Internal processes in hydrological models

A glance at the Meuse basin from space

Bouaziz, L.J.E.

DOI
10.4233/uuid:09d84cc1-27e2-4327-a8c7-207a75952061

Publication date
2021

Document Version
Final published version

Citation (APA)
Bouaziz, L. J. E. (2021). *Internal processes in hydrological models: A glance at the Meuse basin from space*. https://doi.org/10.4233/uuid:09d84cc1-27e2-4327-a8c7-207a75952061

Internal processes in hydrological models

A glance at the Meuse basin from space

Bouaziz, L.J.E.

DOI
10.4233/uuid:09d84cc1-27e2-4327-a8c7-207a75952061

Publication date
2021

Document Version
Final published version

Citation (APA)
Bouaziz, L. J. E. (2021). *Internal processes in hydrological models: A glance at the Meuse basin from space*. https://doi.org/10.4233/uuid:09d84cc1-27e2-4327-a8c7-207a75952061

Internal processes in hydrological models

A glance at the Meuse basin from space

Laurène Bouaziz

INTERNAL PROCESSES IN HYDROLOGICAL MODELS

A GLANCE AT THE MEUSE BASIN FROM SPACE

INTERNAL PROCESSES IN HYDROLOGICAL MODELS

A GLANCE AT THE MEUSE BASIN FROM SPACE

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus Prof.dr.ir. T.H.J.J. van der Hagen;
chair of the Board for Doctorates
to be defended publicly on
Friday 17 September 2021 at 12:30 o'clock

by

Laurène Judith Estelle BOUAZIZ

Master of Science in Civil Engineering,
Delft University of Technology, the Netherlands
born in Les Lilas, France.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Dr. M. Hrachowitz,	Delft University of Technology, promotor
Prof.dr.ir. H.H.G. Savenije,	Delft University of Technology, promotor
Independent members:	
Prof.dr.ir. R. Uijlenhoet,	Delft University of Technology
Prof.dr. V. Andréassian,	INRAE, France
Dr.ir. M.-C. ten Veldhuis,	Delft University of Technology
Dr. B. Arheimer,	SMHI, Sweden
Prof.dr.ir. S. Steele-Dunne,	Delft University of Technology, reserve member
Other member:	
Prof.dr. A.H. Weerts,	Deltares, the Netherlands

Prof.dr. A.H. Weerts has contributed to the preparation of this dissertation.

Keywords: hydrological modeling, Meuse basin, root-zone storage capacity, remote sensing, states and fluxes, intercatchment groundwater flows

Printed by: Ipskamp Printing

Front & Back: Laurène Bouaziz

ISBN 978-94-6421-419-2

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

On ne voit bien qu'avec le cœur. L'essentiel est invisible pour les yeux.

Antoine de Saint-Exupéry

PREFACE

At the time of writing, traveling has become difficult in a world where countries are locking down. While longing for better times ahead to come, I can at least look back at a long and hectic journey over the last four years. It was quite an experience that I would like to share. At the start, I imagined to be sitting on a small raft on a river, when a thick mist came down, obscuring my view of the surroundings. Not knowing where to go, I slowly found my way and made it to the shore. I was drawn towards the forest, as trees loomed out of the mist. The wind was blowing through the canopy when I entered the maze of paths. I was captivated by the complexity, ingenuity and diversity of this world I had joined, with all its hidden treasures rooted under the surface. The density of the forest suddenly vanished and a wide open landscape stretched out ahead. And guided by the lights, I wandered in this beautiful, mysterious and hilly scenery. After a steep ascent, somewhere along my unsettled way, I sat back and enjoyed the view. I hope you will too, it is laid out in front of you.

Laurène Judith Estelle Bouaziz
Delft, January 2021

SUMMARY

Contemplating the Meuse or any other river of the world, one may wonder about the journey of rain in becoming river. This fascinates hydrologists, as they develop theories to understand movement, storage and release of water through the landscape across climates. These theories are translated to hydrological models, which describe the complex reality in a simpler way. Models are then used to predict the hydrological cycle for the nearby or long-term future.

This thesis aims to assist the Dutch Ministry of Infrastructure and Water Management in improving the reliability of hydrological modeling of the Meuse basin for operational and policy applications. Using in-situ and remote-sensing data, the value of representing additional processes in models is explored, as well as the creative use of additional data to improve hydrological predictions.

First, water balance data is used to identify the potential presence of intercatchment groundwater flows (Chapter 3). These underground flow paths cross topographic catchment boundaries and mainly play a role in headwater catchments (< 500 km^2) of the Meuse basin, which are underlain by productive aquifers. Representing this flux as a preferential threshold-initiated process improves low and high flow model performance and increases the consistency between modeled and remote-sensing estimates of actual evaporation.

Besides the importance of quantifying the long-term hydrological partitioning of precipitation into streamflow, evaporation and potentially intercatchment groundwater flows, another key element of the hydrological response is the amount of water available in the root-zone of vegetation. The temporal dynamics of root-zone soil moisture control how much more water can be stored in the soil and how much water is available for transpiration. In Chapter 4, meaningful estimates of root-zone soil moisture are inferred from satellite observations of near-surface soil moisture, by establishing a link between the catchment-scale root-zone storage capacity and the Soil Water Index.

Interestingly, hydrological models with different internal process representations of root-zone soil moisture, evaporation, snow and total storage at the catchment scale may lead to a similar aggregated streamflow response (Chapter 5). This discrepancy implies that models are not necessarily providing the right answers for the right reasons, as they cannot simultaneously be close to reality and different from each other. To circumvent the uncertainty of process representation, which is inherent to hydrological science, the use of multiple model structures is advocated for operational and policy applications. Nonetheless, testing the consistency between modeled hydrological behavior and independent remote-sensing data can foster model developments and lead to creating better models.

Finally, we move beyond the use of historical in-situ and remote-sensing data to predict long-term hydrological behavior of the Meuse basin under projected global warming (Chapter 6). If environmental conditions change, it is likely to also assume ecosystem

adaptation in response to climate change and a potential natural and/or anthropogenic shift in dominant species across the landscape. Non-stationarity in the representation of hydrological systems is introduced in a process-based model with three hydrological response units to account for the spatial variability of hydrological processes. More specifically, we adapt the root-zone storage capacity parameter using the information contained in the projected climate data. This is an important step forward in the great challenge of hydrological predictions under change.

Despite data uncertainties and a lack of data at the required temporal and spatial resolutions, many possibilities are at hand with what is currently available to develop new theories, test and improve hydrological models. Requiring creativity, this is a beautiful challenge to further unravel the mysteries of the hydrological landscape.

SAMENVATTING

Het bewonderen van de Maas of een andere rivier ter wereld wekt nieuwsgierigheid over de afgelegde reis van regen tot het ontstaan van een rivier. Dit fascineert hydrologen, die zich bezighouden met het ontwikkelen van theorieën om berging, verdamping en stroming van water door het landschap te begrijpen. Deze theorieën worden vertaald naar hydrologische modellen, die de complexe werkelijkheid eenvoudiger beschrijven. Vervolgens worden deze modellen gebruikt om de hydrologische cyclus voor de nabije of lange termijn toekomst te voorspellen.

Dit proefschrift heeft als doel om Rijkswaterstaat te ondersteunen bij het verbeteren van de betrouwbaarheid van hydrologische modellen van het Maasstroomgebied voor operationele- en beleidstoepassingen. Met behulp van in-situ en satellietgegevens is er onderzoek gedaan naar het belang van het beschrijven van aanvullende processen en het creatief gebruiken van gegevens om hydrologische voorspellingen te verbeteren.

Ten eerste zijn waterbalansgegevens gebruikt om de mogelijke aanwezigheid van grondwateruitwisselingen tussen stroomgebieden te identificeren (Hoofdstuk 3). Deze ondergrondse stromingen doorkruisen stroomgebiedsgrenzen die zijn afgeleid uit de topografie. De grondwateruitwisselingen spelen vooral een rol in bovenstroomse gebieden van het Maasstroomgebied, gelegen op productieve watervoerende lagen. Het representeren van deze flux als een preferentieel proces, geïnitieerd boven een drempelwaarde, verbetert zowel laag- als hoogwater modelprestaties en verhoogt daarnaast ook de consistentie tussen gemodelleerde en satellietschattingen van werkelijke verdamping.

Naast het belang van het kwantificeren van de lange-termijn hydrologische verdeling van neerslag naar verdamping, afvoer en mogelijke grondwateruitwisselingen, is de hoeveelheid water die beschikbaar is in de wortelzone van vegetatie een ander belangrijk aspect van de hydrologische respons. De temporele dynamiek van bodemvocht in de wortelzone bepaalt hoeveel meer water er in de bodem kan worden opgeslagen en hoeveel water er beschikbaar is voor transpiratie van vegetatie. In Hoofdstuk 4 worden degelijke schattingen van bodemvocht in de wortelzone afgeleid uit satellietwaarnemingen van bodemvocht in de bovenste paar centimeters van het aardoppervlak, door een verband te leggen tussen de wortelzonecapaciteit op stroomgebiedschaal en de *Soil Water Index*.

Interessant is dat hydrologische modellen met verschillende representaties van interne processen, zoals bodemvocht in de wortelzone, verdamping, sneeuw en totale berging, kunnen leiden tot een vergelijkbare afvoerrespons (Hoofdstuk 5). Deze discrepantie houdt in dat modellen mogelijk niet de juiste antwoorden geven om de juiste redenen, aangezien ze niet tegelijkertijd dichtbij de werkelijkheid kunnen staan en van elkaar kunnen verschillen. De inherente onzekerheid van procesrepresentatie in de hydrologie kan expliciet worden meegenomen door meerdere modelstructuren te gebruiken ter ondersteuning van operationele- en beleidstoepassingen. Daarnaast kan het ontwikkelen

van betere modelstructuren ook worden bevorderd door een evaluatie van de consistentie tussen gemodelleerd hydrologisch gedrag en onafhankelijke satellietgegevens.

Tenslotte worden niet alleen historische in-situ- en satellietgegevens gebruikt om het hydrologische gedrag van het Maasstroomgebied te begrijpen en te voorspellen, maar daarnaast ook een lange termijn projectie van klimaatverandering als gevolg van de opwarming van de aarde (Hoofdstuk 6). Als het klimaat verandert, is het aannemelijk dat ecosystemen zich zullen aanpassen aan de nieuwe omstandigheden. Daarnaast kan er een natuurlijke en/of antropogene verschuiving plaatsvinden van dominante soorten in het landschap. We introduceren verandering in de beschrijving van het hydrologische systeem in een model met drie hydrologische responseenheden om rekening te houden met de ruimtelijke variabiliteit van hydrologische processen. In het model wordt de parameter van de wortelzone capaciteit aangepast aan de nieuwe klimaatgegevens. Dit is een belangrijke stap voorwaarts in de geweldige uitdaging van hydrologische voorspellingen onder veranderende omstandigheden.

Ondanks onzekerheden in bestaande data en het gebrek aan meer gegevens op de gewenste temporele en ruimtelijke resolutie, zijn er ontelbare mogelijkheden om met beschikbare gegevens nieuwe theorieën te ontwikkelen, hydrologische modellen te toetsen en te verbeteren. Met de vereiste creativiteit, is dit een mooie uitdaging om de mysteries van het hydrologisch landschap verder te ontraadselen.

NOMENCLATURE

LIST OF ABBREVIATIONS

AMSR2	Advances Microwave Scanning Radiometer 2
CLC	CORINE Land Cover
CSR	U. Texas / Center for Space Research
DMIP	Distributed Model Intercomparison Project
dS2	distributed simple dynamical systems
ECAD	European Climate Assessment Dataset
ECMWF	European Center for Medium-Range Weather Forecasts
FDC	Flow Duration Curve
FUSE	Framework for Understanding Structural Errors
GFZ	GeoForschungsZentrum Potsdam
GLEAM	Global Land Evaporation Amsterdam Model
GR4H	Génie Rural à 4 paramètres Horaire
GRACE	Gravity Recovery and Climate Experiment
HAND	Height Above the Nearest Drainage
HBV	Hydrologiska Byråns Vattenbalansavdelning
HRU	Hydrological Response Unit
IGF	Intercatchment Groundwater Flows
IGME	International Geological Map of Europe
IHME	International Hydrogeological Map of Europe
JPL	Jet Propulsion Laboratory
KGE	Kling-Gupta Efficiency
KNMI	Royal Netherlands Meteorological Institute
LPRM	Land Parameter Retrieval Model
LSASAF	Satellite Application Facility on Land Surface Analysis
MODIS	Moderate Resolution Imaging Spectroradiometer
MSG	Meteosat Second Generation
NAM	NedborAfstrommings Model
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NSE	Nash-Sutcliffe Efficiency
PRESAGES	PREvision et Simulation pour l'Annonce et la Gestion des Etiages Sévères
RC	Runoff coefficient
RCM	Regional Climate Model
RFI	Radio Frequency Interference
SAR	Synthetic Aperture Radar
SMAP	Soil Moisture Active Passive
SPW	Service Public de Wallonie

SSM	Surface Soil Moisture
SWI	Soil Water Index
VHM	Veralgemeend conceptueel Hydrologisch Model
WALRUS	Wageningen Lowland Runoff Simulator

LIST OF SYMBOLS

a, b, S_0	Shape parameters of the streamflow sensitivity of the dS2 model [-]
a_S	Fraction of land surface covered by surface water for WALRUS [-]
a_G	Fraction of land surface not covered by surface water for WALRUS [-]
C	Pseudo-diffusivity coefficient for the calculation of the SWI [L T^{-1}]
C_{cst}	Constant water stress coefficient to estimate E_R of the dS2 model [-]
C_E	Correction factor for E_P (Chapter 5) [-]
C_{IGF}	Constant net intercatchment groundwater flow (IGF$_{net}$) [L T^{-1}]
D	Fraction of preferential recharge from root-zone to slow storage [-]
D_{IGF}	Fraction to IGF$_{net}$ storage [-]
E_A	Total actual evaporation [L T^{-1}]
E_A/P	Evaporative index [-]
E_I	Interception evaporation [L T^{-1}]
E_P	Potential Evaporation [L T^{-1}]
E_P/P	Aridity index [-]
E_R	Evaporation from the root-zone storage [L T^{-1}]
E_W	Sublimation of snow [L T^{-1}]
$E_{NS,Q}$	Nash-Sutcliffe efficiency of the streamflow [-]
$E_{NS,logQ}$	Nash-Sutcliffe efficiency of the logarithm of the streamflow [-]
$E_{NS,FDC,logQ}$	Nash-Sutcliffe efficiency of the FDC of the logarithm of the streamflow [-]
$E_{NS,RC,m}$	Nash-Sutcliffe efficiency of the monthly runoff coefficient [-]
$E_{NS,RC,s}$	Nash-Sutcliffe efficiency of the seasonal runoff coefficient [-]
$E_{NS,RC,w}$	Nash-Sutcliffe efficiency of the weekly runoff coefficient [-]
$E_{KG,Q}$	Kling-Gupta efficiency of the streamflow [-]
F_M	Degree day factor [mm d^{-1} °C^{-1}]
I_b	Baseflow index (Chapter 5) [-]
I_f	Flashiness index (Chapter 4) [-]
I_{max}	Maximum interception capacity [L]
K_F	Characteristic time scale of the fast recession [T]
K_{IGF}	Characteristic time scale of the IGF$_{net}$ [T]
K_{River}	Characteristic time scale of the river flow (overflow model) [T]
K_S	Characteristic time scale of the slow recession [T]
K_n	Gain coefficient of the SWI at time T_n [-]
L	Depth of the reservoir for the calculation of the SWI [L]
L_P	Reduction parameter for potential evaporation [-]
n	Parameter of the Turc-Mezentsev equation [-]
P	Precipitation [L T^{-1}]
P_E	Effective precipitation [L T^{-1}]
P_{erc}	Fraction of the recharge to IGF$_{net}$ [-]
P_M	Snow melt [L T^{-1}]

P_{max}	Maximum percolation rate [L T^{-1}]
P_R	Precipitation entering the root-zone (Chapter 5) [L T^{-1}]
P_R	Rainfall (Chapter 6) [L T^{-1}]
P_S	Snowfall (Chapter 6) [L T^{-1}]
Q	Streamflow [L T^{-1}]
Q/P	Runoff ratio [-]
Q_C	Capillary rise from slow to root-zone storage (Chapter 5) [L T^{-1}]
Q_F	Fast runoff [L T^{-1}]
Q_G	Seepage (up/down) / extraction (Chapter 5) [L T^{-1}]
Q_{IGF}	Net intercatchment groundwater flows [L T^{-1}]
Q_{obs}	Observed streamflow [L T^{-1}]
Q_P	Percolation from root-zone to slow runoff storage (Chapter 5) [L T^{-1}]
Q_R	Flux from root-zone to fast and/or slow runoff storage (Chapter 5) [L T^{-1}]
Q_{River}	Streamflow which ends up in the river [L T^{-1}]
Q_S	Slow runoff [L T^{-1}]
R_C	Capillary rise (Chapter 6) [L T^{-1}]
R_P	Percolation [L T^{-1}]
R_R	Outflow out of root-zone storage [L T^{-1}]
R_{RF}	Recharge to the fast storage [L T^{-1}]
R_{RS}	Recharge to the slow storage [L T^{-1}]
S_D	Storage deficit (Chapter 5) [L]
S_I	Interception storage [L]
S_F	Fast runoff storage [L]
$S_{R,def}$	Annual maximum cumulative deficit in the root-zone storage [L]
$S_{R,max}$	Root-zone storage capacity [L]
S_R	Root-zone storage [L]
$\overline{S}_R$	Relative root-zone storage ($S_R/S_{R,max}$) [-]
S_S	Slow runoff storage [L]
$S_{S,max}$	Maximum capacity of underground storage [L]
S_{SW}	Surface water storage (Chapter 5) [L]
S_T	Total storage [L]
S_{thresh}	Threshold of root-zone storage above which $E_R = E_P$ (Chapter 5) [L]
S_{VQ}	Very quick runoff storage (Chapter 5) [L]
S_W	Snow storage [L]
T_F	Time lag [T]
T	Characteristic time length of the SWI [T]
T_M	Threshold temperature for snow melt [°C]
T_{opt}	Optimal characteristic time length of the SWI [T]
T_T	Threshold temperature to partition precipitation to snow and rain [°C]
W	Moisture content of the lower reservoir of SWI [m^3 m^{-3}]
W_s	Moisture content of surface reservoir of SWI [m^3 m^{-3}]
α	Non-linear coefficient of the fast storage [-]
β	Shape parameter of storage capacity distribution [-]
μ	Threshold of the recharge above which IGF_{net} occurs [L T^{-1}]
ω	Parameter of the Fu equation [-]

CONTENTS

1

INTRODUCTION

Als je fietst met tegenwind, dan merk je dat voortdurend. Als de wind draait en je helpt, merk je dat alleen de eerste minuut. Daarna voelt het normaal.

Ben Tiggelaar - *Wie succesvol wil zijn moet vooral een beetje geluk hebben* (2019)

"*The Meuse is a lady - the French say so, and they should know.*" This is how Bernard Newman begins his book *The lazy Meuse* on his cycling journey along *la* Meuse in 1948, to tell about ancient and modern history lived along her banks. Exactly 70 years later and about a year after starting my PhD journey on the hydrology of the Meuse, I also embarked on a cycling trip along the Meuse. And I couldn't agree more with Bernard Newman, when he says: "*She is, as I hope to show, a fascinating river*".

1.1 RELIABLE HYDROLOGICAL PREDICTIONS FOR THE MEUSE

The Meuse river originates in North-West France and follows its course through Belgium to the Netherlands. The Meuse is a rain-fed river with a large seasonal streamflow variability. People living in the areas around Brussels, Antwerp and Rotterdam drink water withdrawn from the Meuse. Besides providing recreational and ecological functions, it also provides water for industry, agriculture, navigation and electricity generation. Limited water availability during low flows increases the pressure experienced by the users, as recently occurred during the 2018 and 2019 summers. On the other hand, periods of high flow may cause floods with large societal impact as happened in 1993, 1999, and more recently in 2011. Projected climate change may further exacerbate the pressure by increasing extremes, in particular, heat waves, drought and heavy precipitation events (Kovats et al., 2014). Therefore, it is of critical importance to make reliable predictions of high and low flows within the Meuse river basin for operational and policy applications.

The Dutch Ministry of Infrastructure and Water Management currently uses a hydrological model (HBV-96) to predict flows at Monsin, where the Meuse enters the Netherlands. The model functions adequately under normal conditions, but does not perform at the required level under exceptional circumstances and does not optimally make use of the available data. For example, in 2011 widespread flooding caused by rain on snow was not accurately predicted. The model also shows shortcomings during low flow periods or transition periods from low to high flows.

To overcome these limitations, understanding the hydrological processes in the Meuse river basin has sparked the interests of many researchers. Berger (1992) conducted a system analysis of each tributary of the Meuse to set-up models for flood forecasting. An overall description of the physical properties of the Meuse river basin is provided by de Wit (2008a). He describes four main lithological zones to characterize the spatial variability of hydrological properties in the basin. The gentle hills in the French part are underlain by consolidated sedimentary rocks, while the steep Belgian Ardennes are characterized by relatively impermeable metamorphic rocks. Deep groundwater systems in porous chalk layers are found on the Western side of the Meuse in Wallonia (Reggiani and Rientjes, 2010), in contrast to unconsolidated sedimentary rocks in the Northern lowlands. Land use mainly consists of forestry (35%), agriculture (32%), pasture (21%) and urban areas (9%). Since the 18th century, large areas of broadleaved forests have been converted to agricultural land and coniferous plantations in the Walloon region, while these forests largely remained in the French part of the basin (Tu, 2006). Land-use management and climate variability over the last century had a significant impact on the hydrological behavior of the Meuse river basin (de Wit et al., 2001; Booij, 2005; Ashagrie

et al., 2006; Fenicia et al., 2009; Tu, 2006).

More recently, de Boer-Euser (2017) developed a flexible model structure based on topographic and land-use catchment characteristics, which was extensively tested in the catchment of the Ourthe, an important tributary of the Meuse. The study shows the added value of spatially distributing forcing and hydrological processes to improve model performance. Important processes were identified and included in the model structure, e.g. a reduced infiltration capacity when soils are frozen and a very quick runoff component from artificial drainage on agricultural land.

The current research aims to further identify dominant hydrological processes and understand the drivers of spatial variability at the scale of the Meuse river basin. In collaboration with Delft University of Technology, Deltares and the Dutch Ministry of Infrastructure and Water Management, this study contributes to increasing the reliability of low and high flow predictions.

1.2 HYDROLOGICAL MODELING

Hydrological modeling is a well-established field. A considerable amount of the knowledge we currently rely on has been developed more than 30 years ago (e.g. Turc, 1954; Budyko, 1961; Tóth, 1963; Klemeš, 1986; Beven, 1989; Beven and Binley, 1992; Milly, 1994; Dooge, 1997). The numerous studies in the past 30 years have not lead to a unified approach on hydrological modeling, but rather to competing modeling philosophies (Hrachowitz and Clark, 2017) and probably thousands of hydrological models.

Despite differences in model philosophies, models fulfill two main roles. The first is to provide a means for hypotheses testing of the representation of internal processes to understand catchment functioning. This is exploratory science (Beven, 2019a). The second is to provide predictions and simulations for future short- or long-term applications to support decision-making on the impact of changes into the future (Beven, 2019a). Ideally, knowledge gained through science is applied in practice.

In the plethora of available models, the selection of a suitable model is often driven by personal preference and experience instead of detailed model test procedures (Holländer et al., 2009; Clark et al., 2015; Addor and Melsen, 2019). Yet the choice of model structure and parameterization has considerable impacts on the outcome, making models less objective than they might seem at first sight (Melsen et al., 2019). In fact, models usually reflect the system understanding of their developer.

Models consist of simplified representations of the complex interactions of hydrological processes in a heterogeneous landscape. There are three main modeling philosophies. In the **bottom-up** philosophy, high-resolution descriptions of small-scale processes are numerically integrated to a larger scale (Hrachowitz and Clark, 2017). The **top-down** approach, on the other hand, describes emerging parsimonious patterns at the larger-scale from the small-scale natural heterogeneity of the system: "the whole is greater than the sum of the parts" (Aristotle 384-322 BC and Heraclitus 535-475 BC as cited by Savenije and Hrachowitz, 2017). **Empirical** methods derive governing relations from large catchment samples of data.

In the **bottom-up** approach, the so-called physically-based models usually discretize the study area in grid cells, connected through surface and lateral subsurface flow along the river network. The model represents several soil layers and their properties, where water can be stored before it leaves the system as evaporation or runoff. Recent advances in computational power allows modelers to run these models at hyper-resolution (Bierkens et al., 2015), with the associated challenge of identifying parameters for each grid cell. The use of pedo-transfer functions to relate soil and landscape characteristics to model parameter values has shown interesting results (Samaniego et al., 2010, 2017; Imhoff et al., 2020). However, these high-resolution parameter maps may suggest a high level of certainty despite the limited number of field measurements they rely on and the use of pedo-transfer functions derived at the lab-scale and applied to a heterogeneous landscape (Beven et al., 2015).

In contrast, process-based models (or so-called conceptual models) associated with the **top-down** approach are usually developed at the catchment or sub-catchment scale. They typically consist of storage, transmission and release of water to represent the dominant hydrological processes occurring in a specific catchment. These models accommodate features of spatial organization and evolution of the catchment that translate to parsimonious relations (Savenije and Hrachowitz, 2017). They rely on a limited number of effective parameters at the catchment scale, which require calibration but are linked to a physical understanding of catchment functioning. Key parameters such as the root-zone storage capacity can be robustly estimated using water balance data (Gao et al., 2014; de Boer-Euser et al., 2016; Wang-Erlandsson et al., 2016). The root-zone storage capacity may vary with time as catchments themselves are evolving/adapting as living organisms (Nijzink et al., 2016a; Savenije and Hrachowitz, 2017; Hrachowitz et al., 2020). The spatial variability of dominant hydrological processes, based on land-use or physical characteristics, can be implemented by delineating Hydrological Response Units (HRU) within the sub-catchment scale (Savenije, 2010; Fenicia et al., 2016; de Boer-Euser, 2017; Dal Molin et al., 2020). The definition of these HRU to improve the modeling process requires creativity, experience and skill (Savenije, 2009).

Other approaches include the **empirical** relations derived from a large sample of catchment data. The most noteworthy is likely the Budyko framework (Turc, 1954; Mezentsev, 1955; Budyko, 1961) which relates the dryness index (ratio of potential evaporation over precipitation) to the runoff coefficient (Fig. 1.1). Many studies aim to understand why catchments may deviate from the Budyko curve (e.g., Gentine et al., 2012; Li et al., 2014; Le Moine et al., 2007; Bouaziz et al., 2018). Others use the framework to predict the impact of climate or land-use changes on evaporation and streamflow (Berghuijs et al., 2017; Teuling et al., 2019) or to constrain and evaluate models (Hrachowitz et al., 2014; Nijzink et al., 2018).

The similarity of these three approaches is the quest to understand the causes of streamflow variability due to climate and catchment characteristics. This is the holy grail of hydrology, allowing us to design models able to represent hydrological processes across a variety of hydrological settings, at multiple spatiotemporal scales, and under changing environmental conditions (Gupta et al., 2014).

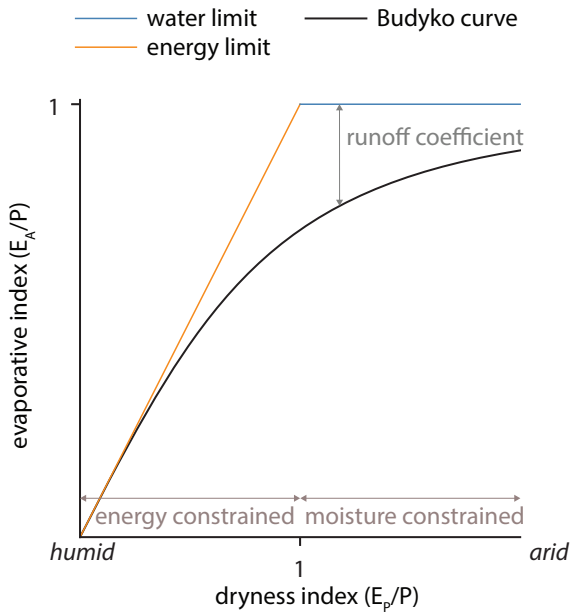

Figure 1.1: Long-term water balance components following the Budyko curve in the non-dimensional representation of the evaporative index as a function of the dryness index

1.3 THE WATER BALANCE

As complex as hydrological models may be, they all rely on conservation of mass. The change in water storage in a catchment ($\mathrm{d}S/\mathrm{d}t$) equals the difference between precipitation (P) and the sum of streamflow (Q_{River}) and actual evaporation (E_{A}), as represented in Fig. 1.2. Over the long term, the change in storage is assumed to be close to zero and the water balance reduces to $P = Q_{\mathrm{River}} + E_{\mathrm{A}}$. The long-term partitioning of precipitation into evaporation and drainage is controlled by the dryness index (defined as the ratio of potential evaporation over precipitation E_{P}/P), according to the Budyko curve (or variants thereof, Turc, 1954; Mezentsev, 1955; Budyko, 1961; Zhang et al., 2004). This is a powerful framework to get a first estimate of the long-term water balance in a catchment (Fig. 1.1).

These are fundamental concepts in hydrological science taught in freshman classes. Yet, this water balance formulation neglects the potential presence of regional groundwater flows across topographic catchment boundaries. Catchments are not closed entities underlain by impervious layers. Groundwater gains or losses through intercatchment groundwater flows are likely related to physical catchment properties including deep permeable layers, catchment size and the presence of complex geological features. Despite the presence of limestone and karstic features underlying significant parts of the Meuse basin, the presence of intercatchment groundwater flows has not been studied. Nevertheless, assuming this process to be negligible may introduce misrepresentation

of the natural system in hydrological models, also affecting our understanding of other processes.

One of these elements is the representation of the root-zone storage capacity ($S_{R,max}$) in hydrological models. It is a key element regulating the partitioning of water fluxes of terrestrial hydrological systems. The root-zone storage capacity represents the maximum amount of soil moisture accessible to the roots of vegetation for transpiration. There is increasing evidence that vegetation adapts to its environment in an efficient way, by developing a root system which ensures access to sufficient water to survive dry periods (Milly, 1994; Gentine et al., 2012). Using the long-term water balance as starting point, we can estimate transpiration as the main evaporation component after deducing interception and potential intercatchment groundwater flows. Next, we can derive annual water deficits experienced by vegetation from the seasonal water balance to estimate the root-zone storage capacity (de Boer-Euser et al., 2016; Nijzink et al., 2016a).

Intercatchment groundwater flows and root-zone storage capacities are two examples of processes that are very difficult to observe or measure directly in the field, especially at the scale typically required for modeling studies. Even if rooting profiles are available, they are sparse snapshots in time and space (Fan et al., 2017) and difficult to translate to root-zone storage capacity due to variations in horizontal and vertical root density and uncertainties in soil profiles (Wang-Erlandsson, 2017). To overcome these limitations, the creative use of in-situ and remote-sensing data is explored.

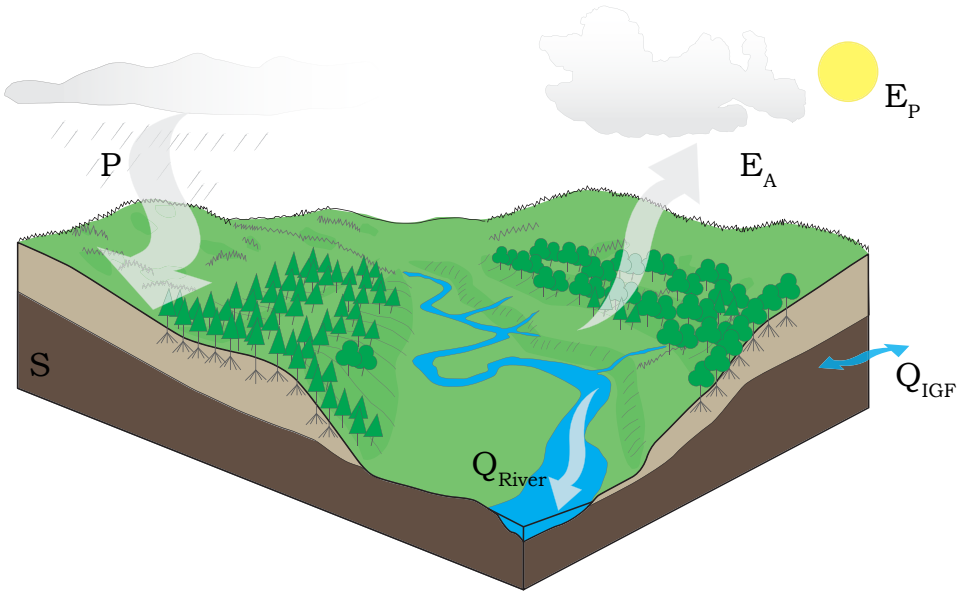

Figure 1.2: Main components of a catchment's water balance (with S for storage, P for precipitation, E_A for actual evaporation, E_P for potential evaporation, Q_{River} for streamflow and Q_{IGF} for intercatchment groundwater flows)

1.4 CREATIVE USE OF IN-SITU AND REMOTE-SENSING DATA

Hydrology is a challenging discipline because it belongs to the field of inexact sciences (Beven, 2019a). We lack knowledge on the subsurface through which water is flowing. We are unable to measure meteorological information at sufficient temporal and spatial scale. The data we use to evaluate models is uncertain and consists at most of several pieces of the puzzle. This implies that models have more degrees of freedom than the evaluation data available to constrain the model parameters, which leads to an ensemble of (more or less) acceptable solutions. This is the equifinality thesis (Beven, 2006).

Enriching our evaluation data through a creative use of in-situ and remote-sensing data in addition to streamflow is key to support or adapt the developed theories in hydrological models (Melsen, 2017). Remote-sensing datasets of e.g. snow, soil moisture, vegetation characteristics, total water storage and evaporation have been successfully used for calibration, evaluation or assimilation in hydrological models (e.g., Crow et al., 2005; Beck et al., 2009; Parajka et al., 2009; Sutanudjaja et al., 2014; Wanders et al., 2014; Silvestro et al., 2015; Leroux et al., 2016; López López et al., 2016, 2017; Rakovec et al., 2016a; Tian, 2007; Nijzink et al., 2018; Gevaert et al., 2018). These datasets are often praised for their spatial information content (even at coarse resolution), although their temporal resolution may be insufficient due to low revisit time, which can be overcome by combining data from multiple satellites.

A drawback of many of these datasets is that the raw measured variables often lack direct hydrological relevance. Near-surface soil moisture is not directly measured; instead, the soil thermal properties (brightness temperature) and/or backscatter properties are measured and used to estimate near-surface soil moisture. For hydrological applications, an additional step is necessary to estimate root-zone soil moisture from near-surface soil moisture. This implies that remote-sensing products often rely on models themselves and can therefore hardly be treated as "reality". However, they can be used as independent data sources, which are useful as reference or benchmark against which to constrain or evaluate the dynamics of modeled states and fluxes. The growing availability of data is surely beneficial to unravel the hydrological functioning of catchments around the world. However, the limitations of these data need to be considered.

Commensurability issues are not only affecting remote-sensing products. Point-scale measurements of precipitation, soil moisture or piezometric levels are often not representative of the catchment-scale inputs or response. Yet, these observations are critical as they provide pieces to understand the hydrological puzzle and they are needed to calibrate and evaluate remote-sensing data products (Vidon, 2015; Burt, T. P., McDonnell, 2015; van Emmerik et al., 2018).

1.5 SYSTEMS UNDER CHANGE

Using in-situ and remote-sensing observations from the past, theories on hydrological functioning are developed and translated to models, that are subsequently used to predict the future. If the future is several days, weeks, months or seasons ahead, it seems acceptable to consider no fundamental changes in the system. However, if projections

are for several years or decades ahead in changing environmental conditions, one may question the validity of assuming a static system (Milly et al., 2008; Merz et al., 2011). Yet, in many climate change studies, models calibrated to historical conditions are used to assess the impact of future change (Booij, 2005; de Wit et al., 2007; Prudhomme et al., 2014; Wanders and Wada, 2015; Gao et al., 2020; Hakala et al., 2019). In doing so, the human interference in system management and the adaptation of ecosystems to a changing environment to maintain crucial hydrological functions are implicitly disregarded (Savenije and Hrachowitz, 2017).

The difficulty is that we lack information on how systems will evolve in response to changing environmental conditions. Especially as land-use and climate change are occurring at unprecedented rates (Hurtt et al., 2011; Ostberg et al., 2015; Gleeson et al., 2020). Moreover, changes in a system usually affect several aspects of hydrological functioning (Seibert and van Meerveld, 2016; Levia et al., 2020). When 18th century broadleaved forests were converted to coniferous plantations in the Belgian Ardennes, not only vegetation changed, but also soil properties and runoff generation mechanisms (Jacquemin et al., 2014). Increasing drought severity and heat stress associated with climate change already affect tree mortality in Europe and could fundamentally alter the composition and structure of forests in many regions (Allen et al., 2010). While the integration and quantification of all these feedback mechanisms of change into models are highly uncertain and perhaps still unfeasible; opportunities are at hand to test the sensitivity of a changing system in response to climate change using readily available data.

1.6 THIS THESIS

To increase our understanding of the spatial and temporal variability of hydrological processes in the Meuse river basin, we benefit from the creative use of additional data to evaluate the theories we develop in our hydrological models. Additionally, a key approach is to combine knowledge from different fields, including groundwater hydrology and remote sensing to answer the main question:

How can we increase the reliability of hydrological models within the Meuse river basin using remote-sensing and in-situ data?

Several research directions are explored to improve hydrological modeling of the Meuse basin (Fig. 1.3). They include identifying and quantifying dominant hydrological processes, increasing our understanding on how to use additional data, evaluating internal process representation between models and putting models to the test under changing climate and land use.

First, a general overview of landscape, climate, hydrological and land-use characteristics of the Meuse river basin is provided in Chapter 2.

In Chapter 3, we identify and quantify the magnitude and spatial variability of inter-catchment groundwater flows in the Meuse basin using water balance data within the Budyko framework, conceptual models and remote-sensing data of actual evaporation.

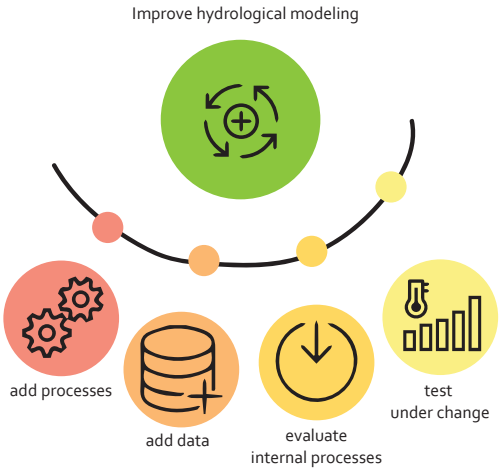

Figure 1.3: Overview of the research directions to improve hydrological modeling

Several options are tested to represent intercatchment groundwater flows as additional process in a model structure.

The water available in the root zone of vegetation for transpiration is a key element to understand the partitioning of precipitation into streamflow and evaporation. Satellites provide worldwide estimates of near-surface water content within the first few centimeters of the soil. In Chapter 4, we hypothesize that root-zone soil moisture can be inferred from remote-sensing near-surface soil moisture estimates through the link with catchment-scale vegetation accessible water storage capacities.

In Chapter 5, we hypothesize that process-based models with similar streamflow performance have similar representations of internal state and flux variables. We identify and quantify the benefits, limitations and uncertainties associated with the use of remote-sensing products to evaluate the plausibility of a suite of different hydrological models, which were calibrated by several research institutes and universities working on the Meuse basin.

In Chapter 6, we propose a methodology to represent potential system changes in process-based hydrological models in response to +2°C global warming using readily available climate change data. We rely on long-term water balance components to estimate potential shifts in the Budyko framework due to potential changes in climate and land use. We implement a larger root-zone storage capacity parameter to represent the adaptive behavior of vegetation in response to increasing water storage deficits during dry periods and evaluate the effect on hydrological predictions.

The concluding Chapter 7 provides a synthesis of the main findings and discusses further opportunities for scientists and practitioners.

2

THE MEUSE RIVER BASIN

I've been following the river, until it joins hands with the sea

The Rolling Stones (1972)

The Meuse river basin, in North-Western Europe, is the area of interest of this thesis. The Meuse is important for drinking water production, navigation, recreation, ecology, industry and agriculture. Reliable short- and long-term predictions of streamflow entering the Netherlands are of high importance, due to relatively short response times and the large seasonal variability of streamflow. General characteristics of the basin are given in this chapter. In the subsequent chapters, additional information is provided on the selection of catchments used for a specific study.

2.1 LANDSCAPE

The Meuse basin upstream of Borgharen, at the border between Belgium and the Netherlands, covers an area of approximately 21,300 km^2 in mainly France and Belgium and can be divided in three main zones. The French Southern part of the basin is characterized by thick soil layers, broad valleys bottoms and gentle slopes, underlain by sedimentary consolidated rock from the Middle and Late Jurassic. Thin soils on relatively impermeable Cambrian metamorphic rock and Early Devonian sandstone dominate the steeper and relatively high Ardennes Massif in Belgium. On the West bank of the Meuse in Wallonia, the lithology is characterized by porous chalk layers with deep groundwater systems. Elevation in the basin ranges between 50 and 700 m (Fig. 2.1).

2.2 HYDRO-CLIMATIC SETTING

The Meuse is a rain-fed river with relatively short response times. During floods, streamflow can rise within several hours due to the steep slopes and impermeable soils of the Ardennes. The strong streamflow seasonality with low summer and high winter flows reflects the seasonality of potential evaporation, as precipitation is relatively uniformly distributed throughout the year. Streamflow was at its lowest during the summer of 1976 with 20 m^3 s^{-1} and at its highest in the winter of 1993 with 3000 m^3 s^{-1}, while mean annual flows at Borgharen are around 250 m^3 s^{-1}. The large storage capacity due to relatively thick soils in the French part of the Meuse basin increases the hydrological memory of the system, implying a strong influence of winter precipitation on streamflow deficits in the subsequent summer (de Wit et al., 2007). Snow is not a major component of the water balance, but snow melt can have a significant influence during specific events (de Wit et al., 2001). For example in 2011, when rain on snow caused widespread flooding in the catchments of the Belgian Ardennes. Mean annual precipitation, potential evaporation and streamflow is approximately 950 mm yr^{-1}, 580 mm yr^{-1} and 407 mm yr^{-1}, implying a runoff ratio of 0.44 and an aridity index of 0.61.

2.3 LAND USE

Land use in the basin consists of 35% forest, 32% agriculture, 21% pasture and 9% urban areas (Fig. 2.1, European Environment Agency, 2018). The large majority of forests in the French part of the basin has been continuously wooded since at least the middle of the 19th century (Cateau et al., 2015). These broadleaved forests consist primarily of European Oak, Sessile Oak and Beech (Institut National de l'Information Géographique et Forestière, 2019). In contrast, only 44% of the 18th century Walloon forests of Belgium has remained from the original broadleaved forest, while 26 % of the original forest has been converted to coniferous plantations (Scots pine, Norway spruce and Douglas-fir) on the poor soils of the Ardennes and 30 % has been cleared for agriculture on high fertility soils in the North West (Kervyn et al., 2018). A relatively higher evaporation water use is expected in these younger, short-rotation exotic plantations in comparison to older, more natural forests.

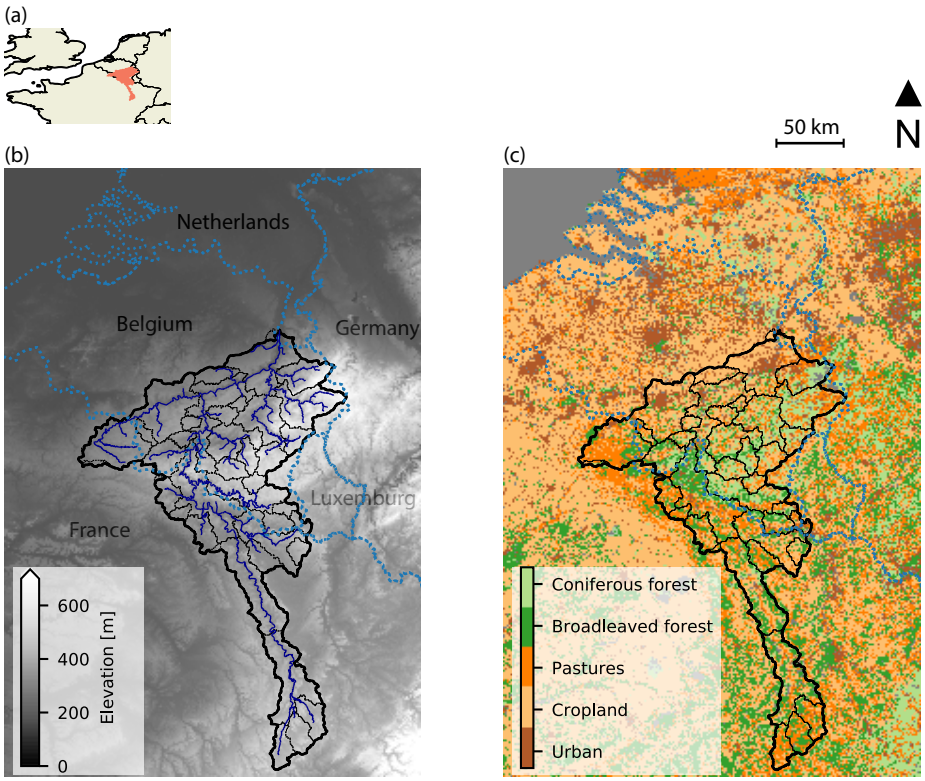

Figure 2.1: **(a)**: Overview map of Meuse basin in North-West Europe. **(b)**: Digital elevation model and outline of the Meuse basin with main catchments. **(c)**: Main land-use types according to CORINE Land Cover (European Environment Agency, 2018).

3

REDRESSING THE BALANCE: QUANTIFYING NET INTERCATCHMENT GROUNDWATER FLOWS

The Meuse had passed by a dozen villages, but now approached her first town. How appalling—thousands of people to see her, and in her bed! At such a thought she promptly gets under it. La Perte de la Meuse it is called. [...] After a couple of miles the Meuse finds this underground travel uncomfortable. After all, she must face many towns in her course: better brazen it out. So she emerges, showing a new confidence as she pushes her way through a tangle of weeds and water flowers.

