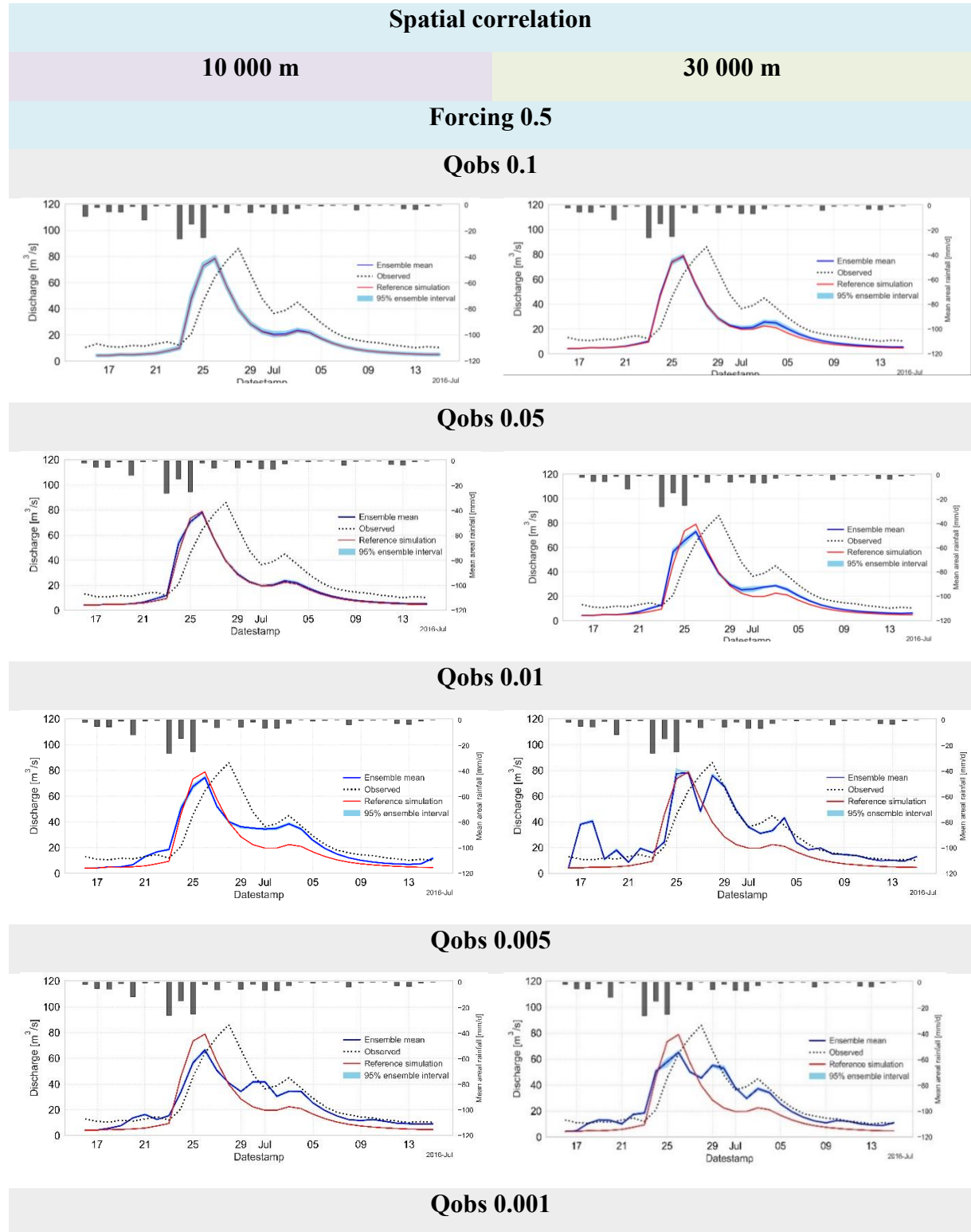
B.3 Annual max and mean precipitation

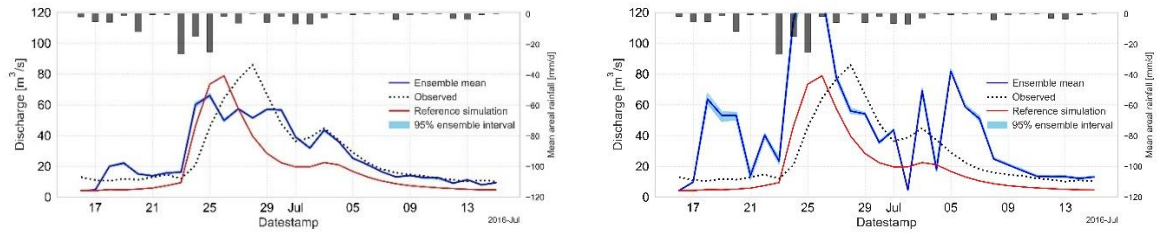




Appendix C. Simulations with DA

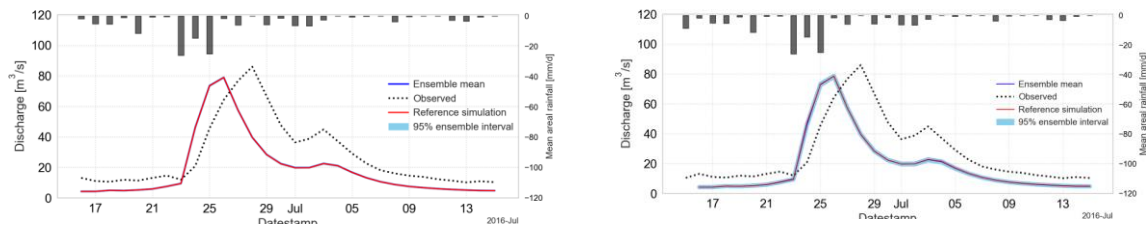
C.1 Sensitivity analysis of the simulations with DA, for a run with 8 ensemble members, with different values of uncertainties for forcing and observations



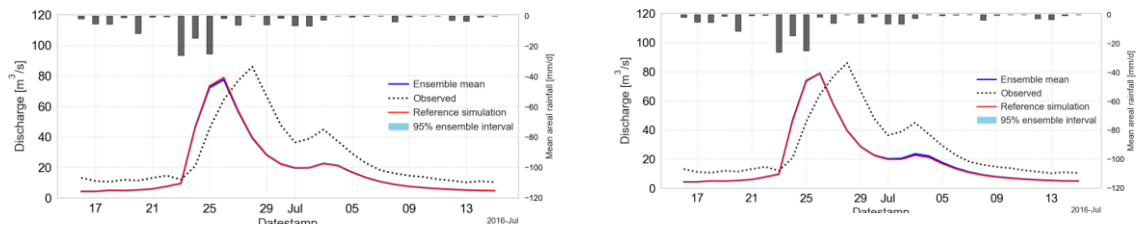


Forcing 0.1

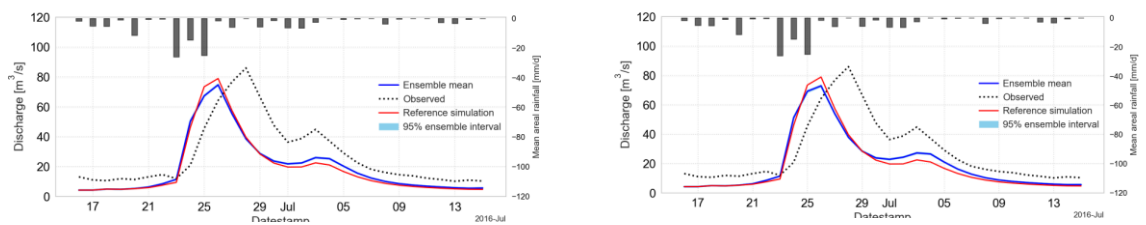
Qobs 0.1



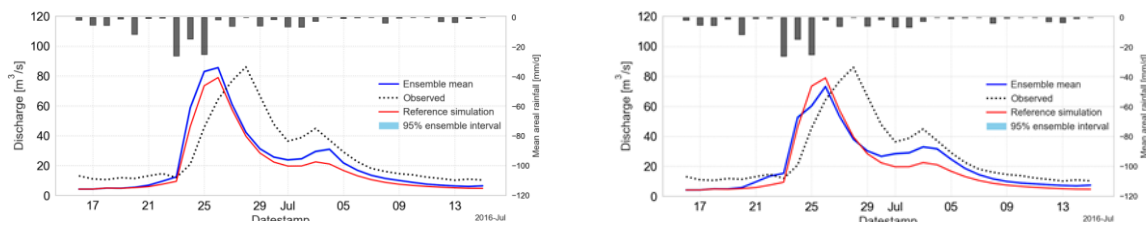
Qobs 0.05



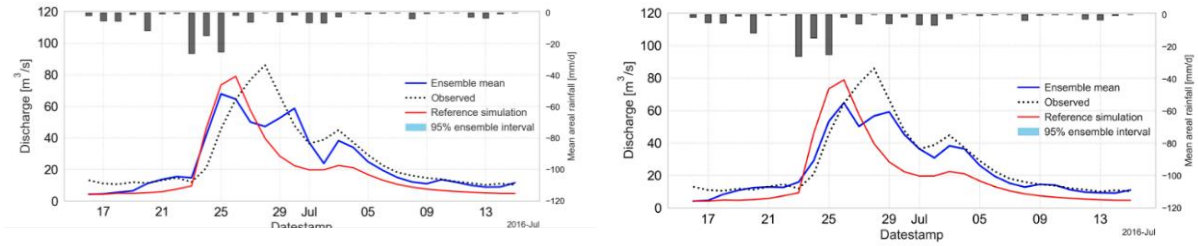