Bernard Newman - *The lazy Meuse* (1948)

Conservation of mass is the main physical law on which all hydrological models rely. The change of water storage in a catchment is equal to the difference between incoming fluxes (precipitation) and outgoing fluxes (e.g. evaporation, streamflow and intercatchment groundwater flows). Many models neglect the presence of intercatchment groundwater flows across topographic divides, despite their potential contribution in the water balance. In this chapter, we use water balance data to identify, quantify and evaluate the potential presence of net intercatchment groundwater flows in the Meuse river basin.

Parts of this chapter have been published in Hydrology and Earth System Sciences (Bouaziz et al., 2018)

SUMMARY

Intercatchment groundwater flows (IGFs), defined as groundwater flows across topographic divides, can occur as regional groundwater flows that bypass headwater streams and only drain into the channel further downstream or directly to the sea. However, groundwater flows can also be diverted to adjacent river basins due to geological features (e.g., faults, dipping beds and highly permeable conduits). Even though intercatchment groundwater flows can be a significant part of the water balance, they are often not considered in hydrological studies. Yet, assuming this process to be negligible may introduce misrepresentation of the natural system in hydrological models, for example in regions with complex geological features. The presence of limestone formations in France and Belgium potentially further exacerbates the importance of intercatchment groundwater flows, and thus brings into question the validity of neglecting intercatchment groundwater flows in the Meuse basin. To isolate and quantify the potential relevance of net intercatchment groundwater flows in this study, we propose a three-step approach that relies on the comparison and analysis of (1) observed water balance data within the Budyko framework, (2) results from a suite of different conceptual hydrological models and (3) remote-sensing-based estimates of actual evaporation. The data of 58 catchments in the Meuse basin provide evidence of the likely presence of significant net intercatchment groundwater flows occurring mainly in small headwater catchments underlain by fractured aquifers. The data suggest that the relative importance of net intercatchment groundwater flows is reduced at the scale of the Meuse basin, as regional groundwater flows are mostly expected to be self-contained in large basins. The analysis further suggests that net intercatchment groundwater flow processes vary over the year and that at the scale of the headwaters, net intercatchment groundwater flows can make up a relatively large proportion of the water balance (on average 10 % of mean annual precipitation) and should be accounted for to prevent overestimating actual evaporation rates.

3.1 INTRODUCTION

Intercatchment groundwater flows are defined as groundwater fluxes crossing topographic divides, implying that precipitation falling in one catchment affects the streamflow in another catchment. A theoretical framework to describe groundwater flows was introduced by Tóth (1963). He classified different systems of groundwater flows, starting from local flow paths, nested in larger intermediate systems, which in turn are nested in regional flow systems. The theory describes that regional groundwater flow paths transport water from small headwaters to the larger and lower elevation basin, meaning that small basins tend to export or import water and large basins are likely self-contained (Schaller and Fan, 2009). This is based on the assumption that regional flow paths occur within surface drainage boundaries at the largest scale; however, systems with dipping sedimentary beds can divert groundwater away from the basin, leading to complications of the above-described theories and to intercatchment groundwater flows between adjacent basins (Schaller and Fan, 2009; Frisbee et al., 2016). Regional flow paths within a basin and between adjacent basins are the subject of this study as they characterize intercatchment groundwater flows.

Large-scale studies and theoretical models can help to understand the link between intercatchment groundwater flows and physical catchment characteristics. Schaller and Fan (2009) assessed the role of topography, aquifer properties, climate and geology on intercatchment groundwater flows. On the continental scale, they found that arid climates favor intercatchment groundwater flows. However, on the regional and basin scale, geology exerts the strongest control on intercatchment groundwater flows. The particularities of the geological systems (e.g., faults, connectivity between faults, subsurface flow conduits) can inhibit expected correlations between the magnitude of intercatchment groundwater flows and physical catchment characteristics (e.g., lithology), as was also pointed out by Le Moine et al. (2007). This highlights the difficulty to generalize the presence of intercatchment groundwater flows based on similarities in climate and topography between catchments.

Intercatchment groundwater flows cannot be directly measured and are therefore difficult to quantify, which can explain why they are often neglected in small catchment studies (Genereux et al., 2002). However, Schaller and Fan (2009) showed that intercatchment groundwater flows can be a significant portion of a basin's water balance across the continental United States; with up to 90 % of flow leaving catchments as groundwater export and up to 50 % of river flow originating from groundwater imported from other basins. Methods to identify and quantify intercatchment groundwater flows in real-world basins either rely on stream chemistry and isotope analyses (Genereux et al., 2002; Genereux and Jordan, 2006; Ajami et al., 2011; Frisbee et al., 2011, 2012, 2016), numerical groundwater flow and transport modeling (Gleeson and Manning, 2008; Welch and Allen, 2012; Ameli et al., 2018), or on water budget analyses (within the Budyko framework) (Genereux et al., 2005; Le Moine et al., 2007, 2008; Schaller and Fan, 2009; Hrachowitz et al., 2014). Depending on the type of solute (Ameli et al., 2017), higher solute concentrations in regional groundwater flows (due to longer residence time) compared to local flow paths can provide evidence for groundwater gains through intercatchment groundwater flows. Water budget analyses, using observed streamflow in real-world

catchments, can in contrast show net gains or losses due to intercatchment groundwater inflow or outflow (Genereux et al., 2002).

Intercatchment groundwater flows impact water quality in higher order streams, the alteration of nonpoint source agricultural pollution, water replenishment in aquifers, the generation and migration of petroleum and mineral deposits and the ecological functioning of a catchment (Ameli et al., 2018), and it is therefore essential to understand intercatchment groundwater flows in spite of the difficulties in quantifying them.

Most conceptual hydrological models, including HBV (Bergström, 1992), TOPMODEL (Beven and Kirkby, 1979), HyMOD (Wagener et al., 2001), SUPERFLEX (Fenicia et al., 2014a), VHM (Willems et al., 2014) and NAM (Nielsen and Hansen, 1973), solely rely on closing the water balance and neglect the possible presence of intercatchment groundwater flows by relating the change in storage over time to the difference between precipitation and the sum of actual evaporation and streamflow. These models assume watertight catchment boundaries derived from surface elevation, an impermeable substratum and no deep subsurface flow bypassing the stream. These assumptions imply the absence of intercatchment groundwater flows. Adding a loss or gain term to represent such intercatchment groundwater flows is often not warranted in models due to limited data availability for calibration (often only streamflow) and the difficulties involved in determining potential and actual evaporation (Beven, 2001; Mouelhi et al., 2006). Conceptual models have several possibilities to adjust the water balance and a "correction" factor on climatic input data has often been favored over an explicit representation of intercatchment groundwater flows. Yet, this common practice may introduce misrepresentation of the natural system in hydrological models, for example in regions with complex geological features (Zhang and Savenije, 2005; Zhang et al., 2005; Reggiani and Rientjes, 2010). In the absence of robust quantitative evidence on the magnitude and temporal variability of intercatchment groundwater flow, the errors introduced by an omission of this process in models is typically compensated for by the actual evaporation term. Examples of conceptual (or empirical) models that explicitly account for net intercatchment groundwater flows include the GR4J empirical model (Perrin et al., 2003) often applied in French catchments, HYDROLOG (Chiew and McMahon, 1990), SMAR (Goswami et al., 2007; Goswami and O'Connor, 2010), mHM (Samaniego et al., 2011) and the flexible model structure used in Hrachowitz et al. (2014).

Including intercatchment groundwater flows in conceptual models has been studied in a large set of French catchments (Le Moine et al., 2007) and results in a more plausible partitioning between evaporation, streamflow and underground fluxes than methods correcting for potential errors in climatic input data or catchment area instead. Isotopic and chemical analyses indicate an intra-annual variability of intercatchment groundwater flow processes (Ajami et al., 2011; Frisbee et al., 2012).

While several studies used extensive tracer and geochemical data or developed detailed flow and transport models to quantify intercatchment groundwater flows, we propose a framework that uses widely available hydrometric observations. Previous research also using water balance data shows that different methods for estimating intercatchment groundwater flows are characterized by different uncertainties. The novelty

Table 3.1: Catchment characteristics. Meteorological and hydrological data are based on data between October 2006 and September 2016. *Fissured denotes the percentage of highly productive fissured aquifers based on the International Hydrogeological Map of Europe, IHME.

Station River	Straimont Vierre	Ste-Marie Semois	Tintigny Semois	Chiny Semois	Membre P Semois	Huccorgne Mehaigne	Yvoir Bocq	Belval Sormonne	Pierrepont Crusnes	V-le-C Aroffe
Area (km^2)	182	143	381	738	1226	305	230	369	207	198
Mean elev. (m)	440	366	405	407	390	158	268	254	340	367
Mean slope (–)	0.067	0.044	0.055	0.060	0.083	0.026	0.064	0.066	0.054	0.060
Forest (%)	34	38	50	47	56	3	16	28	23	48
Pasture (%)	29	26	22	24	18	1.7	14	48	13	20
Urban (%)	8	11	6	6	5	15	10	4	4	1
Crop (%)	29	26	22	22	21	80	60	20	60	30
Hillslopes (%)	7.4	1.5	4.6	6.0	15	0.9	7.5	9.1	6.6	8.4
Fissured* (%)	0	63	27	16	9	16	71	48	94	72
P ($mm\,y^{-1}$)	1176	1041	1110	1152	1183	753	867	1114	939	833
Q_{obs} ($mm\,y^{-1}$)	665	455	570	600	665	206	297	422	337	88
E_P ($mm\,y^{-1}$)	608	615	611	614	611	627	618	620	618	621
Q_{obs}/P (–)	0.57	0.44	0.51	0.52	0.56	0.27	0.34	0.38	0.36	0.11
E_P/P (–)	0.52	0.59	0.55	0.53	0.52	0.83	0.71	0.56	0.66	0.75

of this study is that, here, we aim to limit these uncertainties and to detect and quantify net intercatchment groundwater flows (i.e., $Q_{IGF,in} - Q_{IGF,out}$) in a complementary three-step approach through (1) water budget accounting, (2) testing a set of model concepts and (3) evaluating the results against remote-sensing estimates of actual evaporation. In a proof-of-concept study in the Meuse basin, we test the following hypotheses:

1. Observed water balance data in combination with the Budyko framework can provide robust evidence of the likelihood and spatial variability of net intercatchment groundwater flows.

2. Simple hydrological conceptual models enable to quantify the magnitude and intra-annual variability of net intercatchment groundwater flows over mesoscale catchments and to assess the likelihood that intercatchment groundwater flows occur within a basin or between neighboring basins.

3. Actual evaporation estimates from remote sensing provide additional evidence to support the presence of net intercatchment groundwater flows.

3.2 STUDY AREA

This study uses data from 58 catchments within the Meuse basin (Fig. 3.1) with areas varying between 50 and 16 500 km^2, with a median value of 370 km^2. Mean annual precipitation varies between 750 and 1200 $mm\,yr^{-1}$ and median annual streamflow and potential evaporation in these catchments are approximately 420 and 620 $mm\,yr^{-1}$, respectively.

From the 58 available stations, five stations are available in the Semois River catchment (Fig. 3.2 and Table 3.1) and are studied in more detail along with five additional stations (Fig. 3.1a and Table 3.1).

The Semois catchment upstream of Membre-Pont is interesting because it combines

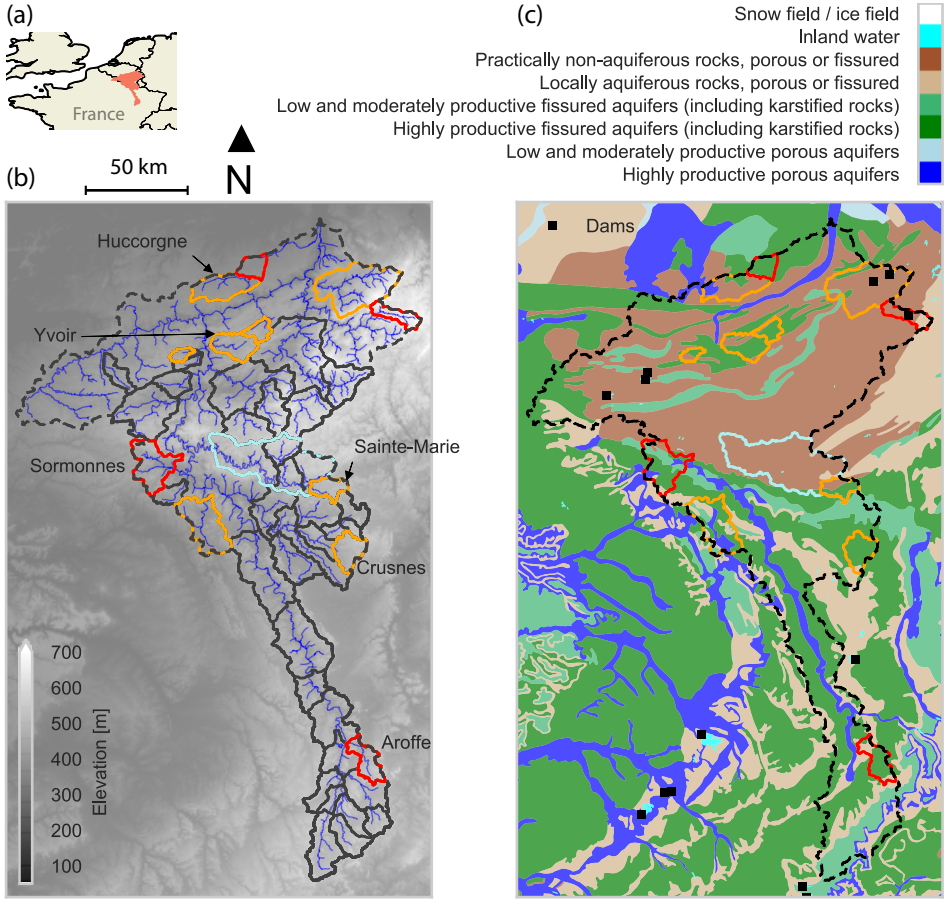

Figure 3.1: **(a)**: Overview map of the Meuse basin. **(b)**: Digital elevation model and outline of the Meuse basin with all catchments (black), catchments plotting beyond the energy limit (red), catchments very close to the energy limit (orange). The location of the Semois catchment at Membre-Pont is indicated in pale turquoise. **(c)**: International Hydrogeological Map of Europe (IHME), location of main dams (black squares, FAO database) and catchments close to (orange) and beyond (red) the energy limit.

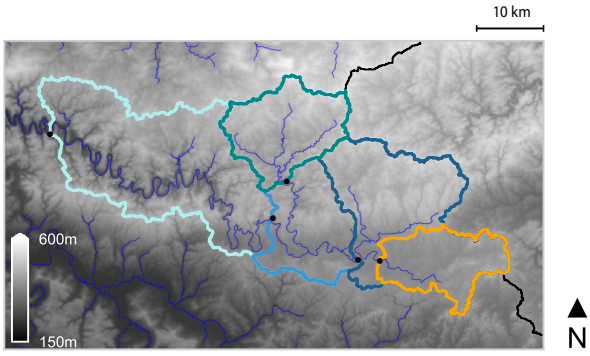

Figure 3.2: Digital elevation model of the Semois catchment and location of the stations from upstream to downstream: Sainte-Marie (orange), Tintigny (dark blue), Chiny (light blue), Membre-Pont (pale turquoise) and additional tributary in the north is the Vierre at Straimont (green). The catchment of Sainte-Marie plots very close to the energy limit as shown in Fig. 3.3.

both the Jurassic and Early Devonian geological horizons: only the upstream catchment of Sainte-Marie consists of marl (and limestone), while further downstream the basin is underlain by relatively impermeable sandstone and schist. In addition, several streamflow stations along the Semois river are available and allow us to detect how net intercatchment groundwater flows (IGF_{net}) evolve as we move further downstream along the same river. Characteristics of the Semois catchments are included in Table 3.1 and a map is provided in Fig. 3.2.

In the French part of the Meuse basin, the tributary of the Aroffe River at Vannes-le-Châtel (198 km^2; Fig. 3.1a) flows underground through limestone deposits towards the Moselle catchment (Fister, 2012). The Aroffe is a typical example of an overflow spring that is activated when the capacity of the conduit is exceeded, while it flows underground to the Moselle the rest of the time. The Aroffe is one of the additional five catchments where IGF_{net} is quantified (Sect. 3.5.2).

3.3 DATA

3.3.1 METEOROLOGICAL AND HYDROLOGICAL DATA

For each catchment, areal averages of precipitation, potential evaporation and observed streamflow (available between 2006 and 2016) are required for the analyses.

Hourly precipitation measurements are interpolated using climatological monthly background grids, using a combination of the HYRAS (Rauthe et al., 2013) and E-OBS (Haylock et al., 2008) datasets and following the method described in van Osnabrugge et al. (2017). Precipitation measurements in Belgium were provided by the Service Public de Wallonie (2018); in France data were retrieved from the Dutch operational forecasting system. Potential evaporation estimates are based on the Makkink formula (Hooghart and Lablans, 1988) and rely on hourly interpolated temperature station data (using a

lapse rate of 6.6×10^{-3} °C m^{-1}) and hourly radiation data from Maastricht (Royal Netherlands Meteorological Institute, 2018). Mean hourly values of precipitation and potential evaporation are derived from the 1200 m resolution gridded data for each catchment.

Observed streamflow data are available at the hourly time step for the stations in Belgium from the Service Public de Wallonie (2018) and at the daily timestep for the stations in France from Banque Hydro (2018). In the Semois catchments, streamflow between March and mid-June 2013 were set to missing because of high observed streamflow with too-limited precipitation amounts.

3.3.2 REMOTE-SENSING-BASED ACTUAL EVAPORATION ESTIMATES

Two products of remote-sensing-based actual evaporation estimates are used for comparison with modeled actual evaporation: the Global Land Evaporation Amsterdam Model (GLEAM, Miralles et al., 2011; Martens et al., 2017) and Satellite Application Facility on Land Surface Analysis daily MSG actual evaporation (LSA SAF, Trigo et al., 2011).

GLEAM v3a calculates actual evaporation based on satellite-observed soil moisture, vegetation optical depth and snow water equivalent, reanalysis air temperature and radiation, and a multi-source precipitation product. GLEAM provides actual evaporation estimates at a spatial resolution of 0.25°and accounts for subgrid heterogeneity by considering three land surface types (bare soil, short vegetation and vegetation with a tall canopy). GLEAM estimates are available for the entire studied period between 2006 and 2016.

LSA SAF daily MSG (Meteosat Second Generation) actual evaporation (hereafter referred to as LSA SAF) includes soil evaporation, interception and transpiration and is calculated by solving the energy balance by combining radiative, land surface, vegetation and meteorological data. Each pixel (3 km × 3 km resolution at nadir) is split into four tiles to represent main land cover types (bare soil, grassland, crops and forests) and the surface energy balance is solved for each tile separately and results in an actual evaporation value per pixel based on the weighted average of the tiles (https://landsaf.ipma.pt/en/products/evapotranspiration/dailymet/, 29 November 2018). LSA SAF estimates are only available for the validation period (2012–2016).

3.4 METHODS

This study consists of three parts aimed to identify, quantify and test for the presence of net intercatchment groundwater flows (IGF_{net}) in the Meuse basin. First, we use long-term observed water balance data in combination with the Budyko framework (Budyko, 1961) to identify catchments with evidence of water losses or gains through IGF_{net}. Second, we use conceptual hydrological models to assess the magnitude and temporal variability of potential IGF_{net} in the Meuse basin and we assume that they are the main cause of water balance discrepancies and thereby neglect uncertainties in forcing data. We model IGF_{net} as independent losses or gains in alternative model concepts and evaluate their magnitude in several catchments of the Meuse basin. To assess if part of the

groundwater flow bypasses the headwater stream to reach the river further downstream, we model the losses or gains in increasingly large catchments along the same tributary. Thirdly, we use actual evaporation from remote-sensing estimates to provide additional evidence for the likelihood and magnitude of IGF_{net}.

3.4.1 IDENTIFYING NET INTERCATCHMENT GROUNDWATER FLOWS FROM OBSERVED DATA SIGNALS

The water balance of a catchment reads as follows:

$$\frac{dS}{dt} = P(t) - Q_{obs}(t) - E_A(t) - Q_{IGF}(t), \tag{3.1}$$

where S is the storage in the catchment, P is the precipitation at time step t, Q_{obs} is the observed streamflow at the catchment outlet, E_A is the actual evaporation and Q_{IGF} is the groundwater net loss (if Q_{IGF} is positive, meaning that the groundwater flow out of the catchment is larger than the flow into the catchment) or net gain (if Q_{IGF} is negative) to the catchment, where all variables represent instantaneous fluxes (in mm h^{-1}).

Intercatchment groundwater flows are often not considered and over a long period (several years), the change in storage is assumed to be zero, and long-term mean precipitation P, actual evaporation E_A and observed streamflow Q_{obs} (in mm yr^{-1}) can be reduced to the following:

$$P = Q_{obs} + E_A \tag{3.2}$$

The Budyko framework (Budyko, 1961) describes the empirical global relation between the long-term evaporative index (E_A/P) and the dryness index (E_P/P, with E_P the long-term mean potential evaporation) and shows that natural catchments show a tendency to plot along the Budyko curve in the theoretical range located in between the energy and water limits. The water limit implies that a catchment cannot evaporate (or streamflow) more water than it receives from precipitation; this implies that catchments with higher streamflow than precipitation plot beyond the water limit (gaining catchments) in the Budyko framework. The energy limit implies that catchments cannot evaporate (E_A) more than the energy available for evaporation (E_P), and therefore catchments where the difference between precipitation and streamflow is larger than potential evaporation are beyond the energy limit (leaky catchments), as shown in Fig. 3.3a. Assuming negligible observation errors, they are likely affected by net intercatchment groundwater inflows (gaining catchments) or outflows (leaky catchments). Andréassian et al. (2012) suggest replacing the axis of the evaporative index ($E_A/P = 1 - Q_{obs}/P$) with the runoff ratio (Q_{obs}/P) in the Budyko framework because gaining catchments would otherwise have a negative evaporative index and because E_A itself is not measured at the catchment scale. We therefore plot each catchment in the nondimensional representation of the runoff ratio (Q_{obs}/P) as a function of the dryness index (E_P/P), hereinafter referred to as the Budyko framework for the sake of convenience, using hydrological years between October 2006 and September 2016 (10 years) with more than 350 days of streamflow data per year.

Catchments show a tendency to plot close to the Budyko curve or other alternative expressions. The Turc–Mezentsev formula (Turc, 1954; Mezentsev, 1955) plots very close

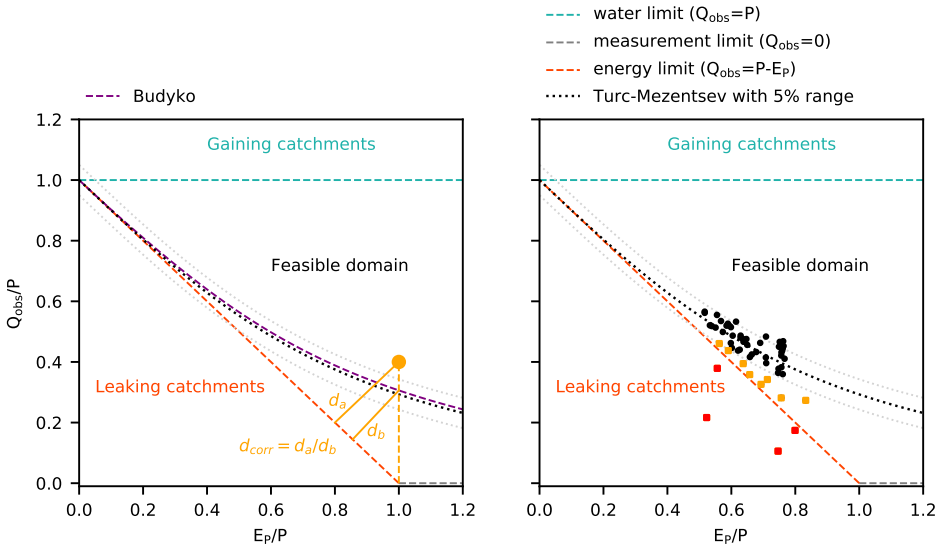

Figure 3.3: **(a)** Dimensionless representation of the runoff ratio (Q_{obs}/P) as a function of the dryness index (E_P/P), referred to as the Budyko framework. The red line is the energy limit ($Q_{obs} = P - E_P$) beyond which catchments are leaking water; the blue line is the water limit ($Q = P$) above which catchments are gaining water; the dark grey line is the measurement limit ($Q = 0$). The domain within these three limits is the theoretical feasible domain. The Turc–Mezentsev and the Budyko curves plot very close to each other. The 5 % uncertainty bound around the Turc–Mezentsev curve is also shown. For each catchment, the ratio of the distance to the energy limit (d_a) over the distance of Turc–Mezentsev to the energy limit (d_b) is used as a proxy for the presence of net intercatchment groundwater flows. **(b)** The catchments of the Meuse basin are located around the Turc–Mezentsev curve (black circles). However, four catchments plot beyond the energy limit (red squares) and eight catchments plot very close to the energy limit and are beyond the lower 5 % range of Turc–Mezentsev (orange squares). In these catchments, we expect net intercatchment groundwater flow losses to occur.

to the Budyko curve (Fig. 3.3) and has often been used in studies of French catchments (Le Moine et al., 2007) and was therefore applied in our analysis. The Turc–Mezentsev formula is the most general function that fulfills the two conditions $Q_{obs} \sim 0$ when $P \ll E_P$ (in very dry, moisture-constrained catchments) and $Q_{obs} \sim P - E_P$ when $P \gg E_P$ (in very wet, energy-constrained catchments) (Turc, 1954; Lebecherel et al., 2013), and transposed to streamflow, it reads as follows:

$$\frac{Q_{obs}}{P} = 1 - \frac{1}{(1 + (\frac{P}{E_P})^n)^{\frac{1}{n}}}, \tag{3.3}$$

in which n is an exponent to estimate. Depending on the value of the parameter n, the Turc–Mezentsev relation occupies the domain from the energy limit to the water limit; Turc (1954) retained a value of $n = 2$. Here we define catchments plotting more than 5 % away from this curve (which implies a narrower range than in Gentine et al., 2012, but wider than in Li et al., 2014) and close to the limits as likely to be affected by IGF_{net}. More specifically, catchments plotting beyond the energy limit and between the energy limit and the lower boundary of the Turc–Mezentsev uncertainty range (Fig. 3.3) potentially

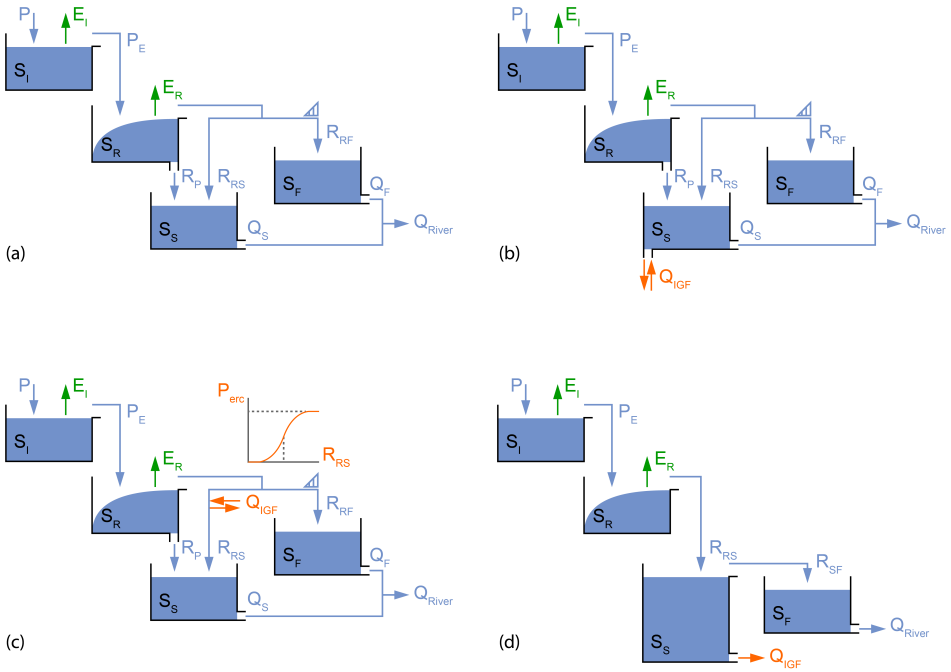

Figure 3.4: Conceptual model schematizations. **(a)** Reference model without net intercatchment groundwater flows. **(b)** Reference model with net constant intercatchment groundwater flows from the slow reservoir. **(c)** Reference model with net preferential intercatchment groundwater flows retrieved from or added to the recharge to the slow reservoir (using an error function that relates the percentage or recharge lost or gained to net intercatchment groundwater flows as a function of the recharge rate). **(d)** Overflow model used for the Aroffe catchment at Vannes-le-Châtel that simulates river flows in the Aroffe only when the underground storage capacity is exceeded. The rest of the time, flows occur underground towards neighboring basins. Here, we define P as precipitation, P_E as effective precipitation, E as evaporation, R as an internal flux and Q as surface or subsurface streamflow (all in mm h^{-1}) and S as storage (in mm). For the subscripts, we define $_I$ as interception, $_R$ as root zone, $_S$ as slow response, $_F$ as fast response and $_P$ as percolation. The parameter P_{erc} defines the maximum percentage of recharge as net intercatchment groundwater flow.

indicate the presence of net subsurface losses. Indeed, catchments that plot very close to the energy limit imply that the difference between precipitation and streamflow approximates the total energy available for evaporation ($P - Q_{obs} \approx E_P$). During dry and/or very warm periods, however, evaporation is constrained by water availability and mean annual actual evaporation is therefore expected to be considerably lower than potential; this in turn means that water must be leaving the catchment through another route to comply with the observed long-term water balance. We hypothesize that water is leaving the catchment through underground pathways.

We consider the shortest distance between each catchment and the energy limit in the Budyko framework as a proxy for the presence of IGF_{net}. The closer a catchment is to the energy limit, the higher the probability of IGF_{net}. We adjust this distance by the shortest distance of the point on the Turc–Mezentsev curve at the catchment's E_P/P

to the energy limit (Fig. 3.3a) because arid catchments have lower runoff ratios and are therefore expected to be further away from the energy limit. Negative distances imply that catchments plot beyond the energy limit.

We then assess if the adjusted distance to the energy limit is correlated with several physical catchment characteristics that may influence the formation of IGF_{net}, including the percentage of highly productive fissured aquifers (including karstified rocks) as provided by the International Hydrogeological Map of Europe (IHME, https://www.bgr.bund.de/ihme1500, last access: 29 November 2018) and International Geological Map of Europe (IGME), catchment area, and percentage of hillslopes (slopes steeper than 13 %, Gharari et al., 2011).

3.4.2 QUANTIFYING NET INTERCATCHMENT GROUNDWATER FLOW PROCESSES USING CONCEPTUAL MODELS

MODELS DESCRIPTION

A reference conceptual model is developed including interception, soil moisture, fast and slow reservoirs, but no IGF_{net} (Fig. 3.4a). This conceptual model is similar to the model used by Fovet et al. (2015) and has 10 calibration parameters. The characteristic timescale of the recession of the slow reservoir is determined with a master recession curve analysis.

Two options are investigated to incorporate IGF_{net} in the reference model. The first one involves a continuous constant groundwater exchange flux (loss or gain) from or to the slow reservoir ($Q_{IGF}(t) = C_{IGF}$), assuming a slowly draining, homogeneous, low-permeability aquifer (Fig. 3.4b). The second relies on preferential permeable pathways, activated above a certain threshold, to lose or gain water (Fig. 3.4c and Appendix A). In the preferential model, part of the recharge is lost or gained (before entering the slow reservoir) when the recharge exceeds a certain threshold. An error function is used to simulate this behavior: $Q_{IGF}(t) = \mathrm{erf}(R_{RS}(t), \mu, m_3) \cdot p_{erc} \cdot R_{RS}(t)$, with $R_{RS}(t)$ the recharge from the root zone storage to the slow reservoir, μ the threshold parameter of the recharge above which IGF_{net} occurs, P_{erc} the maximum fraction of the recharge to IGF_{net} and m_3 a shape parameter of the error function (not calibrated). The constant loss–gain model resembles the one in Hrachowitz et al. (2014) and counts one extra parameter, while the preferential IGF_{net} model has two additional parameters compared to the reference model.

In the catchment of the Aroffe River, water sinks in the karstified limestone after traveling through sandstone and marl deposits and emerges again in the neighboring catchment of the Moselle (which is a tributary of the Rhine River). During peak flows, the conduit capacity is exceeded and water flows in the river bed of the Aroffe (Fister, 2012). To simulate the hydrological functioning of the Aroffe river, an overflow type of model is developed to quantify the losses of this catchment to the neighboring Moselle basin, according to $Q_{IGF}(t) = K_{IGF}^{-1} \cdot S_S(t)$, with K_{IGF} the characteristic timescale of the underground stores (S_S), as shown in Fig. 3.4d.

Parameters, water balance and constitutive equations of all models are provided in

Appendix A and model schematizations are shown in Fig. 3.4. All models are programmed in Python and an implicit Euler time stepping scheme is used to solve the model equations.

MODEL EXPERIMENTS – GENERAL PROCEDURE

The model was run between 1 January 2006 and 31 December 2011, using 2006 as a warm-up year, to explore the parameter space with a Monte Carlo strategy and sampling from uniformed prior parameter distributions (10^5 realizations). This was done at an hourly time step because of the fast processes occurring in the Meuse river basin. Feasible parameter sets are retained based on their simultaneous ability to reproduce high- and low-flow metrics during calibration with Nash–Sutcliffe efficiencies of at least 0.7 for different indicators (Nash–Sutcliffe efficiency of the flows $E_{\mathrm{NS},Q}$ and of the log of the flows $E_{\mathrm{NS},\log Q}$, Nash–Sutcliffe efficiency of the flow duration curve of the log of the flows $E_{\mathrm{NS,FDC},\log Q}$), and to reproduce streamflow volumes at different temporal scales (relative volume error E_{RVE}, Nash–Sutcliffe efficiency of runoff ratios for 6-monthly $E_{\mathrm{NS,RC,6m}}$, monthly $E_{\mathrm{NS,RC,m}}$ and weekly $E_{\mathrm{NS,RC,w}}$ periods). The tested models are evaluated in an independent validation period running from 1 January 2012 to 31 December 2016.

Prior and posterior parameter ranges are provided in Appendix A. The characteristic timescale of the recession of the slow reservoir is estimated with a master recession curve analysis for each catchment (Fenicia et al., 2006). A range of 10 days around the derived value is used as a calibration range to account for nonlinear recession when a constant loss or gain is added to the slow reservoir.

The experiments designed to test the hypotheses of this paper are described in the following sections.

REPRESENTATION: HOW TO REPRESENT NET INTERCATCHMENT GROUNDWATER FLOWS – ZERO, CONSTANT OR PREFERENTIAL FLOWS?

The stations on the Semois River and its tributary (Vierre at Straimont and Semois at Sainte-Marie, Tintigny, Chiny, Membre-Pont, shown in Fig. 3.2) are used to assess three alternative model concepts: the reference model without $\mathrm{IGF_{net}}$, constant $\mathrm{IGF_{net}}$ from or to the slow reservoir and preferential $\mathrm{IGF_{net}}$ from or to the recharge to the slow reservoir. These stations are selected because they also allow us to quantify how $\mathrm{IGF_{net}}$ evolve from upstream to downstream along the same river (Sect. 3.4.2). The five stations are calibrated independently using the three models to quantify the magnitude of $\mathrm{IGF_{net}}$ in the subsequent catchments. The most suitable model structure is determined based on a visual inspection of hydrographs and modeled streamflow regime, a comparison of performance indicators in the validation period, and a comparison between the magnitude of the loss and the distance to the energy limit (long-term mean and annual variability). Additionally, modeled mean annual actual evaporation are compared to Turc–Mezentsev estimates and we assess the shift of the modeled water balance in the Budyko framework when $\mathrm{IGF_{net}}$ are considered versus neglected.

DIRECTION: WHERE DO INTERCATCHMENT GROUNDWATER FLOWS GO?

To test if part of the groundwater flow bypasses the headwater stream to reach the river only further downstream, we model the Semois River catchments (using the experiments described in Sect. 3.4.2) to quantify how the loss–gain term varies as catchment size increases along the same river. Additionally, we looked for examples in the literature located in the Meuse basin to highlight the possible difference between IGF_{net} that is internal to a river basin and IGF_{net} to neighboring river basins.

MAGNITUDE: WHAT IS THE MAGNITUDE OF NET INTERCATCHMENT GROUNDWATER FLOWS AT THE SCALE OF THE MEUSE BASIN?

Several catchments plotting close to or beyond the energy limit (from the analysis described in Sect. 3.4.1) are modeled to quantify the magnitude of potential IGF_{net} at several locations in the Meuse basin. Additional catchments where the magnitude of IGF_{net} is evaluated using the preferential model (because it performed better for the Semois at Sainte-Marie; see the results in Sect. 3.5.2) include the Sormonne at Belval, the Mehaigne at Huccorgne, the Bocq at Yvoir and the Crusnes at Pierrepont (Fig. 3.1a). For the Aroffe at Vannes-le-Châtel, the overflow type of model (Fig. 3.4d) is used to model the loss towards the Moselle basin, based on findings from the literature (Fister, 2012).

3.4.3 EVALUATING NET INTERCATCHMENT GROUNDWATER FLOWS USING REMOTELY-SENSED ACTUAL EVAPORATION ESTIMATES

We test for the presence of IGF_{net} using independent additional data sources. Actual evaporation is a major component of the water balance at the catchment scale, but it is also a great unknown. Reliable estimates of actual evaporation at the catchment scale would allow us to attribute the gap in the water balance to IGF_{net}, assuming minor anthropogenic activities. Global evaporation products are, however, not derived directly from earth observations, but rely on remotely sensed data in combination with models to derive actual evaporation. In this study, we compare two sources of remotely sensed actual evaporation estimates (LSA SAF and GLEAM) with our modeled actual evaporation to test the hypothesis of IGF_{net}.

3.5 RESULTS

3.5.1 IDENTIFICATION OF NET INTERCATCHMENT GROUNDWATER FLOWS AND LINK WITH PHYSICAL CATCHMENT CHARACTERISTICS

The analysis of observed water balances in the Budyko framework shows that relatively small headwater catchments of the Meuse basin (50–700 km^2; Fig. 3.1) plot closest to or beyond the energy limit (Fig. 3.3b); this suggests that these catchments exhibit the highest potential for the presence of net intercatchment groundwater flows (IGF_{net}). Amongst them is the headwater catchment of the Semois at Sainte-Marie (Fig. 3.2), which plots close to the energy limit, suggesting underground losses towards other catchments. The water balance of two catchments in the northeast (Fig. 3.1) might be affected by the

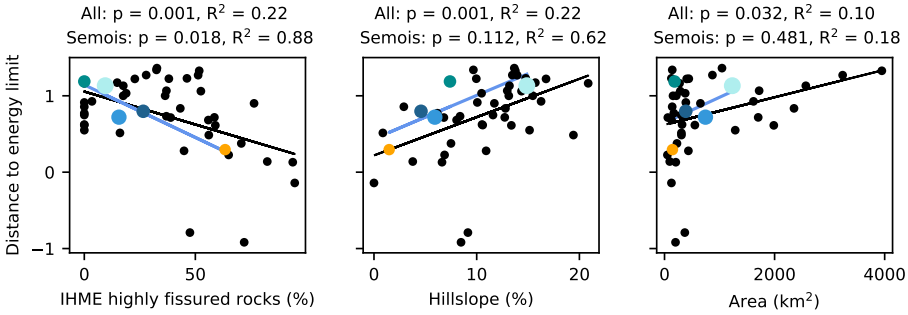

Figure 3.5: Adjusted distance of each catchment to the energy limit in the Budyko framework (as explained in Fig. 3.3) is plotted as a function of several catchment characteristics. This distance is used as a proxy for the presence of net intercatchment groundwater flows. The black line and dots show the correlation for all stations of the Meuse basin and the colored dots (with sizes scaled to catchment area) and blue line display the catchments of the Semois River only. **(a)** Distance to the energy limit as a function of the percentage highly fissured aquifers including karstified rocks based on the International Hydrogeological Map of Europe (IHME), indicating larger net losses as the percentage of highly fissured aquifers increases because of lower (or negative) values of the distance to the energy limit. **(b)** Distance to the energy limit as a function of percentage of hillslopes defined as slopes above 13 % and representative for the competition between surface and subsurface drainage. **(c)** Distance to the energy limit as a function of catchment areas of the main tributaries (up to 4000 km^2).

presence of dams (FAO, 2016) and the two catchments are therefore left out of further analyses. The net losses calculated with long-term observed streamflow, precipitation and Turc–Mezentsev estimates of actual evaporation in these headwater catchments range between 70 mm yr^{-1} (for the Semois at Sainte-Marie, which corresponds to 7 % of mean annual precipitation), and 260 mm yr^{-1} (for the Aroffe catchment at Vannes-le-Châtel, which is 31 % of annual precipitation), with a median of 100 mm yr^{-1} (or 12 % of median annual precipitation). The distance of the Aroffe catchment to the energy limit is negative (the catchment plots beyond the energy limit) and approximately three times larger than the (positive) distance of the Semois at Sainte-Marie.

The catchments of the Meuse basin show a significant trend ($p = 0.001$ and $R^2 = 0.22$) indicating more losses from the catchment (negative or shorter distance to the energy limit) as the percentage of highly productive fissured aquifers increases, as shown in Fig. 4.10a. Intercatchment groundwater flows in the Meuse basin are therefore likely to occur in catchments with highly productive fissured aquifers, including karstified rocks (see the IHME hydrogeological map in Fig. 3.1b). These productive aquifers are characterized by limestone, marl or chalk lithologies (IGME). Karstification processes may cause "piracy" routes to develop (Hartmann et al., 2014) and therefore be at the origin of IGF_{net}.

We use the percentage of hillslopes in a catchment (defined as areas with a slope steeper than 13 %, Gharari et al., 2011) as a proxy for how well the drainage network is defined from the surface and relate it to the potential presence of IGF_{net} (through the distance to the energy limit) as shown in Fig. 4.10b. The data show a significant trend

($p = 0.001$ and $R^2 = 0.22$), indicating fewer losses from the catchment (larger distance towards the energy limit) as the percentage hillslope increases. The underlying idea is that surface topography displays the result of a competition between surface and subsurface flows. Catchments dominated by steep valleys, as encountered in the Ardennes, clearly show their drainage network at the surface. The steeper the catchment, the higher the relative importance of lateral flow through a subsurface preferential path network to the channel or stream. On the other hand, catchments lying on permeable lithologies as chalk and middle Jurassic limestones may be dominated by rivers cutting through relatively flat plateaus and may hide an underground network of subsurface flow paths from the surface (Le Moine, 2008). The flatter the catchment, the higher the potential importance of an underground flow network and therefore of subsurface losses or gains. In the Meuse basin, IGF_{net} is therefore likely to occur in catchments dominated by a relatively flat topography.

We also tested the hypothesis that part of the groundwater flow bypasses the channel to reach the river only further downstream by correlating the distance of each catchment to the energy limit (as a proxy for the presence of IGF_{net}) to the catchment area, for the main tributaries of the Meuse basin (Fig. 4.10c). We expected the presence of IGF_{net} to be reduced as catchment size increases, and although this trend is significant ($p = 0.032$ and $R^2 = 0.10$), the correlation is weak. The data shown in Fig. 4.10c, however, suggest that evidence for IGF_{net} is highest in small catchments (with areas less than 500 km^2) and much less pronounced in larger downstream catchments, although there are also small catchments with little evidence of it. This is likely related to the variability of local geological features underlying these small catchments.

3.5.2 VARIABILITY OF NET INTERCATCHMENT GROUNDWATER FLOWS ACROSS THE MEUSE BASIN

REPRESENTATION: A PREFERENTIAL MODEL TO REPRESENT NET INTERCATCHMENT GROUNDWATER FLOWS

The reference (without IGF_{net}), constant and preferential IGF_{net} models are calibrated on subsequent catchments along the Semois river. In the following sections, the models are evaluated based on (1) performance indicators during the validation period and visual inspection of the hydrographs and seasonal behavior, (2) the magnitude of modeled IGF_{net}, and (3) modeled actual evaporation.

PERFORMANCE INDICATORS AND VISUAL INSPECTION OF THE HYDROGRAPHS

Performance indicators of the feasible realizations of the three models in the Semois catchments during the calibration and validation period are shown in Fig. 3.6. The preferential model shows an improvement in high- and low-flow indicators, and in modeled runoff ratio in the Semois catchment at Sainte-Marie compared to the constant and zero IGF_{net} models, whereas in the other catchments of the Semois River, performance indicators are similar for the three models. Nash–Sutcliffe efficiencies of daily flows ($E_{NS,Q}$) and log of the flows ($E_{NS,\log Q}$) increase when the reference model (no IGF_{net})

is extended with a constant IGF_{net} term and increase even more when a preferential IGF_{net} term is included in the catchment upstream of Sainte-Marie. This also applies for the Nash–Sutcliffe efficiency applied on monthly and weekly runoff ratio ($E_{NS,RC,m}$ and $E_{NS,RC,w}$). On the other hand, all performance indicators for the Vierre at Straimont (Fig. 3.6) show similar results for the three models. Adding an exchange term in this sandstone-dominated catchment (constant or preferential) does not lead to an improved performance. This behavior also characterizes the other catchments at Tintigny, Chiny and Membre-Pont (Fig. 3.6).

A visual inspection of the in 2014 modeled and observed hydrographs at Sainte-Marie (Fig. 3.7) shows a decrease in modeled winter peak flows at the beginning of the year and an increase of modeled peak flows after the dry season (October) for the preferential model compared to the zero IGF_{net} model, which better approximates observed behavior. Although this behavior might vary throughout the years, a higher performance of the preferential model in reproducing the observed streamflow regime is also visible in Fig. 3.8. Including preferential IGF_{net} in the model reduces the mean overestimation of 9 mm month^{-1} at the beginning of the year and the underestimation of 11 mm month^{-1} in October and November, simulated by the zero IGF_{net} model, to respectively 0.5 and 3 mm month^{-1} on average. This implies that the error is reduced by 94 % at the beginning of the year and by 73 % in October and November. The improved simulation of the seasonal behavior indicates a better representation of the underlying processes and the resulting partitioning of water fluxes.