Qobs 0.01



Qobs 0.005

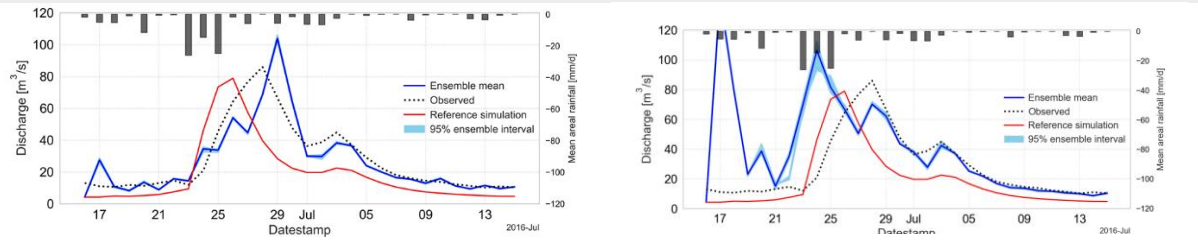


Qobs 0.001

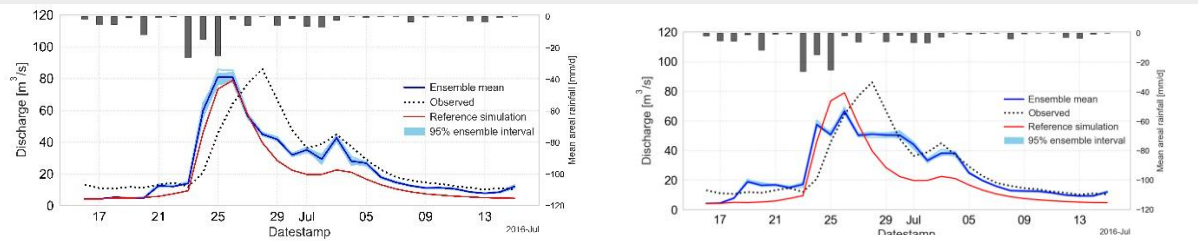


Forcing 1

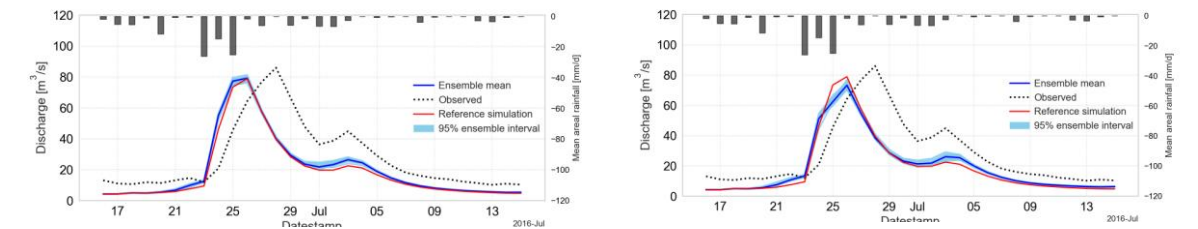
Qobs 0.001



Qobs 0.01

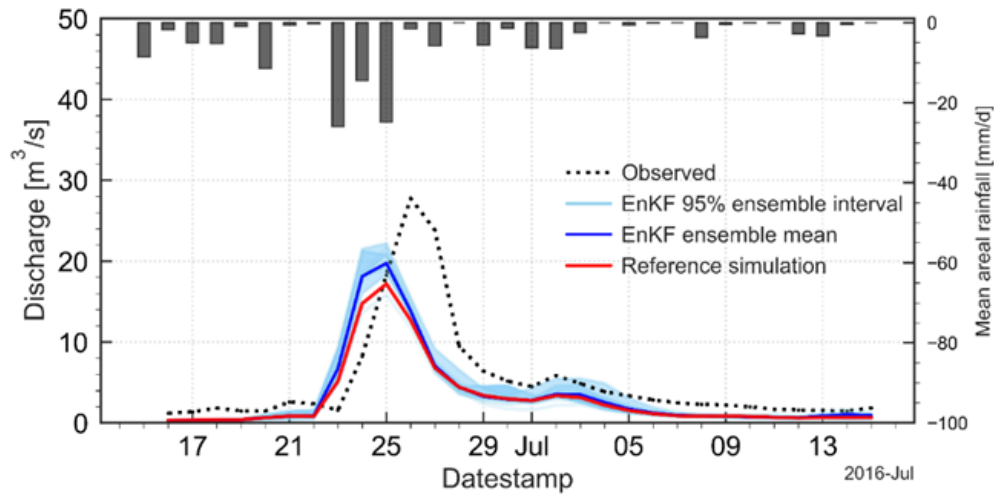


Qobs 0.1

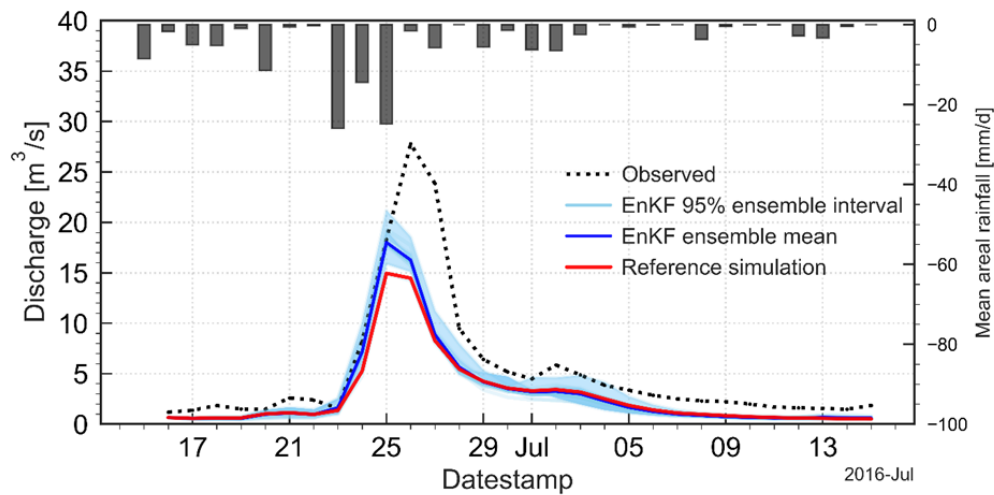


C.2 Discharge prediction results with DA in different station locations – experiment 1