GROUNDWATER NET LOSSES OR GAINS IN THE SEMOIS CATCHMENT

In the catchment upstream of Sainte-Marie, a median annual loss term of 17 % and 20 % of observed streamflow (corresponding to 77 and 90 mm yr^{-1}) is modeled by the feasible realizations of the preferential and constant IGF_{net} model, respectively, as shown in Fig. 3.9a, b. The magnitude of IGF_{net} decreases in the catchments further downstream on the Semois River. At the catchment outlet (Membre-Pont) and in the Vierre tributary, the magnitude of IGF_{net} is centered around zero. The range of IGF_{net} is larger for the constant model compared to the preferential model. For the preferential model, IGF_{net} approximates a value of zero for all other catchments than Sainte-Marie. In the constant model, median values of IGF_{net} are positive (losses), but some realizations imply a slight gain. Additionally, Fig. 3.9c, d show that the magnitude of IGF_{net} decreases as the distance to the energy limit increases. This means that as catchments plot closer to the Budyko curve (and further away from the energy limit), we see the relative importance of IGF_{net} decreasing, which is in line with expectations.

EFFECT ON ACTUAL EVAPORATION

Turc–Mezentsev estimates of actual evaporation are compared with modeled mean annual actual evaporation of the feasible realizations of the three models in all Semois stations in Fig. 3.10. Including (constant or preferential) IGF_{net} in the catchment of Sainte-Marie leads to median annual actual evaporation rates close to Turc–Mezentsev estimates, whereas the reference model leads to 10 % higher actual evaporation rates

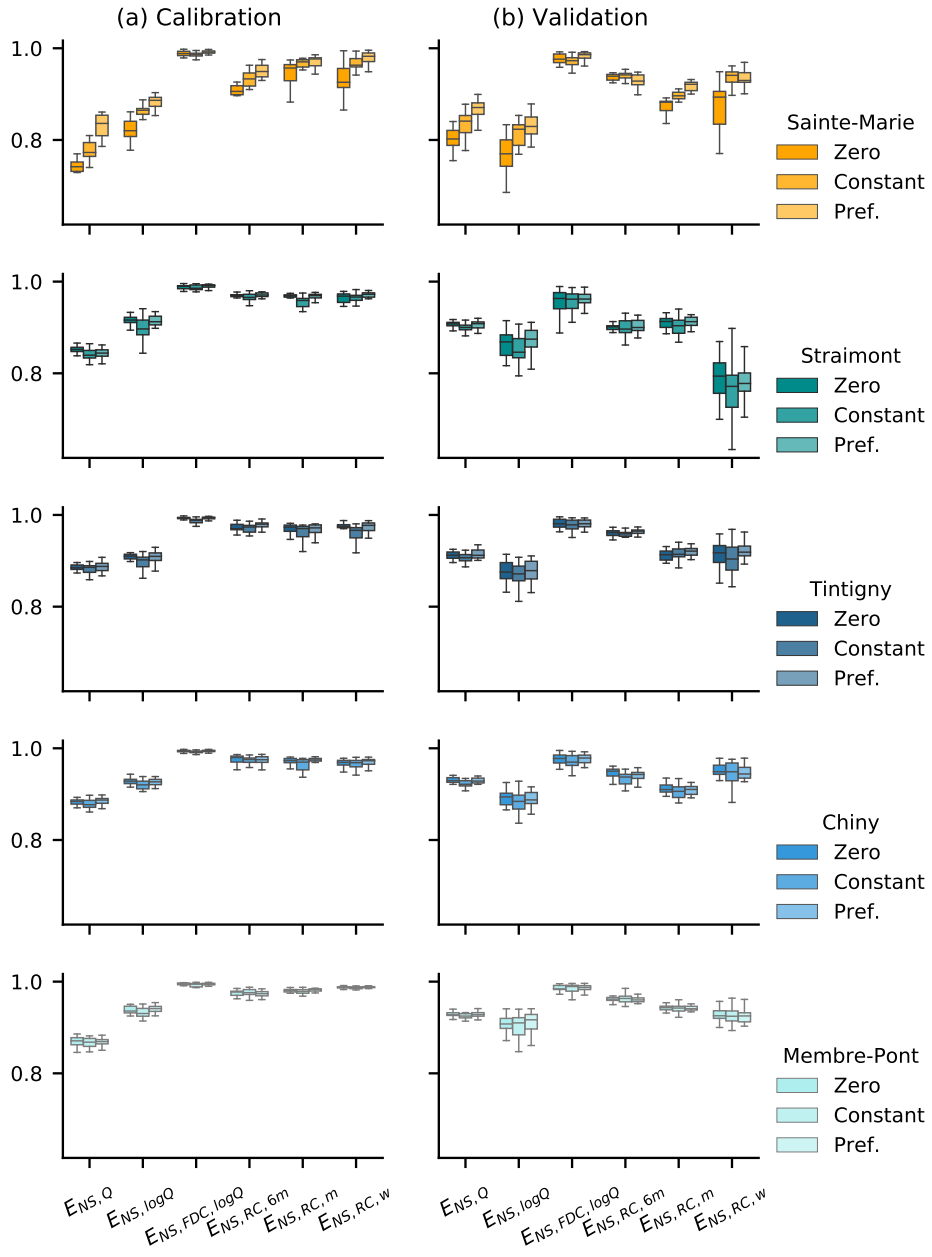

Figure 3.6: Performance indicators during the calibration (2007–2011, **a**) and the validation period (2012–2016, **b**) for the zero, constant and preferential models for the Semois at Sainte-Marie, the Vierre at Straimont, the Semois at Tintigny, the Semois at Chiny and the Semois at Membre-Pont. Including net intercatchment groundwater flows leads to an improved performance in the catchment of Sainte-Marie but not in the other catchments of the Semois.

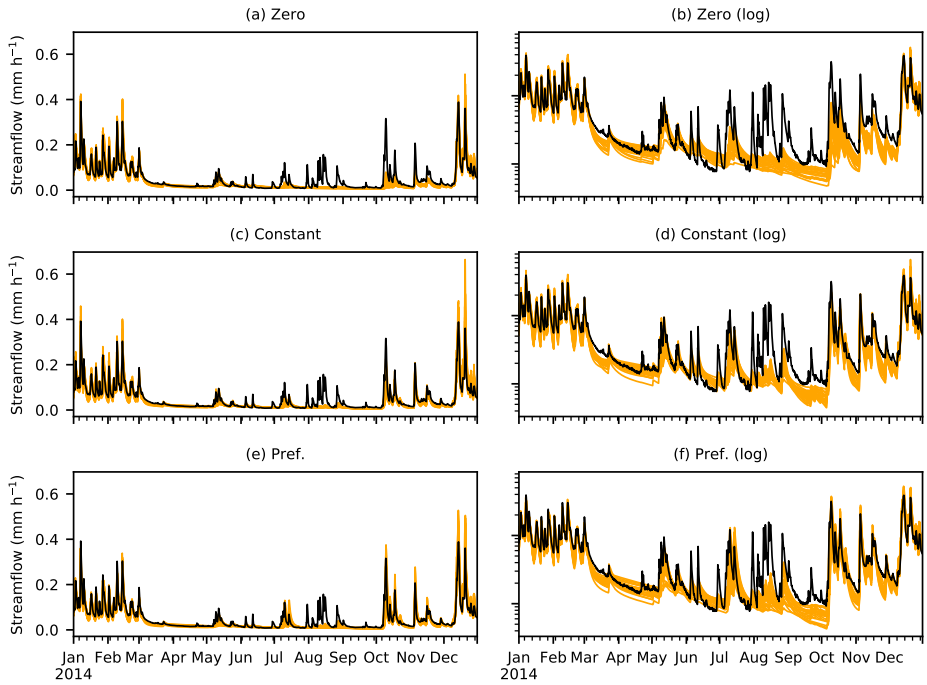

Figure 3.7: Observed (black line) and feasible realizations of modeled hydrographs (orange) in the catchment of the Semois at Sainte-Marie in 2014 for the three models (**a, b**: zero, **c, d**: constant and **e, f**: preferential model) on a normal **(a, c, e)** and log **(b, d, f)** scale. Including net intercatchment groundwater flows leads to lower simulated winter streamflow (January–March) and higher streamflow in the wetting-up period (October–November).

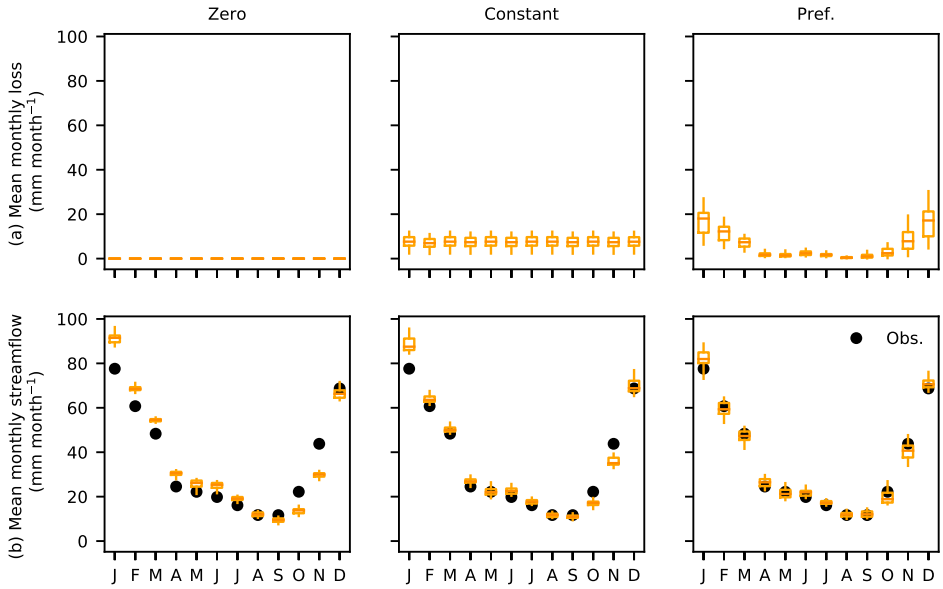

Figure 3.8: **(a)** Mean monthly loss between 2007 and 2016 for the feasible model realizations in the Semois catchment at Sainte-Marie. **(b)** Mean monthly streamflow between 2007 and 2016 for the feasible model realizations (orange) and observations (black dots) for the three models at Sainte-Marie. The preferential model leads to better performances with lower simulated streamflow in the first half-year and higher streamflow in the wetting-up period (October–November).

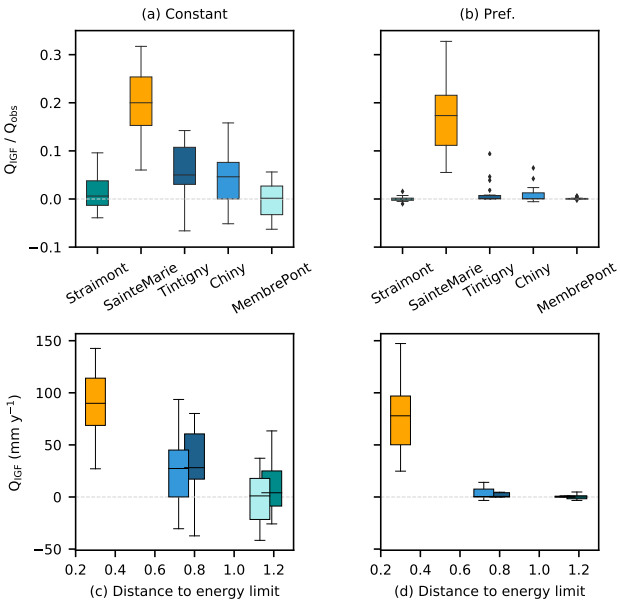

Figure 3.9: **(a, b)** Ratio of modeled net loss over observed streamflow for the Semois stations for the period 2007–2016 (positive values indicate a net loss, whereas negative values indicate a net gain) for the constant model **(a)** and the preferential model **(b)**. **(c, d)** Mean annual net intercatchment groundwater flow rates (for the feasible model realizations) as a function of the observed distance to the energy limit for the catchments of the Semois river (same color code as **a, b**) for the constant model **(c)** and preferential model **(d)**. Both models show a decrease in net intercatchment groundwater flows as the distance to the energy limit increases. The streamflow observations of all Semois stations are provided by the Service Public de Wallonie.

($535\,\mathrm{mm\,yr^{-1}}$ for the preferential model versus $590\,\mathrm{mm\,yr^{-1}}$ for the zero IGF_{net} model). The reference model compensates for the absence of an intercatchment groundwater flow term by increasing actual evaporation rates to reproduce observed flow volumes. For the majority of the other catchments, the effect of adding IGF_{net} on modeled actual evaporation rates is less pronounced, but still visible.

When using observed river streamflow and neglecting IGF_{net}, the catchment of Sainte-Marie plots close to the energy limit in the Budyko framework; however, when IGF_{net} is modeled and added to river flows, the catchment of Sainte-Marie plots close to the Turc–Mezentsev curve, as shown in Fig. 3.11. This shift in the Budyko framework occurs because we acknowledge that part of the streamflow produced from the catchment bypasses the measuring gauge in the river. Including IGF_{net} in the representation of the system results in a higher degree of plausibility, based on the Budyko framework. The shift is most obvious for the catchment of Sainte-Marie, although it also occurs in the other catchments.

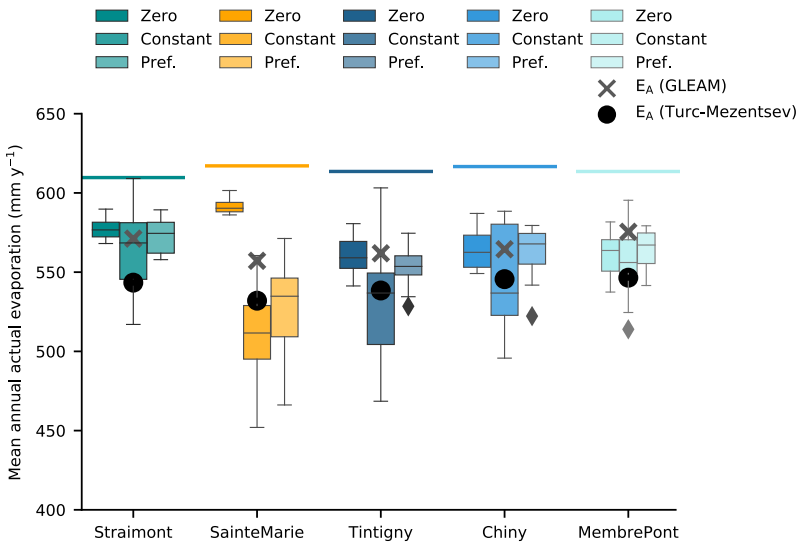

Figure 3.10: Best realizations of modeled mean annual actual evaporation in all stations in the Semois catchment for the three models (zero, constant and preferential models) during 2007–2016. Colored horizontal lines indicate mean annual potential evaporation used as forcing. Estimates of actual evaporation from the Turc–Mezentsev curve are shown as black dots and GLEAM estimates are shown as grey crosses. In the catchment of Sainte-Marie, the reference model without net intercatchment groundwater flows overestimates actual evaporation compared to the other two models and Turc–Mezentsev estimates.

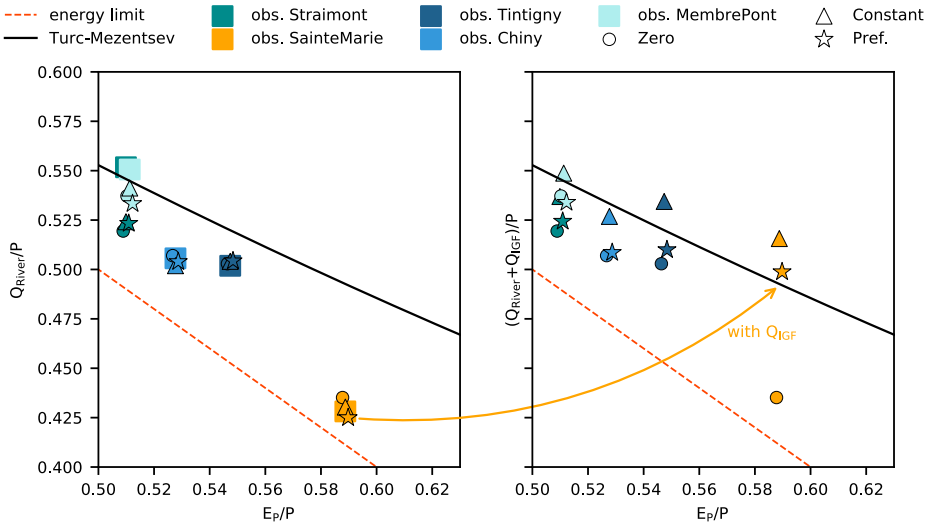

Figure 3.11: **(a)** Dimensionless representation of Q_{river}/P as a function of E_P/P. Long-term observed values between 2007 and 2016 are shown together with modeled river flows (runoff from fast and slow reservoirs) using the three models for all stations of the Semois River. **(b)** Dimensionless representation of $(Q_{river}+Q_{IGF})/P$ as a function of E_P/P. In this plot, we acknowledge that part of the groundwater bypasses the gauging station and we consider this flow in addition to the river flow. For the catchment of Sainte-Marie, we see a shift towards the Turc–Mezentsev curve when net intercatchment groundwater flows are acknowledged.

DIRECTION: GROUNDWATER BYPASS ROUTES VERSUS INTERCATCHMENT GROUNDWATER FLOWS TO EXTERNAL BASINS

The magnitude of modeled IGF_{net} decreases from nested upstream to downstream catchments along the Semois River (Fig. 3.9a,b), which is an indication that "losses" modeled at Sainte-Marie are internal to the catchment of the Semois. Losses in the upstream catchment of Sainte-Marie reappear as additional groundwater inflows in the downstream parts of the Semois, thereby reducing the IGF_{net} from upstream to downstream.

In contrast, experiments previously conducted in the Aroffe River catchment (Fister, 2012; Martin and Zany, a) revealed the presence of groundwater flows, leaving the Meuse basin towards the Moselle catchment (which is part of the Rhine basin). Losses from the Meuse basin also occur along the northern boundary of the tributary of the Geer River catchment (Reggiani and Rientjes, 2010). Additionally, downstream of the village of Bazoilles, the Meuse flows underground during a large part of the year, leaving its surface bed empty, before emerging again at Noncourt, just upstream of Neufchâteau (in the upstream part of the Meuse basin); this is referred to as "la Perte de la Meuse" (translation: the Loss of the Meuse) (Newman, 1949; Martin and Zany, b). This variety of processes highlights the contrast between stations that are losing water to neighboring catchments (Aroffe to the Rhine) and catchments that are losing water to themselves further downstream.

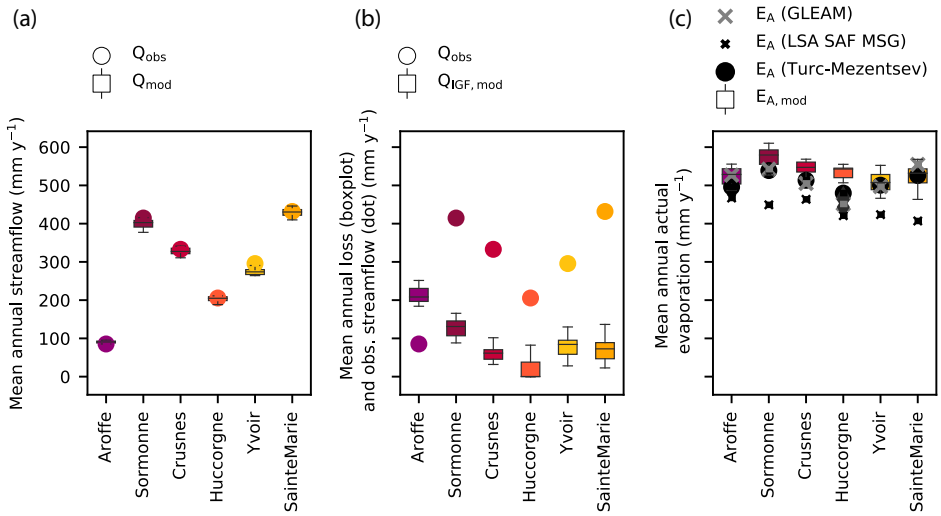

Figure 3.12: Water balance components of additional modeled catchments over the period 2007–2016. **(a)** Modeled (box plot) and observed (dot) mean annual streamflow overlap well. **(b)** Modeled mean annual net loss (box plot) and observed mean annual streamflow (dot), showing the large proportion of net intercatchment groundwater flows especially in the Aroffe catchment. **(c)** Modeled actual evaporation (box plot), GLEAM actual evaporation (grey cross), LSA SAF actual evaporation (black cross) and Turc–Mezentsev estimates (black dot). It should be noted that LSA SAF estimates are only available during the validation period (2012–2016). Model results overlap relatively well with GLEAM and Turc–Mezentsev estimates of actual evaporation, but LSA SAF estimates are lower.

MAGNITUDE: QUANTIFICATION OF NET INTERCATCHMENT GROUNDWATER FLOWS AT THE SCALE OF THE MEUSE BASIN

The magnitude of IGF_{net} is assessed in several other catchments of the Meuse basin that plot close to or beyond the energy limit (annotated catchments in Fig. 3.1a). The preferential model is used to assess the magnitude of IGF_{net} because it performed better in the catchment of Sainte-Marie. In the Aroffe catchment, an overflow type of model is applied to represent the functioning of the system based on a priori available knowledge (Sect. 3.4.2). Modeled mean annual flows between 2007 and 2016 overlap well with observations, as shown in Fig. 3.12a. The ratio of mean annual net intercatchment groundwater flows to observed streamflow is always positive (indicating a loss). Modeled losses can be substantial compared to observed streamflow, as shown in Fig. 3.12b. In the Aroffe, the median loss rate (of 208 mm yr^{-1}) is approximately 2.5 times higher than observed river flows (85 mm yr^{-1}). Median values of annual loss rates over observed streamflow range from 0.1 % to 32 % (0.3 to 130 mm yr^{-1}) in the other catchments. Modeled actual evaporation is close to or slightly overestimates Turc–Mezentsev estimates (Fig. 3.12c), showing that the models are able to reproduce the observed long-term water balance in a meaningful way.

At the scale of the Meuse basin, intercatchment groundwater flow processes play only a small role because they occur in relatively small catchments and because part of these losses may be internal to the Meuse basin. However, IGF_{net} occurring at the scale of headwater catchments make up a considerable part of the water balance (on average 10 % and up to 25 % of mean annual precipitation), which in many current models is wrongly attributed to actual evaporation.

3.5.3 EVALUATION AGAINST ACTUAL EVAPORATION FROM REMOTE SENSING

GLEAM estimates of mean annual actual evaporation approximate or slightly overestimate (< 5 %) modeled and Turc–Mezentsev estimates of actual evaporation, as shown in Figs. 3.10 and 3.12c, whereas estimates from a land surface modeling approach, such as LSA SAF data, are considerably lower (between 400 and 470 mm yr^{-1}, Fig. 3.12c) in the studied catchments. While the difference in both products highlights the uncertainty in remote-sensing-based estimates of actual evaporation, it also shows that actual evaporation might even be less than that resulting from our models, which might imply even larger magnitudes of losses due to IGF_{net}. The simple conceptualization of soil-moisture-constrained evaporation used in our models, which does not account for a temperature-based stress function, might lead to an overestimation of transpiration. Thus, being arguably conservative modeled estimates, the low estimates of LSA SAF evaporation lend further credibility to evidence suggesting the presence of considerable IGF_{net}.

3.6 DISCUSSION

3.6.1 IMPLICATIONS

In this study, we question in three steps the validity of neglecting intercatchment groundwater flows in catchment-scale hydrological studies. In the Meuse basin, the potential presence of net intercatchment groundwater flows is detected from observed water balance data in relatively small headwater catchments (< 500 km^2) and is much less pronounced in larger downstream catchments (Fig. 3.1). In the theory advanced by Tóth (1963), regional groundwater flows occur from the headwaters to the bottom of the basin. This implies that headwater catchments may export water through groundwater flow paths into the river further downstream, thereby increasing the groundwater contributions in larger downstream catchments; this suggests a variability of dominant hydrological (subsurface) processes across spatial scales, as also demonstrated by Frisbee et al. (2011). Schaller and Fan (2009) found that the largest magnitudes of intercatchment groundwater flow occur at catchment size near 100 km^2, which also results from our analysis (Fig. 4.10c). Catchment size might not be the most important control (as compared to geology or topography), but it is to some extent a proxy of the position of a catchment in the landscape. We could show that the largest evidence for intercatchment groundwater flows occurs in small headwater catchments, whereas in the lowlands, where the proportion of larger catchments is also higher, net intercatchment flow is of less relevance: the further downstream the catchment is situated, the more the losses that occurred upstream become accounted for. Schaller and Fan (2009) also report that efficient aquifers favor intercatchment groundwater flows. In the Meuse basin, the identified headwater catchments are relatively flat and underlain by highly productive and fissured aquifers (Fig. 4.10b,c), where karstification processes might be at the origin of underground exchange flow paths between catchments. While previous research (Gleeson and Manning, 2008; Ameli et al., 2018) showed that more intercatchment groundwater flows are to be expected with increasing catchment slope, these studies assume a homogeneous subsurface, which is not the case in the Meuse basin. The relatively weak correlations between physical catchment characteristics and intercatchment groundwater flows shown in Fig. 4.10 can be explained by the high spatial variability of intercatchment groundwater flows due to local geological features that overrule theoretical relations at the basin scale, as also argued by Genereux et al. (2002), Schaller and Fan (2009) and Frisbee et al. (2016).

We make one of the first steps to bridge the gap between regional groundwater models in which topographic catchment boundaries are not considered and lumped conceptual hydrological models that treat catchments as well-defined impermeable entities, by adding an additional flux in conceptual models to represent net intercatchment groundwater flows. We model net intercatchment groundwater flows as preferential fluxes, occurring when recharge exceeds a threshold, to represent the filling of underground stores before intercatchment flow paths are activated (Fig. 3.4c), rather than as constant matrix flow. Interestingly, we show that accounting for net preferential intercatchment groundwater flows improves not only low-flow performance indicators, but also high-flow simulations (Fig. 3.6). The increased performance achieved with the pref-

erential model during both high and low flows suggests that the relative importance of intercatchment groundwater flow processes change throughout the year, as also found by Frisbee et al. (2012) based on a chemical and isotopic analysis. Ajami et al. (2011) also suggest that local, intermediate and regional groundwater flow paths are active during winter, while mainly local groundwater flow paths are active during summer. The ratio of net intercatchment groundwater exports over total streamflow ($Q_{IGF}/(Q_{IGF} + Q_{river})$) is about 70 % in the Aroffe catchment (where the flow is diverted into the neighboring Moselle river) and is on average 17 % in the other catchments; these values are within the range provided by Schaller and Fan (2009).

3.6.2 LIMITATIONS

In this work, we rely on the empirical organizing principle provided by the Turc–Mezentsev or Budyko curves (Turc, 1954; Mezentsev, 1955; Budyko, 1961) and assume that catchments of the Meuse basin plotting close to or beyond the energy limit (Fig. 3.3) may be subject to losses due to net intercatchment groundwater flows. Changing vegetation, climate and human interactions might, however, also be at the origin of catchments deviating from the Budyko curve (Velde et al., 2014; Berghuijs et al., 2014). The location of each catchment within the Budyko framework is also subject to uncertainties in the data used to calculate long-term mean precipitation, streamflow and potential evaporation. Data uncertainties can originate from the spatial interpolation of the precipitation, the choice of a potential evaporation formula, errors in streamflow measurements or in catchment delineation, or the presence of unknown anthropogenic activities affecting the water balance. The Budyko framework relies on long-term mean precipitation, potential evaporation and streamflow data as well as the underlying assumption that changes in total catchment water storage (for example in lakes, wetlands or groundwater) can be treated as negligible, i.e $dS \sim 0$ [$L\,T^{-1}$]. However, the assumption of $dS \sim 0$ [$L\,T^{-1}$] is unlikely to hold over shorter time periods, which implies that the framework cannot be used in a meaningful way to estimate short-term differences (e.g., annual) in intercatchment groundwater flows. In spite of these shortcomings, the three-step approach of this study, which combines different perspectives and data to estimate net intercatchment groundwater flows, allowed us to plausibly attribute deficits in the observed water balance to the potential presence of net intercatchment groundwater flows.

We treated intercatchment groundwater flows as independent net losses or gains in lumped conceptual catchment models, without explicitly connecting the loss of one catchment to the gain of another. By modeling several stations along the same tributary (the Semois), we hypothesized that the loss in the headwater catchment at Sainte-Marie might bypass the channel to reach the river only further downstream, implying an "internal" loss within the river system; but other configurations of groundwater flows in this area might lead to similar results. Additionally, we found evidence in literature (Fister, 2012) that the Aroffe catchment flows underground to the Moselle catchment (a tributary of the Rhine), but we could not relate the flow out of the relatively small Aroffe catchment (198 km^2) to its emergence in the much larger Moselle catchment near Toul (3338 km^2) due to the difference in catchment area. Interestingly, in the recent geological past (250 000 years ago), the upstream catchment of the Moselle at Toul was flow-

ing through the Meuse valley before it changed course to join the Rhine basin (de Wit, 2008b). Subsurface flow paths connecting both catchments may therefore still remain from these earlier geological times.

We use independent data sources of remotely sensed actual evaporation estimates to quantify the overestimation of actual evaporation modeled when intercatchment groundwater flows are neglected. Both global actual evaporation products (GLEAM and LSA SAF) rely on different models and remotely sensed data and provide relatively large differences in mean annual values (up to 150 mm yr^{-1}), highlighting the large uncertainty in estimating actual evaporation. While GLEAM actual evaporation estimates approximate our model results and Turc–Mezentsev estimates, LSA SAF estimates indicate lower evaporation rates (Figs. 3.10, 3.12c), potentially indicating an underestimation of actual evaporation in this area, or the even larger importance of losses due to net intercatchment groundwater flows in the studied catchments.

3.7 CONCLUSION

This proof-of-concept study in the Meuse basin shows strong evidence that we can identify net intercatchment groundwater flow processes from analyzing the long-term observed water balance of a catchment. The results suggest that intercatchment groundwater flows mainly play a role in headwater catchments (< 500 km^2) with productive aquifers. In these catchments, we then use simple conceptual models to show that a net groundwater loss occurs when recharge exceeds a threshold. This preferential net loss term represents the filling of underground stores before intercatchment flow paths are activated, and ranges between 0 and 208 mm yr^{-1} (0 % and 25 % of annual precipitation) with an average of 100 mm yr^{-1} (10 % of mean annual precipitation) in the studied catchments (Fig. 3.12b). Some of these underground flow paths may lead to downstream catchments along the same river (regional groundwater flow paths), while others may lead to neighboring river basins (diverted groundwater flows due to the presence of geological features), which explains why these net losses can be considerable at the headwater catchment scale and negligible at the scale of larger catchments (modeled net intercatchment groundwater flows reduced to zero at the most downstream station of the Semois tributary). These findings therefore highlight that dominant streamflow generation processes vary across spatial scales. Additionally, errors in simulating the seasonal behavior are reduced by more than 70 % with the preferential model (Fig. 3.8b), this suggests a pronounced intra-annual variability of the magnitude of net intercatchment groundwater flow processes. Neglecting net intercatchment groundwater flows in conceptual models may still result in high performances of streamflow simulation; however, it comes at the cost of overestimating actual evaporation rates to compensate for this lack (Fig. 3.10). Including net intercatchment groundwater flow processes in models can considerably increase the correspondence between modeled actual evaporation and remote-sensing estimates, and this provides additional evidence for the presence and magnitude of net intercatchment groundwater flows.

4

GETTING TO THE ROOT OF THE T-VALUE: LINKING THE ROOT-ZONE STORAGE CAPACITY WITH THE SOIL WATER INDEX AT THE CATCHMENT SCALE

'Nature is a living whole,' he [Humboldt] later said, not a 'dead aggregate'.

Andrea Wulf - *The Invention of Nature* (2015)

In the previous chapter, we used the long-term water balance to evaluate the presence of intercatchment groundwater flow processes across the Meuse river basin. Once the different components of the long-term water balance in a specific catchment are identified, they can be used to estimate the water storage accessible to the roots of vegetation for transpiration. This root-zone storage capacity ensures the long-term survival of vegetation to overcome droughts. The spatiotemporal dynamics of root-zone soil moisture can be inferred from remote-sensing estimates of near-surface soil moisture through the Soil Water Index. The aim of this chapter is to link the catchment-scale root-zone storage capacity to the Soil Water Index.

Parts of this chapter have been published in Water Resources Research (Bouaziz et al., 2020)

SUMMARY

The spatio-temporal dynamics of water volumes stored in the unsaturated root-zone are a key control on the response of terrestrial hydrological systems. Robust, catchment-scale root-zone soil moisture estimates are thus critical for reliable predictions of river flow, groundwater recharge or evaporation. Satellites provide estimates of near-surface soil moisture that can be used to approximate the moisture content in the entire unsaturated root-zone through the Soil Water Index (SWI). The characteristic time length (T, in days), as only parameter in the SWI approach, characterizes the temporal variability of soil moisture. The factors controlling T are typically assumed to be related to soil properties and climate, however, no clear link has so far been established. In this study, we hypothesize that optimal T values (T_{opt}) are linked to the interplay of precipitation and evaporation during dry periods, thus, to catchment-scale vegetation-accessible water storage capacities in the unsaturated root-zone. We identify T_{opt} by matching modeled time series of root-zone soil moisture from a calibrated process-based hydrological model to SWI from several satellite-based near-surface soil moisture products in 16 contrasting catchments in the Meuse river basin. T_{opt} values are strongly and positively correlated with vegetation-accessible water volumes that can be stored in the root-zone, here estimated for each study catchment both as model calibration parameter and from a water balance approach. Differences in T_{opt} across catchments are also explained by land cover (% agriculture), soil texture (% silt) and streamflow signatures (flashiness index).

4.1 INTRODUCTION

Catchment-scale estimates of water volumes stored in the unsaturated root-zone are a key element regulating the partitioning of water fluxes in terrestrial hydrological systems (Savenije and Hrachowitz, 2017). The spatio-temporal dynamics of this moisture content control how much additional water can be stored in the soil of the unsaturated root-zone and how much water is thus available for plant transpiration. Capillary forces in soils have the ability to retain water against gravity to delay drainage. Conversely, excess water that cannot be held against gravity is released from soils as lateral flow and/or groundwater recharge to eventually generate the runoff response in streams. For reliable estimates of streamflow, accurate estimates of soil moisture contents in the unsaturated root-zone are thus required (Blöschl and Zehe, 2005).

However, in-situ soil moisture observations are typically not available at sufficient spatio-temporal scales and resolutions, except for in a handful of small experimental catchments (e.g., Bogena et al., 2010). An increasing number of studies has therefore previously explored the value of globally available remotely-sensed soil moisture estimates for calibration and evaluation of or assimilation in different types of hydrological models (e.g., Crow et al., 2005; Beck et al., 2009; Parajka et al., 2009; Sutanudjaja et al., 2014; Wanders et al., 2014; Silvestro et al., 2015; Leroux et al., 2016; López López et al., 2016, 2017; Rakovec et al., 2016a; Tian, 2007; Nijzink et al., 2018; Gevaert et al., 2018).

The use of remotely-sensed soil moisture products has proven to be relevant for a variety of hydrological applications, however, a key issue is that the raw products often lack direct hydrological relevance. Depending on the satellite mission, the soil moisture estimates are generally limited to the upper-most few centimeters of the soil. This "near-surface soil moisture" is in itself uninformative to quantify water release to streams, which is regulated by the integrated moisture content over the entire unsaturated root-zone. Establishing a systematic quantitative link between near-surface soil moisture and the hydrologically relevant soil moisture in the unsaturated root-zone is therefore of critical importance, but remains challenging (McCabe et al., 2017; Sheffield et al., 2018; Blöschl et al., 2019).

Analytical, statistical or modelling methods have been proposed to estimate soil moisture in the unsaturated root-zone by smoothing and delaying the near-surface soil moisture signal (Ragab, 1995; Entekhabi et al., 1994; Wagner et al., 1999; Mahmood and Hubbard, 2007; Sabater et al., 2007; Manfreda et al., 2014). For example, Wagner et al. (1999) proposes a simplified two-layer model to convolve the near-surface soil moisture signal to a Soil Water Index (SWI) using an exponential filter. The underlying assumption implies that the water fluxes from the surface through the entire root-zone is proportional to the difference in soil moisture between both. Stroud (1999) and Albergel et al. (2008) reformulate the exponential filter to a recursive filter to handle data more easily than the original formulation. Despite its simplicity and lack of explicit link to physical processes, SWI provides operationally useful estimates of aggregated soil moisture content in the soil profile (Ceballos et al., 2005; Albergel et al., 2008; Brocca et al., 2010a, 2011; Albergel et al., 2012; Ford et al., 2014).

The proposed recursive filter for the SWI calculation requires a single parameter, the

characteristic time length T, expressed in unit of time (usually days), which is proportional to the ratio of the depth of the reservoir below the surface and a pseudo-diffusivity coefficient. The larger the value of T, the more smoothing and delaying of the near-surface soil moisture signal occurs in the estimation of the root-zone soil moisture. The parameter T has been interpreted to be a lumped surrogate for multiple interacting processes influencing soil moisture dynamics, such as soil depth, evaporation, streamflow, soil hydraulic properties or vegetation cover (Ceballos et al., 2005).

Optimal values of the characteristic time lengths T_{opt} are often derived by relating satellite-derived SWI to observed and/or modeled soil moisture time series integrated to different depths (Wagner et al., 1999; Albergel et al., 2008; Ceballos et al., 2005; Brocca et al., 2011; de Lange et al., 2008; Beck et al., 2009; Ford et al., 2014; Paulik et al., 2014). T_{opt} values increase with increasing depth of soil moisture measurements or deeper soil moisture profiles (up to 100 cm), implying less temporal variability of soil moisture at increasing depths (Wagner et al., 1999; Albergel et al., 2008; Ceballos et al., 2005; Paulik et al., 2014; Wang et al., 2017).

To infer SWI from near-surface soil moisture in areas where no in-situ soil moisture observations are available, it is important to understand the factors controlling the characteristic time length T. Previous studies assessed either soil properties and/or climate as main controls to explain the variability of T_{opt} values.

The role of climate as a control on T_{opt} was assessed by Albergel et al. (2008) and Wang et al. (2017). Albergel et al. (2008) suggest that a weak relation with climate may exist, with lower values of T_{opt} in areas with higher evaporative demand and less frequent but more intense precipitation. Wang et al. (2017) did not find a correlation with mean annual potential evaporation. However, they found a negative correlation with mean annual precipitation, suggesting that areas with mean annual precipitation larger than 500 mm yr^{-1} have smaller values of T_{opt} due to stronger hydraulic connections between the surface and deeper layers.

Ceballos et al. (2005), de Lange et al. (2008) and Wang et al. (2017) found that soil texture is an important control on T_{opt}. Sandy soils enable fast drainage and low water retention as compared to clayey soils, therefore resulting in a low temporal persistence of soil moisture in the system. This implies a stronger similarity between the near-surface and deeper soil moisture and therefore lower T_{opt} values for sandy than for clayey soils (Ceballos et al., 2005; Wang et al., 2017). On the other hand, de Lange et al. (2008) report higher T_{opt} values for sandy than clayey soils. Albergel et al. (2008) and Paulik et al. (2014) did not find a correlation between T_{opt} and fractions of clay and sand.

These contradicting findings hinder our efforts to estimate T-values that represent the integrated soil moisture content in the unsaturated root-zone. As a result of the inconclusive role of soil properties and climate, T_{opt} values of 20 days are often assumed from literature to represent soil moisture in the first 100 cm of the soil (Wagner et al., 1999). However, the vegetation accessible water storage, which is a key variable in hydrological applications as it controls the partitioning between drainage and evaporation, does not necessarily correspond to soil moisture in the first 100 cm of the soil.

Despite the important hydrological role of the root-zone storage capacity, it is diffi-

cult to observe it at the catchment scale. Heterogeneity within catchments impedes our ability to integrate soil and plant root properties beyond point-scale measurements. The root-zone storage capacity is, therefore, often estimated from calibration of a hydrological model, which includes a representation of this storage capacity in the dynamic part of the unsaturated root-zone. However, the derived catchment representative value of root-zone storage capacity may be subject to equifinality, even if additional data is used to constrain the calibration, making it difficult to assess its plausibility (de Boer-Euser et al., 2016).

Recent studies have demonstrated that root-zone storage capacities can be estimated from the amount of water accessible to vegetation for transpiration (Gao et al., 2014; Wang-Erlandsson et al., 2016; Nijzink et al., 2016a; de Boer-Euser et al., 2016). The underlying assumption is that vegetation creates a buffer large enough to fulfill evaporative demand and overcome dry spells with a certain return period to ensure their long-term survival (Eagleson and Tellers, 1982; Milly, 1994). The interplay between water supply through precipitation and evaporative water demand, accumulated over dry periods, enables us to estimate catchment-scale vegetation-accessible water storage capacities in the unsaturated root-zone (Kleidon and Heimann, 1998; Schymanski et al., 2008; Donohue et al., 2012; Gentine et al., 2012; Gao et al., 2014; Nijzink et al., 2016a).

The overarching aim of our study is to establish a quantitative link between optimal T-values and catchment-scale vegetation accessible water storage capacities. This allows us to make informed decisions on optimal T-values to meaningfully use near-surface soil moisture data in hydrological models. We first identify T_{opt} values that maximize the correlation between time series of SWI derived from several remotely-sensed products and modeled root-zone soil moisture using a process-based hydrological model. We test this in 16 contrasting catchments of the Meuse river basin (Fig. 4.1 and Table 4.1) to highlight the large variability of T_{opt} values and therefore the need to adequately estimate T_{opt}. We then test the relation between T_{opt} and root-zone water storage capacities derived as calibrated model parameter. This is useful to emphasis the role of T_{opt} in providing root-zone soil moisture estimates consistent with their representation in hydrological applications. Finally, we use a well-established method relying on the interplay between precipitation and evaporation to independently estimate catchment-scale vegetation accessible water storage capacities (Milly, 1994; Nijzink et al., 2016b; de Boer-Euser et al., 2016) and test for the relation with T_{opt} to answer our overarching hypothesis.

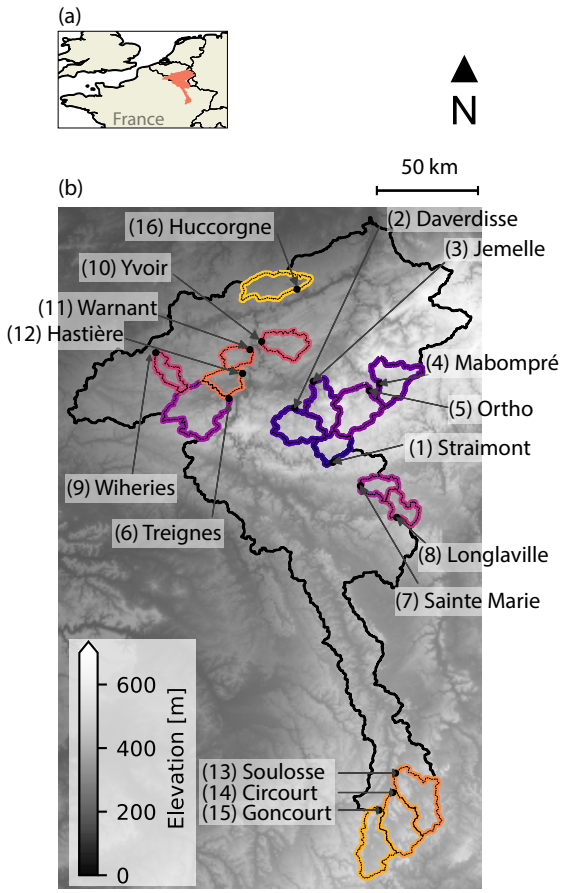

Figure 4.1: (a) Outline of the Meuse river basin upstream of Borgharen at the border between Belgium and the Netherlands, (b) Digital Elevation Model and 16 studied catchments with areas varying between 127 and 551 km^2.

4.2 DATA

4.2.1 SATELLITE-BASED NEAR-SURFACE SOIL MOISTURE PRODUCTS

The satellite-based near-surface soil moisture products used in this study included both, passive and active microwave products, for L, C and X bands at spatial resolutions of 100 m x 100 m and 1 km x 1 km, respectively, as provided by VanderSat (http://docs.vandersat.com/index.html, https://patents.google.com/patent/WO2017216186A1/en) and Copernicus (Sentinel-1 based Surface Soil Moisture SSM1km, Bauer-Marschallinger et al. (2019)), as well as at spatial resolutions of 9 km x 9 km or 25 km x 25 km as provided by the National Aeronautics and Space Administration (NASA, Owe et al. (2008); Entekhabi et al. (2016)), as detailed in Table 4.2.

Copernicus SSM1km applies a change detection method on Sentinel-1 C-SAR backscatter values to derive relative soil moisture in percentage saturation. VanderSat products are based on the Land Parameter Retrieval Model (LPRM, van der Schalie et al. (2016)) method to estimate near-surface soil moisture, using descending overpasses at 6 AM for the Soil Moisture Active Passive satellite (SMAP) and 1.30 AM for the Advanced Microwave Scanning Radiometer 2 (AMSR2). The NASA SMAP Level-3 SPL3SMP-E product (Entekhabi et al., 2016) is based on a composite of ascending and descending overpasses and uses the single channel algorithm to estimate soil moisture (Entekhabi et al., 2014). The NASA AMSR2 soil moisture products are based on the LPRM algorithm (Owe et al., 2008) and are available for both descending and ascending overpasses. Brocca et al. (2011) mention that some studies find higher correlations between ascending overpasses and in-situ measurements, while others favor night-time descending overpasses to take advantage of the reduced difference between surface and canopy temperature, making it relevant to assess both ascending and descending overpasses in this study.

The data of all used products were spatially averaged in the 16 catchments of the Meuse basin for the study period 05/2016 until 12/2017. Radio frequency interferences affect soil moisture estimations of NASA AMSR2 C1 and C2 bands and Copernicus Sentinel-1 C-band Synthetic Aperture Radar (SAR) products for 01/2015 until 05/2016 and this period was, therefore, left out from the analyses. Whenever mean daily catchment temperature (Sect. 4.2.3) dropped below $1°C$, near-surface soil moisture estimates were set to missing to avoid potential biases resulting from frozen soils.

4.2.2 SOIL WATER INDEX (SWI)

The Soil Water Index (SWI) approach estimates time series of soil moisture in deeper layers from antecedent satellite-based near-surface soil moisture estimates. The approach relies on a two-layer water balance model, where the upper layer represents the near-surface soil moisture and the bottom layer is a deeper reservoir with length L only in contact with the surface layer. The soil moisture content in the deeper layer is fed by infiltration from the upper layer and is therefore explained by past dynamics, where more recent events have a stronger influence. This model assumes that the water flux from the surface through the entire deeper layer is proportional to the difference in soil moisture between the two layers, as shown by Eq. 4.1.

Table 4.1: Mean annual precipitation (P), potential evaporation (E_P), streamflow (Q), aridity index (E_P/P), runoff ratio (Q/P), flashiness index (I_f) for hydrological years 2006-2011; and catchments characteristics including percentage forest (for.) and agriculture cover (agr.), percentage of highly productive and fissured aquifers (fiss.) based on the International Hydrogeological Map of Europe (IHME, https://www.bgr.bund.de/), percentage clay, sand and silt (Ballabio et al., 2016), catchment area, drainage density (dd), mean slope, of study catchments with IDs sorted on aridity index from the most to the least humid catchment.

(ID) Station	P	E_P	Q	E_P/P	Q/P	I_f	for.	agr.	fiss.	clay	sand	silt	area	dd	slope
	mm y^{-1}			%	%	%	%	%	%	%	%	%	km^2	km^{-1}	%
(1) Straimont	1187	574	621	50	52	17	34	30	0	22	16	62	183	0.36	6.7
(2) Daverdisse	1142	568	555	51	49	15	57	28	0	21	19	61	303	0.32	7.7
(3) Jemelle	1005	563	477	56	47	18	68	20	0	20	25	55	274	0.40	9.4
(4) Mabompre	985	572	460	59	47	16	46	28	0	21	24	55	319	0.47	7.4
(5) Ortho	990	574	470	60	47	14	40	33	0	22	21	58	387	0.37	7.3
(6) Treignes	985	579	398	60	41	28	54	27	0	22	21	57	551	0.33	6.6
(7) Sainte-Marie	1043	600	418	60	40	36	38	26	63	23	17	60	144	0.28	4.4
(8) Longlaville	971	608	440	65	45	39	20	28	18	28	20	52	154	0.16	6.9
(9) Wiheries	889	590	356	67	40	30	19	43	0	25	15	59	140	0.33	4.9
(10) Yvoir	865	577	264	68	31	13	16	60	71	24	13	63	226	0.33	6.4
(11) Warnant	819	586	275	72	34	12	20	64	56	24	15	61	127	0.29	6.2
(12) Hastiere	802	582	285	73	36	32	41	40	0	24	21	55	169	0.40	5.4
(13) Soulosse	831	642	331	79	39	35	30	26	38	34	15	51	441	0.30	6.3
(14) Circourt	823	642	312	80	38	42	41	12	23	33	18	49	403	0.30	7.3
(15) Goncourt	815	646	295	81	36	40	19	23	19	34	14	52	376	0.27	5.6
(16) Huccorgne	737	593	181	82	25	19	3	80	16	21	9	70	307	0.25	2.6

Table 4.2: Satellite surface soil moisture products with associated IDs (where d indicates downscaled high-resolution product), provider, sensor, band (where C1 and C2 corresponds to 6.9 and 7.3 GHz, respectively), method or product name, overpass, resolution and range of observation count per catchment between 05/2016 and 12/2017 (Nr. obs.).

ID	Provider	Sensor	Band	Method/ Product	Overpass	Resol. km^2	Nr. obs.
S1-CSAR	Copernicus	Sentinel-1	C-SAR	SSM1km	6AM 6PM	1x1	199-290
SMAP-L-am_d	VanderSat	SMAP	L	LPRM	6 AM	0.1 x 0.1	297-415
SMAP-L	NASA	SMAP	L	SPL3SMP-E	6AM 6PM	9 x 9	316-345
AMSR2-X-am_d	VanderSat	AMSR2	X	LPRM	1.30 AM	0.1 x 0.1	511-572
AMSR2-X-am	NASA	AMSR2	X	LPRM	1.30 AM	25 x 25	508-556
AMSR2-X-pm	NASA	AMSR2	X	LPRM	1.30 PM	25 x 25	495-570
AMSR2-C1-am_d	VanderSat	AMSR2	C1	LPRM	1.30 AM	0.1 x 0.1	511-571
AMSR2-C1-am	NASA	AMSR2	C1	LPRM	1.30 AM	25 x 25	508-556
AMSR2-C1-pm	NASA	AMSR2	C1	LPRM	1.30 PM	25 x 25	495-570
AMSR2-C2-am	NASA	AMSR2	C2	LPRM	1.30 AM	25 x 25	508-556
AMSR2-C2-pm	NASA	AMSR2	C2	LPRM	1.30 PM	25 x 25	495-570

$$L \cdot \frac{\mathrm{d}W(t)}{\mathrm{d}t} = C \cdot (W_s(t) - W(t)) \tag{4.1}$$

With W the moisture content in the lower reservoir ($m^3\ m^{-3}$ or percentage saturation), L the depth of the reservoir [m], W_s the surface soil moisture content ($m^3\ m^{-3}$ or percentage saturation) and C represents a pseudo-diffusivity coefficient dependent on soil moisture and is assumed constant [m d^{-1}]. The equation is solved by assuming a constant parameter T [d], which increases with increasing soil depth or decreasing pseudo-diffusivity constant (Eq. 4.2).

$$T = \frac{L}{C} \tag{4.2}$$

As only parameter in the SWI, the T-value is referred to as the characteristic time length. More smoothing and delaying of the near-surface soil moisture signal occurs as the T-value increases.

Stroud (1999) and Albergel et al. (2008) solve the differential equation in a recursive form to handle the irregular time steps of satellite near-surface soil moisture data more easily than the original exponential filter proposed by Wagner et al. (1999). Details are provided in Appendix B.

$$SWI(t_n) = SWI(t_{n-1}) + K_n \cdot (SSM(t_n) - SWI(t_{n-1})) \tag{4.3}$$

Where SWI, the Soil Water Index at time t_n, has replaced the continuous W and SSM, the near-surface soil moisture estimate at time t_i, has replaced the continuous W_s (all in $m^3\ m^{-3}$ or percentage saturation).

The gain K_n [-] at time t_n is given by the following recursive formula:

$$K_n = \frac{K_{n-1}}{K_{n-1} + e^{-\frac{(t_n - t_{n-1})}{T}}} \tag{4.4}$$

With initial values $SWI(t_0) = SSM(t_0)$ and $K_0 = 1$, and where t_n and t_{n-1} are the observation times of the current and previous SSM observation in Julian days. The gain K_n ranges between [0,1]. When many observations during the characteristic time length T are available, the gain will be small, meaning that the prior value will only be changed slightly towards the new observation. On the other hand, when no data has been received in quite some time relative to T, the gain will be large implying that the new estimate of SWI will converge towards the value of the new observation. Therefore, the T-value dictates how strongly previous near-surface soil moisture observations influence the current SWI.

For each of the products and study catchments, we calculate time series of daily catchment average SWI from near-surface soil moisture for values of T varying between 1 and 100 days (with time step of 1 day), according to the recursive formulation (Eq. 4.3) for the period 05/2016 - 12/2017. Spin-up effects are reduced by starting the SWI calculation at the start of 2016.

4.2.3 METEOROLOGICAL AND STREAMFLOW DATA

Hourly precipitation data from stations of Meteo France and the Service Public de Wallonie (Service Public de Wallonie (2018), http://voies-hydrauliques.wallonie.be/opencms/opencms/fr/hydro/Archive/annuaires/index.html), are spatially interpolated using climatological monthly background grids (van Osnabrugge et al., 2017) and then spatially averaged over the catchments for a calibration (2006-2011) and evaluation (2012-2017) period. Potential evaporation is estimated based on the Makkink equation (Hooghart and Lablans, 1988) and relies on hourly interpolated temperature station data (provided by the Service Public de Wallonie and retrieved from the archive of the Dutch operational flood forecasting system) and radiation grids from the Satellite Application Facility on Land Surface Analysis (LSA SAF, Trigo et al. (2011)). Daily observed river streamflow for stations in France (IDs 8, 13-15) are retrieved from Hydro Banque (Banque Hydro (2018), http://hydro.eaufrance.fr/) and are provided by the Service Public de Wallonie for Belgium (IDs 1-7, 9-12, 16).

Streamflow data between March and mid-June 2013 were discarded from the analysis due to implausibly high streamflow compared to observed precipitation amounts at all stations except in the Vair at Soulosse-sous-Saint-Élophe (ID13), Mouzon at Circourt-sur-Mouzon (ID14) and Meuse at Goncourt (ID15) (Bouaziz et al., 2018).

4.3 METHODS

Sect. 4.3.1 describes the set-up of a process-based lumped hydrological model to compute soil moisture time series. Sect. 4.3.2 details the methodology to derive water balance estimates of root-zone storage capacities. Subsequently in Sect. 4.3.3, the optimal characteristic time lengths (T_{opt}) in the 16 study catchments are derived using the retained set of feasible model-generated time series of daily soil moisture contents in the unsaturated root-zones (S_R). In Sect. 4.3.4, we then test for a relation between root-zone water storage capacity ($S_{R,max}$) inferred from the set of calibrated model parameters $S_{R,max}$ retained as feasible for each study catchment and from a water balance approach. We also assess how the interactions between soil texture, land cover, hydrometeorological variables, streamflow signatures, geological features and topographical indices explain the variability in T_{opt} between catchments.

4.3.1 HYDROLOGICAL MODEL

A process-based lumped hydrological model (Fig. 4.2a) is set-up and calibrated for each of the 16 study catchments to estimate the dynamics of catchment-scale, daily soil moisture content in the unsaturated root-zone. The model consists of four storage components, including an interception reservoir (S_I), a reservoir representing the unsaturated root-zone (S_R), a fast responding (S_F) as well as a slow responding reservoir (S_S), representing the groundwater, Fig. 4.2b. The storage components are linked through water fluxes, which also include losses to or gains from deep groundwater, which can be a significant factor in the water balance in the headwaters of the Meuse basin (Bouaziz et al., 2018). This model was selected because of the satisfying performance achieved in catch-

ments of the Meuse (Bouaziz et al., 2018). The model has 12 calibration parameters and resembles in its core formulation FLEX-type model concepts, in the past successfully used in a wide range of environments (Fenicia et al., 2006, 2008; Hrachowitz et al., 2014; Euser et al., 2015; Fovet et al., 2015; Nijzink et al., 2016b,a; Mostbauer et al., 2018).

The model was run at hourly time steps which were aggregated to daily for model calibration, post-calibration evaluation and further analyses. After a one-year warm-up period in 2006, the model was calibrated for the 01/2007 to 12/2011 period, based on a multi-objective calibration strategy (Hulsman et al., 2019). The parameter space was explored with a Monte Carlo approach, sampling 10^5 realizations from uniform prior parameter distributions. Similar to Bouaziz et al. (2018), feasible parameter sets were retained based on their ability to simultaneously and adequately reproduce six different aspects of the observed hydrograph including daily flows (Q), the logarithm of the daily flows (log(Q)), the logarithm of the flow duration curves (FDC, log(Q)) as well as the time-series of weekly (RC,w), monthly (RC,m) and seasonal (RC,s) runoff ratios. This calibration approach was followed to limit uncertainties in medium- to long-term partitioning between drainage and evaporative fluxes and to thus approximate at least longer-term conservation of energy (Hrachowitz and Clark, 2017). All the six above described modeled variables were evaluated against their observed values based on their associated Nash-Sutcliffe Efficiencies (E_{NS}) as objective functions. Feasible parameter sets were retained when E_{NS} scores of all six variables were at least above the 90th percentile best value. For evaluation, the model was tested without further calibration in an independent period from 01/2012 to 12/2017 based on the same performance metrics as above. All relevant model equations as well as prior and posterior parameter ranges are provided in Appendix A.

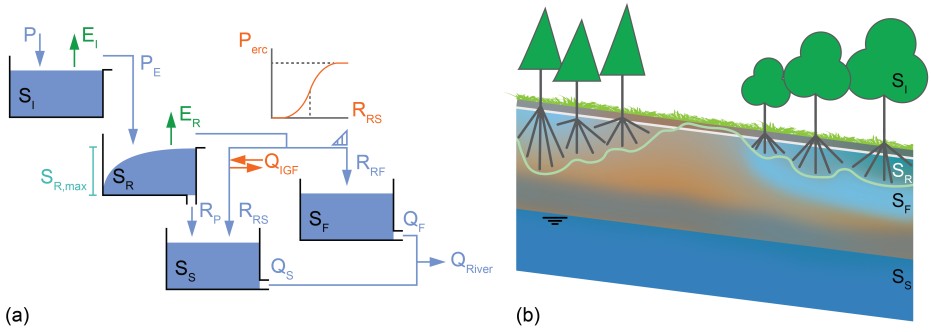

Figure 4.2: (a) Hydrological model with root-zone soil moisture S_R (mm, variable in time) and root-zone water storage capacity $S_{R,max}$ (mm, calibrated and constant in time) (b) associated perceptual model, where the area above the green curve represents the root-zone storage capacity. Both S_R and $S_{R,max}$ are spatially heterogeneous but aggregated to lumped effective values for each individual catchment. We define P as precipitation, P_E as effective precipitation, E as evaporation, R as an internal flux, Q as surface or subsurface drainage (all in mm d^{-1}) and S as storage (in mm). For the subscripts, we define $_I$ as interception, $_R$ as unsaturated root-zone, $_S$ as slow response, $_F$ as fast response, $_P$ as percolation, $_{IGF}$ as net groundwater losses or gains. The parameter P_{erc} defines the maximum percentage of recharge as net groundwater losses or gains.

4.3.2 WATER-BALANCE-DERIVED ROOT-ZONE STORAGE CAPACITY

While hydrological models estimate root-zone storage capacities through an inverse process of calibration with associated risk of equifinality, the water balance approach directly infers root-zone storage capacities from hydrometeorological data. The absence of calibration considerably limits computational power and calculation time of the water balance approach. Daily time series of precipitation and potential evaporation are used as inputs in combination with an estimate of long-term runoff ratio, while a thorough calibration of a hydrological model would also require daily time series of streamflow to constrain the model parameters.

More importantly, there is increasing evidence that catchment-scale root-zone water storage capacities $S_{R,max}$ can be robustly and readily estimated following this water balance approach (Gao et al., 2014; Wang-Erlandsson et al., 2016; Nijzink et al., 2016a; de Boer-Euser et al., 2016). This root-zone storage capacity reflects the integrated interactions between atmospheric water supply and vegetation water demand as controlled by energy supply (i.e. potential evaporation) and vegetation type within a specific spatial domain (here: catchment). It should be clear that the root-zone storage capacity is not necessarily related to root depth but rather to root density as it reflects the pore volume within the influence area of roots (Schenk and Jackson, 2002; Gentine et al., 2012; de Boer-Euser et al., 2016). The underlying assumption implies that vegetation adapts its storage to overcome dry spells with certain return periods (Milly, 1994; Kleidon and Heimann, 1998; Donohue et al., 2012; Gentine et al., 2012).

The approach requires to estimate daily actual transpiration, as it depletes the root-zone storage during dry periods. We first derive the long-term actual transpiration from the water balance of the catchment:

$$E_R \approx P - E_I - Q_{River} - Q_{IGF} \tag{4.5}$$

with long-term annual mean actual transpiration E_R, precipitation P, interception evaporation E_I, streamflow Q_{River} and potential intercatchment groundwater flow losses Q_{IGF}, all provided in mm yr^{-1}.

An interception reservoir is simulated to quantify effective precipitation P_E that reaches the soil, interception evaporation and storage, as a function of the interception storage capacity (I_{max}; see Appendix A for the detailed equations). Due to the lack of more detailed information, we quantified the effect of different interception storage capacities in a sensitivity analysis, i.e. I_{max} = 0.5, 1.0, 2.0 and 3.0 mm.

If the difference between mean annual precipitation and streamflow exceeds potential evaporation, catchment are likely affected by deep groundwater losses, as they were shown to be significant in several catchments of the Meuse (Bouaziz et al., 2018). We estimate mean annual intercatchment groundwater flow losses with the Budyko/Turc-Mezentsev framework (Turc, 1954; Mezentsev, 1955; Budyko, 1961). Long-term mean annual actual transpiration can then be determined from mean annual effective precipitation, streamflow and where applicable groundwater losses (Eq. 4.5).

Daily actual transpiration is subsequently scaled to the daily signal of potential evaporation after removal of the interception evaporation. This scaling allows us to introduce seasonality in actual transpiration.

The storage deficits are then calculated by accumulating the difference between effective precipitation and transpiration assuming an "infinite-reservoir". The initial deficit is assumed to be zero at the start of the calculation, i.e. end of the wet period. The deficit increases when transpiration exceeds effective precipitation during summer until it becomes zero again when all excess water is assumed to drain away as direct runoff. The annual maximum cumulative deficit ($S_{R,def}$) between the time where the deficit equals zero until the time where the total deficit returns to zero is illustrated in Fig. 4.3 for two consecutive years. The maximum storage deficit decreases as the maximum interception capacity increases because more water is intercepted and less goes to transpiration. However, the magnitude of the fluctuations due to interception is minor relative to the magnitude of the storage deficits, as shown in Fig. 4.3.

The annual maxima of the storage deficit $S_{R,def}$ are fitted to the extreme value distribution of Gumbel. Following Gao et al. (2014) and Nijzink et al. (2016a), the 20 years drought return period is used to estimate the water-balance-derived root-zone moisture capacity $S_{R,max}$ for each catchment using hydrometeorological data between 2006 and 2011. The detailed equations for the calculation are provided in Appendix C.

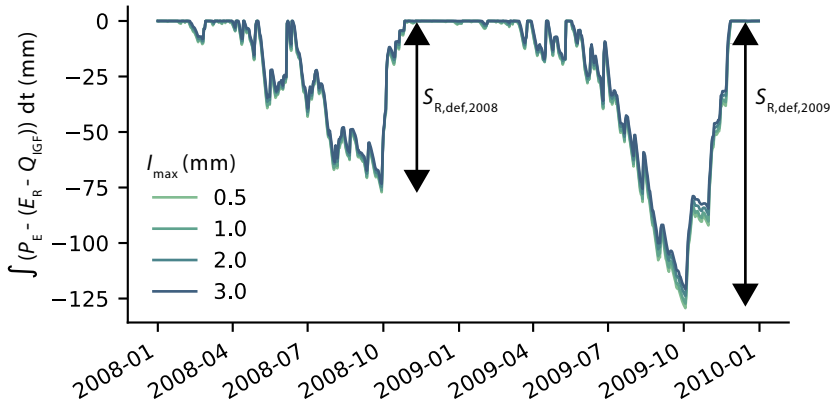

Figure 4.3: Annual maximum cumulative storage deficits $S_{R,def}$ calculated from estimated daily effective precipitation, transpiration and groundwater losses for two consecutive years and several values of interception capacity (I_{max}) in the Hermeton catchment at Hastière (ID12).

4.3.3 IDENTIFYING THE OPTIMAL CHARACTERISTIC TIME LENGTH (T_{opt})

Spearman rank correlations are calculated between time series of daily SWI for T values ranging from 1 to 100 days and modeled time series of daily soil moisture content in the unsaturated root-zone S_R (from Sect. 4.3.1; Fig. 4.2a,b). This calculation is done for each catchment, each parameter set kept as feasible and each satellite soil moisture product

for the period 05/2016 - 12/2017, when all products are available. The optimal T value maximizes the median Spearman rank correlation between SWI and S_R across the feasible model realizations and provides the best representation of the moisture content in the unsaturated root-zone. We use a variety of remotely-sensed products to understand how much of the variability in T_{opt} is related to the product itself. For comparison, the Spearman rank correlation between raw and thus unfiltered values of the near-surface soil moisture SSM (note: $\lim_{T \to 0} SWI(t) = SSM(t)$) and the modeled root-zone soil moisture content S_R is also calculated for each product and catchment. Also note that the strong seasonal cycle of both variables, SWI and S_R, may excessively inflate Spearman rank correlation coefficients, however, with less than 2 years of data, it remains problematic to meaningfully discount seasonality from the time series.

4.3.4 UNDERSTANDING CONTROLS OF THE OPTIMAL CHARACTERISTIC TIME LENGTH (T_{opt})

We first test the relation between T_{opt} and root-zone storage capacities derived as calibrated parameters for the studied catchments by calculating Spearman rank coefficients for each remotely-sensed product. Catchments with relatively small water storage capacities are expected to show a high variability of integrated soil moisture from one time step to another, while catchments with relatively large storage capacities are likely to show a more damped response. Indeed, a small water storage capacity is likely to fill through precipitation and empty through evaporation and drainage more rapidly than a large water storage capacity, leading to an increased variability in time. This is also in line with Eq. 4.2, where T is proportional to the depth of the reservoir below the surface.

We then test the relation between T_{opt} and root-zone storage capacities derived from the water balance approach using meteorological and streamflow data for each catchment. We also compare calibrated and water-balance-derived root-zone water storage capacities between each other.

Finally, we go beyond our main hypothesis and also test the link between T_{opt} and hydrometeorological variables, land cover, soil texture, geological features, topographic indices and streamflow characteristics (provided in Table 4.1) to further explain the observed differences in T_{opt} values between catchments. Similarly to Wang et al. (2017), we consider mean annual precipitation and potential evaporation. Additionally, we assess the link with runoff ratio, aridity index, and percentage of forest and agriculture. We also assess the relation between T_{opt} and soil texture (clay, sand and silt percentages), as it has been the subject of several studies (de Lange et al., 2008; Albergel et al., 2008; Wang et al., 2017) and to the presence of highly productive aquifers including karstified rocks (based on the International Hydrogeological Map of Europe, IHME, http://www.brg.bund.de/). As topography is related to climate, land cover and soil characteristics, especially in natural landscapes (Savenije and Hrachowitz, 2017), we include catchment size, drainage density and mean slope as potential predictors to explain the variability in T_{opt} between catchments. While T_{opt} characterizes the temporal variability of soil moisture, the flashiness index (I_f) is a measure of the responsiveness of a catchment in terms of the variability of its streamflow from one time step to another (Fenicia

et al., 2016). We therefore test for a relation between the flashiness index and T_{opt}, as slow dynamics in root-zone soil moisture can be expected to result in a smooth runoff response.

4.4 RESULTS

4.4.1 MODEL EVALUATION

The calibration and evaluation performance, in terms of the six objective functions, have rather similar performances in both periods except for some modest divergence of the models' ability to reproduce runoff ratios for the evaluation period in several catchments (Fig. 4.4). $E_{NS,Q}$ during evaluation ranges between 0.51 and 0.93 for all catchments with a mean of 0.83, while $E_{NS,RC,m}$ range between 0.52 and 0.97 with a mean of 0.82. The model slightly overestimates peak flows (Fig. 4.5a,d), but reproduces flow duration curves (Fig. 4.5b,e) and monthly runoff ratios (Fig. 4.5c,f) relatively well.

Overall, the models can reproduce high- and low flow metrics quite well and exhibit plausible long-term partitioning of water fluxes into streamflow and actual evaporation (i.e. $E_A = E_I + E_R$), according to runoff ratio, enabling us to use soil moisture S_R and $S_{R,max}$ for the subsequent analyses.

4.4.2 VARIABILITY OF IDENTIFIED T_{opt}

Highest Spearman rank correlation coefficients r between time series of daily SWI and time series of modeled root-zone soil moisture S_R are obtained for T-values between 6 and 32 days (median of 18 days) per soil moisture product for one illustrative catchment (see Fig. 4.6). The maximum Spearman rank correlation achieved with the SWI is always higher than with the near-surface soil moisture SSM (Fig. 4.6), as also found by Paulik et al. (2014). Especially for S1-CSAR, the Spearman rank correlation from SSM to SWI increases on average (over all catchments) from 0.54 to 0.83, as opposed to an increase from 0.75 to 0.88 for SMAP-L-am_d, possibly related to higher variance of the SSM signal of S1-CSAR compared to SMAP-L-am_d. Introducing noise to near-surface soil moisture time series leads to a larger increase in Spearman rank correlations from SSM to SWI and slightly higher T_{opt} values, suggesting noise filtering by the Soil Water Index. The narrow band in Fig. 4.6 suggests that Spearman rank correlations are not very sensitive to the uncertainty in model parameters, implying relatively similar root-zone soil moisture dynamics between parameter sets.

Fig. 4.7 shows modeled daily root-zone moisture S_R for each parameter set kept as feasible, near-surface soil moisture SSM and SWI as inferred from a selection of satellite products using the associated T_{opt} (see Fig. 4.6) for two of the 16 study catchments. Higher values of T_{opt} lead to more smoothing and delaying of the original near-surface soil moisture signal (Fig. 4.7b-e versus 4.7g-j showing the effect of T_{opt} of 6-26 days versus 2-6 days) and can therefore reproduce daily fluctuations of the modeled soil moisture S_R reasonably well (Spearman rank correlations $r > 0.82$ for the best performing products S1-CSAR and SMAP L-band products versus $r > 0.61$ for the poorer perform-

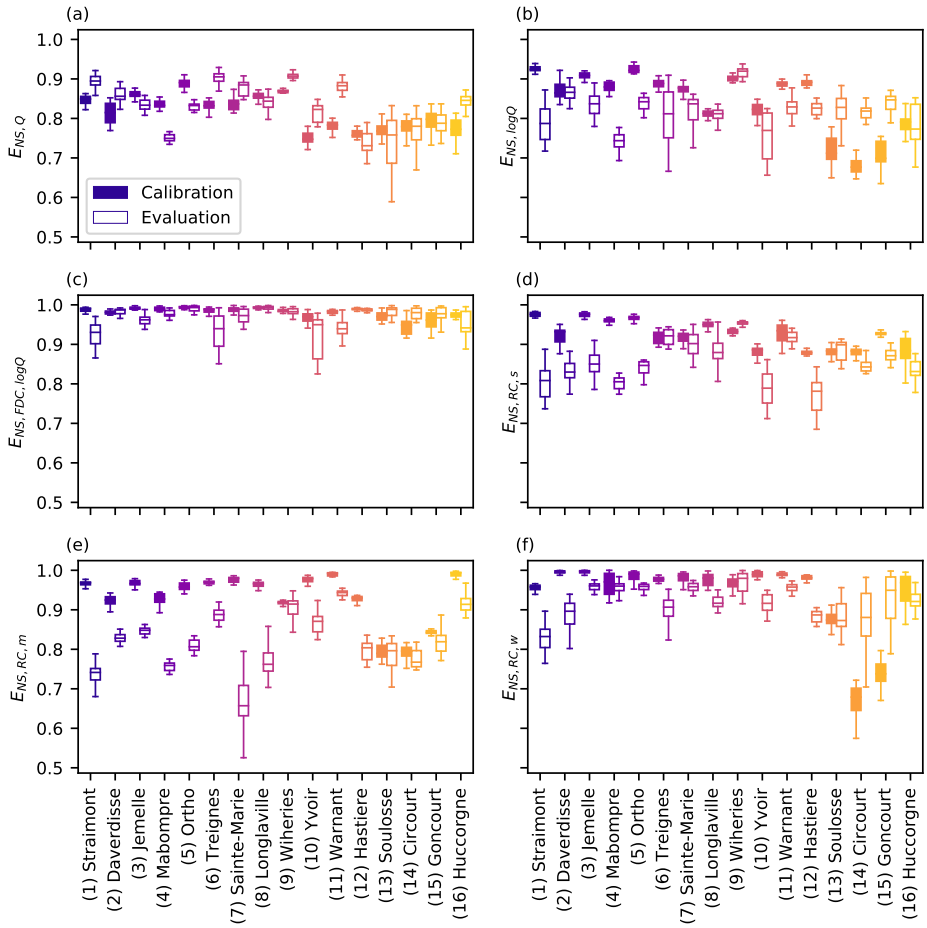

Figure 4.4: Calibration (filled boxplot) and evaluation (no fill) performances of Nash-Sutcliffe efficiencies of the: (a) flows; (b) logarithm of the flows; (c) flow duration curve of the logarithm of the flows; (d) seasonal runoff ratio; (e) monthly runoff ratio; (f) weekly runoff ratio for the 16 study catchments (colors as in Fig. 4.1)

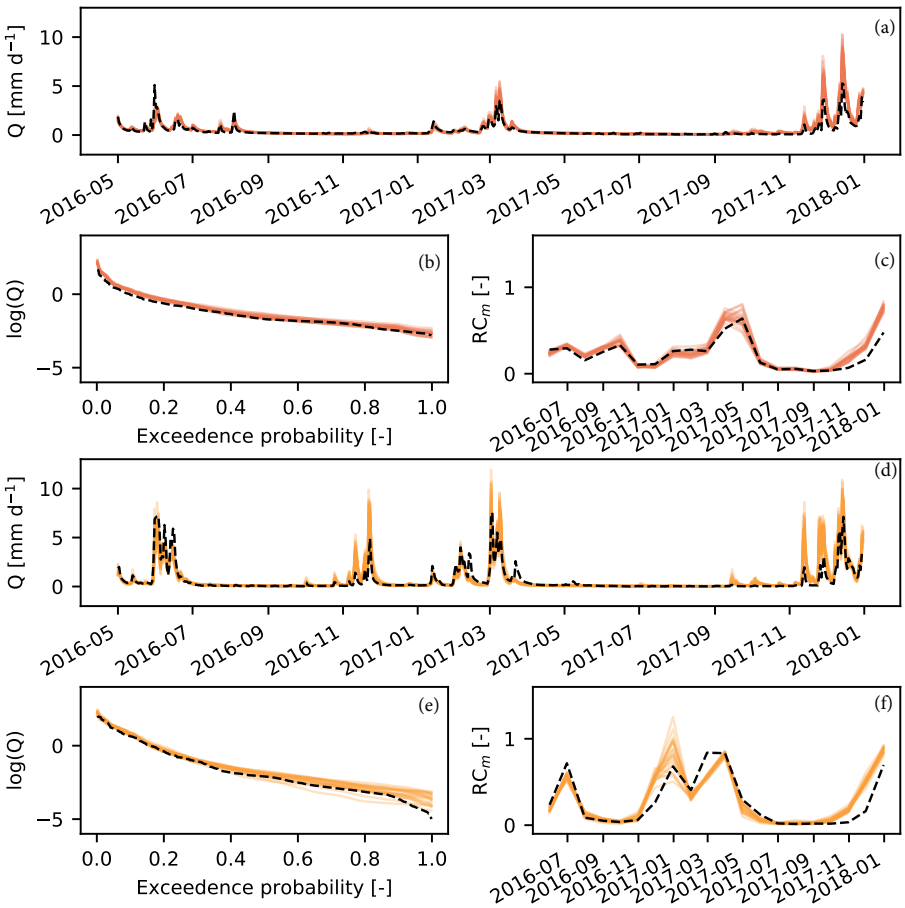

Figure 4.5: Observed (dashed black lines) and feasible modeled (colored lines) hydrographs for the period with available SSM data (05/2016-12/2017), and associated flow duration curves of the log of the flows and monthly runoff ratios for (a-c) the Hermeton at Hastière (ID12) and (d-f) the Mouzon at Circourt-sur-Mouzon (ID14) (colors as in Fig. 4.1).

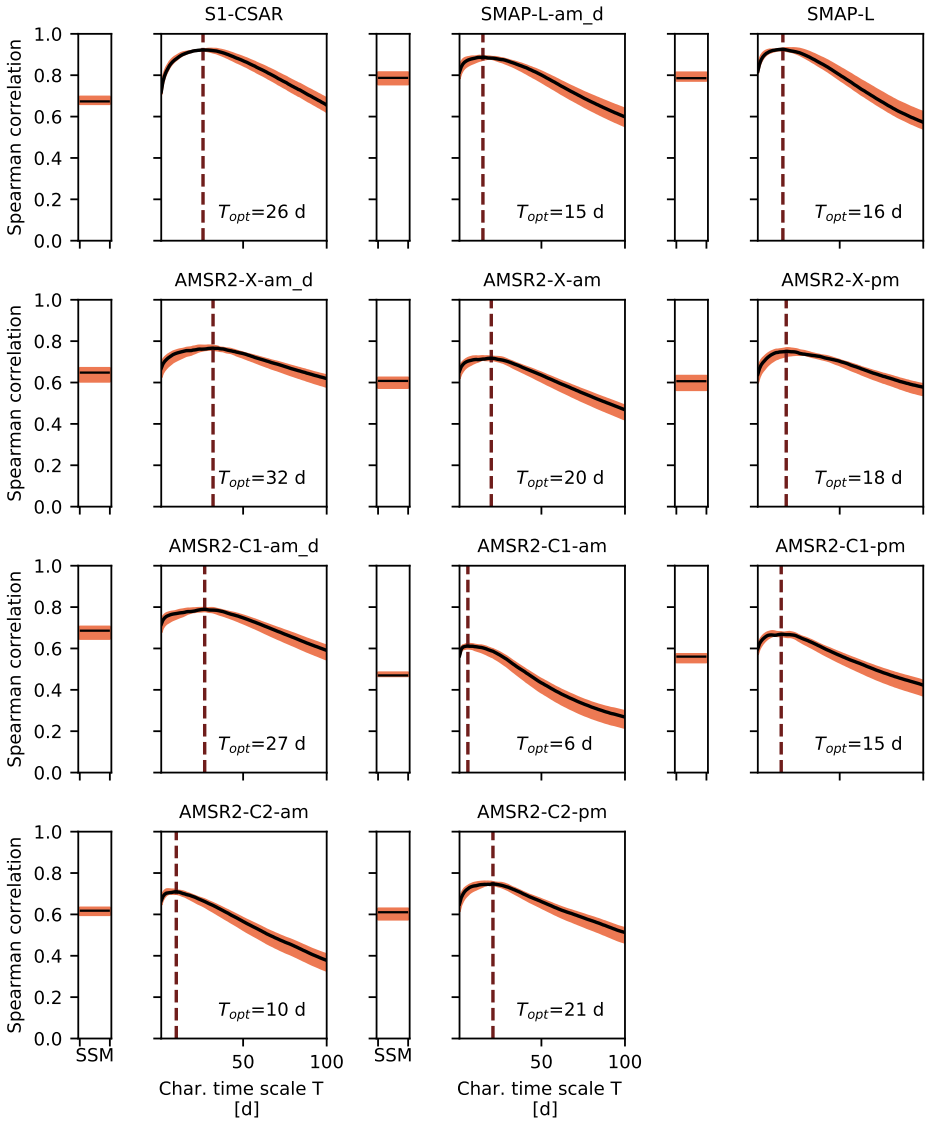

Figure 4.6: Spearman rank correlations coefficients between time series of modeled daily root-zone soil moisture S_R [mm] and SWI [-] for different values of the characteristic time scale T [d] in the Hermeton at Hastière (ID12) for the period 05/2016 to 12/2017 for each soil moisture product. In the narrow subplots on the left side of each larger subplot, the correlation of the time series of daily near-surface soil moisture SSM and modeled root-zone soil moisture S_R is shown. The black line and colored band represent the median value and associated 25^{th} to 75^{th} percentiles of S_R from the ensemble of parameter sets retained as feasible. The vertical dashed line indicates the optimal value of the characteristic time scale T_{opt} [d] here defined to be at the highest correlation of the median value.

ing AMSR2-C1-am product). For the S1-CSAR product in the Hermeton catchment at Hastière (ID12), the variance decreases from 0.066 for near-surface soil moisture SSM to 0.034 for SWI, while a reduction of variance from 0.065 to 0.046 can be seen for the Mouzon at Circourt-sur-Mouzon (ID14).

During the wetting phase in the catchment of the Hermeton at Hastière (ID12), S1-CSAR (Fig. 4.7b) and SMAP-L (Fig. 4.7d) have the highest visual similarity with modeled root-zone moisture S_R (Fig. 4.7a), while the drying phase behavior is best reproduced by SMAP-L-am_d (Fig. 4.7c) and SMAP-L (Fig. 4.7d). Noise in AMSR2-C1-am occurs in the Hermeton at Hastière (ID12), possibly caused by radio frequency interference (RFI) in the area. More spread in SSM of S1-CSAR is also observed compared to SSM of the L-band products (Fig. 4.7b and 4.7g versus 4.7c-d and 4.7h-i).

Note that in the month following May 22th of 2016, a series of heavy and long-lasting precipitation events occurred over France and Belgium, with total monthly volumes between 146 and 236 mm month^{-1} over the study catchments. Despite the large spatial scale of these events, most satellite products only show a limited increase in soil moisture, except S1-CSAR, in particular in the Mouzon at Circourt-sur-Mouzon (Fig. 4.7g). The underlying reason remains unclear.

T_{opt} varies per product (Figs. 4.6 and 4.8a) with lowest 5^{th} and highest 95^{th} percentiles of 1 and 98 days (median of 17.5 days), with product AMSR2-C1-am showing the lowest T_{opt} ($5/95^{th}$ percentiles of 1 and 30 days with a median of 4 days) and product S1-CSAR the highest values ($5/95^{th}$ percentiles of 6 and 98 days with a median of 33 days). The higher T_{opt} of the only radar (active microwave) S1-CSAR product amongst radiometers (passive microwave) are likely related to the different sensing techniques, the larger initial variance of the SSM data compared to variances of the passive products and to the larger sampling intervals (see Table 4.2). de Lange et al. (2008) investigated the influence of satellite sampling intervals on the T parameter and suggested that for larger sampling intervals, soil dynamics are less well reflected, therefore, higher values of the characteristic time scale are expected.

Similarly, T_{opt} varies across the study catchments (Fig. 4.8b), with lowest 5^{th} and highest 95^{th} percentiles of 1 and 89 days (median of 17 days). The catchment of the Mehaigne at Huccorgne (ID16) exhibits the largest values ($5/95^{th}$ percentiles of 33.5 and 89 days with a median of 58 days), this catchment is characterized by 80% agriculture cover, relatively low runoff ratio (25%) and low flashiness index (19%), see Table 4.1. The Bocq at Yvoir (ID10) and the Molignée at Warnant (ID11), catchments with relatively similar characteristics, also show high outlier values of T_{opt} (Fig. 4.8b). The three most upstream catchments (IDs 13-15) exhibit the lowest T_{opt} values ($5/95^{th}$ percentiles of 1 and 6 days with a median of approximately 3 days) and these catchments are, on the other hand, more responsive in time (relatively high flashiness indices varying between 35-42%) and have relatively low agriculture cover (12-26%), Table 4.1.

In general, the strongest correlations with median Spearman rank correlation coefficients r between the daily times series of SWI and S_R of 0.84, 0.88 and 0.91, respectively, can be found for the S1-CSAR and L-band (SMAP-L-am_d, SMAP-L) products, while the lowest correlations are observed for AMSR2-C1-am with a median of r=0.63 (Fig. 4.8c).

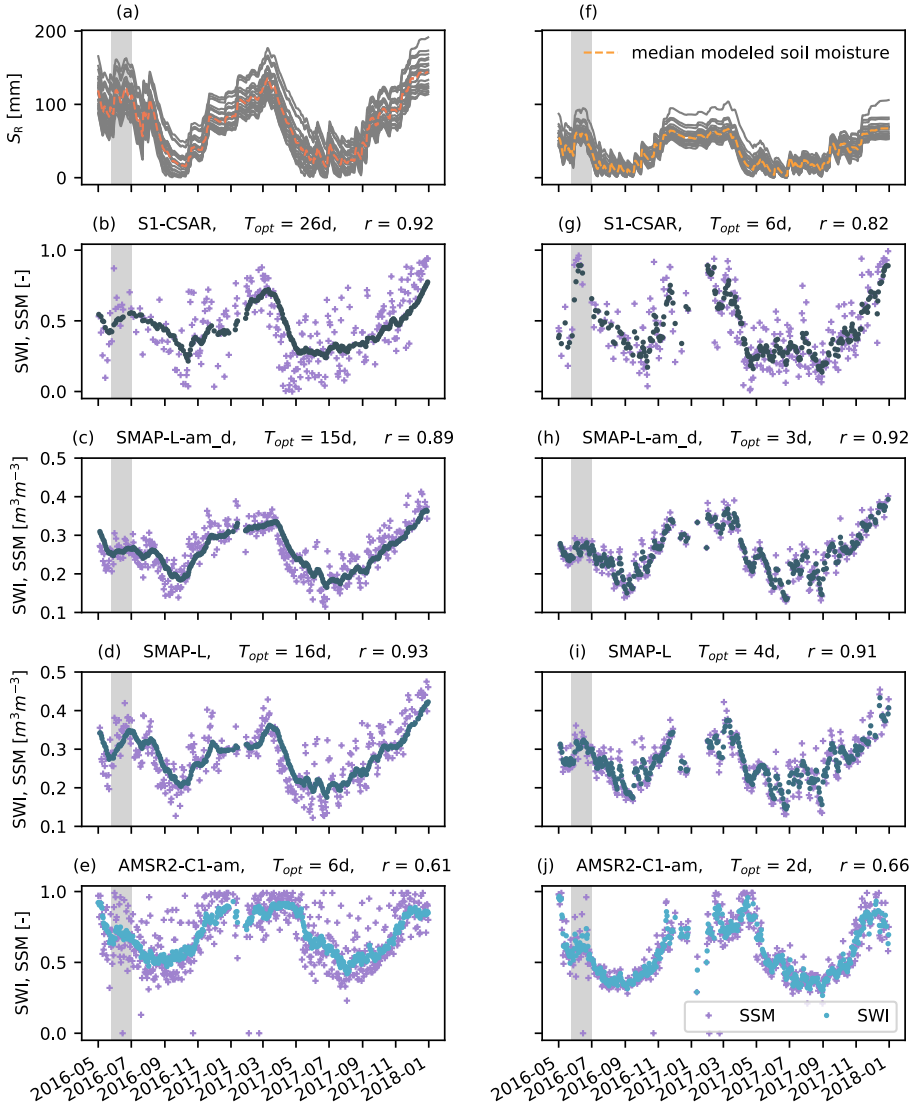

Figure 4.7: Time series of (a,f) modeled root-zone soil moisture S_R for all model parameter sets retainted as feasible, (b-e and g-j) SSM and SWI based on the respective T_{opt} values for a selection of products with highest performance (S1-CSAR, SMAP-L-am_d and SMAP-L) and lowest performance (AMSR2-C1-am) for (a-e) the Hermeton catchment at Hastière (ID12) and (f-j) the Mouzon catchment at Circourt-sur-Mouzon (ID14). The grey shaded area in summer 2016 indicates a series of very wet events. Units vary between soil moisture products as they are either provided as volumetric weigths [m^3 m^{-3}] or percentage saturation [-]. The blue shades used for each soil moisture product follow the color scheme of Fig. 4.8a,c.

A higher sensitivity to soil moisture is expected for L-band retrievals over C- and X-band as a result of a higher penetration depth, a higher sensitivity of the dielectric constant to soil moisture and an increased transmissivity of vegetation. SAR provides information at a higher spatial resolution compared to passive radiometers (Entekhabi et al., 2010). The reason explaining the lower correlations of AMSR2-C1-am is likely related to RFI. The AMSR2 low-resolution ascending products (overpass at 1.30 PM) show a higher degree of similarity with modeled root-zone water storage dynamics than the descending overpasses (Fig. 4.8c). This is in line with findings from Brocca et al. (2011), despite the benefit of reduced difference between surface and canopy temperature of night-time overpasses.

Similarly, lowest median correlations (r varying between 0.62 and 0.72) are found in the wettest, steep and forested Ardennes catchments (IDs 1-5; Fig. 4.8d) and highest (0.80-0.84) in the catchments with highest aridity index (IDs 13-16), see Table 4.1. Accuracies of soil moisture retrieval are indeed affected by a complex topography and high vegetation density (Brocca et al., 2017). It should be noted that the comparison of Spearman rank correlation is based on a different number of observations for each product (Table 4.2).

4.4.3 INFLUENCE OF CATCHMENT CHARACTERISTICS ON T_{opt}

Root-zone storage capacities $S_{R,max}$ were available as calibrated parameters from the hydrological model and from the water balance approach described in Sect. 4.3.2 and Appendix C. We calculate the 20-year return period root-zone water storage capacity for several values of the interception capacity I_{max}. Both approaches provide broadly consistent values with a Spearman rank correlation coefficient of r=0.67 with p=0.004 (and Pearson correlation of 0.86 with p=2 x 10^{-5}) for median values of $S_{R,max}$ retained as feasible and median $S_{R,max}$ values estimated from the water balance approach (Fig. 4.9).

High (median r >0.80) and significant positive (p <0.05) correlations are found between T_{opt} values of all soil moisture products and the calibrated root-zone storage capacities $S_{R,max}$ (Table 4.3a and Fig. 4.10a for a selection of products). Spearman rank correlations between T_{opt} values and water-balance-derived root-zone storage capacities are less strong but still statistically significant (p <0.05) for most soil moisture products (Table 4.3a and Fig. 4.10b). The correlations are strongest for the S1-CSAR product (r=0.65), while the weakest link was found with the AMSR2-C1-am product (r=0.36). Similarly, T_{opt} in the AMSR2-C1-am product is less sensitive to $S_{R,max}$ ($dT_{opt}/dS_{R,max}$ = 0.23 d mm^{-1}), possibly because of RFI, while the S1-CSAR product exhibits the highest sensitivity ($dT_{opt}/dS_{R,max}$ = 0.52 d mm^{-1}), which is likely related to the larger spread in T_{opt} values (Fig. 4.8a, Fig. 4.10b). In addition, all products suggest that T_{opt} is insensitive to $S_{R,max}$ at $S_{R,max}$ values below 100 mm, this could be related to the already very low T_{opt} (< 6 days) in these catchments. The vertical error bars in Fig. 4.10 suggest that T_{opt} is not very sensitive to the uncertainty in model parameters. While previous studies (Albergel et al., 2008; Paulik et al., 2014) showed an increase in T_{opt} for increasing depth of the soil profile, in line with Eq. 4.2, we explicitly make the link with root-zone storage capacity derived from streamflow, precipitation and evaporation data.

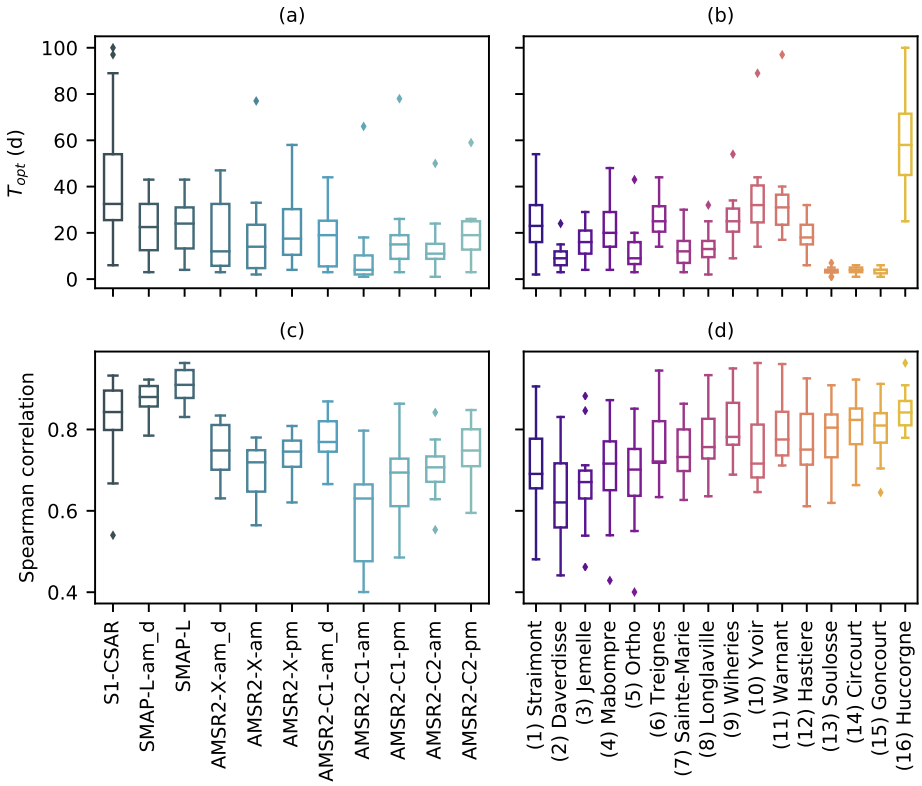

Figure 4.8: (a) T_{opt} range across the 16 study catchments per product. (b) T_{opt} range across all products per catchment. (c) Spearman rank correlations coefficients between SWI using the optimal T value and modeled time series of root-zone soil moisture S_{R} for the period 05/2016 to 12/2017 per product for all 16 catchments. (d) Spearman rank correlations coefficients between SWI using the optimal T value and time series of modeled root-zone soil moisture S_{R} for the period 05/2016 to 12/2017 per catchment for all products.