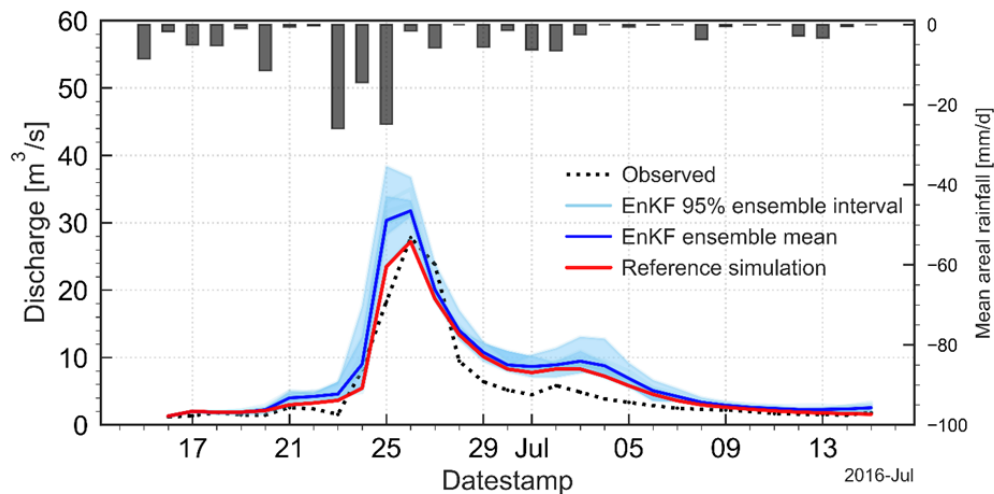
River station – Bilk



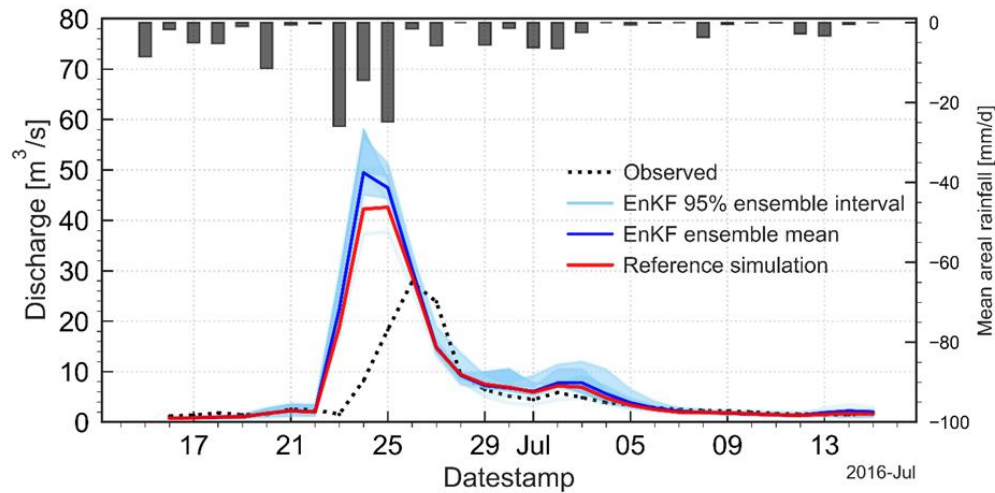
River station – Gronau



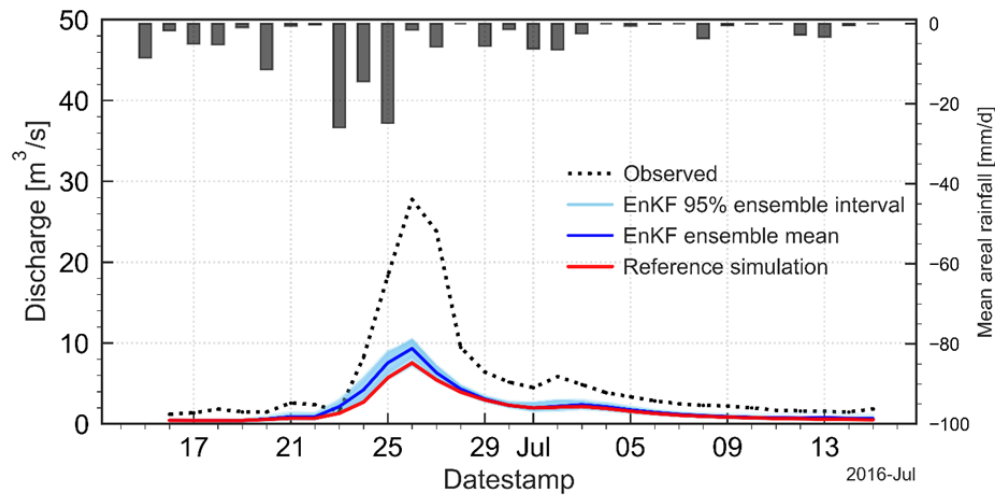
River station – Lagel