We found no significant correlations between T_{opt} and aridity index, runoff ratio or mean annual precipitation and potential evaporation (Table 4.3a). Instead, the interplay between accumulated precipitation and actual evaporation during dry periods represented by the root-zone storage capacity $S_{\text{R,max}}$, shows a significant positive correlation with T_{opt}.

T_{opt} values are inversely correlated with the flashiness indices I_{f} of the study catchments (Table 4.3a, Fig. 4.10c), implying that a more flashy streamflow (high flashiness index) indicates lower T_{opt} and, by extension, lower $S_{\text{R,max}}$ and therefore higher temporal variability in soil moisture.

We found significant positive correlations between T_{opt} and percentage agriculture (Table 4.3b, Fig. 4.10d), suggesting a low temporal variability of soil moisture in agriculture-dominated catchments, also implying higher $S_{\text{R,max}}$ values in these catchments, which

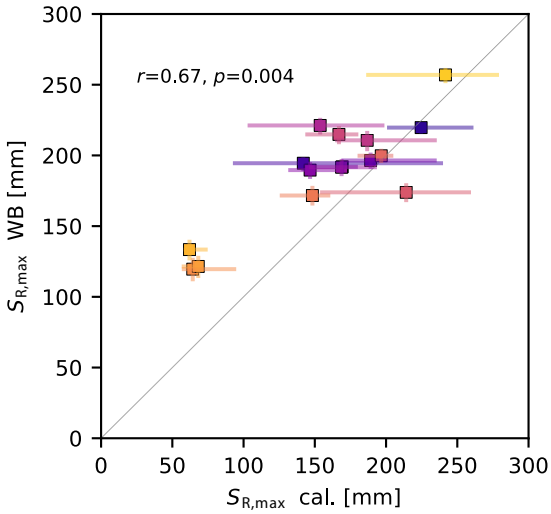

Figure 4.9: Relation between water-balance-derived (WB) and calibrated (cal.) $S_{R,max}$ from all 16 study catchments. The symbols indicate the median $S_{R,max}$ for each method, the horizontal error bars indicate the range of feasible $S_{R,max}$ values ($25/75^{th}$ percentiles) from the model and vertical error bars indicate the min-max range obtained from several I_{max} values (colors as in Fig. 4.1).

are related to low runoff ratios and, therefore, high evaporation rates (Table 4.1). Interestingly, T_{opt} is positively related to percentage silt in a catchment (Table 4.3b, Fig. 4.10e), and therefore also to the percentage of agriculture, which is likely related to the fertility and suitability of silt for growing crops as it promotes water retention and air circulation. No significant relations are found between T_{opt} and the percentages of clay and sand (Table 4.1 and 4.3).

No significant relations between T_{opt} and geological features (percentage of highly productive aquifers) or topographic indicators (catchment area, drainage density and mean slope) are found (see Table 4.3b).

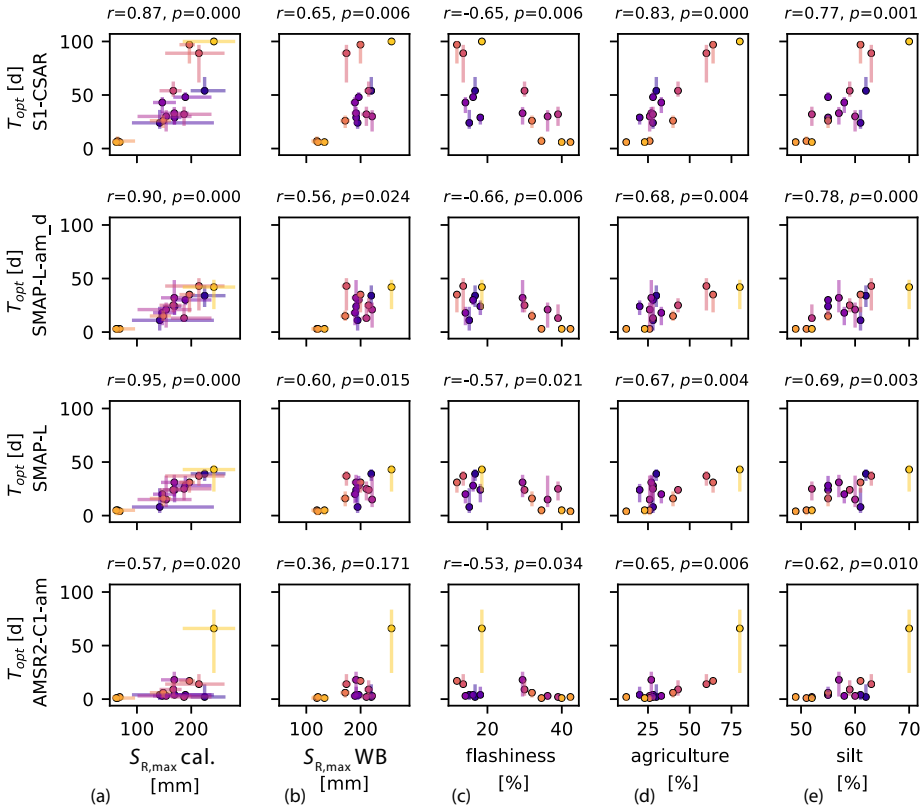

Figure 4.10: Optimal value of the characteristic time length T_{opt} (with vertical bars associated with the $25/75^{th}$ percentiles of T_{opt} from parameter sets kept as feasible) as a function of (a) calibrated root-zone water storage capacities ($S_{\text{R,max}}$ cal., median and $25/75^{th}$ percentiles of parameter sets kept as feasible), (b) water-balance-derived root-zone water storage capacities ($S_{\text{R,max}}$ WB, median and min-max range based on I_{max} values for a 20-year return period), (c) flashiness index, (d) percentage agriculture and (e) percentage silt for the 16 studied catchments for a selection of products (colors as in Fig. 4.1). Spearman rank coefficients r with associated p-values are calculated for the median values.

Table 4.3: Spearman rank correlation coefficients between the optimal values of the characteristic time length T_{opt} [d] per product and (a) the calibrated (cal.) and water-balance-derived (WB) root-zone water storage capacities $S_{R,max}$, mean annual precipitation (P) and potential evaporation (E_P), aridity index (E_P/P), runoff ratio (Q/P), flashiness index (I_f), and (b) percentage of forest cover (for.), percentage of agricultural cover (agr.), percentage of highly productive and fissured aquifers (fiss., based on IHME), percentage clay, sand and silt, and catchment area, drainage density (dd) and mean slope.

(a)

Product	$S_{R,max}$ (cal.)	$S_{R,max}$ (WB)	P	E_P	E_P/P	Q/P	I_f
S1-CSAR	0.87**	0.65*	-0.01	-0.32	-0.11	-0.16	-0.65*
SMAP-L-am_d	0.90**	0.56*	0.04	-0.43	-0.18	-0.17	-0.66*
SMAP-L	0.95**	0.60*	0.05	-0.39	-0.21	-0.05	-0.57*
AMSR2-X-am_d	0.80**	0.44	-0.34	-0.16	0.16	-0.44	-0.40
AMSR2-X-am	0.81**	0.56*	-0.24	-0.18	0.07	-0.38	-0.43
AMSR2-X-pm	0.92**	0.54*	-0.12	-0.40	-0.08	-0.22	-0.62*
AMSR2-C1-am_d	0.61*	0.37	-0.11	-0.39	-0.12	-0.14	-0.44
AMSR2-C1-am	0.57*	0.36	-0.20	-0.38	0.02	-0.33	-0.53*
AMSR2-C1-pm	0.85**	0.60*	0.02	-0.44	-0.19	-0.12	-0.60*
AMSR2-C2-am	0.90**	0.56*	0.01	-0.50*	-0.20	-0.11	-0.70*
AMSR2-C2-pm	0.82**	0.55*	-0.17	-0.28	-0.01	-0.31	-0.46

(b)

Product	for.	agr.	fiss.	clay	sand	silt	area	dd	slope
S1-CSAR	-0.45	0.83**	-0.01	-0.38	-0.29	0.77**	-0.42	0.07	-0.27
SMAP-L-am_d	-0.28	0.68*	-0.01	-0.49	-0.21	0.78**	-0.34	0.19	-0.21
SMAP-L	-0.30	0.67*	-0.13	-0.47	-0.15	0.69*	-0.28	0.11	-0.12
AMSR2-X-am_d	-0.45	0.80**	-0.02	-0.20	-0.26	0.59*	-0.44	0.09	-0.42
AMSR2-X-am	-0.41	0.82**	0.01	-0.24	-0.26	0.66*	-0.51*	-0.00	-0.44
AMSR2-X-pm	-0.30	0.78**	-0.09	-0.44	-0.17	0.68*	-0.39	0.22	-0.21
AMSR2-C1-am_d	-0.07	0.67*	-0.36	-0.33	0.03	0.51*	-0.37	0.33	-0.26
AMSR2-C1-am	-0.07	0.65*	-0.15	-0.48	-0.06	0.62*	-0.24	0.16	-0.25
AMSR2-C1-pm	-0.24	0.73*	-0.17	-0.51*	-0.25	0.83**	-0.37	0.16	-0.23
AMSR2-C2-am	-0.24	0.73*	-0.14	-0.54*	-0.20	0.78**	-0.32	0.18	-0.08
AMSR2-C2-pm	-0.31	0.74*	-0.05	-0.38	-0.19	0.66*	-0.40	0.12	-0.35

* $p < 0.05$

** $p < 0.001$

4.5 DISCUSSION

4.5.1 IMPLICATIONS

We found a median T_{opt} of 17 days, which is close to the often used value of 20 days (Wagner et al., 1999), however, we show that T_{opt} significantly varies between catchments ($5/95^{th}$ percentiles of 1 and 98 days) and products (highest T_{opt} for S1-CSAR). This is in line with the variability reported by other studies. Ceballos et al. (2005) found values of approximately 50 days to represent the soil profile between 0 and 100 cm based on soil moisture measurements in agricultural fields in Spain. de Lange et al. (2008) mentions a general characteristic time length of 20 days based on observed and modeled soil moisture series, while Beck et al. (2009) found a T_{opt} value of 5 days for the calculation of SWI in Australian catchments.

Previously, T_{opt} was shown to increase with increasing depth of the soil profile at the point-scale (Wagner et al., 1999; Albergel et al., 2008; Ceballos et al., 2005; Paulik et al., 2014; Wang et al., 2017). However, T_{opt} was not linked to a defined storage in the subsurface at the catchment scale. Here, we explicitly show the increase of T_{opt} with increasing estimates of catchment-scale root-zone water storage capacities $S_{R,max}$. For modeling applications and data assimilation of satellite soil moisture products in streamflow forecasting applications (Brocca et al., 2010b; Wanders et al., 2014; Laiolo et al., 2016; López López et al., 2017; Loizu et al., 2018), our results suggest that suitable values of T can be inferred from a calibrated model or based on estimates of root-zone water storage capacities. These estimates can readily be determined using available hydrometeorological data, instead of selecting a standard value of 20 days based on the work of Wagner et al. (1999).

Albergel et al. (2008) and Wang et al. (2017) suggest that T values are low in areas with high evaporative demand and less frequent but more intense precipitation. The results from our water balance approach rather suggest that these are typically areas where vegetation needs to overcome long dry spells and, therefore, with relatively high root-zone water storage capacities and T_{opt}. This is an important conclusion as it shows that the interplay between precipitation and evaporation is the main climatic driver that controls T, and not the precipitation and evaporation individually, as often tested.

The highest T_{opt} values occur in agricultural-dominated catchments, which is in line with the relatively high values of T_{opt} (40-60 days) reported by Ceballos et al. (2005) in agricultural fields in Spain. In the Meuse, these catchments are characterized by low flashiness indices, low runoff ratios and, therefore, high actual evaporation, resulting in a small temporal variability of soil moisture (high T_{opt}) and large root-zone water storage capacities $S_{R,max}$. The high agricultural cover also coincides with a relatively high silt percentage, related to the fertility and suitability of silt for growing crops due to the high water holding capacity. Soils with fine texture promote high water retention, slow drainage and, therefore higher T_{opt} than soils with coarser textures, as also mentioned by Ceballos et al. (2005). These findings highlight the interactions between soil properties (% silt), the interplay between precipitation and evaporation ($S_{R,max}$), land cover (% agriculture) and streamflow signatures (I_f) to characterize soil moisture behavior and

estimate T_{opt} to infer root-zone from near-surface soil moisture.

SMAP-L-am_d, SMAP-L and S1-CSAR derived times series of SWI show the highest similarity with modeled root-zone soil moisture S_R. The longer wavelength of L-band compared to C- and X-bands allows for a deeper soil penetration and a higher sensitivity to soil moisture. The difference in Spearman rank correlations between modeled root-zone soil moisture and SWI for the high- and low-resolution products is, therefore, less than between bands (Fig. 4.8c). The advantage of S1-CSAR is the high spatial resolution to capture small-scale changes, easily missed out by coarse resolution sensors (Bauer-Marschallinger et al., 2019). During large scale and high intensity precipitation events from end of May until end of June 2016, the S1-CSAR product also shows the expected sharp increase in soil moisture, which is, for unknown reasons, missed by or relatively limited in most passive microwave products (Fig. 4.7).

The variability of T_{opt} values is related to the soil moisture product itself. The larger variance of the near-surface soil moisture signal of S1-CSAR leads to a larger range of T_{opt} values (Fig. 4.8a). Paulik et al. (2014) show a stronger agreement between in-situ measurements and SWI compared to SSM, which could be related to noise being filtered out by the SWI. Our findings further suggest that T_{opt} values are likely to increase with increasing noise in the observations.

Both the remote sensing and the hydrological communities can benefit from our analysis, as it provides guidance for hydrologists to meaningfully infer root-zone soil moisture from near-surface soil moisture for hydrological applications, while it clarifies the behavior of T for several near-surface soil moisture products and its relation with hydrometeorological variables, soil texture, land cover and streamflow dynamics for the remote sensing community .

4.5.2 LIMITATIONS

In previous studies, T_{opt} was often estimated using observed soil moisture time series. However, these point observations often fail to represent the spatial heterogeneity at the catchment scale. In this study, we instead calibrate a process-based hydrological model (Fig. 4.2a) against observed streamflow to derive T_{opt} by relating the time series of SWI to modeled daily root-zone soil moisture S_R (Fig. 4.6 and 4.7). However, it implies that the model provides a meaningful representation of true soil moisture at catchment scale. As we are interested in root-zone soil moisture and because T is a function of the depth of the layer below the surface (Eq. 4.2), we related T_{opt} to root-zone water storage capacities $S_{R,max}$ estimated as a calibration parameter and, in an independent way, from a water balance approach. The water balance approach, however, also relies on assumptions related to estimation of the return period, interception evaporation, actual evaporation, groundwater losses and periods where infiltration takes place and deficits start to accumulate. Yet, the important additional information used in both methods is the streamflow data, from which actual evaporation can be estimated and, by extension, the water storage capacity in the root-zone, which influences root-zone soil moisture behavior.

As the selected catchments are relatively similar in size and all located in a temperate

climate zone, an interesting next step would be to assess the relation between T_{opt} and $S_{R,max}$, land cover and soil texture in a larger variety of catchments, differing in size and climate zone. We expect the positive correlation between T_{opt} and $S_{R,max}$ to be transferable to other areas, as the water balance approach has successfully been applied in a variety of climate zones (New Zealand in de Boer-Euser et al. (2016), Australia in Donohue et al. (2012) and United States in Gentine et al. (2012) and Gao et al. (2014)) or even globally (Wang-Erlandsson et al., 2016). The positive relation between agricultural cover, silt percentage and T_{opt} found in the studied catchments might not apply in other areas, in contrast to the expected relation between T_{opt} and flashiness index. Furthermore, the approach could be tested using several other models. However, if they provide a plausible long-term partitioning of water fluxes to evaporation and drainage from the root-zone soil moisture, we do not expect our results to be significantly affected by model selection. Additionally, our results with respect to differences in Spearman rank correlations between products could be affected if ten years of satellite near-surface soil moisture data based on SMAP and Sentinel-1 would be available, instead of the less than two years of data available here.

4.6 CONCLUSION

The Soil Water Index (SWI) enables us to infer root-zone soil moisture from satellite-based near-surface soil moisture, using an estimate of the characteristic time length (T) of the water flux from the surface to the root-zone. Estimating T has so far proven difficult as no clear link with climate or soil has been established. Using a process-based lumped hydrological model calibrated on streamflow data, we identified optimal T values (T_{opt}) that lead to the highest correlation between SWI and modeled root-zone soil moisture S_R in the Meuse basin (Fig. 4.6 and 4.7). While the median T_{opt} value of 17 days approximates the often used standard value of 20 days, T_{opt} significantly varies between catchments and soil moisture products ($5/95^{th}$ percentiles of 1 and 98 days, Fig. 4.8a,b). In the past, T has been conceptually linked to an undefined storage volume in the subsurface. We now show for 16 contrasting catchments in the Meuse river basin that T is strongly and positively related with the root-zone water storage capacity (Fig. 4.10). This catchment-scale vegetation accessible water storage capacity can be readily and robustly estimated based on water balance data. Our key finding implies that the interplay between precipitation and evaporation during dry periods, which regulate the size of the storage capacity in the unsaturated root-zone, is the main driver controlling T_{opt} in temperate climates. Such a clear link between T_{opt} and hydrometeorological variables opens the opportunity to generate meaningful estimates of water contents in the root-zone from globally available remotely-sensed near-surface soil moisture data. These are of critical importance for hydrological and meteorological applications, as root-zone soil moisture controls how much water is available for plant transpiration and, therefore, the partitioning of precipitation to streamflow and evaporation.

5

PEEKING BEHIND THE SCENES OF STREAMFLOW PERFORMANCE: COMPARING INTERNAL PROCESS REPRESENTATION

L'opinion publique, c'est sacré : pas d'affolement, surtout pas d'affolement.

Albert Camus - *La Peste* (1947)

In the previous chapters, we quantified the long-term water balance and the root-zone soil moisture dynamics in several catchments of the Meuse river basin. Remotely-sensed estimates of evaporation and near-surface soil moisture were used as additional sources of information to increase our understanding of hydrological processes. However, streamflow is often the only variable used to constrain and evaluate hydrological models. Here, we test if twelve process-based models with similar streamflow performance have a similar representation of internal processes. Using expert knowledge in combination with remotely-sensed estimates of snow cover, evaporation, soil moisture and total storage anomalies, we test the plausibility of model behavior.

Parts of this chapter have been published in Hydrology and Earth System Sciences (Bouaziz et al., 2021)

SUMMARY

Streamflow is often the only variable used to evaluate hydrological models. In a previous international comparison study, eight research groups followed an identical protocol to calibrate twelve hydrological models using observed streamflow of catchments within the Meuse basin. In the current study, we quantify the differences in five states and fluxes of these twelve process-based models with similar streamflow performance, in a systematic and comprehensive way. Next, we assess model behavior plausibility by ranking the models for a set of criteria using streamflow and remote-sensing data of evaporation, snow cover, soil moisture and total storage anomalies. We found substantial dissimilarities between models for annual interception and seasonal evaporation rates, the annual number of days with water stored as snow, the mean annual maximum snow storage and the size of the root-zone storage capacity. These differences in internal process representation imply that these models cannot all simultaneously be close to reality. Modeled annual evaporation rates are consistent with GLEAM estimates. However, there is a large uncertainty in modeled and remote-sensing annual interception. Substantial differences are also found between MODIS and modeled number of days with snow storage. Models with relatively small root-zone storage capacities and without root water uptake reduction under dry conditions tend to have an empty root-zone storage for several days each summer, while this is not suggested by remote-sensing data of evaporation, soil moisture and vegetation indices. On the other hand, models with relatively large root-zone storage capacities tend to overestimate very dry total storage anomalies of GRACE. None of the models is systematically consistent with the information available from all different (remote-sensing) data sources. Yet, we did not reject models given the uncertainties in these data sources and their changing relevance for the system under investigation.

5.1 INTRODUCTION

Hydrological models are valuable tools for short-term forecasting of river flows, long-term predictions for strategic water management planning but also to develop a better understanding of the complex interactions of water storage and release processes at the catchment scale. In spite of the wide variety of existing hydrological models, they mostly include similar functionalities of storage, transmission and release of water to represent the dominant hydrological processes of a particular river basin (Fenicia et al., 2011), differing mostly only in the detail of their parameterizations (Gupta et al., 2012, 2014; Hrachowitz and Clark, 2017).

In all of these models, each individual model component constitutes a separate hypothesis of how water moves through that specific part of the system. Frequently, the individual hypotheses remain untested. Instead only the model output, i.e. the aggregated response of these multiple hypotheses, is confronted with data of the aggregated response of a catchment to atmospheric forcing. Countless applications of different hydrological models in many different regions across the world over the last decades have shown that these models often provide relatively robust estimates of streamflow dynamics, for both calibration and evaluation periods. However, various combinations of different untested individual hypotheses, can and do lead to similar aggregated outputs, i.e. model equifinality (Beven, 2006; Clark et al., 2016).

To be useful for any of the above applications, it is thus of critical importance that not only the aggregated but also the individual behaviors of the respective hypotheses are consistent with their real-world equivalents. Given the complexity and heterogeneity of natural systems together with the general lack of suitable observations, this remains a major challenge in hydrology (e.g., Jakeman and Hornberger, 1993; Beven, 2000; Gupta et al., 2008; Andréassian et al., 2012).

Studies have addressed the issue by constraining the parameters of specific models through the use of additional data sources besides streamflow. Beven and Kirkby (1979); Güntner et al. (1999) and Blazkova et al. (2002) mapped saturated contributing areas during field surveys to constrain model parameters, while patterns of water tables in piezometers were used by Seibert et al. (1997); Lamb et al. (1998) and Blazkova et al. (2002). Other sources include satellite-based total water storage anomalies (e.g., Winsemius et al., 2006; Werth and Güntner, 2010; Yassin et al., 2017), evaporation (e.g., Livneh and Lettenmaier, 2012; Rakovec et al., 2016a; Bouaziz et al., 2018; Demirel et al., 2018; Hulsman et al., 2019), near-surface soil moisture (e.g., Franks et al., 1998; Brocca et al., 2010a; Sutanudjaja et al., 2014; Adnan et al., 2016; Kunnath-Poovakka et al., 2016; López López et al., 2017; Bouaziz et al., 2020), snow cover information (e.g., Gao et al., 2017; Bennett et al., 2019; Riboust et al., 2019), or a combination of these variables (e.g., Nijzink et al., 2018; Dembélé et al., 2020). Reflecting the results of many studies, Rakovec et al. (2016b) showed that streamflow is necessary but not sufficient to constrain model components to warrant partitioning of incoming precipitation to storage, evaporation and drainage.

Hydrological simulations are, however, not only affected by model parameter uncertainty, but also by the selection of a model structure and its parameterization (i.e. the

choice of equations). Modeling efforts over the last four decades have led to a wide variety of hydrological models providing flexibility to test competing modeling philosophies, from spatially lumped model representations of the system to high-resolution small-scale processes numerically integrated to the catchment scale (Hrachowitz and Clark, 2017). Haddeland et al. (2011) and Schewe et al. (2014) compared global hydrological models and found that differences between models are a major source of uncertainty. Nonetheless, model selection is often driven by personal preference and experience of individual modelers rather than detailed model test procedures (Holländer et al., 2009; Clark et al., 2015; Addor and Melsen, 2019).

A suite of comparison experiments tested and explored differences between alternative modeling structures and parameterizations (Perrin et al., 2001; Reed et al., 2004; Duan et al., 2006; Holländer et al., 2009; Knoben et al., 2020). However, these studies mostly restricted themselves to analyses of the models' skills to reproduce streamflow ("aggregated hypothesis"), with little consideration for the model internal processes ("individual hypotheses"). The Framework for Understanding Structural Errors (FUSE) was one of the first initiatives towards a more comprehensive assessment of model structural errors, with special consideration given to individual hypotheses (Clark et al., 2008).

Subsequent efforts towards more rigorous testing of competing model hypotheses, partially based on internal processes include Smith et al. (2012a,b) who tested multiple models for their ability to reproduce in-situ soil moisture observations as part of the Distributed Model Intercomparison Project 2 (DMIP2). They found that only two out of sixteen models provided reasonable estimates of soil moisture. In a similar effort, Koch et al. (2016) and Orth et al. (2015) also compared modeled soil moisture to in-situ observations of soil moisture for a range of hydrological models in different environments. In contrast, Fenicia et al. (2008) and Hrachowitz et al. (2014) used groundwater observations to test individual components of their models. There are actually relatively few studies that comprehensively quantified differences in internal process representation by simultaneously analyzing multiple models and multiple state and flux variables.

Here, in this model comparison study, we instead use globally available remote-sensing data to evaluate five different model state and flux variables of twelve process-based models with similar overall streamflow performance, which are calibrated by several research groups following an identical protocol. The calibration on streamflow was conducted in our previous study (de Boer-Euser et al., 2017), in which we compared models using hourly streamflow observations, leaving the modeled response of internal processes unused.

In a direct follow-up of the above study, we here hypothesize that process-based models with similar overall streamflow performance rely on similar representations of their internal states and fluxes. We test our hypothesis by simultaneously quantifying the differences in the magnitudes and dynamics of five internal state and flux variables of twelve models, in a comprehensive and systematic way. Our primary aim is to test if models calibrated to streamflow with similar high-performance levels in reproducing streamflow, follow similar pathways to do so, i.e. represent the system in a similar way. A secondary objective is to evaluate the plausibility of model behavior by introducing a

set of "soft" measures based on expert knowledge in combination with remote-sensing data of evaporation, snow cover, soil moisture and total water storage anomalies.

5.2 STUDY AREA

We test our hypothesis using data from three catchments in the Belgian Ardennes, the Ourthe upstream of Tabreux (ID1), the nested Ourthe Orientale upstream of Mabompré (ID2) and the Semois upstream of Membre-Pont (ID3), as shown in Fig. 6.1a,b.

The rain-fed Ourthe River at Tabreux (ID1) is fast-responding due to shallow soils and steep slopes in the catchment. Agriculture is the main land cover (27 % crops and 21 % pasture), followed by 46 % forestry and 6 % urban cover in an area of 1607 km^2 and an elevation ranging between 107 m and 663 m (European Environment Agency, 2000; de Boer-Euser et al., 2017). Mean annual precipitation, potential evaporation and streamflow are 979 mm yr^{-1}, 730 mm yr^{-1} and 433 mm yr^{-1} respectively for the period 2001–2017.

Nested within the Ourthe catchment (ID1), the Ourthe Orientale upstream of Mabompré (ID2) is characterized by a narrow elevation range from 294 m to 662 m, with 65 % of the catchment falling within a 100 m elevation band, making this catchment suitable to analyze snow processes modeled by lumped models. The Ourthe Orientale upstream of Mabompré has an area of 317 km^2 which corresponds to 20 % of the Ourthe area upstream of Tabreux and has similar land cover fractions. Mean annual precipitation, potential evaporation and streamflow for the period 2001–2017 are also relatively similar with 1052 mm yr^{-1}, 720 mm yr^{-1} and 462 mm yr^{-1}, respectively.

Forest is the main land cover in the Semois upstream of Membre-Pont (ID3) with 56 %, followed by agriculture (18 % pasture and 21 % crop) and 5 % urban cover. The Semois upstream of Membre-Pont is 24 % smaller than the Ourthe upstream of Tabreux with 1226 km^2 and elevation ranges between 176 m and 569 m. Mean annual precipitation, potential evaporation and streamflow are respectively 38 %, 4 % and 46 % higher in the Semois at Membre-Pont with 1352 mm yr^{-1}, 759 mm yr^{-1} and 634 mm yr^{-1}.

5.3 DATA

5.3.1 HYDROLOGICAL AND METEOROLOGICAL DATA

Hourly precipitation gauge data are provided by the Service Public de Wallonie (Service Public de Wallonie, 2018) and are spatially interpolated using Thiessen polygons. Daily minimum and maximum temperatures are retrieved from the 0.25° resolution gridded E-OBS dataset (Haylock et al., 2008) and disaggregated to hourly values by linear interpolation using the timing of daily minimum and maximum radiation at Maastricht (Royal Netherlands Meteorological Institute, 2018). Daily potential evaporation is calculated from daily minimum and maximum temperatures using the Hargreaves formula (Hargreaves and Samani, 1985) and is disaggregated to hourly values using a sine function during the day and no evaporation at night. We use the same forcing for 2000–2010 as

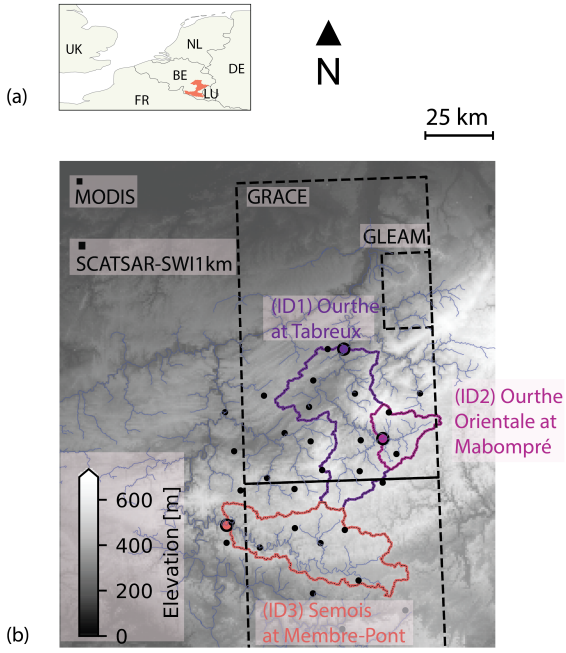

Figure 5.1: **(a)** Location of the study catchments in Belgium, northwestern Europe. **(b)** Digital elevation model and catchments of the Ourthe upstream of Tabreux (ID1), Ourthe Orientale upstream of Mabompré (ID2) and Semois upstream of Membre-Pont (ID3). Pixel size of GRACE, GLEAM, MODIS and SCATSAR-SWI1km. Colored dots are the streamflow gauging locations and black dots are the precipitation stations.

in the previous comparison study (de Boer-Euser et al., 2017) and follow the same approach to extend the meteorological dataset for the period 2011–2017. Uncertainty in meteorological data is not explicitly considered, but our primary aim is to compare the models forced with identical data. Observed hourly streamflow data for the Ourthe at Tabreux, Ourthe Orientale at Mabompré and Semois at Membre-Pont are provided by the Service Public de Wallonie for the period 2000–2017.

5.3.2 REMOTE-SENSING DATA

GLEAM EVAPORATION

The Global Land Evaporation Amsterdam Model (GLEAM, Miralles et al., 2011; Martens et al., 2017) provides daily estimates of land evaporation by maximizing the information recovery on evaporation contained in climate and environmental satellite observations. The Priestley and Taylor (1972) equation is used to calculate potential evaporation for bare soil, short canopy and tall canopy land fractions. Actual evaporation is the sum of interception and potential evaporation reduced by a stress factor. This evaporative stress factor is based on microwave observations of vegetation optical depth and estimates of root-zone soil moisture in a multi-layer water balance model. Interception evaporation is estimated separately using a Gash analytical model and only depends on precipitation and vegetation characteristics. GLEAM v3.3a relies on reanalysis net radiation and air temperature from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 data, satellite and gauge-based precipitation, satellite-based vegetation optical depth, soil moisture and snow water equivalent. The data are available at 0.25° resolution (Fig. 6.1b) and account for subgrid heterogeneity by considering three land cover types. We spatially average GLEAM interception and total actual evaporation estimates over the Ourthe catchment upstream of Tabreux for the period 2001–2017.

MODIS SNOW COVER

The Moderate Resolution Imaging Spectroradiometer (MODIS) AQUA (MYD10A1, version 6) and TERRA (MOD10A1, version 6) satellites provide daily maps of the areal fraction of snow cover per 500 m × 500 m cell (Fig. 6.1b) based on the Normalized Difference Snow Index (Hall and Riggs, 2016a,b). For each day, AQUA and TERRA observations are merged into a single observation by taking the mean fraction of snow cover per day. The percentage of cells with a fractional snow cover larger than zero and fraction of cells without missing data (i.e. due to cloud cover) for the catchment of the Ourthe Orientale upstream of Mabompré is calculated for each day. For this study, we disregard observations during the summer months (JJA, when temperatures did not drop below 4°C) and only use daily observations in which at least 40 % of the catchment area has snow cover retrievals not affected by clouds, implying that we have 1463 valid daily observations of mean fractional snow cover. This corresponds to 32 % of all observations of the Ourthe Orientale catchment upstream of Mabompré between 2001 and 2017.

SCATSAR-SWI1KM SOIL WATER INDEX

SCATSAR-SWI1km is a daily product of soil water content relative to saturation at a 1 km × 1 km resolution (Fig. 6.1b) obtained by fusing spatio-temporally complementary radar sensors (Bauer-Marschallinger et al., 2018). Estimates of the moisture content relative to saturation at various depths in the soil, referred to as Soil Water Index (SWI), are obtained through temporal filtering of the 25 km METOP ASCAT near-surface soil moisture (Wagner et al., 2013) and 1 km Sentinel-1 near-surface soil moisture (Bauer-Marschallinger et al., 2018). The Soil Water Index features as single parameter the characteristic time length T (Wagner et al., 1999; Albergel et al., 2008). The T-value is required to convert near-surface soil moisture observations to estimates of root-zone soil moisture. The T-value increases with increasing root-zone storage capacities (Bouaziz et al., 2020), resulting in more smoothing and delaying of the near-surface soil moisture signal. The Copernicus Global Land Service (2019) provides the Soil Water Index for T-values of 2, 5, 10, 15, 20, 40, 60 and 100 days. Since Sentinel-1 was launched in 2014, the Soil Water Index is available for the period 2015–2017. We calculate the mean soil moisture over all SCATSAR-SWI1km pixels within the Ourthe upstream of Tabreux for the available period.

5

GRACE TOTAL WATER STORAGE ANOMALIES

The Gravity Recovery and Climate Experiment (GRACE, Swenson and Wahr, 2006; Swenson, 2012) twin satellites, launched in March 2002, measure the Earth's gravity field changes by calculating the changes in the distance between the two satellites as they move one behind the other in the same orbital plane. Monthly total water storage anomalies (in mm) relative to the 2004–2009 time-mean baseline are provided at a spatial sampling of 1° (approximately 78 km x 110 km at the latitude of the study region, Fig. 6.1b) by three centers: U. Texas / Center for Space Research (CSR), GeoForschungsZentrum Potsdam (GFZ) and Jet Propulsion Laboratory (JPL). These centers apply different processing strategies which lead to variations in the gravity fields. These gravity fields require smoothing of the noise induced by attenuated short wavelength. The spatial smoothing decreases the already coarse GRACE resolution even further through signal "leakage" of one location to surrounding areas (Bonin and Chambers, 2013), which increases the uncertainty especially at the relatively small scale of our study catchments. We apply the scaling coefficients provided by NASA to restore some of the signal loss due to processing of GRACE observations (Landerer and Swenson, 2012). The data of the three processing centers are each spatially averaged over the catchments of the Ourthe upstream of Tabreux and the Semois upstream of Membre-Pont for the period April 2002 to February 2017.

5.3.3 DATA UNCERTAINTY

The hydrological evaluation data are all subject to uncertainties (Beven, 2019a). Streamflow is not measured directly but depends on water level measurements and a rating curve. Westerberg et al. (2016) quantify a median streamflow uncertainty of ±12 %, ±24 % and ±34 % for average, high and low streamflow conditions, respectively, using a Monte Carlo sampling approach of multiple feasible rating curve for 43 UK catchments. We

sample from these uncertainty ranges to transform the streamflow observations (100 realizations). We then quantify signature uncertainty originating from streamflow data uncertainty using the 100 sampled time series for a selection of streamflow signatures (Sect. 5.4.2). The 5-95th uncertainty bounds of median annual streamflow, baseflow and flashiness indices result in ±11 %, ±9 % and ±12 %, respectively. These magnitudes are similar to those reported by Westerberg et al. (2016).

GLEAM evaporation estimates are inferred from models and forcing data which are all affected by uncertainty. Yet, uncertainty estimates of GLEAM evaporation are not available. However, GLEAM evaporation was evaluated against FLUXNET data by Miralles et al. (2011). For the nearby station of Lonzee in Belgium, they report similar annual rates and a high correlation coefficient of 0.91 between the daily time series. GLEAM mean annual evaporation was compared to the ensemble mean of five evaporation datasets in Miralles et al. (2016) and shows higher than average values in Europe (of approximately 60 mm yr^{-1} or 10 % of mean annual rates for our study area). The partitioning of evaporation in different components (transpiration, interception and soil evaporation) differs substantially between different evaporation datasets, as shown by Miralles et al. (2016). GLEAM interception currently only considers tall vegetation and underestimates in-situ data (Zhong et al., 2020) and is ~50 % lower than estimates from other datasets (Miralles et al., 2016). These uncertainties underline that GLEAM (and other remote-sensing data) cannot be considered as a reliable representation of real-world quantities. However, the comparison of daily dynamics and absolute values of this independent data source with modeled results is still valuable to detect potential outliers and understand their behavior. Besides, the different methods used to estimate potential evaporation of GLEAM and our model forcing should not impede us from testing the consistency between the resulting actual evaporation (Oudin et al., 2005).

Most frequent errors within the MODIS snow cover products are due to cloud/snow discrimination problems. Daily MODIS snow maps have an accuracy of approximately 93 % at the *pixel scale*, with lower accuracy in forested areas, complex terrain and when snow is thin and ephemeral and higher accuracy in agricultural areas (Hall and Riggs, 2007). However, here, MODIS data is used to estimate the number of days with snow at the *catchment scale*. We expect lower classification errors at the catchment scale as it would require many pixels to be misclassified at the same time. For each day and each pixel of valid MODIS observations, we sample from a binomial distribution with a probability of 93 % that MODIS is correct when the pixel is classified as snow and assume a higher probability of 99 % that MODIS is correct when the pixel is classified as no-snow to prevent overestimating snow for days without snow (Ault et al., 2006; Parajka and Blöschl, 2006). We repeat the experiment for 1000 times in a Monte Carlo procedure. This results in less than ±2 % uncertainty in the number of days when MODIS observes snow at the catchment scale.

The soil water content relative to saturation of SCATSAR-SWI1km is estimated from observed radar backscatter through a change detection approach, which interprets changes in backscatter as changes in soil moisture, while other surface properties are assumed static (Wagner et al., 1999). The degree of saturation of the near-surface is given in relative units from 0 % (dry reference) to 100 % (wet reference) and is converted to deeper

layers through an exponential filter called the Soil Water Index. The smoothing and delaying effect of the Soil Water Index narrows the range of the near-surface degree of saturation. Therefore, data matching techniques are often used to rescale satellite data to match the variability of modeled or observed data (Brocca et al., 2010a), which suggests the difficulty to meaningfully compare the *range* of modeled and remote-sensing estimates of root-zone soil moisture content relative to saturation. However, the *dynamics* of SCATSAR-SWI1km data have been evaluated against in-situ stations of the International Soil Moisture Network, despite commensurability issues of comparing in-situ point measurements and areal satellite data. Spearman rank correlation coefficients of 0.56 are reported for T-values up to 15 days and 0.43 for T-value of 100 days (Bauer-Marschallinger, 2020).

GRACE estimates of total water storage anomalies suffer from signal degradation due to measurement errors and noise. Filtering approaches are applied to reduce these errors, but induce leakage of signal from surrounding areas. The uncertainty decreases as the size of the region under consideration increases. However, time series of a single pixel may still be used to compare dynamics and amplitudes of total water storage anomalies despite possible large uncertainty (Landerer and Swenson, 2012). We estimate an uncertainty in total water storage anomalies of ~18 mm in the pixels of our catchments by combining measurement and leakage errors in quadrature, which are both provided for each grid location (Landerer and Swenson, 2012).

5.4 METHODS

5.4.1 MODELS AND PROTOCOL

Eight research groups (Wageningen University, Université de Lorraine, Leuven University, Delft University of Technology, Deltares, Irstea (now INRAE), Eawag and Flanders Hydraulics Research) participated in the comparison experiment and applied one or several hydrological models (Fig. 5.2). The models include WALRUS (Wageningen Lowland Runoff Simulator, Brauer et al., 2014a,b), PRESAGES (PREvision et Simulation pour l'Annonce et la Gestion des Etiages Sévères, Lang et al., 2006), VHM (Veralgemeend conceptueel Hydrologisch Model, Willems, 2014), FLEX-Topo which was still under development when it was calibrated for our previous study (Savenije, 2010; de Boer-Euser et al., 2017; de Boer-Euser, 2017), a distributed version of the HBV model (Hydrologiska Byråns Vattenbalansavdelning, Lindström et al., 1997), SUPERFLEX M2 to M5 models (Fenicia et al., 2011, 2014b), dS2 (distributed simple dynamical systems, Buitink et al., 2020), GR4H (Génie Rural à 4 paramètres Horaire, Mathevet, 2005; Coron et al., 2017, 2019) combined with the CemaNeige snow module (Valéry et al., 2014) and NAM (NedborAfstrommings Model, Nielsen and Hansen, 1973). Main differences and similarities between models in terms of snow processes, root-zone storage, total storage and evaporation processes are summarized in Tables 5.1-5.3.

In our previous study (de Boer-Euser et al., 2017), we defined a modeling protocol to limit the degrees of freedom in the modeling decisions of the individual participants (Ceola et al., 2015), allowing us to meaningfully compare the model results. The proto-

col involved forcing the models with the same input data and calibrating them for the same time period, using the same objective functions. However, participants were free to choose a parameter search method, as we considered it to be part of the modelers experience with the model, even if this would make comparison less straightforward. The models were previously calibrated using streamflow of the Ourthe at Tabreux (ID1) for the period 2004 to 2007, using 2003 as a warm-up year (de Boer-Euser et al., 2017). The Nash-Sutcliffe efficiencies of the streamflow and the logarithms of the streamflow were simultaneously used as objective functions to select an ensemble of feasible parameter sets to account for parameter uncertainty and ensure a balance between the models' ability to reproduce both high and low flows. The temporal and spatial transferability of the models was tested by evaluating the models in a pre- and post-calibration period and by applying the calibrated model parameter sets to nested and neighboring catchments, including catchment ID2 and ID3 (Klemeš, 1986). Results thereof are presented in de Boer-Euser et al. (2017).

In the current study, we run the calibrated models for an additional period from 2011 to 2017 for the Ourthe at Tabreux (ID1), the Ourthe Orientale at Mabompré (ID2) and the Semois at Membre-Pont (ID3). The modeling groups have provided simulation results for each catchment in terms of streamflow, groundwater losses/gains, interception evaporation, root-zone evaporation (transpiration and soil evaporation), total actual evaporation, snow storage, root-zone storage and total storage as a sum of all model storage volumes (Table 5.2) at an hourly time step for the total period 2001–2017. We compare these modeled states and fluxes and evaluate them against their remote-sensing equivalents as further explained in Sects. 5.4.2 and 5.4.3.

5.4.2 MODEL EVALUATION: WATER BALANCE

All models are evaluated in terms of the long-term water balance, which indicates the partitioning between drainage and evaporative fluxes and allows us to assess long-term conservation of water and energy. We compare mean annual streamflow with observations and mean annual actual evaporation and interception evaporation with GLEAM estimates for the Ourthe at Tabreux during the evaluation period 2008–2017. A detailed description of streamflow performance for specific events (low and high flows, snowmelt event, transition from dry to wet period) has been detailed in the previous paper (de Boer-Euser et al., 2017). In the current study, differences in streamflow dynamics are briefly summarized by assessing observed and modeled baseflow indices (I_b, van Dijk, 2010) and flashiness indices (I_f, Fenicia et al., 2016), as these are representative of the partitioning of drainage into fast and slow responses. Seasonal dynamics of actual evaporation over potential evaporation and runoff ratios during winter (Oct-Mar) and summer (Apr-Sep) are compared between models.

5.4.3 MODEL EVALUATION: INTERNAL STATES

We compare modeled snow storage, root-zone soil moisture and total storage between models and with remote sensing estimates of MODIS snow cover, SCATSAR-SWI1km Soil Water Index and GRACE total storage anomalies, respectively, as shown in Tables 5.2-5.3.

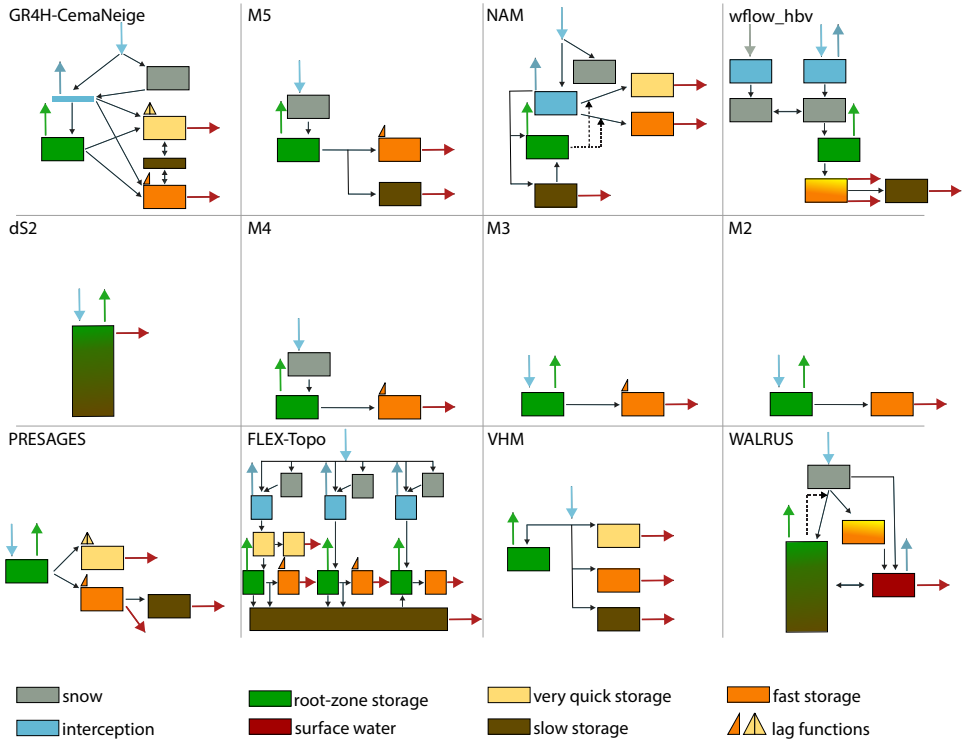

Figure 5.2: Simplified schematic overview of 12 model structures (adapted from de Boer-Euser et al., 2017) with the aim to highlight similarities and differences between the models. Solid arrows indicate fluxes between stores, while dashed arrows indicate the influence of a state to a flux. Colored arrows indicate incoming or outgoing fluxes, whereas black arrows indicate internal fluxes. The narrow blue rectangle in GR4H indicates the presence of an interception module without interception storage capacity (Table 5.3). Storages with a color gradient indicate the combination of several components in one reservoir. FLEX-Topo consists of three hydrological response units connected through a shared slow reservoir and wflow_hbv is a distributed model.

SNOW DAYS

As most models used in this study are lumped, it is not possible to spatially evaluate modeled snow cover versus MODIS snow cover. However, we can classify each day in a binary way according to the occurrence of snow, based on a threshold for the percentage of cells in the catchment where snow cover is detected. MODIS snow cover observations are classified as days with and without snow using thresholds of both 10 and 15 % of snow-covered cells in the catchment to be counted as a day with snow, in a sensitivity analysis. For each model, snow days are distinguished from non-snow days whenever the water stored as snow is above 0.05 mm to account for numerical rounding. For each model (and each retained parameter set), we then compare if modeled snow coincides with 'truly' observed snow by MODIS, for each day with a valid MODIS observation. We create a confusion matrix with counts of true positives when observations and model results agree on the presence of snow (hits), false positives when the model indicates

the presence of snow but this is not observed by MODIS (false alarms), false negatives when the model misses the presence of snow observed by MODIS (miss) and true negatives when observations and model results agree on the absence of snow (correct rejections). This allows us to identify the trade-off between, on the one hand, the miss rate between model and remote-sensing observation, as the ratio of misses over actual positives (number of days when snow is observed by MODIS) and, on the other hand, the false discovery rate as the ratio of false alarms over predicted positives (number of days when snow is modeled). We also compare annual maximum snow storage and number of days with snow between the seven models with a snow module (GR4H, M5, NAM, wflow_hbv, M4, FLEX-Topo, WALRUS). The snow analysis is performed in the catchment of the Ourthe Orientale upstream of Mabompré as it features the narrowest elevation range among the study catchments (i.e. 294-662 m a.s.l. versus 108-662 m for the Ourthe upstream of Tabreux) and thus plausibly permits a lumped representation of the snow component.

ROOT-ZONE SOIL MOISTURE

We compare the range of the relative root-zone soil moisture $\overline{S}_R$ ($\overline{S}_R = S_R/S_{R,max}$, Table 5.1) between models for the period in which SCATSAR-SWI1km is available (2015–2017). Time series of catchment-scale root-zone soil moisture are available for all models except WALRUS and dS2 as these models have a combined soil reservoir (Fig. 5.2). The dS2 model only relies on the sensitivity of streamflow to changes in total storage. In WALRUS, the state of the soil reservoir (which includes the root zone) is expressed as a storage deficit and is therefore not bound by an upper limit (Table 5.2). Root-zone storage capacities ($S_{R,max}$, mm) are available as calibration parameter for all other models. We relate the range in relative root-zone soil moisture to the maximum root-zone storage capacity $S_{R,max}$, because we expect models with small root-zone storage capacities $S_{R,max}$ to entirely utilize the available storage, through complete drying and saturation.

We then compare the similarity of the dynamics of modeled time series of the relative root-zone soil moisture with remotely sensed SCATSAR-SWI1km Soil Water Index for several values of the characteristic time length parameter (T in days). The T-value has previously been positively correlated with root-zone storage capacity, assuming a high temporal variability of root-zone soil moisture and therefore a low T-value for small root-zone storage capacities $S_{R,max}$ (Bouaziz et al., 2020). For each model and feasible realization, we identify the T-value that yields the highest Spearman rank correlation between modeled root-zone soil moisture and Soil Water Index. We then relate the optimal T-value to the root-zone storage capacity $S_{R,max}$. This analysis enables us to identify potential differences in the representation and the dynamics of root-zone storage between models.

TOTAL STORAGE ANOMALIES

For each model, we calculate time series of total storage (Table 5.2) and mean monthly total storage anomalies relative to the 2004-2009 time-mean baseline for comparison with GRACE estimates for the Ourthe upstream of Tabreux (ID1) and the Semois upstream of Membre-Pont (ID3). Both catchments coincide with two neighboring GRACE

cells, allowing us to test how well the models reproduce the observed spatial variability. We further relate the modeled range of total storage (maximum minus minimum total storage over the time series) to Spearman rank correlation coefficients between modeled and GRACE estimates of total storage anomalies.

5.4.4 INTERACTIONS BETWEEN STORAGE AND FLUXES DURING DRY PERIODS

The impact of a relatively large (> 200 mm) versus relatively small (< 150 mm) root-zone storage capacity on actual evaporation, streamflow and total storage is assessed during a dry period in September 2016 by selecting two representative models with high streamflow model performance (GR4H and M5). The plausibility of the hydrological response of these two model representations is evaluated against remote-sensing estimates of root-zone soil moisture and actual evaporation.

5.4.5 PLAUSIBILITY OF PROCESS REPRESENTATIONS

The models are subsequently ranked and evaluated in terms of their consistency with observed streamflow, remote-sensing data and expert knowledge with due consideration of the uncertainty in the evaluation data, as detailed in Sect. 5.3.3. We summarize our main findings by evaluating the models in terms of their deviations around median annual streamflow, flashiness and baseflow indices, median annual actual evaporation and interception compared to GLEAM estimates, the number of days with snow over valid MODIS observations, the number of days per year with empty root-zone storage and the very dry total storage anomalies compared to GRACE estimates.

5.5 RESULTS

5.5.1 WATER BALANCE

STREAMFLOW

All models show high Nash-Sutcliffe Efficiencies of the streamflow and the logarithm of the streamflow ($E_{NS,Q}$ and $E_{NS,logQ}$) with median values of above 0.7 for the post-calibration evaluation period 2008–2017 (Fig. 5.3a and Table 5.2 for the calculation of the Euclidean distances). The interannual variability of streamflow agrees strongly with observations for each model in the period 2008–2017 (Fig. 5.3b). The difference between modeled and observed median streamflow varies between -5.6 % and 5.6 % and the difference in total range varies between -9.6 % and 20 %. This is in line with our results in the previous paper, in which we also showed that all models perform well in terms of commonly used metrics (de Boer-Euser et al., 2017). However, there are differences in the partitioning of fast and slow runoff, as shown by the flashiness and baseflow indices (I_f and I_b) in Fig. 5.3c. Largest underestimation of the flashiness index occurs for M2 and dS2 (~20 %), while FLEX-Topo shows the highest overestimation (26 %). FLEX-Topo and WALRUS underestimate the baseflow index most (41 % and 70 % respectively), while

GR4H and M5 show the highest overestimation (15 % and 21 % respectively). There is a strong similarity between modeled and observed hydrographs for one of the best performing models M5, as quantified by its low Euclidean distance (Fig. 5.3d and Table 5.2). The GR4H model is the only model which includes deep groundwater losses, but they are very limited and represent only 1.6 % of total modeled streamflow of the Ourthe at Tabreux, or approximately 7 mm yr^{-1}.

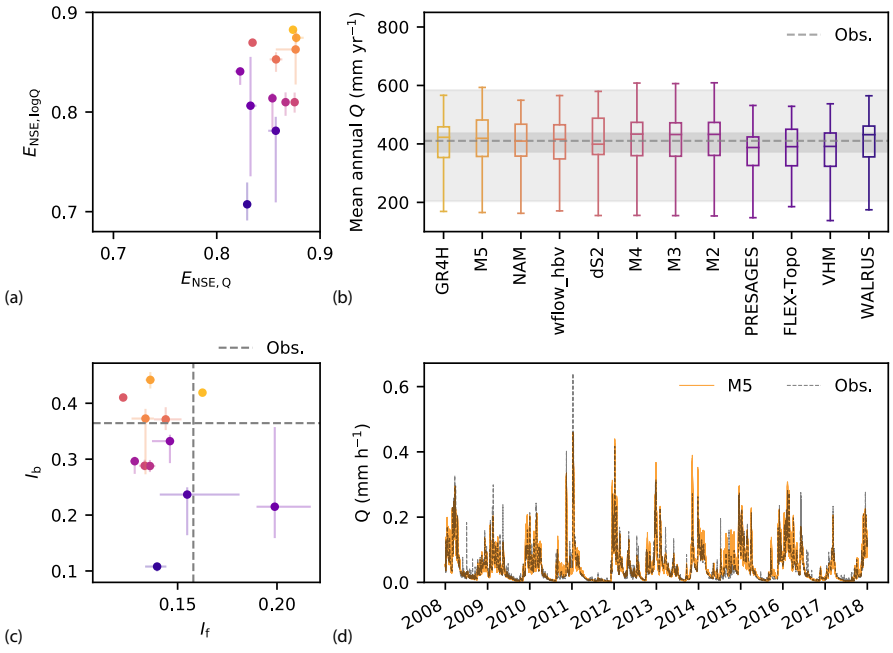

Figure 5.3: Evaluation of modeled streamflow performance for the Ourthe at Tabreux for the period 2008–2017. **(a)** Nash-Sutcliffe Efficiencies of the streamflow $E_{NS,Q}$ and the logarithm of the streamflow $E_{NS,logQ}$ (median, 25/75th percentiles across parameter sets). **(b)** Modeled mean annual streamflow for hydrological years between 2008–2017 across feasible parameter sets. The models are ranked from the highest to the lowest performance according to the Euclidean distance of streamflow performance (see Table 5.2). The dashed line and grey shaded areas show median, 25/75th and minimum-maximum range of observed mean annual streamflow. **(c)** Baseflow index I_b as a function of the flashiness index I_f (median, 25/75th percentiles across parameter sets). Observed values are shown by the grey dashed lines. **(d)** Observed and modeled hydrographs of model M5 with high streamflow model performance (low Euclidean distance).

ACTUAL EVAPORATION

Modeled median annual actual evaporation E_A (computed as the sum of soil evaporation, transpiration, (separate) interception evaporation and, if applicable, sublimation, Table 5.3) for hydrological years between October 2008 and September 2017 varies between 507 and 707 mm yr^{-1} across models, with a median of 522 mm yr^{-1}, which is approximately 10 % lower than the GLEAM estimate of 578 mm yr^{-1}, as shown in Fig. 5.4a. Annual actual evaporation of the VHM model is very high compared to the other mod-

els, with a median of 707 mm yr^{-1} and approximates potential evaporation (median of 732 mm yr^{-1}). Calibration of the VHM model is meant to follow a manual stepwise procedure including the closure of the water balance during the identification of soil moisture processes (Willems, 2014). However, in the automatic calibration prescribed by the current protocol, this step was not performed, which explains the unusual high actual evaporation in spite of relatively similar annual streamflow compared to the other models, as there is no closure of the water balance (Fig. 5.3a).

Interception evaporation is included in four models, with GR4H showing the lowest annual interception evaporation of 100 mm yr^{-1} (19 % of E_A or 10 % of P), FLEX-Topo and wflow_hbv having relatively similar amounts of approximately 250 mm yr^{-1} (~45 % of E_A or 26 % of P) and NAM having the highest annual interception evaporation of 340 mm yr^{-1} (65 % of E_A or 36 % of P), as shown in Fig. 5.4a. Differences are related to the presence and maximum size of the interception storage (I_{max}), as shown in Table 5.3. GLEAM interception estimates of 189 mm yr^{-1} are almost twice as high as GR4H estimates, 25 % lower than FLEX-Topo and wflow_hbv, and 44 % lower than NAM values, suggesting a large uncertainty in the contribution of interception and transpiration to actual evaporation. For comparison, measurements of the fraction of interception evaporation over precipitation in forested areas vary significantly depending on the site location, with estimates of 37 % for a Douglas fir stand in the Netherlands (Cisneros Vaca et al., 2018), 27 %, 32 % and 42 % for three coniferous forests of Great Britain (Gash et al., 1980) and 50 % for a forest in Puerto Rico (Schellekens et al., 1999) and are difficult to extrapolate to other catchments due to the heterogeneity and complexity of natural systems.

GLEAM estimates of actual evaporation show relatively high evaporation rates in winter and are never reduced to zero in summer, as opposed to modeled M5 estimates, as shown in Fig. 5.4b. GLEAM actual evaporation minus the separately calculated interception is 94 % of potential evaporation, implying almost no water-limited conditions, as opposed to our models in which actual evaporation in summer (Apr–Sep) is, due to water stress, reduced to approximately 73 % of potential evaporation on average for all models except VHM (Fig. 5.4c). Larger differences between models occur in the ratio E_A/E_P during winter (Oct–Mar), when FLEX-Topo, wflow_hbv and VHM show E_A/E_P ratios close to unity, and dS2 the lowest values of $E_A/E_P \sim 0.75$ as shown in Fig. 5.4c. The dS2 model differs from all other models as it relies on a year-round constant water stress coefficient (C_{cst}), independent of water supply, while the stress coefficient depends on root-zone soil moisture content in all other models (Table 5.3).

Most models slightly overestimate summer runoff ratios with values between 0.22 and 0.26 which are very close to the observed value of 0.22, as shown in Fig. 5.4d. During winter, runoff ratios vary between 0.55 and 0.71, which is close to the observed value of 0.66. This implies a relatively high level of agreement between models in reproducing the medium- to long-term partitioning of precipitation into evaporation and drainage and thus in approximating at least long-term conservation of energy (Hrachowitz and Clark, 2017).

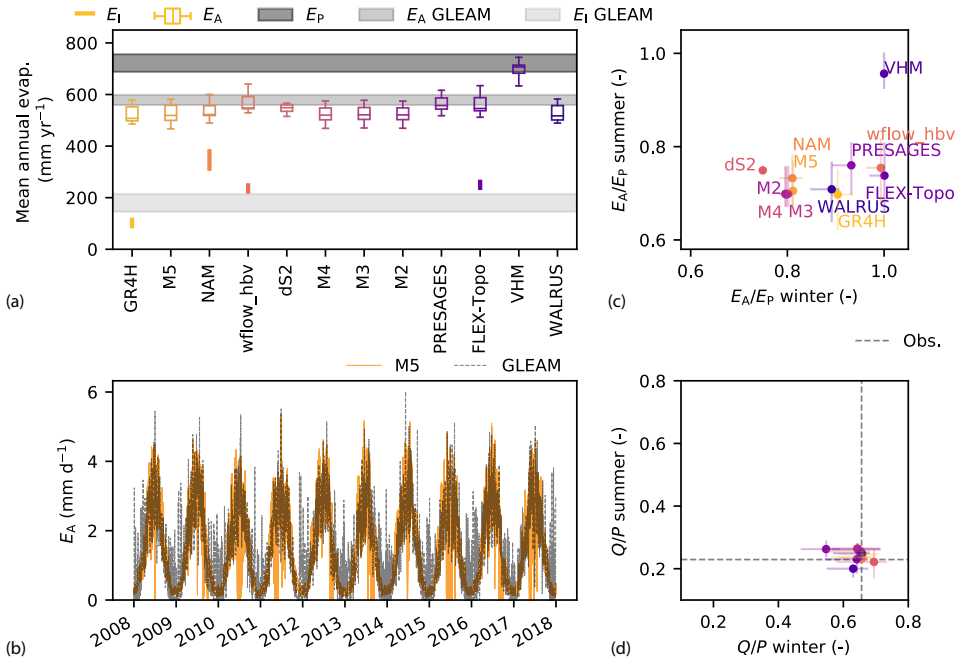

Figure 5.4: Evaluation of modeled evaporation for the Ourthe upstream of Tabreux for the period 2008–2017. **(a)** Modeled mean annual actual evaporation E_A and minimum-maximum range of mean annual interception evaporation E_I for hydrological years between 2008–2017 across feasible parameter sets. The dark grey shaded area shows the range of potential evaporation E_P used as input for the models. The light grey shaded area shows GLEAM actual and interception evaporation. **(b)** Daily actual evaporation from GLEAM and modeled by the M5 model. **(c)** Summer against winter E_A/E_P ratios for each model (median and 25/75th percentiles across parameter sets). **(d)** Summer against winter runoff ratio Q/P for each model (median and 25/75th percentiles across parameter sets), plotted on the same scale. The dashed grey lines indicate the observed median runoff ratios in summer and winter.

5.5.2 INTERNAL MODEL STATES

SNOW DAYS

MODIS snow cover is detected over most of the catchment area for some time each year between November 2001 and November 2017, except for the periods of November 2006 to March 2007 and November 2007 to March 2008, when snow is detected in less than half of the catchment cells, as shown in Fig. 5.5a. The number and magnitude of modeled snow storage events varies between models (Fig. 5.5b). The modeled number of snow days per hydrological year varies from ~28 days for FLEX-Topo, WALRUS and wflow_hbv to ~62 days for GR4H and ~90 days for NAM, M4 and M5, as shown in Fig. 5.5c. The variability in median annual maximum snow storage varies from 3 mm for wflow_hbv and ~5-6 mm for FLEX-Topo and WALRUS to ~10 mm for GR4H, M4, M5 and 15 mm for NAM. We further evaluate the plausibility of these modeled snow processes by comparing modeled and observed snow cover, for days when a valid MODIS observation is available.

The presence of snow modeled by FLEX-Topo, wflow_hbv and WALRUS coincides for 92 % with the presence of snow observed by MODIS. However, these models fail to model snow for ~62 % of days when MODIS reports the presence of snow, implying that these models miss many observed snow days, but when they predict snow, it was also observed (Fig. 5.5d).

NAM, M4 and M5, on the other hand, predict the presence of snow which coincides with observed snow by MODIS in ~68 % of the positive predictions, implying a relatively high probability of false alarm snow prediction of ~32 %. However, they miss only ~29 % of actual positive snow days observed by MODIS (Fig. 5.5d). This suggests that these models miss fewer observed snow days, but they also overpredict snow days numbers, which could be related to the use of a single temperature threshold to distinguish between snow and rain, as opposed to a temperature interval in the other models (Table 5.2).

GR4H is in between the two previously mentioned model categories, with a snow prediction which coincides with observed snow by MODIS in 79 % of the positive predictions and therefore only 21 % of false alarms. The model misses 42 % of actual positive snow days observed by MODIS. GR4H therefore shows a more balanced trade-off between the number of false alarms and the amount of observed snow events missed. This is illustrated in Fig. 5.5d.

With an increased threshold to distinguish snow days in MODIS, from 10 % to 15 % of cells in the catchment with a detected snow cover (Fig. 5.5d and Fig. 5.5e respectively), we decrease the number of observed snow days. For all models, this leads to an increase in the ratio of false alarms over predicted snow days but also a decrease of the ratio of missed events over actual observed snow days by MODIS. However, as all models are similarly affected by the change in threshold, our findings on the differences in performance between models show little sensitivity to this threshold.

Despite the large variability in snow response between models, snow processes are represented by a degree-hour method in all models, suggesting a high sensitivity of the snow response to the snow process parameterization (Table 5.2).

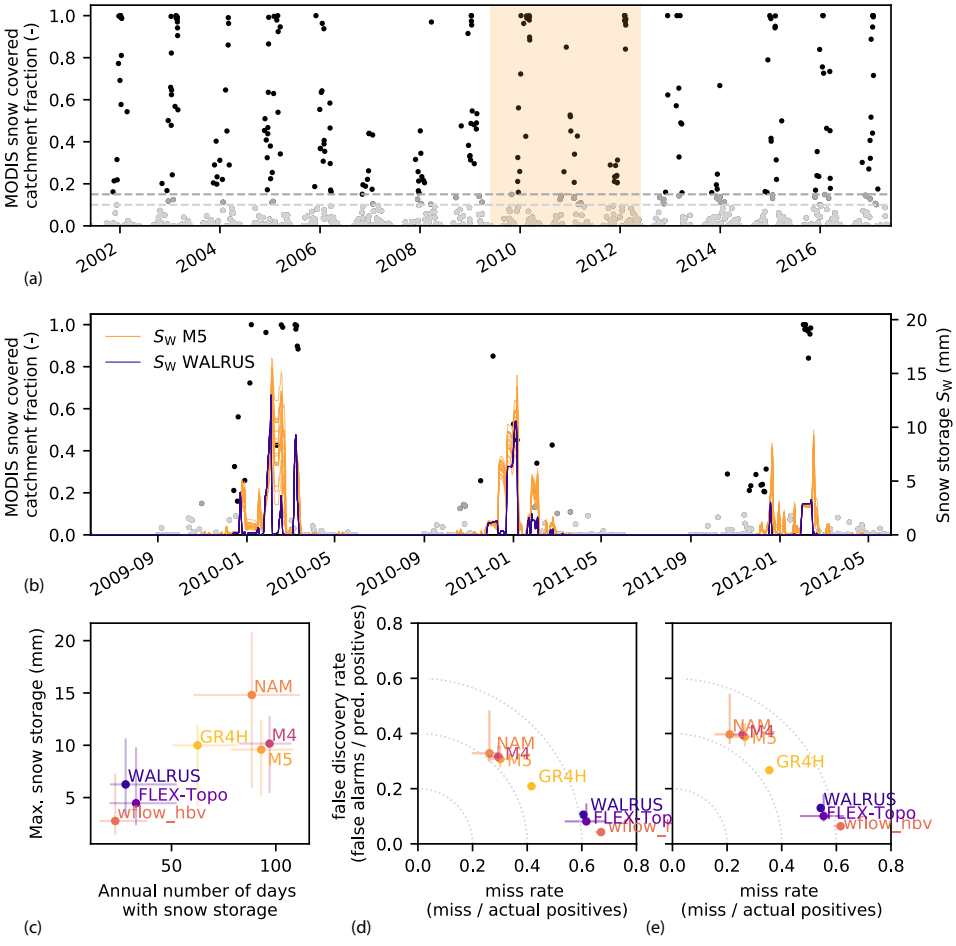

Figure 5.5: **(a)** Fraction of cells with a MODIS areal fraction snow cover greater than zero in the Ourthe Orientale upstream of Mabompré for the period 2001–2017. MODIS data are available once every three days on average. The dashed lines show the two thresholds of 10 % and 15 % selected to distinguish snow days. **(b)** Modeled snow storage for two contrasting models M5 and WALRUS for the light orange shaded period. **(c)** Median annual maximum snow storage as a function of number of days per year with snow. Light (yellowish) colors indicate models with higher performance (lower Euclidean distances). The vertical and horizontal error bars indicate the 25/75th percentiles over time and feasible parameter sets. **(d,e)** Two-dimensional representation of the false discovery rate as a function of the miss rate, when applying a threshold of **(d)** 10 % and **(e)** 15 % of cells within the catchment with snow cover greater than zero. In this representation, the perfect model would be at the origin (0 % misses and 0 % false alarms).The dotted lines show the distance from the origin. The vertical and horizontal error bars indicate the uncertainty within feasible parameter sets.

ROOT-ZONE SOIL MOISTURE

Vegetation-accessible water volumes that can be stored in the root zone largely control the long-term partitioning of precipitation into evaporation and drainage. Most hydrological models include a representation of this root-zone storage capacity $S_{R,max}$, which is estimated through calibration (Table 5.2). The calibrated root-zone storage capacities vary between 74 mm and 277 mm across studied models. The root-zone soil moisture content relative to saturation of models with relatively large root-zone storage capacities (here defined as $S_{R,max} > 200$ mm) tends to never fully dry out (>0.20) and saturate (<0.94) as opposed to models with lower root-zone storage capacities ($S_{R,max} < 150$ mm), in which the storage tends to either dry out completely and/or fully saturate (Fig. 5.6a). If the vegetation-accessible water storage dries out, this will lead to water stress and reduced transpiration. On the other hand, if the root-zone storage is saturated, no more water can be stored, resulting in fast drainage. The size of the root-zone storage capacity is therefore a key control of the hydrological response, allowing us to explain some of the observed variability between models.

We compare the dynamics of modeled and remote-sensing estimates of root-zone soil moisture by calculating Spearman rank correlations between modeled root-zone soil moisture and remote-sensing estimates of the Soil Water Index for the available T-values of 2, 5, 15, 20, 40, 60 and 100 days. As the T-value increases, the Soil Water Index is more smoothed and delayed. For each model realization, we identify the T-value which yields the highest Spearman rank correlation between Soil Water Index and modeled root-zone soil moisture (Fig. 5.6b). The optimal T-value increases with the size of the calibrated root-zone storage capacity and varies between 15 and 60 days. A small root-zone storage capacity is indeed likely to fill through precipitation and empty through evaporation and drainage more rapidly than a large water storage capacity, leading to a higher temporal relative soil moisture variability. The mismatch between the relatively high root-zone storage capacities of VHM ($S_{R,max} \sim 200$ mm) in relation to the relatively low optimal T-values of 20 days is likely related to the unclosed water balance (Sect. 5.5.1). The similarity between modeled root-zone soil moisture and Soil Water Index with optimal T-values is high, as implied by Spearman rank correlations varying between 0.88 and 0.90 across models. However, the disparity in optimal T-values between models underlines the different temporal representations of root-zone soil moisture content across models, implying that all these models cannot simultaneously provide a plausible representation of the catchment-scale vegetation-accessible water content.

TOTAL STORAGE ANOMALIES

Total water storage anomalies obtained from GRACE are compared to the storage as simulated by the models, showing relatively similar seasonal patterns, as illustrated in Fig. 5.7a for model M5. GRACE total storage anomalies of the Semois upstream of Membre-Pont and the Ourthe upstream of Tabreux are mainly represented by two neighboring cells (Fig. 6.1b), allowing us to test how models represent the observed spatial variability. The range of anomalies in the Semois upstream of Membre-Pont is larger than in the Ourthe upstream of Tabreux, implying 18 %, 3 % and 7 % less summer and 19 %, 19 % and 10 % more winter storage in the Semois upstream of Membre-Pont for each of the

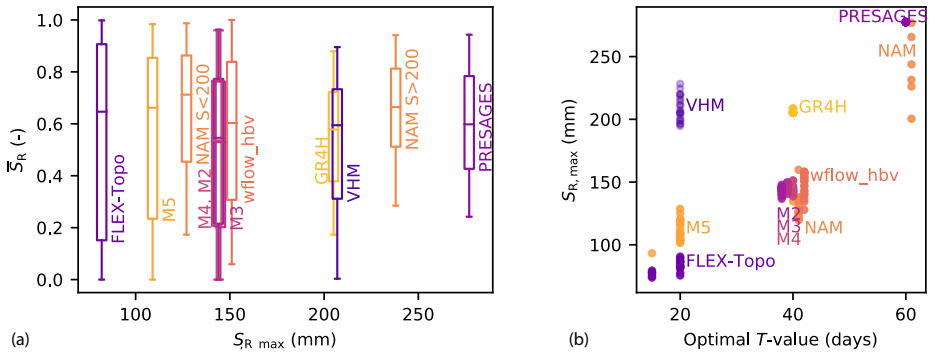

Figure 5.6: **(a)** Range of relative root-zone soil moisture $\overline{S}_R$ in the Ourthe upstream of Tabreux for the period 2015–2017 as a function of the median root-zone storage capacity ($S_{R,max}$) across parameter sets. The feasible parameters for NAM are split in two groups due to the large variability of $S_{R,max}$ (subsets with $S_{R,max}$ of ~130 mm and ~240 mm). **(b)** Root-zone storage capacity $S_{R,max}$ as a function of the optimal T-value for each model realization. Optimal T-values are derived at the highest Spearman rank correlation between Soil Water Index and modeled root-zone soil moisture.

three GRACE processing centers (Fig. 5.7b). Median precipitation is also 37 % higher in the Semois upstream of Membre-Pont than in the Ourthe upstream of Tabreux during winter months (Oct-Mar), but relatively similar during summer months (Apr-Sep). This difference in precipitation potentially leads to a wider range of modeled anomalies in the Semois upstream of Membre-Pont than in the Ourthe upstream of Tabreux for all models, as shown in Fig. 5.7c. This implies that all models reproduce the spatial variability between both catchments observed by GRACE. As the models were calibrated for the Ourthe at Tabreux and parameter sets were transferred to the Semois upstream of Membre-Pont, the forcing data is the main difference to explain the modeled spatial variability.

The models are also able to represent the observed temporal dynamics of total storage anomalies, as suggested by Spearman rank correlation coefficients ranging between 0.62 and 0.80 for the Ourthe upstream of Tabreux (Fig. 5.7d). There is, however, no relation between the Spearman rank correlations of the anomalies and the total modeled storage range (difference between maximum and minimum values), as shown in Fig. 5.7d. PRESAGES, WALRUS, VHM and dS2 have the largest ranges of total modeled storage, varying between 260 and 280 mm and are also characterized by a relatively large root-zone storage capacities (PRESAGES and VHM) or no separate root zone (WALRUS and dS2), while the total storage range of all other models is between 200 and 220 mm. The similarity in total storage range between most models is likely related to the identical forcing data and the similarity in the long-term partitioning of precipitation into drainage and evaporation (Sect. 5.5.1). However, the absolute values of total storage during a specific event or the partitioning in internal storage components may vary between models (Sect. 5.5.3).

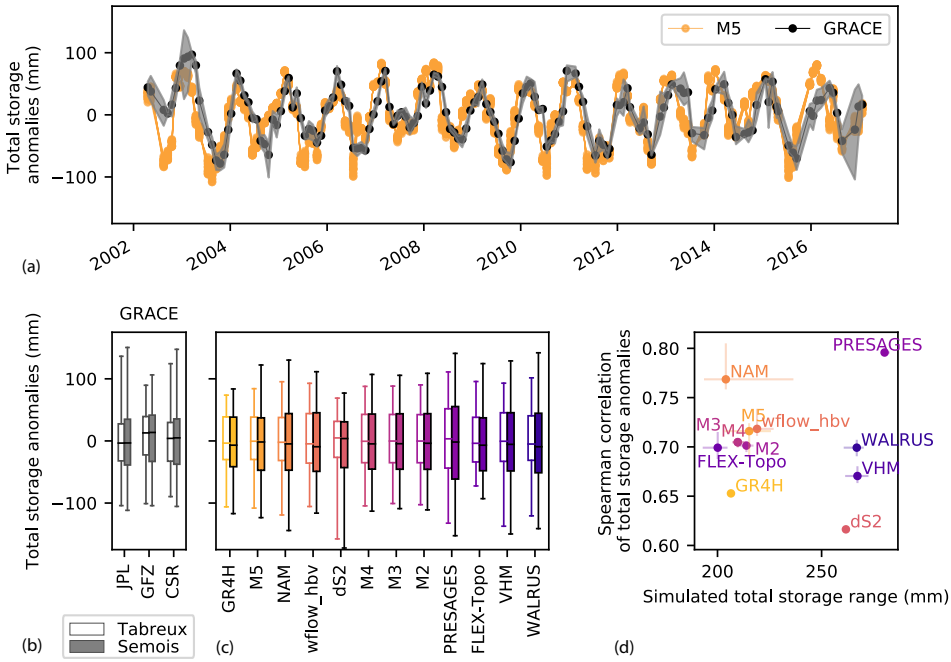

Figure 5.7: **(a)** Total storage anomalies modeled by M5 and compared to GRACE for the Ourthe upstream of Tabreux. The grey band shows the variability in total storage anomalies of the three processing centers. **(b)** Range of GRACE total storage anomalies for the three processing centers for the Semois upstream of Membre-Pont compared to the Ourthe upstream of Tabreux for the period 2001–2017. **(c)** Modeled total storage anomalies for both catchments. **(d)** Spearman rank correlations between GRACE and modeled total storage anomalies as a function of the range of modeled total storage for the Ourthe upstream of Tabreux.

5.5.3 INTERACTIONS BETWEEN STORAGE AND FLUXES DURING DRY PERIODS

As previously seen in Fig. 5.6a, the relative root-zone soil moisture content of the GR4H model is always above 0.2 for the three years for which SCATSAR-SWI1km data are available, as opposed to M5 which fully dries out for some time during the summers of 2015–2017. The Normalized Difference Vegetation Index of MODIS (NDVI, Didan, 2015a,b) also does not show a sharp decrease during these periods (Fig. 5.8a,b). Actual evaporation in M5 is strongly reduced during these dry soil moisture periods unlike GR4H, as shown in Fig. 5.8c,d. When zooming into the dry period around September 2016, Fig. 5.8e,f shows median relative root-zone soil moisture in GR4H of ~0.24 versus ~0.01 for M5, while SCATSAR-SWI1km has a higher median value of ~0.55 (for both optimal T-values of 20 and 40 days). The dryness of root-zone soil moisture in M5 leads to median daily evaporation of 0.8 mm d^{-1} against 1.3 mm d^{-1} for GR4H and prolonged periods of almost zero evaporation in M5 (e.g. 31/08–03/09, 09/09–15/09 and 22/09–30/09), while this neither occurs in GR4H nor in GLEAM actual evaporation, as shown in Fig. 5.8g,h. Despite the high streamflow performance of model M5 (Fig. 5.3, Table 5.2), it is unlikely that transpiration is reduced to almost zero for several days in a row each

summer in a catchment where approximately half of the area is covered by forests. This is also not supported by the remote-sensing data of soil moisture, NDVI and evaporation. High streamflow performances, therefore, do not warrant the plausibility of internal process representation. Despite the dried-out root-zone storage in M5, there is still water available in the slow storage to sustain a baseflow close to observed values, as shown in Fig. 5.8j,l. The streamflow responses of GR4H and M5 are both close to observations (Fig. 5.8i,j) in spite of differences in storage and evaporation, suggesting different internal process representations for a similar aggregated streamflow response during a low flow period.

5.5.4 PLAUSIBILITY OF PROCESS REPRESENTATIONS

The models are ranked and evaluated for a selection of criteria using observed streamflow, remote-sensing data and expert knowledge (Fig. 5.9). All models deviate less than ±6 % from observed median annual streamflow (Fig. 5.9a), which is less than the estimated uncertainty of 11 % (Sect. 5.3.3). In contrast, the modeled flashiness and baseflow indices of most models deviate more than the estimated uncertainty (Fig. 5.9b,c). FLEX-Topo is the only model with a clear overestimation of the flashiness index, which relates to the calibration aim of having a flashy model to reproduce small summer peaks (de Boer-Euser et al., 2017).

Modeled median annual total actual evaporation deviates by approximately -10 % from GLEAM estimates, except for the +22 % overestimation of the VHM model due to the issue of the unclosed water balance, as shown in Fig. 5.9d. These results are consistent with the evaluation study of GLEAM compared to other evaporation products (Miralles et al., 2016) which reports higher than average values for GLEAM in Europe (~+10 % at our latitude).

Four models explicitly account for interception with a separate module. Median annual interception rates deviate substantially from GLEAM estimates (-47 % to +80 %) as shown in Fig. 5.9e. There is a high uncertainty in the partitioning of evaporation into different components in evaporation products and GLEAM likely underestimates interception rates (Miralles et al., 2016; Zhong et al., 2020). Therefore, we consider a large uncertainty of +50 % to evaluate and rank the models. The GR4H interception is lower than GLEAM estimates. However, an interception storage was recently included in an hourly GR model (GR5H), to better represent the interception processes (Ficchì et al., 2019; Thirel et al., 2020).

All models substantially underestimate the number of days when snow is observed by MODIS at the catchment scale for all valid MODIS observations (cloud cover < 40 % and excluding summer months), as shown in Fig. 5.9f. Yet, we estimate a low uncertainty of less than 2 % around this number (Sect. 5.3.3). The NAM, M4 and M5 models are closest to MODIS estimates, but they are characterized by high false alarm rates (Fig. 5.5d), which implies a mismatch in the modeled and observed days with snow for valid MODIS observations. Based on expert knowledge (Royal Meteorological Institute Belgium, 2015) and the trade-off between miss rate and false discovery rate (Fig. 5.5d,e), we expect the annual number of days with snow storage to be between 28 and 62 days yr^{-1} as modeled

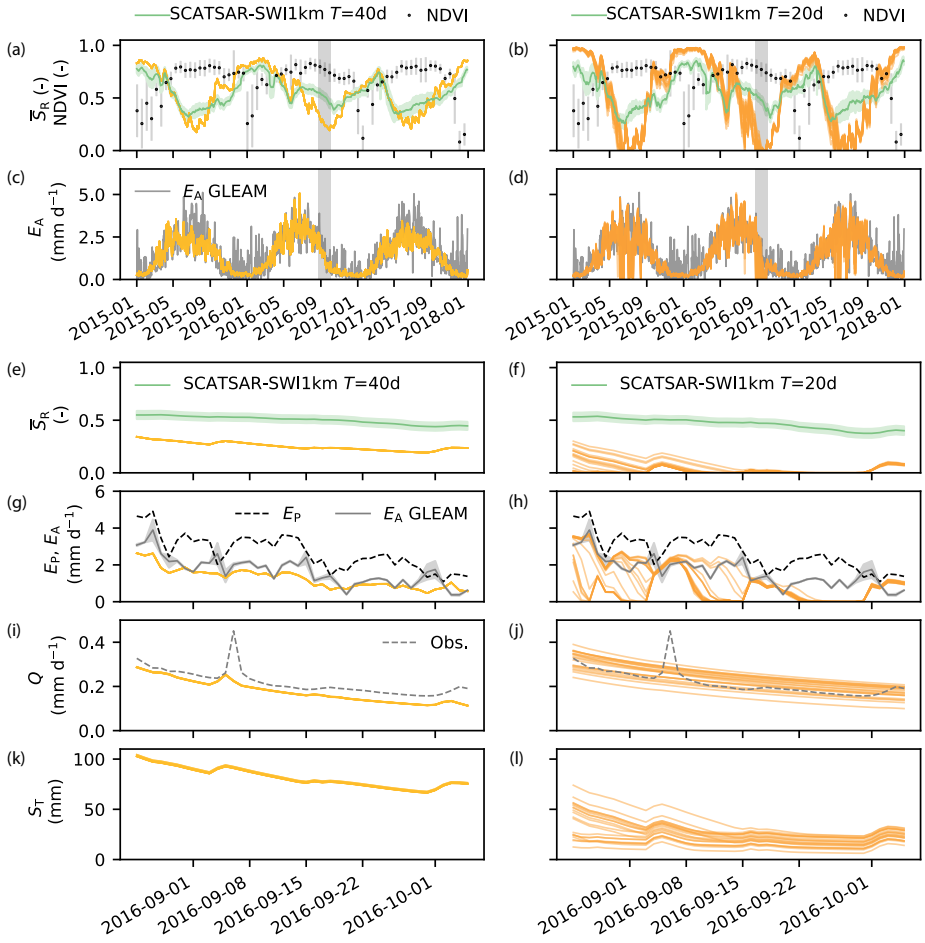

Figure 5.8: **(a,b)** Modeled relative root-zone soil moisture $\overline{S}_R$, SCATSAR-SWI1km Soil Water Index with optimal T-value and NDVI for the period 2015–2017 for GR4H (yellow) and M5 (orange) respectively. The error bars and bands show the standard deviation of the remote-sensing data within the catchment area. **(c,d)** Actual evaporation E_A by GR4H and M5 for the period 2015–2017, showing a large reduction of evaporation during summer for M5 unlike GR4H and GLEAM actual evaporation. **(e,f)** Zoomed-in modeled $\overline{S}_R$ and SCATSAR-SWI1km root-zone soil moisture for the grey shaded period of September 2016 in (a,b,c,d). **(g,h)** Potential, modeled and GLEAM actual evaporation, **(i,j)** Modeled and observed streamflow Q, **(k,l)** Total storage S_T for the September 2016 dry period. The narrow uncertainty band of the GR4H model is related to its converging parameter search method.

by wflow_hbv, WALRUS, FLEX-Topo and GR4H, whereas the ~90 days yr^{-1} of NAM, M4 and M5 seems too high.

The FLEX-Topo and M2 to M5 models are characterized by an empty root-zone storage for approximately 10 days yr^{-1} ($\overline{S}_R < 1\,\%$) as shown in Fig. 5.9g. These models have in common that evaporation from the root-zone occurs at potential rate and is not (or

hardly) reduced when soils are becoming dry until the point where the storage is empty. This is the case for models with very low or absence of the evaporation reduction parameter L_P. This behavior is not supported by the remote-sensing data of evaporation, soil moisture and NDVI (Sect. 5.5.3), nor by theory on root water uptake reduction under dry conditions (Feddes et al., 1978). The additional slow groundwater reservoir added in model M5 compared to M2-M4 leads to a smaller root-zone storage capacity as the available storage is partitioned into the root-zone storage and the additional groundwater store. The smaller root-zone storage capacity of model M5 exacerbates the number of annual days with empty storage. This highlights the complex interactions in internal dynamics even in parsimonious lumped models with similar mean annual streamflow performance.

Catchments with relatively large root-zone storage capacities underestimate GRACE estimates of very dry storage anomalies most (Figs. 5.6 and 5.9h). The uncertainty of GRACE is represented by the estimates of the three processing centers and the ~18 mm uncertainty estimate mentioned in Sect. 5.3.3. FLEX-Topo has a low root-zone storage capacity and is the only model which overestimates the very dry storage anomalies. Models with root-zone storage capacities of around 110 mm to 150 mm show the most consistent behavior with GRACE estimates of very dry storage anomalies.

5.6 DISCUSSION

5.6.1 IMPLICATIONS

While streamflow alone may be used to evaluate hydrological models, we subsequently use these models to understand internal states and fluxes in current and future conditions (Alcamo et al., 2003; Hagemann et al., 2013; Beck et al., 2017) or to make operational streamflow predictions (e.g. HBV and GR types of models are used by the Dutch and French forecasting services). Our findings show that similar streamflow responses obtained by models calibrated according to an identical protocol rely on different internal process representations. While not unexpected, it implies that we might get the right answers but for the wrong reasons (Kirchner, 2006), as these models cannot at the same time all be right and different from each other (Beven, 2006).

Almost all models show a similar long-term partitioning of precipitation into drainage and evaporation, as they are forced and constrained by the same data, also leading to relatively similar volumes of total storage. However, the partitioning of total storage in several internal storage components differs between models, resulting in distinct runoff responses as expressed by the baseflow and flashiness indices.

None of the models is systematically consistent with the information available from streamflow observations, remote-sensing data and expert knowledge. However, some processes either play a limited role on the overall water balance or can be compensated by other processes. Snow occurs every year but is not a major component of the streamflow regime (de Wit et al., 2007), interception evaporation can be compensated by root-zone evaporation, and very dry periods only occur for several weeks per year when streamflow is already very low. There is also a large uncertainty in each of the data

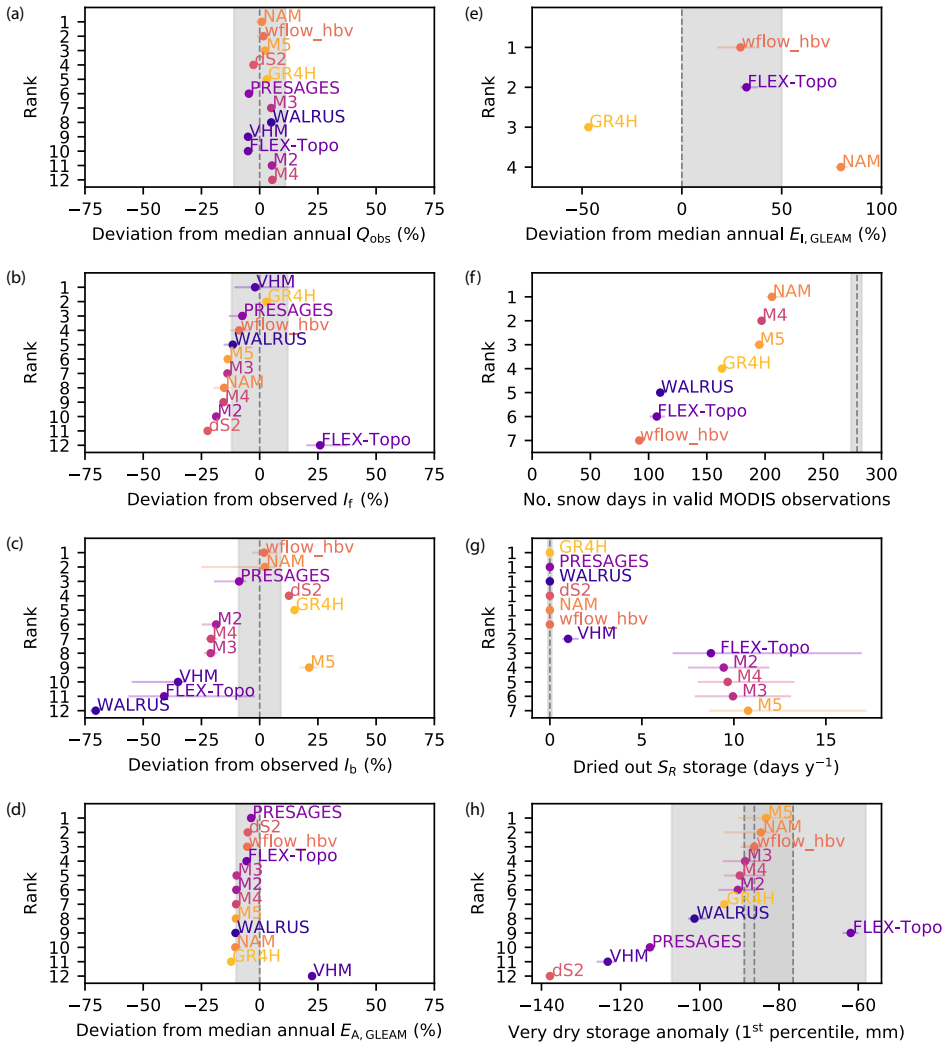

Figure 5.9: Ranking and evaluation of model behavior for a selection of criteria based on observed streamflow, remote-sensing data and expert knowledge. The grey shaded areas are soft indications of more plausible behavior based on uncertainty estimates and expert knowledge. Model ranks as a function of the: **(a)** deviation from observed median annual streamflow; **(b)** deviation from the flashiness index I_f; **(c)** deviation from the baseflow index I_b; **(d)** deviation from median annual GLEAM actual evaporation; **(e)** deviation from median annual GLEAM interception for models with a separate interception module; **(f)** number of days with snow cover for valid MODIS observations between 2001-2017, for models with a snow module; **(g)** annual number of days when the root-zone storage is dry (filled with less than 1 % of its capacity); **(h)** deviation from the 1st percentile of GRACE total storage anomalies for the three centers. The error bars show the 25-75th range across the ensemble of feasible parameter sets. Results are for the Ourthe catchment upstream of Tabreux, except for the snow analysis.

sources, which makes us reluctant to use them to determine hard thresholds to reject models. Instead, we ranked the models for a selection of "soft" criteria and found that NAM, wflow_hbv and PRESAGES are overall most consistent with the evaluation data, with median ranks of 2-3. While an overall ranking may be useful for practitioners, modelers benefit more from the specific ranking for each criterion to detect specific model deficiencies that could be improved in the model structure. An overall ranking is only a mere indication, which should be interpreted carefully due to uncertainty in the evaluation data and the applied calibration strategy.