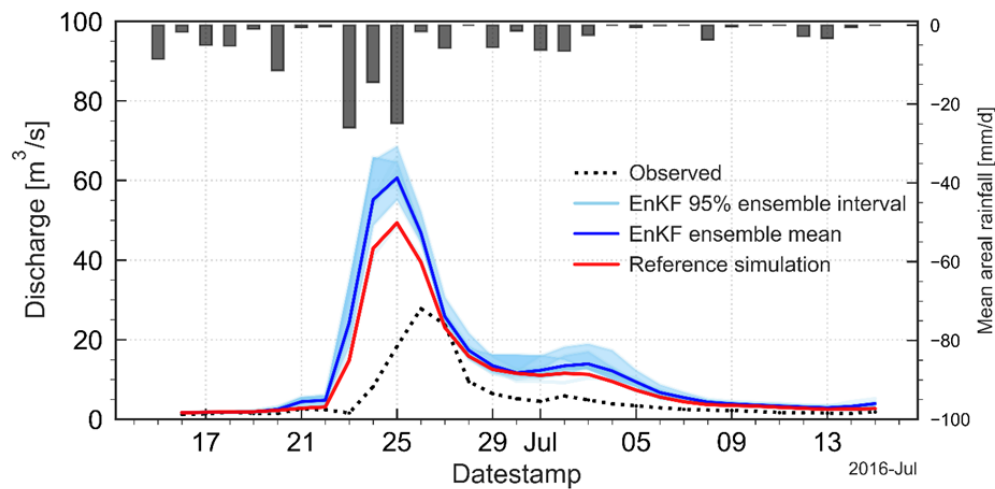
River station – Ohne



River station – Osterwald

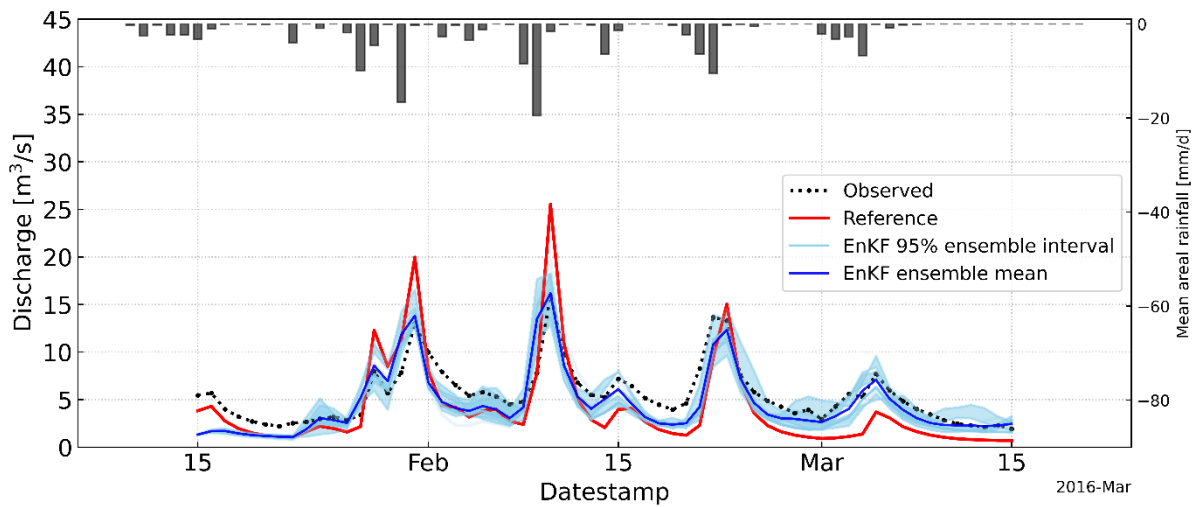


River Station – Wehr Neuenhaus