The presence of interception or a slow storage (absent in M2-M4 but added in M5) affects the representation of other internal processes, including transpiration and/or root-zone soil moisture, implying that individual internal model components are altered by the presence/absence of other potentially compensating processes. Adding an additional internal model component changes the internal representation of water storage and fluxes through the system, which should be kept in mind if model parameters were to be fixed in alternative model structures. Furthermore, model improvements through additional process components and/or adapted parameterization should not only be evaluated in terms of the aggregated response, but also in the partitioning of fluxes and storages through the system (e.g. does the groundwater component improve the baseflow index at the expense of the availability of root-zone soil moisture during dry periods?). Models should be confronted with expert knowledge, e.g. on the occurrence of days with water stress or snow storage, to assess the plausibility of internal states and fluxes (Gharari et al., 2014; Hrachowitz et al., 2014; van Emmerik et al., 2015).

Applying these models to a future, more extreme climate in the same region might lead to contrasting insights regarding impacts of climate change, as also shown by studies of Hagemann et al. (2013), Melsen et al. (2018) and de Niel et al. (2019) in which model structures may lead to different signs of change of mean streamflow. Using one model or the other to assess the effect of rising temperatures on snow could lead to very different timescales of snow storage decline. Vegetation already experiences more intense water stress in some models compared to others and this would be exacerbated in more extreme drought scenarios (Melsen and Guse, 2019). More intense precipitation events could affect interception evaporation and therefore water availability in the root-zone differently from one model to another. Beyond model structure, the experience each modeler has with its model and associated calibration procedure to constrain model parameters may also impact the simulation results (Melsen et al., 2019).

Our findings should, therefore, encourage modelers to use multiple data sources for model calibration and evaluation, as already suggested by many other studies (Samaniego et al., 2010; Rakovec et al., 2016a; Koch et al., 2018; Stisen et al., 2018; Nijzink et al., 2018; Veldkamp et al., 2018; Dembélé et al., 2020). Remote-sensing estimates of soil moisture, evaporation and total storage anomalies are available at the global scale and in spite of potential biases with models, the temporal dynamics are useful to constrain our models (McCabe et al., 2017; Sheffield et al., 2018). Additionally, it seems essential to support decision-makers by studies relying on multi-model and multi-parameter systems, as also suggested by Haddeland et al. (2011) and Schewe et al. (2014), to reveal uncertainties inherent to the heterogeneous hydrological world (Beven, 2006; Savenije, 2010;

Samaniego et al., 2010; Hrachowitz and Clark, 2017).

This study is the result of a joint research effort of scientists and practitioners gathering each year in Liège at the International Meuse Symposium to exchange interdisciplinary and intersectoral knowledge related to the Meuse basin. Although coordination of large international teams may be challenging, international studies favor a close collaboration between scientists and practitioners that can learn from each other to accelerate modeling advances (Archfield et al., 2015). Another advantage of comparing modeling results of several research groups is to quickly detect small mistakes in the modeling process, including shifts in the time series or using forcing data of one catchment to model another catchment. While hydrograph characteristics were the main focus of the previous study (de Boer-Euser et al., 2017), we gain distinct insights on the plausibility of model behavior by evaluating additional facets of internal process representation using remote-sensing data.

5.6.2 LIMITATIONS AND KNOWLEDGE GAPS

Many aspects of the hydrological response remain unknown and can hardly be evaluated against observations. While in-situ observations of snow, evaporation or soil moisture are rarely available at sufficient spatio-temporal scale, remote-sensing estimates have the advantage of high spatial resolution and worldwide spatial coverage, though they often rely on models themselves and are affected by high and often unknown uncertainty. Comparing models with these independent observations is valuable to evaluate their consistency and detect outliers. However, these observations cannot be considered as representative of the truth as they rely on many assumptions themselves, hindering "real" hypotheses testing. The ratio of actual over potential evaporation as a result of water stress at the catchment scale, therefore, remains highly uncertain (Coenders-Gerrits et al., 2014; Mianabadi et al., 2019). While areal fractions of snow cover can be estimated by MODIS, the presence of clouds limits the usability of the data and knowledge of catchment-scale snow water equivalent is lacking. If remote-sensing estimates of near-surface relative soil moisture are available, root-zone water content remains uncertain and while GRACE provides estimates of total storage anomalies, we lack knowledge on absolute total water storage. The spatial variability and the temporal dynamics of these remote-sensing products provide useful, additional, independent information to understand the hydrological puzzle, but certainly not all the answers to evaluate the states typically included in process-based models. Measurements are, therefore, of crucial importance to increase our understanding of hydrological processes at the catchment scale, which in turn will improve the quality of remote-sensing products and model development (Vidon, 2015; Burt, T. P., McDonnell, 2015; van Emmerik et al., 2018).

The evaluation of model behavior is conditional on the calibration procedure, which was freely chosen by the individual contributing institutes. The use of different or more calibration objectives and in-depth uncertainty estimation (Beven and Binley, 1992) may have resulted in different conclusions in terms of the plausibility of the behavior of each model.

We performed a thorough analysis of twelve models, five variables and three catch-

ments. We deliberately chose to limit the number of study catchments to balance depth with breadth, allowing us to dive into process-relevant insights.

5.7 CONCLUSIONS

Similar streamflow performance of process-based models, calibrated following an identical protocol, relies on different internal process representations. Most models are relatively similar in terms of the long-term partitioning of precipitation into drainage and evaporation. However, the partitioning between transpiration and interception, snow processes and the representation of root-zone soil moisture varies significantly between models, suggesting variability of water storage and release through the catchment. The comparison of modeled states and fluxes with remote-sensing estimates of evaporation, root-zone soil moisture and vegetation indices suggests that models with relatively small root-zone storage capacities and without reduction in root water uptake during dry conditions lead to unrealistic drying-out of the root-zone storage and significant reduction of evaporative fluxes each summer. Expert knowledge in combination with remote-sensing data further allows us to "softly" evaluate the plausibility of model behavior by ranking them for a set of criteria. Even if none of the models is systematically consistent with the available data, we did not formally reject specific models due to the uncertainty in the evaluation data and their changing relevance for the studied catchments. The dissimilarity in internal process representations between models implies that they are not necessarily providing the right answers for the right reasons, as they cannot simultaneously be close to reality and different from each other. While the consequences for streamflow may be limited for the historical data, the differences may exacerbate for more extreme conditions or climate change scenarios. Considering the uncertainty of process representation behind the scenes of streamflow performance and our lack of knowledge and observations on these internal processes, we invite modelers to evaluate their models using multiple variables, we encourage more experimental research, and highlight the value of multi-model multi-parameter studies to support decision making.

Table 5.1: Description of symbols used for fluxes, storages and parameters in Tables 5.2 and 5.3

Symbol	unit	Description
Fluxes		
E_P	mm h^{-1}	Potential evaporation
E_I	mm h^{-1}	Interception evaporation
E_R	mm h^{-1}	Transpiration and soil evaporation
E_W	mm h^{-1}	Sublimation
E_A	mm h^{-1}	Total actual evaporation (sum of soil evaporation, transpiration, (separate) interception and, if applicable, sublimation)
P	mm h^{-1}	Precipitation
P_R	mm h^{-1}	Precipitation entering the root-zone storage (after snow and/or interception if present or fraction/total precipitation)
Q	mm h^{-1}	Streamflow
Q_R	mm h^{-1}	Flux from root-zone to fast and/or slow runoff storage
Q_P	mm h^{-1}	Percolation flux from root-zone storage to slow runoff storage
Q_C	mm h^{-1}	Capillary flux from slow runoff storage to root-zone storage
Q_G	mm h^{-1}	Seepage (up/down) / extraction
Storages		
S_T	mm	Total storage
S_W	mm	Snow storage
S_I	mm	Interception storage
S_R	mm	Root-zone storage
$\overline{S_R}$	-	Relative root-zone storage ($S_R/S_{R,max}$)
S_D	mm	Storage deficit
S_{VQ}	mm	Very quick runoff storage
S_F	mm	Fast runoff storage
S_S	mm	Slow runoff storage
S_{SW}	mm	Surface water storage
Parameters		
C_E	-	Correction factor for E_P
I_{max}	mm	Maximum interception capacity
$S_{R,max}$	mm	Maximum root-zone storage capacity
S_{thresh}	mm	Threshold of root-zone storage above which $E_R = E_P$
L_P	-	Threshold of relative root-zone storage above which $E_R = E_P$
C_{cst}	-	Constant water stress coefficient to estimate E_R
a, b, S_0	-	Parameters describing the shape of the streamflow sensitivity
a_S	-	Fraction of land surface covered by surface water
a_G	-	Fraction of land surface not covered by surface water

Table 5.2: Number of calibrated model parameters, spatial distribution, and model performance calculated for the period 2008–2017 with the Euclidean distance where a value of 0 would indicate a perfect model. Main characteristics describing snow storage, root-zone storage and total storage per model. Notations are defined in Table 5.1.

	GR4H	*M5*	*NAM*	*wflow_hbv*	*dS2*	*M4*	*M3*	*M2*	*PRESAGES*	*FLEX-Topo*	*VHM*	*WALRUS*
Number of calibrated parameters	4	9	12	9	4	7	6	5	6	20	12	3
Lumped (L) / Semi-distributed (S) / Distributed (D)	L	L	L	D	L	L	L	L	L	S	L	L
Euclidean distance $\sqrt{(1-E_{\mathrm{NS,Q}})^2+(1-E_{\mathrm{NS,logQ}})^2}$	0.17	0.18	0.18	0.20	0.21	0.23	0.23	0.24	0.24	0.26	0.26	0.34
Snow storage S_{W} (compared to MODIS snow cover)												
Snow module	✓	✓	✓	✓	-	✓	-	-	-	✓	-	✓
Degree-hour method	✓	✓	✓	✓	-	✓	-	-	-	✓	-	✓
Elevation zones	✓	-	-	✓	-	-	-	-	-	✓	-	-
Temperature interval for rainfall and snow	✓	-	-	✓	-	-	-	-	-	✓	-	✓
Melt factor constant in time	-	✓	-	✓	-	✓	-	-	-	✓	-	✓
Melt factor ~ snow storage	✓	-	✓	-	-	-	-	-	-	-	-	-
Refreezing of liquid water	-	-	✓	✓	-	-	-	-	-	-	-	-
Sublimation	-	-	-	-	-	-	-	-	-	✓	-	-
Calibration snow parameters	-	✓	✓	-	-	✓	-	-	-	✓	-	-
Root-zone storage S_{R} (compared to SCATSAR-SWI1km Soil Water Index)												
Separate root-zone module with capacity $S_{\mathrm{R,max}}$	✓	✓	✓	✓	-	✓	✓	✓	✓	✓	✓	-
$\frac{\mathrm{d}S_{\mathrm{R}}}{\mathrm{d}t} = P_{\mathrm{R}} - E_{\mathrm{R}}$	-	-	-	-	-	-	-	-	-	-	✓	-
$\frac{\mathrm{d}S_{\mathrm{R}}}{\mathrm{d}t} = P_{\mathrm{R}} - E_{\mathrm{R}} + Q_{\mathrm{C}}$	-	-	✓	-	-	-	-	-	-	-	-	-
$\frac{\mathrm{d}S_{\mathrm{R}}}{\mathrm{d}t} = P_{\mathrm{R}} - E_{\mathrm{R}} - Q_{\mathrm{R}}$	✓	✓	-	✓	-	✓	✓	✓	✓	-	-	-
$\frac{\mathrm{d}S_{\mathrm{R}}}{\mathrm{d}t} = P_{\mathrm{R}} - E_{\mathrm{R}} - Q_{\mathrm{R}} - Q_{\mathrm{P}} + Q_{\mathrm{C}}$	-	-	-	-	-	-	-	-	-	✓	-	-
Total storage S_{T} (anomalies are compared to GRACE total storage anomalies)												
$S_{\mathrm{T}} = -S_{\mathrm{D}} \cdot a_{\mathrm{G}} + S_{\mathrm{F}} \cdot a_{\mathrm{G}} + S_{\mathrm{SW}} \cdot a_{\mathrm{S}}$	-	-	-	-	-	-	-	-	-	-	-	✓
$S_{\mathrm{T}}(Q) = \frac{1}{a}\frac{1}{1-b} \cdot Q^{1-b} + S_0$	-	-	-	-	✓	-	-	-	-	-	-	-
$S_{\mathrm{T}} = S_{\mathrm{R}} + S_{\mathrm{F}}$	-	-	-	-	-	-	✓	✓	-	-	-	-
$S_{\mathrm{T}} = S_{\mathrm{W}} + S_{\mathrm{R}} + S_{\mathrm{F}}$	-	-	-	-	-	✓	-	-	-	-	-	-
$S_{\mathrm{T}} = S_{\mathrm{W}} + S_{\mathrm{R}} + S_{\mathrm{F}} + S_{\mathrm{S}}$	-	✓	-	-	-	-	-	-	-	-	-	-
$S_{\mathrm{T}} = S_{\mathrm{R}} + S_{\mathrm{VQ}} + S_{\mathrm{F}} + S_{\mathrm{S}}$	-	-	-	-	-	-	-	-	✓	-	✓	-
$S_{\mathrm{T}} = S_{\mathrm{W}} + S_{\mathrm{R}} + S_{\mathrm{VQ}} + S_{\mathrm{F}} + S_{\mathrm{S}}$	✓	-	-	-	-	-	-	-	-	-	-	-
$S_{\mathrm{T}} = S_{\mathrm{I}} + S_{\mathrm{W}} + S_{\mathrm{R}} + S_{\mathrm{F}} + S_{\mathrm{S}}$	-	-	-	✓	-	-	-	-	-	-	-	-
$S_{\mathrm{T}} = S_{\mathrm{I}} + S_{\mathrm{W}} + S_{\mathrm{R}} + S_{\mathrm{VQ}} + S_{\mathrm{F}} + S_{\mathrm{S}}$	-	-	✓	-	-	-	-	-	-	✓	-	-

Table 5.3: Main characteristics describing evaporation processes per model (with ✓[1] indicates $L_P = 1$ and ✓[2] indicates $E_I = 0$). Notations are defined in Table 5.1.

	GR4H	*M5*	*NAM*	*wflow_hbv*	*dS2*	*M4*	*M3*	*M2*	*PRESAGES*	*FLEX-Topo*	*VHM*	*WALRUS*
Correction factor for potential evaporation	-	✓	-	-	-	✓	✓	✓	-	-	-	-
Interception evaporation E_I	✓	-	✓	✓	-	-	-	-	-	✓	-	-
Maximum interception storage I_{max}	-	-	✓	✓	-	-	-	-	-	✓	-	-
$I_{max} \sim 1.1 - 3.4$ mm	-	-	-	-	-	-	-	-	-	✓	-	-
$I_{max} \sim 1.4 - 2.9$ mm	-	-	-	✓	-	-	-	-	-	-	-	-
$I_{max} \sim 5.3 - 6.9$ mm	-	-	✓	-	-	-	-	-	-	-	-	-
$E_I = \begin{cases} E_P, & \text{if } S_I > 0. \\ 0, & \text{otherwise.} \end{cases}$	-	-	✓	✓	-	-	-	-	-	✓	-	-
$E_I = \begin{cases} E_P, & \text{if } P > E_P. \\ P, & \text{otherwise.} \end{cases}$	✓	-	-	-	-	-	-	-	-	-	-	-
Transpiration and soil evaporation E_R	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
$E_R = E_P \cdot C_{cst}$	-	-	-	-	✓	-	-	-	-	-	-	-
$E_R = E_P \cdot \frac{\overline{S}_R \cdot (2 - \overline{S}_R)}{1 + E_P / S_{R,max} \cdot (2 - \overline{S}_R)}$	✓	-	-	-	-	-	-	-	✓	-	-	-
$E_R = E_P \cdot C_E \cdot \frac{\overline{S}_R \cdot (1 + m_1)}{\overline{S}_R + m_1}$, with $m_1 = 10^{-2}$	-	✓	-	-	-	✓	✓	✓	-	-	-	-
$E_R = \begin{cases} (E_P - E_I) \cdot \frac{\overline{S}_R}{L_P}, & \text{if } \overline{S}_R < L_P. \\ E_P - E_I, & \text{otherwise.} \end{cases}$	-	-	✓[1]	✓	-	-	-	-	-	✓	✓[2]	-
$E_R = E_P \cdot f(S_d)$	-	-	-	-	-	-	-	-	-	-	-	✓
Total actual evaporation E_A	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
$E_A = E_R$	-	✓	-	-	✓	✓	✓	✓	✓	-	✓	✓
$E_A = E_R + E_I$	✓	-	✓	✓	-	-	-	-	-	-	-	-
$E_A = E_R + E_I + E_W$	-	-	-	-	-	-	-	-	-	✓	-	-

6

ADAPTING ECOSYSTEMS: TESTING THE IMPACT OF TIME-VARIANT HYDROLOGICAL MODELS IN RESPONSE TO CLIMATE CHANGE

Hij zegt dat de toekomst belangrijker is omdat die nog moet komen, zodat je daar nog iets aan kunt doen.

Ilja Leonard Pfeijffer - *Grand Hotel Europa* (2018)

In the previous chapters, in-situ and remote-sensing data are used to increase our understanding of hydrological behavior for historical conditions. If predicting present-day hydrological functioning is already challenging, it is even more so under changing environmental conditions. Ecosystems are likely to adapt in response to climate change and other species might become dominant, both under natural and anthropogenic influence. In this chapter, we introduce non-stationarity in the representation of hydrological systems for a long-term projection of global warming. Using the projected changes in magnitude and seasonality of hydro-climatic variables, we test the sensitivity of the hydrological response to adapting the root-zone storage capacity parameter in a process-based hydrological model and changing dominant land-use characteristics by trading space-for-time.

Parts of this chapter are in review in Hydrology and Earth System Sciences Discussions (Bouaziz et al., 2021)

SUMMARY

To predict future hydrological behavior in a changing world, often use is made of models calibrated on the past, disregarding that hydrological systems, hence model parameters, will change as well. Yet, ecosystems likely adjust their root-zone storage capacity, which is the key parameter of any hydrological system, in response to climate change. In addition, other species might become dominant, both under natural and anthropogenic influence. In this study, we propose a top-down approach, which directly uses projected climate data to estimate how vegetation adapts its root-zone storage capacity at the catchment scale in response to changes in magnitude and seasonality of hydro-climatic variables. Additionally, the Budyko characteristics of different dominant ecosystems in sub-catchments are used to simulate the hydrological behavior of potential future land-use change, in a space-for-time exchange. We hypothesize that changes in the predicted hydrological response as a result of 2K global warming are more pronounced when explicitly considering changes in the sub-surface system properties induced by vegetation adaptation to changing environmental conditions. We test our hypothesis in the Meuse basin in four scenarios designed to predict the hydrological response to 2K global warming in comparison to current-day conditions using a process-based hydrological model with (a) a stationary system, i.e. no changes in the root-zone storage capacity of vegetation and historical land use, (b) an adapted root-zone storage capacity in response to a changing climate but with historical land use, and (c,d) an adapted root-zone storage capacity considering two hypothetical changes in land use from coniferous plantations/agriculture towards broadleaved forest and vice-versa. We found that the larger root-zone storage capacities (+34 %) in response to a more pronounced seasonality with drier summers under 2K global warming strongly alter seasonal patterns of the hydrological response, with an overall increase in mean annual evaporation (+4 %), a decrease in recharge (-6 %) and a decrease in streamflow (-7 %), compared to predictions with a stationary system. By integrating a time-dynamic representation of changing vegetation properties in hydrological models, we move towards more reliable hydrological predictions under change.

6.1 INTRODUCTION

Hydrological models are required to provide robust short-term hydrological forecasts and long-term predictions of the impact of natural and human-induced change on the hydrological response. Common practice is to predict the future using a hydrological model calibrated to the past (Vaze et al., 2010; Blöschl and Montanari, 2010; Peel and Blöschl, 2011; Coron, 2013; Seibert and van Meerveld, 2016). For the near future, it seems acceptable to assume no fundamental change in the hydrological system, although we know that ecosystems, the manager of the hydrological system, have the capacity to adapt to climatic change (Savenije and Hrachowitz, 2017). For longer term predictions, it is therefore not correct to assume an unchanged system within a changing world. This raises the question on the robustness of hydrological predictions, especially in the context of climate change (Coron et al., 2012).

For example, Merz et al. (2011) clearly shows the non-stationarity of hydrological model parameters when calibrating 273 Austrian catchments in subsequent 5-years periods between 1976 and 2006. Being the core parameter of any hydrological system, Merz et al. (2011) report almost a doubling of the root-zone storage capacity and this gradual increase is assumed to be related to changing climatic conditions, such as increased evaporation and drier conditions in the more recent years. The temporal variability of model parameters could also be attributed to uncertainties in input and model structure or inadequate calibration strategies. However, the observed trends in model parameters are also likely to reflect transient catchment conditions over the historical period.

Under continued global warming, precipitation and temperature extremes are expected to further increase and the hydrological cycle is likely to further accelerate (Allen et al., 2010; Kovats et al., 2014). In addition, natural land cover change and anthropogenic activities of land-cover change and land-use management can substantially alter a catchment's water balance (Brown et al., 2005; Wagener, 2007; Fenicia et al., 2009; Jaramillo and Destouni, 2014; Nijzink et al., 2016b; Hrachowitz et al., 2020; Levia et al., 2020). Considering the unprecedented speed of change, Milly et al. (2008) declared that stationarity is dead and no longer should serve as a default assumption in water management. He advocates the development of methods that quantify the non-stationarity of relevant hydrological variables.

However, understanding and representing non-stationarity is challenging due to the complex interactions and associated feedback between climate, vegetation, soils, ecosystems and humans (Seibert and van Meerveld, 2016). The main methods to understand how changes in hydrological functioning relate to changes in catchment characteristics rely on paired watershed studies and hydrological modeling (Andréassian et al., 2003). In many modeling studies, a selection of one or more parameters are changed using values from literature in combination with adapted land-cover maps to (partly) reflect the characteristics of the altered system (Mao and Cherkauer, 2009; Buytaert and Beven, 2009; Pomeroy et al., 2012; Gao et al., 2015). Alternatively, Duethmann et al. (2020) uses satellite observations of vegetation indices to improve the representation of the surface resistance dynamics to calculate reference evaporation used in conceptual hydrological models over a historical record. A similar approach is applied by Fenicia et al. (2009) to

account for changes in evaporation as a result of land-use management changes in the Meuse basin.

While these approaches are valuable to test the sensitivity of change on the hydrological response (Seibert and van Meerveld, 2016), they require an understanding of how catchment characteristics relate to model parameters. Yet, there is considerable uncertainty in a priori parameter estimation and the use of regionalization approaches (Wagener, 2007). Besides, the required data (e.g. future land-use maps or vegetation indices) may not be available in the context of climate change impact assessment (Duethmann et al., 2020). Instead, a way forward may be to develop robust top-down modeling approaches based on optimality principles by considering the co-evolution of soils, vegetation and climate in a holistic way (Blöschl and Montanari, 2010).

As complex and heterogeneous as landscapes may be across a diversity of climates, the long-term hydrological partitioning of a catchment is governed by a surprisingly simple and predictable relation, which relies on the available water and energy for evaporation (Turc, 1954; Mezentsev, 1955; Budyko, 1961; Fu, 1981; Zhang et al., 2004). The Budyko hypothesis, as often referred to, describes that mean annual evaporation over precipitation (E_A/P) is mainly controlled by the aridity index, defined as the ratio of mean annual potential evaporation over precipitation (E_P/P). However, Troch et al. (2013) found catchments to deviate from the Budyko hypothesis when exchanging climates across different catchments in a modeling experiment. Their results suggest that long-term hydrological partitioning results from the co-evolution of catchment properties and climate characteristics, including not only the aridity index but also climate seasonality, topography, vegetation and soils.

The combination of these other factors influencing the water balance partitioning besides the aridity index are explicitly considered in the ω parameter of the parametric description of the Budyko hypothesis (Fu, 1981; Zhang et al., 2004). Deviations from the Budyko curve suggest that different vegetation develops in different climates, along a different ω curve. If climate changes, catchments are likely not only to shift horizontally in the Budyko space as a result of a changing aridity index, but also vertically as a result of a changing vegetation cover (Jaramillo and Destouni, 2014). Vertical shifts within the Budyko space can also be related to vegetation-CO_2 interactions, e.g. CO_2 fertilization and improved water-use efficiency as a result of increasing CO_2 levels (Keenan et al., 2013; van der Velde et al., 2014; van Der Sleen et al., 2015; Ukkola et al., 2016; Jaramillo et al., 2018).

The interdependence of climate seasonality, aridity index and vegetation to match the expectation from the Budyko curve was also demonstrated by Gentine et al. (2012); Donohue et al. (2012). Vegetation tends to efficiently adapt its root-zone storage capacity to satisfy canopy water demand. This implies that vegetation creates a larger buffer to survive dry spells when seasonal water supply and demand are out of phase, than in a climate where demand and supply are in phase (Milly, 1994; Schymanski et al., 2008; Gerrits et al., 2009; Gentine et al., 2012; Gao et al., 2014). The root-zone storage capacity is, therefore, the key element regulating the partitioning of water fluxes in many terrestrial hydrological systems. In addition, not only natural changes to the environment, but

also human interference with vegetation affect transpiration water demand and hence the root-zone storage capacity (Nijzink et al., 2016b; Hrachowitz et al., 2020).

Detailed observations of rooting-systems are very scarce in time and space and difficult to integrate to the catchment scale due to heterogeneity of landscapes (de Boer-Euser et al., 2016; Hrachowitz et al., 2020). Instead, the catchment-scale root-zone storage capacity is often estimated through calibration of a hydrological model. Other methods rely on optimality principles that maximize net primary production or carbon gain (Kleidon, 2004; Guswa, 2008; Speich et al., 2018). Alternatively, there is increasing evidence that the catchment-scale root-zone storage capacity can be robustly and directly estimated from annual water deficits using water balance data (Gao et al., 2014; de Boer-Euser et al., 2016; Wang-Erlandsson et al., 2016; Nijzink et al., 2016b; Bouaziz et al., 2020; Hrachowitz et al., 2020). However, it remains unclear how vegetation may adapt its root-zone storage capacity to climate change and how these changes affect future hydrological behavior.

The objective of this study in the Meuse basin (Western Europe) is to quantify the sensitivity of the hydrological response to potential changes in the root-zone storage capacity of vegetation in combination with land-use changes as a result of 2K global warming. Using the Budyko framework, we first estimate changes in the long-term hydrological partitioning. To evaluate the effect of land-use change under future conditions, we exchange space-for-time by connecting the spatially variable ω parameter of the Budyko curve to different land uses. We then use water balance data to estimate how the root-zone storage capacity may adapt to increasing seasonal water deficits under climate change.

We hypothesize that changes in the predicted hydrological response as a result of 2K global warming in comparison to current-day conditions are more pronounced when explicitly considering an adapted root-zone storage capacity to reflect changes in the magnitude and seasonality of hydro-climatic variables as well as potential land-use changes. We test our hypothesis using a process-based hydrological model and compare the difference in hydrological response when assuming (a) a stationary system without changes in the root-zone storage capacity and historical land use, with three non-stationary systems involving (b) an adapted root-zone storage capacity in response to climate change but no changes in land use, and (c,d) an adapted root-zone storage capacity and two hypothetical land-use change scenarios.

6.2 STUDY AREA

The study is performed in the Meuse basin upstream of Borgharen (Fig. 6.1). The forests in the French part of the basin are mainly characterized as "old growth", here defined as forested area which has been continuously wooded since at least the middle of the 19th century (Cateau et al., 2015). These broadleaved forests consist primarily of European Oak, Sessile Oak and Beech (Institut National de l'Information Géographique et Forestière, 2019). In contrast, only 44 % of the 18th century Walloon forests of Belgium has remained from the original broadleaved forest, the rest being cleared for agriculture

on high fertility soils in the North West (30 %) or converted to coniferous plantations (Scots pine, Norway spruce and Douglas-fir) on the poor soils of the Ardennes (26 %, Kervyn et al., 2018). The status of "old growth" forest does not exclude human disturbances, but assumes a relatively limited impact. Soils are less disturbed and their structure and biochemical composition have been preserved for several centuries. This favors a high degree of biodiversity, which is a key element for the resilience of forest ecosystems to perturbations. In contrast, recent short-rotation plantations lack many of these characteristics. Particularly thick canopy plantations, such as the spruce and Douglas-fir, significantly alter the typical biodiversity of forests. Additionally, relatively higher evaporation water use is expected in these recent, short-rotation exotic plantations in comparison to older, more natural forests (Fenicia et al., 2009).

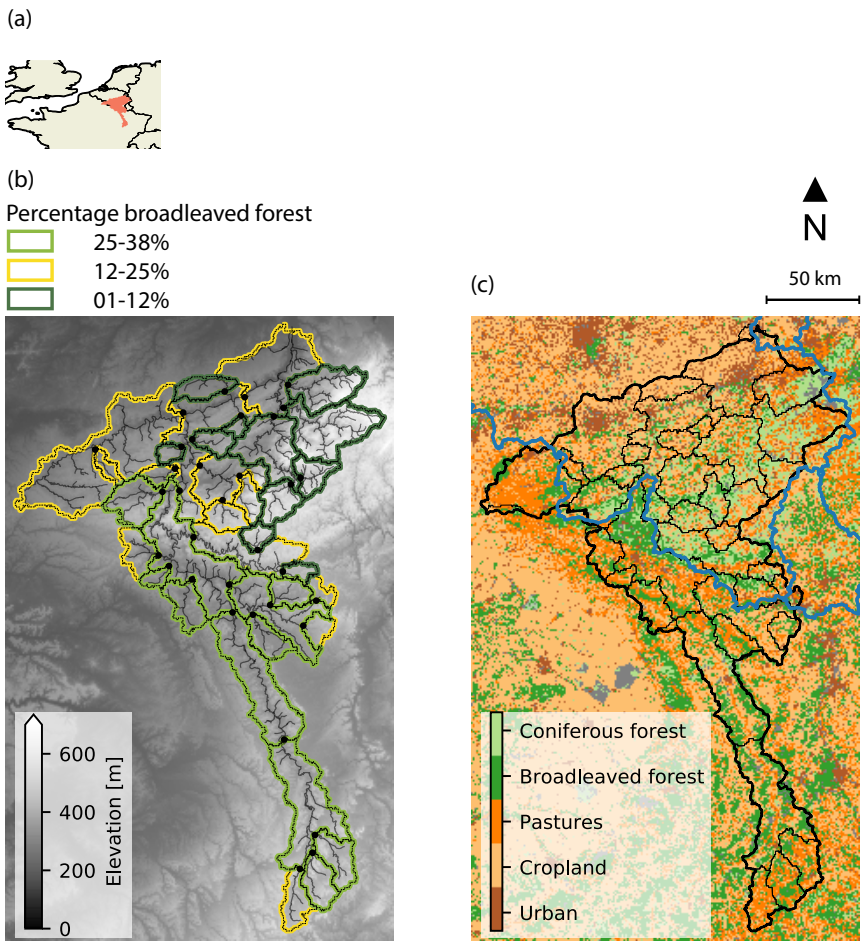

Figure 6.1: **(a)** Location of the Meuse basin in northwestern Europe. **(b)** Elevation in the basin and categorization of catchments according to their areal percentage of broadleaved forest. **(c)** Main land-use types according to CORINE Land Cover (European Environment Agency, 2018).

6.3 DATA

6.3.1 OBSERVED HISTORICAL E-OBS CLIMATE DATA

The E-OBS dataset (v20.0e) includes daily precipitation, temperature and radiation fields for the period 1980-2018 at a 25 km^2 resolution (Cornes et al., 2018). The data are based on station data collated by the European Climate Assessment Dataset (ECA&D) initiative. Temperature is downscaled using the digital elevation model and a fixed lapse rate of 0.0065°C m^{-1}. Potential evaporation is estimated using the Makkink formula (Hooghart and Lablans, 1988). There is a relatively large underestimation of precipitation (> 20 %) in the E-OBS dataset in the center of the basin when compared to an operational dataset, which is based on local precipitation data provided by the Service Public de Wallonie for the period 2005-2017 (Bouaziz et al., 2020). A monthly bias-correction factor is applied to improve the consistency between both datasets.

6.3.2 SIMULATED HISTORICAL AND 2K CLIMATE DATA

To study the impact of 2K global warming on the hydrological response of the Meuse basin, we use climate simulations of the historical period 1979-2018 and a 2K global warming simulation, provided by the Royal Netherlands Meteorological Institute (KNMI). The simulations are generated with the regional climate model KNMI-RACMO2 (van Meijgaard et al., 2008) at 12 km x 12 km resolution. RACMO2 uses the land surface scheme HTESSEL (Balsamo et al., 2009), which employs four soil layers with a total depth of 2.9 m. Each land-grid cell includes separate tiles for high and low vegetation (16 vegetation types), bare soil, snow and intercepted water, for which the energy and water balances are solved individually.

The historical simulation uses ERA5 reanalysis data (Hersbach et al., 2020) as initial- and lateral boundary conditions. The 2K simulation is a so-called pseudo-global warming (PGW) simulation (e.g. Schär et al., 1996; Attema et al., 2014; Prein et al., 2017; Brogli et al., 2019), which is an alternative method to generate high-resolution climate change information. Instead of downscaling global climate model (GCM) projections, the historical period is re-simulated, but set against a warmer climate background by adding perturbations to the ERA5 initial- and boundary conditions. The perturbations represent the change in the mean climate state in a globally 2K warmer world, derived from a large initial condition GCM ensemble (Aalbers et al., 2018). The method minimizes biases in the mean climate state of the historical simulation, guaranties a realistic atmospheric circulation under both historical and 'future' conditions and increases the signal-to-noise ratio of the climate change response. A full description of the dataset is provided in Aalbers et al. (2021).

6.3.3 STREAMFLOW

Streamflow data is available for 35 catchments nested within the Meuse basin upstream of Borgharen for the period 2005-2017 (Fig. 6.1, Service Public de Wallonie, 2018; Banque Hydro, 2018). The streamflow at Borgharen is a constructed time series which sums the

observed streamflow of the Meuse at St Pieter and of the Albert Canal at Kanne to represent the total flow from the tributaries before part of it is extracted in the Albert Canal (de Wit et al., 2007).

6.4 METHODS

To quantify the importance of reflecting ecosystem adaptation in hydrological models in response to climate change, the following stepwise approach is designed: (1) estimate the long-term runoff coefficient in a 2K warmer world from movements in the Budyko space as a result of a shift in aridity index and a potential shift in dominant land-use from broadleaved forests to coniferous plantation and agriculture and vice-versa by trading space-for-time in the Meuse basin; (2) estimate how the root-zone storage capacity adapts in response to a more pronounced seasonality with drier summers and changing dominant land use using the observed historical and estimated long-term runoff coefficient in a 2K warmer world with potential changes in land use; (3) calibrate a hydrological model with observed historical E-OBS climate data to represent current-day hydrological conditions; (4) test if the historical climate data simulated by the regional climate model leads to a plausible representation of current-day hydrological conditions; (5) run the hydrological model with the 2K climate data in four scenarios describing (a) a stationary system with historical root-zone storage capacity and historical land use, (b) an adapted root-zone storage capacity in response to a changing climate but a historical land use, (c,d) an adapted root-zone storage capacity and a shift in dominant land use; and finally (6) compare the change in hydrological response between 2K and historical conditions for these four scenarios.

6.4.1 CHANGING CLIMATE, VEGETATION AND LAND USE

LONG-TERM WATER BALANCE FRAMEWORK FOR ESTIMATING THE CHANGE IN RUNOFF COEFFICIENT

The long-term partitioning of precipitation (P) into evaporation (E_A) and streamflow (Q) is mainly controlled by the long-term aridity index (ratio of potential evaporation over precipitation, E_P/P), according to the Budyko hypothesis. To account for additional factors that influence the long-term hydrological partitioning, Fu (1981) introduces a parameter ω to encapsulate the combined influences of climate, soils, vegetation and topography (Equation 6.1).

$$\frac{E_A}{P} = 1 - \frac{Q}{P} = 1 + \frac{E_P}{P} - \left(1 + \left(\frac{E_P}{P}\right)^{\omega}\right)^{1/\omega} \tag{6.1}$$

We solve Equation 6.1 to determine the value of ω for each of the 35 catchments of the Meuse basin for observed historical conditions for the period 2005 to 2017 (ω_{obs}), using the meteorological E-OBS data (P_{obs} and $E_{P,obs}$) and observed streamflow (Q_{obs}). Assuming only a change in long-term mean climate conditions, i.e. aridity index, a catchment will move along its ω_{obs}-parameterized curve from its original position (p_1) to a

new position (p_2) due to the horizontal shift in aridity index ($\Delta E_P/\Delta P$, Fig. 6.2a). Here, we use the simulated historical and 2K climate data to determine how the change in potential evaporation ($\Delta E_P = E_{P,2K} - E_{P,hist}$) and precipitation ($\Delta P = P_{2K} - P_{hist}$) lead to a change in aridity index (Equation 6.2) and therefore in actual evaporation ($\Delta E_A = E_{A,2K} - E_{A,hist}$) and streamflow ($\Delta Q = Q_{2K} - Q_{obs}$), using Equation 6.1.

$$\left(\frac{E_P}{P}\right)_{2K} = \frac{E_{P,obs} + \Delta E_P}{P_{obs} + \Delta P} \tag{6.2}$$

However, land cover and vegetation are likely to also change in response to a changing climate, introducing an additional vertical shift ($\Delta\omega$) toward a position (p_3) on a different ω_{change} curve (Fig. 6.2a). A downward vertical shift from ω_{obs} to ω_{change} indicates less water use for evaporation, as opposed to an upward shift for higher evaporative water use. These vertical shifts in ω-values represent changes in drivers other than aridity index, including e.g. land cover, tree species, forest age, biomass growth and water use efficiency (Jaramillo et al., 2018).

To test the sensitivity of the hydrological response to a change in ω in addition to a change in aridity index, we consider two scenarios. The catchments with relatively high percentages of broadleaved forests (25-38% as in the French part of the basin) receive the ω-values of catchments with relatively low percentages of broadleaved forests (1-12% as mainly in the Belgian Ardennes) and vice-versa (Fig. 6.1b). We denote $\omega_{broadleaved}$ for the catchments with relatively high percentages of broadleaved forests and $\omega_{coniferous}$ for the catchments where broadleaved forests were largely converted to coniferous plantations or agriculture. These scenarios are meant as a sensitivity analysis in the spirit of trading space-for-time (Singh et al., 2011) to evaluate the effect of potential future land-use management on the overall water balance.

When converting broadleaved forest to coniferous plantations, we expect an increase in water use for evaporation and therefore a vertical upward shift in ω-values, as opposed to a downward shift when converting coniferous plantations to more natural broadleaved forests. The described vertical and horizontal movements in the Budyko space are used to estimate the projected long-term runoff coefficients ($(Q/P)_{2K}$, Equation 6.3) as a result of, both, climate change but no changes in vegetation cover (ω_{obs}), and climate change in combination with changes in vegetation cover (by swapping $\omega_{broadleaved}$ values to $\omega_{coniferous}$ for a selection of catchments and vice-versa). The projected runoff coefficients are subsequently used to estimate changes in the root-zone storage capacity parameter (Sect. 6.4.1).

$$\left(\frac{Q}{P}\right)_{2K,\omega} = \frac{Q_{obs} + \Delta Q}{P_{obs} + \Delta P} = -\left(\left(\frac{E_P}{P}\right)_{2K} - \left(1 + \left(\frac{E_P}{P}\right)^{\omega}_{2K}\right)^{1/\omega}\right) \tag{6.3}$$

SEASONAL WATER BALANCE FOR ESTIMATING THE CHANGE IN ROOT-ZONE STORAGE CAPACITY $S_{R,max}$

The root-zone storage capacity represents the maximum volume of water which can be held in pores of unsaturated soil and which is accessible to roots of vegetation for transpiration. It is a key element controlling the hydrological response of hydrological systems. The long-term partitioning of precipitation into streamflow and evaporation in a changed climate can only match expectations as estimated from movements in the Budyko space (Sect. 6.4.1) if we consider that vegetation has adapted its root-zone storage capacity to offset hydro-climatic seasonality, by creating a buffer large enough to overcome dry spells (Gentine et al., 2012; Donohue et al., 2012; Gao et al., 2014). This is the main assumption underlying the water balance method to estimate the root-zone storage capacity at the catchment scale (Gao et al., 2014; Nijzink et al., 2016b; de Boer-Euser et al., 2016; Wang-Erlandsson et al., 2016; Bouaziz et al., 2020; Hrachowitz et al., 2020).

The water balance method requires daily time series of precipitation, potential evaporation and a long-term runoff coefficient to estimate transpiration, as it depletes the root-zone storage during dry spells. Annual water deficits ($S_{R,def}$) stored in the root-zone of vegetation to fulfill canopy water demand for transpiration are estimated on a daily time step as the cumulative sum of daily effective precipitation (P_E) minus transpiration (E_R).

First, effective precipitation, i.e. the amount of precipitation that reaches the soil after interception evaporation (E_I), is estimated by solving the water balance of a canopy storage (S_I) with maximum interception storage capacity (I_{max}, here taken as 2.0 mm), according to Equation 6.4.

$$P_E(t) = P(t) - E_I(t) - \frac{dS_I(t)}{dt} \tag{6.4}$$

Next, the long-term transpiration $\overline{E}_R$ is estimated from the long-term water balance, using mean annual streamflow and effective precipitation ($\overline{Q}$ and $\overline{P}_E$, all in mm yr^{-1}, Equation 6.5), assuming negligible changes in storage and intercatchment groundwater flows (catchments where $\overline{E}_A = \overline{P} - \overline{Q} < \overline{E}_P$).

$$\overline{E}_R \approx \overline{P}_E - \overline{Q} \tag{6.5}$$

The long-term transpiration $\overline{E}_R$ is subsequently scaled to daily transpiration estimates E_R, using the daily signal of potential evaporation minus interception evaporation, according to Equation 6.6 (Nijzink et al., 2016b; Bouaziz et al., 2020).

$$E_R(t) = (E_P(t) - E_I(t)) \cdot \frac{\overline{E}_R}{(\overline{E}_P - \overline{E}_I)} \tag{6.6}$$

The maximum annual storage deficits can then be derived from the cumulative difference of effective precipitation (P_E) and transpiration (E_R), assuming an "infinite" stor-

age, according to Equation 6.7 and illustrated in Fig. 6.2b. For each year, $S_{R,def}$ represents the amount of water accessible to the roots of vegetation for transpiration during a dry period. Storage deficits are assumed to be zero at the end of the wet period (T_0, here April) and increase when transpiration exceeds effective precipitation during dry periods, until they become zero again (T_1) when excess precipitation is assumed to drain away as direct runoff or recharge.

$$S_{R,def}(t) = \min \int_{T_0}^{T_1} (P_E(t) - E_R(t))\mathrm{d}t \tag{6.7}$$

By fitting the annual maximum storage deficits to the extreme value distribution of Gumbel, the root-zone storage capacity at the catchment scale $S_{R,max}$ can be derived for various return periods. Previous studies used a return period of 20 years for forested areas, meaning that forests develop root systems to survive droughts with a return period of ~20 years (Nijzink et al., 2016b; de Boer-Euser et al., 2016; Hrachowitz et al., 2020). The root-zone storage capacity of cropland and grasslands is assumed to correspond to deficits with a lower return period of ~2 years (Wang-Erlandsson et al., 2016). It should be noted that the methodology assumes that vegetation taps its water from the unsaturated zone and not from the groundwater.

Using the above described methodology, we determine several sets of $S_{R,max}$ values for each of the 35 catchments of the Meuse basin to represent the historical and adapted root-zone storage capacity in response to a changing climate and changing/historical land use, using historical climate observations (E-OBS) and the historical and 2K climate simulations (Table 6.1).

$S_{R,max,A}$: HISTORICAL ROOT-ZONE STORAGE CAPACITY FROM HISTORICAL LAND USE AND OBSERVED HISTORICAL CLIMATE

The first set is $S_{R,max,A}$, which represents the historical meteorological and land-use conditions, derived from observed historical E-OBS data (P_{obs}, $E_{P,obs}$ for the period 1980-2018) and observed streamflow data (Q_{obs} for the period 2005-2017). $S_{R,max,A}$ is used as parameter for three model runs, each forced with a different dataset: historical E-OBS observations, simulated historical and 2K climate data (Sect. 6.4.4).

In this study, we assume that the observed E-OBS historical climate data is the best available estimate of current-day climate conditions and use this data to estimate historical root-zone storage capacities $S_{R,max,A}$ and to calibrate the hydrological model (Sect. 6.4.3). The simulated historical climate data is also required to enable a fair comparison with the simulated 2K climate data, as they are both generated with the regional climate model. Despite potential biases in the climate model simulations compared to the observed historical data (here, E-OBS), we do not apply a formal bias-correction of the climate data which may alter the relations between variables in climate models (Ehret et al., 2012). Instead, we force the hydrological model with the native simulated historical climate data, but use the previously determined $S_{R,max,A}$ parameter. An alternative

approach would have been to estimate the root-zone storage capacities using the simulated historical climate data, to directly correct for potential biases in the climate data in the estimation of the root-zone storage capacity parameter but with the downside of affecting spatial patterns across catchments.