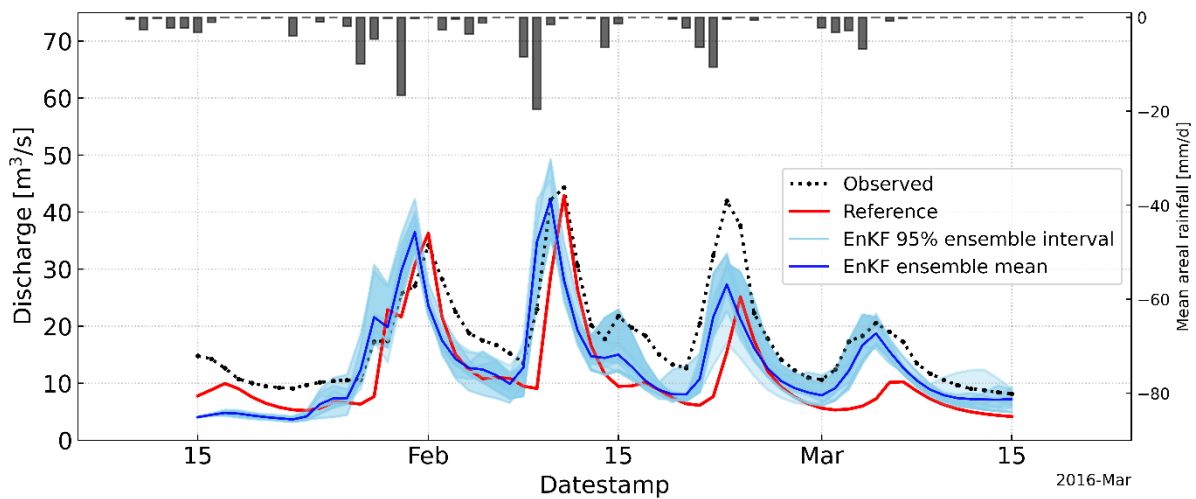


C.3 Discharge prediction results with DA in different station locations – experiment 2

River station – Bilk



River station – Lage1



River station – Ohne