$S_{R,max,B}$: ADAPTED ROOT-ZONE STORAGE CAPACITY FROM HISTORICAL LAND USE AND 2K CLIMATE

We then estimate the root-zone storage capacity $S_{R,max,B}$ based on the 2K climate and historical land use to reflect vegetation adaptation to changing climatic conditions such as differences in seasonality, aridity index (Equation 6.2) and the resulting runoff coefficient (Equation 6.3), but under the assumption that the vegetation cover remains unchanged. To account for differences in the observed and simulated historical climate data, $S_{R,max,B}$ is determined by imposing the difference in storage deficits derived from the 2K and historical climate simulations ($S_{R,def,2K} - S_{R,def,hist}$) on the observed storage deficit derived with E-OBS data $S_{R,def,obs}$, as shown in Table 6.1.

$S_{R,max,C}$: ADAPTED ROOT-ZONE STORAGE CAPACITY FROM LAND-USE CONVERSION TO BROADLEAVED FOREST AND 2K CLIMATE

Subsequently, the root-zone storage capacity is estimated for the 2K climate under two land-use change scenarios, considering that if climate changes, a different vegetation cover might become dominant under natural and anthropogenic influence (Table 6.1). Making use of a space-for-time exchange, we connect the spatially variable ω parameter of the Budyko curve to different land-use categories and use these to evaluate future land-use scenarios. Here, the catchments are categorized according to the areal fraction of broadleaved forest in a catchment (Fig. 6.1b). In the first scenario, land use in the catchments with mainly coniferous plantations and agriculture (as mainly in the Belgian Ardennes, Sect. 6.2 and Fig. 6.1) is assumed to be converted to broadleaved forest, using sampled $\omega_{broadleaved}$ values of catchments within the French part of the basin to estimate the 2K runoff coefficient with Equation 6.3. The sampling is performed because the variability in ω-values in each category is also influenced by other factors besides the dominant presence of broadleaved forest. The resulting $S_{R,max,C}$ thus represents an adapted root-zone storage capacity in response to climate change and land-use conversion to broadleaved forest.

$S_{R,max,D}$: ADAPTED ROOT-ZONE STORAGE CAPACITY FROM LAND-USE CONVERSION TO CONIFEROUS PLANTATION AND AGRICULTURE AND 2K CLIMATE

Similarly, the adapted root-zone storage capacity $S_{R,max,D}$ is estimated for the 2K climate in a land-use scenario where the broadleaved forest in the French part of the basin are converted to coniferous plantations and agriculture, using sampled $\omega_{coniferous}$ values of catchments in the Belgian Ardennes.

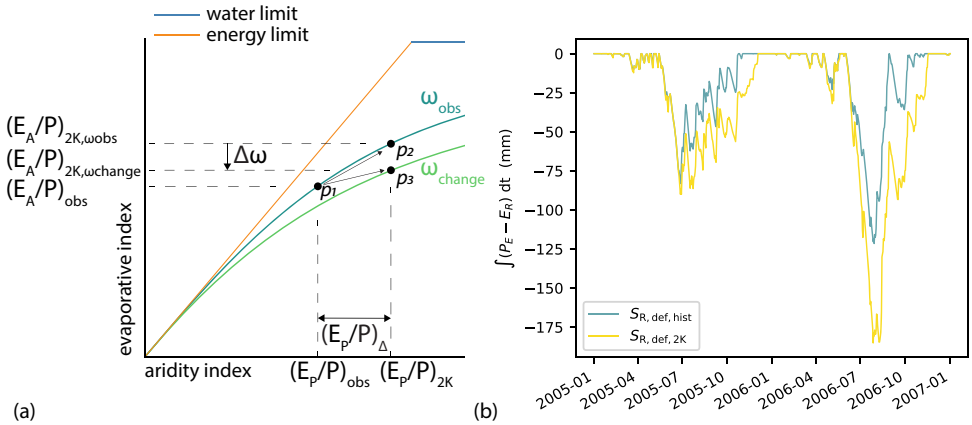

Figure 6.2: **(a)** Representation of the Budyko space, which shows the evaporative index (E_A/P) as a function of the aridity index (E_P/P) and the water and energy limit. A catchment with aridity index $(E_P/P)_{obs}$ and evaporative index $(E_A/P)_{obs}$, derived from observed historical data, plots at location p_1 on the parametric Budyko curve with ω_{obs}. A movement in the Budyko space towards p_2 along the ω_{obs} curve is shown as a result of a change in aridity index $(E_P/P)_{\Delta}$ towards a projected $(E_A/P)_{2K,\omega obs}$ associated with aridity $(E_P/P)_{2K}$. An additional vertical shift $\Delta\omega$ towards a location p_3 on a ω_{change} curve is shown if additional factors (e.g. land use) are projected to change besides aridity index. Here, the represented downward shift in ω reduces the change in evaporative index to $(E_A/P)_{2K,\omega change}$. **(b)** Cumulative storage deficits ($S_{R,def}$) derived from effective precipitation (P_E) and transpiration (E_R) using the simulated historical and 2K climate data. Estimates of transpiration (E_R) are derived from long-term water balance projections as a result of movements within the Budyko framework in response to climate and potential land-use changes.

Table 6.1: Root-zone storage capacity $S_{R,max}$ description and symbols, derived from long-term transpiration and storage deficits calculations for observed historical E-OBS data (P_{obs}) and simulated historical (P_{hist}) and 2K climate data (P_{2K}) for historical land use (ω_{obs}) and land-use change scenarios ($\omega_{broadleaved}$ and $\omega_{coniferous}$). The overline symbol is omitted from P, Q and E_R to increase readability.

Description	$S_{R,max}$ [mm]	Long-term transpiration E_R [mm yr^{-1}] (Eq. 6.5)	Storage deficit $S_{R,def}$ [mm] (Eq. 6.7)
Observed historical E-OBS historical land use (ω_{obs})	$S_{R,max,A}$	$P_{E,obs} - Q_{obs}$	$S_{R,def,obs}$
Simulated historical climate historical land use (ω_{obs}) (historical runoff coefficient)	-	$P_{E,hist} - Q_{obs}/P_{obs} \cdot P_{hist}$	$S_{R,def,hist}$
2K climate historical land use (ω_{obs})	$S_{R,max,B}$	$P_{E,2K} - (Q/P)_{2K,B} \cdot P_{2K}$ (Eq. 6.3)	$\max(\lvert S_{R,def,obs} + \min(0, S_{R,def,2K,B} - S_{R,def,hist})\rvert)$
2K climate broadleaved land use ($\omega_{broadleaved}$)	$S_{R,max,C}$	$P_{E,2K} - (Q/P)_{2K,C} \cdot P_{2K}$ (Eq. 6.3)	$\max(\lvert S_{R,def,obs} + \min(0, S_{R,def,2K,C} - S_{R,def,hist})\rvert)$
2K climate coniferous land use ($\omega_{coniferous}$)	$S_{R,max,D}$	$P_{E,2K} - (Q/P)_{2K,D} \cdot P_{2K}$ (Eq. 6.3)	$\max(\lvert S_{R,def,obs} + \min(0, S_{R,def,2K,D} - S_{R,def,hist})\rvert)$

6.4.2 WFLOW_FLEX-TOPO HYDROLOGICAL MODEL

The wflow_FLEX-Topo model (de Boer-Euser, 2017; Schellekens et al., 2020) is a fully distributed process-based model, which uses different model structures for a selection of Hydrological Response Units (HRUs) to represent the spatial variability of hydrological processes. Here, we develop a model with three HRUs for wetlands, hillslopes and plateaus connected through their groundwater storage (schematized in Fig. 6.3 and model equations in Appendix A; Savenije, 2010; de Boer-Euser, 2017). Thresholds of 5.9 m for the Height Above the Nearest Drainage (HAND Rennó et al., 2008) and 0.129 for slope are used to delineate the three HRUs (Gharari et al., 2011) using the MERIT hydro dataset at ~60 m x 90 m resolution (Yamazaki et al., 2019). Given the high proportion of forest on hillslope and of agriculture on plateau, we here associated hillslope with forest and agriculture with plateau, using the CORINE land cover data (European Environment Agency, 2018). The areal fraction of each HRU are then derived for each cell at the model resolution of 0.00833° (or ~600 m x 900 m). The model includes snow, interception, root-zone, fast and slow storage components. Streamflow is routed through the upscaled river network at the model resolution (Eilander et al., 2020) with the kinematic wave approach. Similar implementations of that model were previously successfully used in a wide variety of environments (e.g. Gharari et al., 2013; Gao et al., 2014; Nijzink et al., 2016b; de Boer-Euser, 2017; Hulsman et al., 2021; Bouaziz et al., 2021).

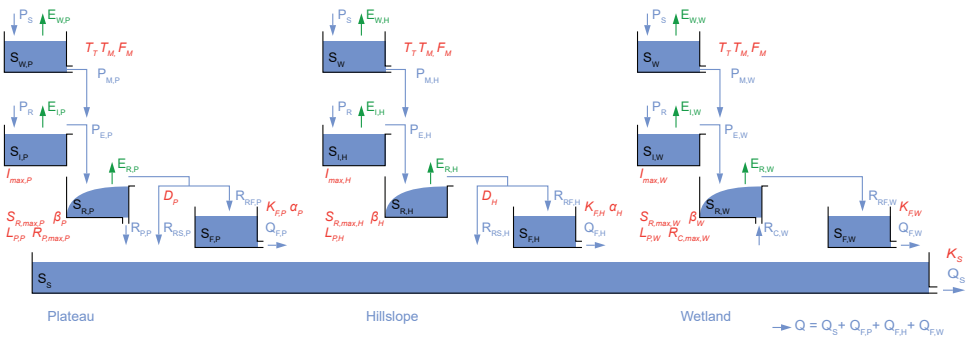

Figure 6.3: Schematic representation of the wflow_FLEX-Topo model with three HRUs for plateau, hillslope and wetland connected through their groundwater storage. The model includes storages for snow S_W, interception S_I, the root-zone S_R, a fast runoff component S_F and groundwater S_S [mm]. The total streamflow Q [mm d^{-1}] is the sum of fast runoff Q_F from the three HRUs and groundwater runoff Q_S. Evaporation [mm d^{-1}] occurs from the snow storage (E_W), the interception storage (E_I) and the root-zone storage (E_R). Main parameters for snow processes include a threshold temperature T_T [$°C$] to distinguish precipitation P falling as rain P_R or snow P_S, a threshold temperature for melt T_M [mm d^{-1}] and a degree-day factor F_M [mm d^{-1} $°C^{-1}$]. For each HRU, other parameters include a maximum interception capacity I_{max} [mm], a maximum root-zone storage capacity $S_{R,max}$ [mm], a shape factor β [-], a transpiration water stress factor L_P [-], a factor for the fraction of preferential groundwater recharge D [-], a recession coefficient for the fast storage K_F [d^{-1}] and a combined recession for the slow storage K_S [d^{-1}]. Parameters specific to plateau, hillslope and wetland include a maximum percolation rate $R_{P,max,P}$ [mm d^{-1}], a non-linear coefficient for fast runoff α_P and α_H [-], and a maximum capillary rise rate $R_{C,max,W}$ [mm d^{-1}]. Effective precipitation is denoted as P_E, fluxes between two stores are denoted as R with subscripts for the stores, and subscripts P, H and W are used to distinguish between the three HRUs.

6.4.3 MODEL CALIBRATION AND EVALUATION

CALIBRATION AND EVALUATION USING THE OBSERVED HISTORICAL E-OBS CLIMATE DATA

The wflow_FLEX-Topo model is calibrated using streamflow at Borgharen and the observed historical E-OBS meteorological forcing data for the period 2007-2011, using 2005-2006 as warm-up years. The observed historical E-OBS dataset is used for calibration of the model as it is assumed to most closely represent current-day conditions. The parameter space is explored with a Monte Carlo strategy, sampling 10000 realizations from uniform prior parameter distributions (Appendix A). The limited number of samples is due to the high computational resources required to run the distributed model. However, our aim is not to find the "optimal" parameter set, but rather to retain an ensemble of plausible parameter sets based on a multiobjective calibration strategy (Hulsman et al., 2019). To best reflect different aspects of the hydrograph, including high flows, low flows and medium-term partitioning of precipitation into drainage and evaporation, parameter sets are selected based on their ability to simultaneously and adequately represent four objective functions, including the Nash-Sutcliffe efficiencies of streamflow, the logarithm of streamflow and, monthly runoff coefficients as well as the Kling-Gupta efficiency of streamflow. Only parameter sets that exceed a performance threshold of 0.9 for each metric are retained as feasible. The root-zone storage capacity parameter $S_{R,max,A}$ is a fixed parameter, which is derived from annual maximum storage deficits with a return period of 2 years for the wetland and plateaus HRUs and 20 years for the hillslopes HRU (Sect. 6.4.1). Next, model performance is evaluated in the 2012-2017 post-calibration period using the same performance metrics, a visual inspection of the hydrographs and the mean monthly streamflow regime. The performance metrics are also evaluated for the 34 remaining nested sub-catchments.

EVALUATION USING THE SIMULATED HISTORICAL CLIMATE DATA

The performance of the calibrated model for the ensemble of retained parameter sets is also evaluated when the model is forced with the simulated historical climate data, using $S_{R,max,A}$ for the root-zone storage capacity parameter. This is the reference historical run against which the relative effect of 2K global warming is evaluated for different scenarios (Fig. 6.4 and Sect. 6.4.4). In addition, we evaluate the performance of the calibrated model forced with the simulated historical climate data but with a root-zone storage capacity parameter derived directly from this data. While this alternative approach enables to correct for potential biases in the simulated historical climate data directly in the estimation of the root-zone storage capacity parameter, it may also affect the spatial patterns of this parameter across catchments.

6.4.4 HYDROLOGICAL CHANGE EVALUATION

We then force the calibrated wflow_FLEX-Topo model for the ensemble of retained parameter sets with the 2K climate data in four scenarios each using a different root-zone storage capacity parameter to represent either stationary or adapted conditions in response to a changing climate and land use (Fig. 6.4 and Sect. 6.4.1). The difference between the modelled historical hydrological response (1980-2018) and the hydrolog-

ical responses predicted by each of the four model scenarios based on the 2K climate is evaluated in terms of runoff coefficient, evaporative index, annual statistics (runoff coefficient, evaporative index, mean, maximum, minimum 7-days streamflow and median volume deficit below the 90^{th} percentile reference streamflow), and monthly patterns of flux and state variables (streamflow, evaporation, root-zone storage, groundwater storage) for a hypothetical 38-year period.

SCENARIO $2K_A$: HISTORICAL LAND USE AND HISTORICAL ROOT-ZONE STORAGE CAPACITY ($S_{R,max,A}$)

In scenario $2K_A$ (Fig. 6.4), we assume an unchanged land use and that vegetation has not adapted its root-zone storage capacity to the aridity and seasonality of the 2K climate. This scenario implies stationarity of model parameters by using $S_{R,max,A}$ in both the historical and 2K runs, a common assumption of many climate change impact assessment studies (Booij, 2005; de Wit et al., 2007; Prudhomme et al., 2014; Hakala et al., 2019; Brunner et al., 2019; Gao et al., 2020; Rottler et al., 2020). This is the benchmark scenario against which we compare the hydrological response considering non-stationarity of the system, as in the following three scenarios.

SCENARIO $2K_B$: HISTORICAL LAND USE AND 2K CLIMATE ADAPTED ROOT-ZONE STORAGE CAPACITY ($S_{R,max,B}$)

In scenario $2K_B$ (Fig. 6.4), we again assume an unchanged land use (ω_{obs}). However, we assume that vegetation has adapted its root-zone storage capacity to the aridity and seasonality of the 2K climate conditions by selecting $S_{R,max,B}$ as parameter for the 2K model run, while the historical $S_{R,max,A}$ is used as parameter in the historical run.

SCENARIO $2K_C$: LAND-USE CONVERSION TO BROADLEAVED FOREST AND 2K CLIMATE ADAPTED ROOT-ZONE STORAGE CAPACITY ($S_{R,max,C}$)

In scenario $2K_C$ (Fig. 6.4), we adapt the root-zone storage capacity to the changing aridity index and seasonality of the 2K climate. Additionally, we assume a change in vegetation cover for the catchments located mainly in the Belgian Ardennes and dominated by coniferous plantation and agriculture to a land use of broadleaved forest as in the French part of the basin. For this purpose, $S_{R,max,C}$ is used as parameter in the model run forced with the 2K climate, while $S_{R,max,A}$ is used as parameter in the historical run.

SCENARIO $2K_D$: LAND-USE CONVERSION TO CONIFEROUS PLANTATION AND AGRICULTURE AND 2K CLIMATE ADAPTED ROOT-ZONE STORAGE CAPACITY ($S_{R,max,D}$)

In scenario $2K_D$ (Fig. 6.4), the approach of scenario $2K_C$ is repeated. However, now the broadleaved forest in the French catchments are assumed to be converted to coniferous plantations or agriculture as in the Belgian Ardennes. The parameter $S_{R,max,D}$ is used in the model run forced with the 2K climate, while $S_{R,max,A}$ is used in the historical run.

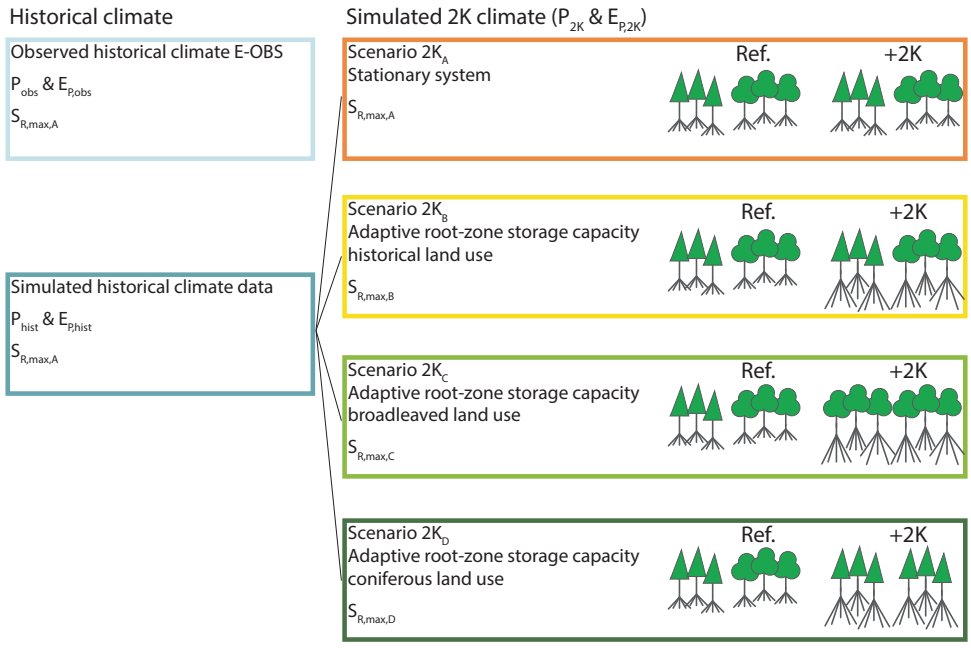

Figure 6.4: Model scenarios using the observed historical and the simulated historical and 2K climate data. The model is calibrated using observed E-OBS data and the historical root-zone storage capacity $S_{R,max,A}$. The model is then forced with the simulated historical climate data using $S_{R,max,A}$ as root-zone storage capacity parameter. We then define four scenarios to compare the change in hydrological response to 2K global warming in comparison to historical conditions for the ensemble of feasible parameter sets. In scenario $2K_A$, we assume an unchanged system (no changes in land use, nor root-zone storage capacity). In scenario $2K_B$, we assume that vegetation has adapted its root-zone storage capacity to the 2K climate, but no changes in land use. In scenario $2K_C$, we test the combination of an adapted root-zone storage capacity in response to the changed climate and a hypothetical conversion of coniferous plantations and agriculture to broadleaved forests in part of the catchment. A similar but reversed approach in land-use changes is assumed in scenario $2K_D$.

6.5 RESULTS

6.5.1 ADAPTED ROOT-ZONE STORAGE CAPACITY $S_{R,max}$ FROM LONG-TERM AND SEASONAL WATER BALANCES AND CHANGING LAND USE

LONG-TERM WATER BALANCE CHARACTERISTICS ACROSS CATCHMENTS

In solving the parametric Budyko curve (Equation 6.1) for the 35 catchments of the Meuse basin using historical E-OBS data and observed streamflow (Fig. 6.5a), we found that ω_{obs} values tend to be lower (median of 2.43 ± 0.48) for catchments with relatively high percentages of broadleaved forests (25-38 % as in the French part of the basin) as compared to catchments with relatively low percentages of broadleaved forests (1-12 % as in the Belgian part of the catchment) with median ω-values of 3.04 ± 0.54, as shown in Fig. 6.5b. Higher values of ω for a same aridity index indicate more water use for evaporation, which is likely related to the increased water use of relatively young coniferous

plantations and agriculture as opposed to older broadleaved forests (Fenicia et al., 2009; Teuling et al., 2019).

$S_{R,max,A}$ FROM HISTORICAL LAND USE AND HISTORICAL CLIMATE

The root-zone storage capacity $S_{R,max,A}$ derived with observed historical E-OBS climate data and observed streamflow is estimated at values of 101 ± 17 mm and 169 ± 24 mm across all study catchments for a 2 year and 20 year return period, respectively (Fig. 6.5c).

If instead the simulated historical climate data is used to derive the root-zone storage capacity, this results in slightly higher values with 110 ± 18 mm and 180 ± 28 mm for 2 and 20 year return periods, respectively. This overestimation of about +7 % is due to the higher precipitation (on average +9 %) in the simulated historical climate data compared to the observed E-OBS historical data, which leads to relatively lower runoff coefficients and therefore larger evaporative indices and storage deficits in the water balance calculation of the root-zone storage capacity.

$S_{R,max,B}$ FROM HISTORICAL LAND USE AND 2K CLIMATE

The adapted root-zone storage capacity $S_{R,max,B}$, in response to changing climate conditions and an unchanged land use, strongly increases with respect to historical conditions ($S_{R,max,A}$) with estimated values of 129 ± 18 mm (+28 %) and 227 ± 27 mm (+34 %) for return periods of 2 year and 20 year, respectively (Fig. 6.5c). This strong increase is explained by larger storage deficits during summer due to an increase of about +10 % in summer potential evaporation in the 2K climate and, therefore, a more pronounced seasonality (Fig. 6.2b). In contrast, the change in aridity index between the historical and 2K climate simulations is relatively small with a median of +0.01 across all study catchments. This can be explained by a simultaneous increase in mean annual precipitation (+5 %) and potential evaporation (+7 %) on average over the basin area in the 2K climate compared to the simulated historical climate data. This increase in precipitation mostly occurs during the winter half year (Nov-Apr). In contrast, there is a relatively large variability in precipitation change in summer, characterized by years with wetter and drier summers.

$S_{R,max,C}$ FROM 2K CLIMATE AND ADAPTED LAND USE TO BROADLEAVED FOREST

The adapted root-zone storage capacity $S_{R,max,C}$, in response to changing climate conditions and a land-use conversion from coniferous plantation and agriculture to broadleaved forest, results in estimated values of 125 ± 17 mm and 219 ± 27 mm for return periods of 2 year and 20 year, respectively (Fig. 6.5c). These values are almost similar to $S_{R,max,B}$, with a difference of about -3 %. This small decrease is in line with the expected reduced water use of broadleaved forests compared to coniferous plantations.

$S_{R,max,D}$ FROM 2K CLIMATE AND ADAPTED LAND USE TO CONIFEROUS PLANTATIONS AND AGRICULTURE

In contrast, the root-zone storage capacity $S_{R,max,D}$, in response to changing climate conditions and a conversion of broadleaved forest to coniferous plantation, result in esti-

mated values of 140 ± 22 mm and 243 ± 35 mm for return periods of 2 year and 20 year, respectively (Fig. 6.5c). This corresponds to an increase of +9 % and +7 % for both return periods in comparison with $S_{R,max,B}$, which does not consider additional land-use changes.

The difference in root-zone storage capacity between the 2K and historical climate simulations as a result of a changing climate (aridity and seasonality) is larger (+58 mm or +34 % for a return period of 20 years) than the difference between root-zone storage capacity for a changing climate and additional changes in land use (-8 mm or -3 % for $S_{R,max,C}$ and +16 mm or +7 % for $S_{R,max,D}$). This indicates that with the assumed land-use change in scenarios $2K_C$ and $2K_D$, the strong increase in water demand during summer as a result of a more pronounced seasonality has greater impact on the estimation of the root-zone storage capacity than a change in ω-values. However, note that land use is changed in only part of the catchment for both land-use change scenarios and that it is plausible to assume that more pronounced changes in land use will reinforce the observed effects.

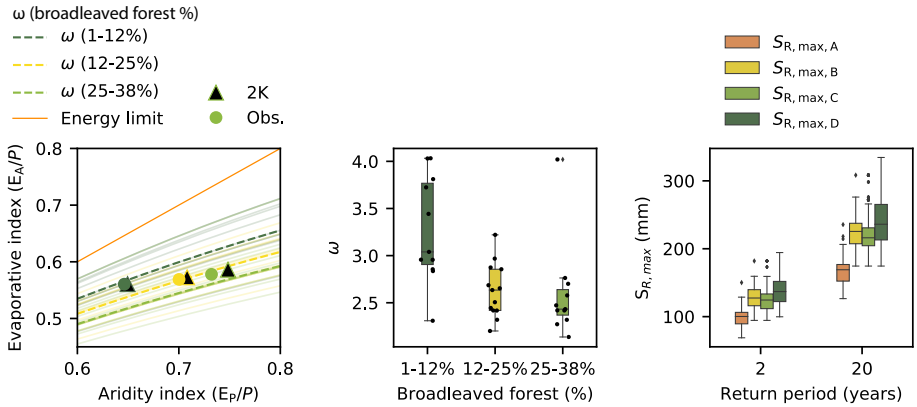

Figure 6.5: **(a)** Budyko space with parametric ω_{obs} curves for each of the 35 catchments of the Meuse basin, categorized according to their percentage of broadleaved forest. The dashes curves represent the median ω_{obs} curves for each category. The change in aridity index from historical to 2K climate conditions along each parameterized ω_{obs} curve is also shown for the median of the three categories. **(b)** Parameterized ω_{obs} values for each of the 35 catchments of the Meuse basin, categorized according to their percentage of broadleaved forest. **(c)** Range of root-zone storage capacities across the 35 catchments of the Meuse basin for the four scenarios. $S_{R,max,A}$ represents the root-zone storage capacity for historical conditions. $S_{R,max,B}$ represents an adapted root-zone storage capacity in response to the 2K climate but no land-use change. In the estimation of $S_{R,max,C}$, catchments with a low percentage of broadleaved forest (1-12%) receive ω-values sampled from catchments with a high percentage of broadleaved forest (25-38%), to represent changes in land use towards a conversion to broadleaved forest. A similar but reversed approach is applied for the estimation of $S_{R,max,D}$.

6.5.2 MODEL EVALUATION (HISTORICAL PERIOD)

MODEL FORCED WITH OBSERVED HISTORICAL CLIMATE DATA

The ensemble of parameter sets retained as feasible after calibration mimics the observed hydrograph at Borgharen relatively well for the evaluation period (Fig. 6.6a). Also the seasonal streamflow regime is relatively well reproduced by the model, except for a slight underestimation of about -9 % in the first half year (Fig. 6.6b). The four objective functions show a relatively similar performance during calibration and evaluation with median values of approximately 0.93 and 0.78 at Borgharen and for the ensemble of nested catchments of the Meuse, respectively (Fig. 6.7a,b).

MODEL FORCED WITH SIMULATED HISTORICAL CLIMATE DATA

When the calibrated model is instead forced with the simulated historical climate data, peaks are slightly overestimated in comparison to the model run forced with the observed historical E-OBS data (Fig. 6.6c). This is due to the on average +9 % overestimation of precipitation in the simulated historical climate data compared to the observed historical E-OBS climate data. This precipitation overestimation results in an overestimation of about +12 % of modeled mean monthly streamflow during the wettest months (Fig. 6.6d). The streamflow model performance at Borgharen slightly decreases when the simulated historical climate data is used instead of E-OBS, but median values across the ensemble of feasible parameter sets are still above 0.77 for each of the objective functions (Fig. 6.7c). Although a decrease in model performance is found in a few nested catchments, the performance in the ensemble of nested catchments of the Meuse remains relatively high with median values of around 0.67 (Fig. 6.7d). The results of the model run forced with the simulated historical data and with the root-zone storage capacity parameter derived directly from this data show a relatively similar behavior (not shown).

The calibrated model forced with the simulated historical climate data shows a plausible behavior with respect to observed streamflow and is also close to the performance achieved with the observed historical E-OBS climate data. This is important because the effect of the 2K climate on the hydrological response is evaluated with respect to the model run forced with the simulated historical climate data, as they are both generated with the regional climate model. Therefore, the relatively high model performance in the evaluation period enable us to use the retained parameter sets from the calibration with E-OBS data for the subsequent analyses with the simulated historical and 2K climate data.

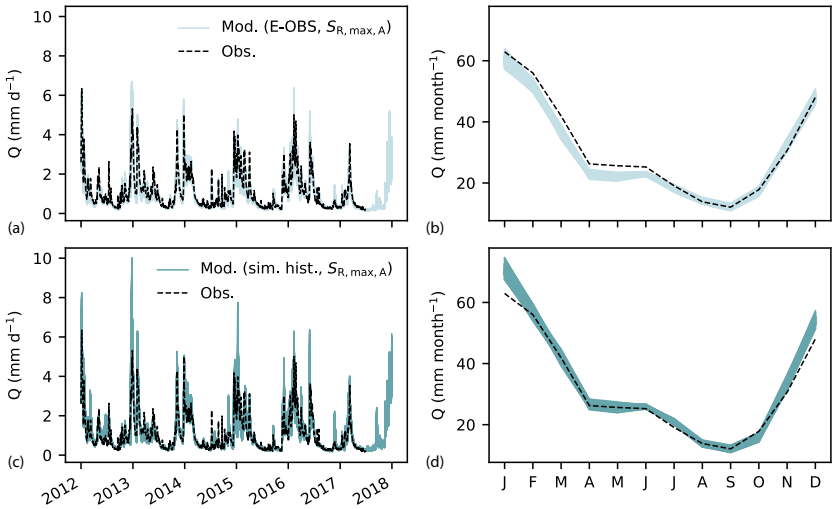

Figure 6.6: Observed and modeled hydrographs and mean monthly streamflow at Borgharen for the ensemble of parameter sets retained as feasible after calibration when the model is: **(a,b)** forced with E-OBS historical data and using $S_{R,max,A}$ as model parameter, and **(c,d)** forced with the simulated historical climate data using $S_{R,max,A}$ as model parameter.

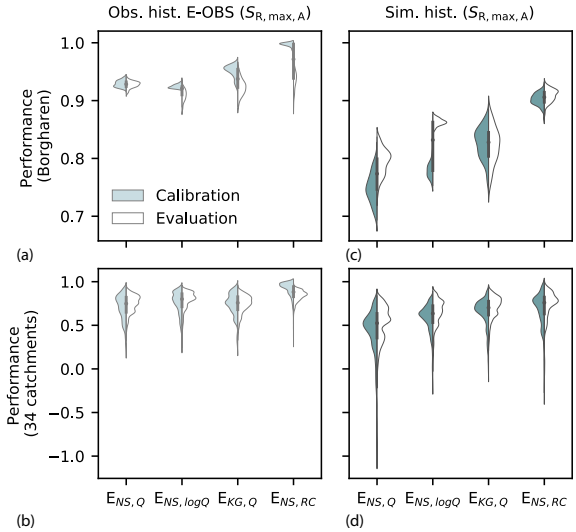

Figure 6.7: Streamflow model performance during calibration and evaluation for the four objective functions when the model is forced with (**a,b**) observed historical E-OBS data and (**c,d**) simulated historical climate data at (**a,c**) Borgharen and (**b,d**) for the ensemble of nested catchments in the Meuse basin. The four objective functions are the Nash-Sutcliffe efficiencies of streamflow, logarithm of streamflow and monthly runoff coefficient ($E_{NS,Q}$, $E_{NS,logQ}$, $E_{NS,RC}$) as well as the Kling-Gupta efficiency of streamflow ($E_{KG,Q}$). Note the different y-axis between rows.

6.5.3 HYDROLOGICAL CHANGE EVALUATION (2K WARMER CLIMATE)

SCENARIO $2K_A$: STATIONARITY WITH HISTORICAL LAND USE AND HISTORICAL ROOT-ZONE STORAGE CAPACITY ($S_{R,max,A}$)

In the $2K_A$ scenario, representing a stationary system with identical parameters in the historical and 2K climate, runoff coefficients are projected to increase with a median of +3 %, the evaporative index (E_A/P) decreases with a median of -2 % and mean annual streamflow increases with a median of +7 %. Maximum annual streamflow is also projected to increase with a median of about +5 %, while the median change in annual minimum of 7-days mean streamflow remains close to zero. The median annual deficit volume below the 90^{th} percentile historical streamflow increases with +10 %, as shown in Fig. 6.8.

Streamflow is projected to increase from December until August with +8 % and decrease between September and November with -7 %. In the months where evaporation demand exceeds precipitation, the root-zone soil moisture decreases, with a maximum of -22 % in September. Actual evaporation increases throughout the year with +3 % except in July and August (-4 %) when the availability of water in the root-zone of vegetation is not sufficient to supply canopy water demand. Recharge to the groundwater storage increases with approximately +5 % in all months except November, as shown in Fig. 6.9.

SCENARIO $2K_B$: HISTORICAL LAND USE AND ADAPTED ROOT-ZONE STORAGE CAPACITY ($S_{R,max,B}$)

Changes are substantially different in the $2K_B$ scenario which considers that the root-zone storage capacity of vegetation has adapted to the change in aridity and seasonality of the 2K climate. Runoff coefficients are instead projected to decrease with a median of -2 %, while the evaporative index increases with a median of +2 % and the median change of mean annual streamflow is close to zero (Fig. 6.8). Also the median change of, both, annual maximum streamflow and minimum 7-days mean streamflow remain close to zero. However, there is a substantial increase of +38 % in median annual deficit volume below the 90^{th} percentile historical streamflow. This result suggests that while the minimum streamflow remains relatively similar, the length of the low flow period strongly increases if we consider that the root-zone storage capacity has adapted to the 2K climate (Fig. 6.8).

Seasonal changes indicate a decrease in streamflow of -19% between September and November, which is longer and more pronounced than in the $2K_A$ scenario (Fig. 6.9). The root-zone soil moisture increases throughout the year with an average of +34 % due to the larger root-zone storage capacities. Actual evaporation is no longer reduced as a result of moisture stress in the root-zone and strongly increases with approximately +7 % from May to October to supply canopy water demand. The increase in evaporation during summer strongly reduces the groundwater recharge with -5 % from October to February (Fig. 6.9).

SCENARIO $2K_C$: LAND-USE CONVERSION TO BROADLEAVED FOREST AND ADAPTED ROOT-ZONE STORAGE CAPACITY ($S_{R,max,C}$)

The predicted hydrological response in the $2K_C$ scenario is very similar to the response of the $2K_B$ scenario, despite considering additional changes in the root-zone storage capacity as a result of a land-use conversion from coniferous plantations and agriculture to broadleaved forest (Figs. 6.8 and 6.9). This is in line with the limited differences in root-zone storage capacities of approximately +3 % between both scenarios (Sect. 6.5.1).

SCENARIO $2K_D$: LAND-USE CONVERSION TO CONIFEROUS PLANTATIONS AND AGRICULTURE AND ADAPTED ROOT-ZONE STORAGE CAPACITY ($S_{R,max,D}$)

In contrast, the change in hydrological response is most pronounced for the scenario $S_{R,max,D}$, which considers land-use conversion of the broadleaved forests in the French part of the basin to coniferous plantations and agriculture (Figs. 6.8 and 6.9). Runoff coefficients decrease with a median of -4 %, while the evaporative index increases with a median value of +4 % and mean annual streamflow decreases with a median of -2 %. If the median change in streamflow extremes remains relatively close to zero, there is a strong increase of +54 % in the median annual deficit volume, suggesting a strong increase in the length of the low flow period (Fig. 6.8).

Streamflow decreases from August to January with an average of -23 % and evaporation strongly increases from May to October with an average of +9 %. This increased evaporation during summer further reduces recharge from October to February with -7 % (Fig. 6.9). In comparison with the hydrological response of scenario $2K_B$, the additional land-use conversion in scenario $2K_D$ results in relatively similar patterns of change but with an additional +2 % increase in evaporation, -2 % decrease in streamflow and -2 % decrease in recharge on average throughout the year.

STATIONARY VERSUS ADAPTIVE ECOSYSTEMS

There is a difference of -7 % in the change of mean annual streamflow between the scenarios $2K_B$, $2K_C$, $2K_D$ with adaptive ecosystems and the stationary $2K_A$ scenario. Additionally, the scenarios with adaptive ecosystems show a more pronounced decrease in streamflow from September to January and a delay in the occurrence of the lowest streamflow from September to October. The change in mean annual actual evaporation is approximately +4 % higher in the scenarios with adaptive ecosystems and the increase mainly occurs between May and October. Instead of a year-round increase in recharge in the $2K_A$ scenario, there is a decrease in winter recharge in the three other scenarios, resulting in a mean annual difference of -6 % between the scenarios with ecosystem adaption and the stationary scenario $2K_A$. Hence, the hydrological response in the 2K climate of the stationary scenario $2K_A$ is substantially different from the responses of the three scenarios $2K_B$, $2K_C$, $2K_D$, which consider a change in the root-zone storage capacity to reflect ecosystem adaptation in response to climate change.

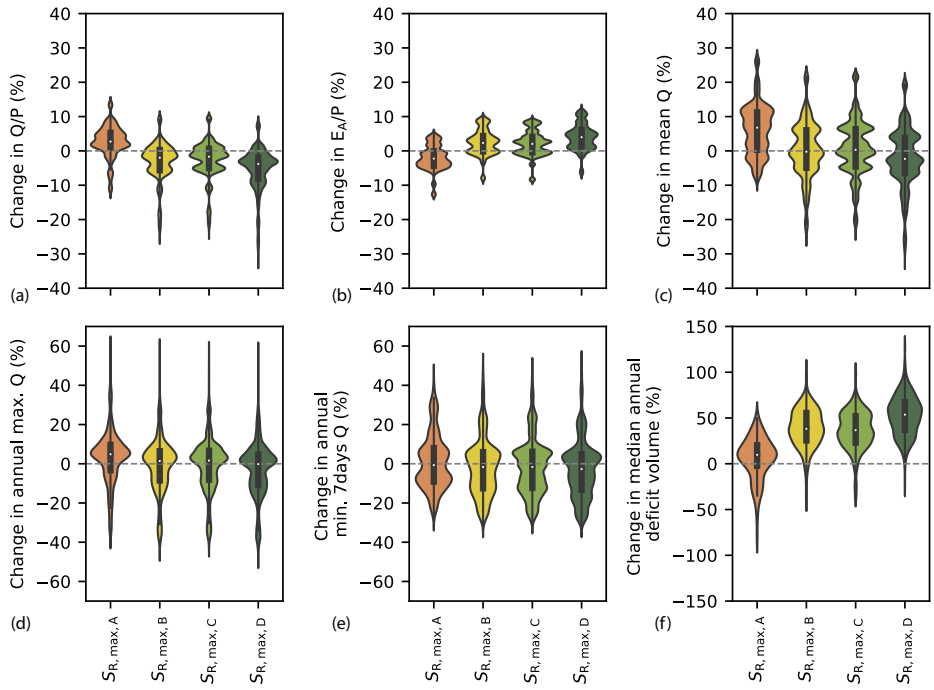

Figure 6.8: Percentage change in annual hydrological response indicators between the 2K and historical model runs for the four scenarios, each based on different assumptions for the root-zone storage capacity parameter $S_{R,max}$. Percentage change in **(a)** runoff coefficient Q/P, **(b)** evaporative index E_A/P, **(c)** mean annual streamflow, **(d)** mean annual maximum streamflow, **(e)** minimum annual 7-days mean streamflow, **(f)** median annual deficit volume below the reference 90th percentile streamflow.

6.6 DISCUSSION

6.6.1 IMPLICATIONS

The hydrological response under 2K global warming with respect to historical conditions shows distinct patterns of change if we explicitly consider the non-stationarity of climate-vegetation interactions in a process-based hydrological model. We implement a dynamic root-zone storage capacity parameter, which is directly inferred from long-term and seasonal water balances of historical observations in combination with simulated historical and future climate conditions. A time-dynamic parameterization of the root-zone storage capacity was previously introduced by Nijzink et al. (2016b) in the context of deforestation, while it was implemented by Speich et al. (2020) in the context of climate change. In the latter study, forest growth in response to climate change leads to a six times higher reduction of streamflow if a dynamic representation of, both, the Leaf Area Index and the root-zone storage capacity is implemented as opposed to a study in which only the Leaf Area Index varies (Schattan et al., 2013). These results are more pronounced than our findings but point towards the same direction of change. While Speich

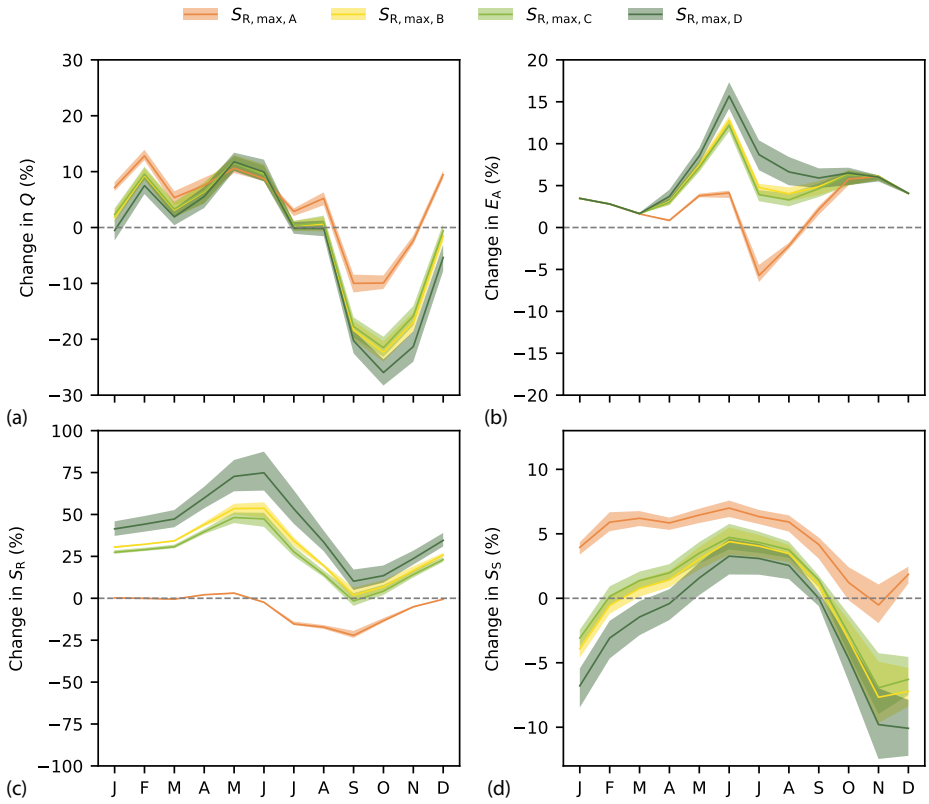

Figure 6.9: Percentage change in mean monthly hydrological response of several flux and state variables between the 2K and historical model runs for the four scenarios, each based on different assumptions for the root-zone storage capacity parameter $S_{R,max}$. Percentage change in mean monthly **(a)** streamflow Q, **(b)** actual evaporation E_A, **(c)** root-zone storage S_R, **(d)** groundwater storage S_S.

et al. (2020) combine a forest landscape model with a hydrological model to simultaneously represent the spatio-temporal forest and water balance dynamics, we rely on a simpler approach of movements in the Budyko framework to include potential land-use change.

The concept of trading space-for-time, which uses space as a proxy for time (Singh et al., 2011) could be further explored by selecting a region outside the Meuse basin with a current climate similar to the projected climate. This approach is also commonly referred to as climate-analogue mapping, i.e. statistical techniques to quantify the similarity between the future climate of a given location and the current climate of another location (Rohat et al., 2018; Bastin et al., 2019; Fitzpatrick and Dunn, 2019). Finding a climate analogue for future projections in present conditions, may allow us to estimate future ω or root-zone storage capacity values in a region where the future climate may resemble today's climate elsewhere. These methods are intuitive but not straightforward,

as they rely on the selection and combination of relevant climate variables and their relation with vegetation, despite non-linear vegetation responses to climate change (Reu et al., 2014).

In comparing several scenarios for the root-zone storage capacity parameter, we include some form of system representation uncertainty, which improves our understanding in the modeled changes by placing them in a broader context (Blöschl and Montanari, 2010). Actual evaporation in the study catchments is projected to decrease if the historical root-zone storage capacity is used as a result of moisture stress in the root-zone of vegetation. In contrast, the increased water demand during summer is met when we assume that vegetation has adapted its root-zone storage capacity. This is an indication of disagreements among model representations on processes that become relevant in the future, in line with findings of Magand et al. (2015); Melsen and Guse (2021).

The impact of climate change on low flows in the Meuse basin has been previously studied by de Wit et al. (2007). Using simulations from regional climate models which project wetter winters and drier summers, they question if the increase in winter precipitation reduces the occurrence of summer low flows due to an increase in groundwater recharge. However, they were unable to address this question with their model due to its poor low-flow performance. Our results indicate an increase in groundwater recharge during winter if the historical root-zone storage capacity parameter is used, as opposed to a decrease for the models with a time dynamic root-zone storage capacity, as a result of an increased water demand for evaporation during summer. This further emphasizes the major impact of vegetation in regulating the water cycle (Luo et al., 2020; Wang et al., 2020).

The land surface scheme HTESSEL, that is used in the regional climate model RACMO2 to generate the historical and 2K climate simulations, assumes, as most land surface models, a fixed root-zone storage capacity. Ideally, this discrepancy between the land surface model and the hydrological model could be reduced by updating the adapted root-zone storage capacity from one model to the other in several iteration steps, thereby including soil moisture - atmosphere feedback mechanisms.

6.6.2 LIMITATIONS AND KNOWLEDGE GAPS

Our study relies on the assumption that vegetation has had the time to adapt its root-zone storage capacity in a changing climate. Yet, considering the unprecedented scale and rate of change (Gleeson et al., 2020), it is unclear how ecosystems will cope with climate change, also considering the impact of storms, heatwaves, fires and biotic infestations as a result of water stress on forest ecosystems (Lebourgeois and Mérian, 2011; Allen et al., 2010; Latte et al., 2017). Additionally, when exposed to different environmental conditions, ecosystems may adapt their behavior by reducing or increasing their water use to the water availability (Zhang et al., 2020). Similarly, direct human interventions, such as the conversion of natural forests to fast-growing monoculture plantations in many parts of the world has significantly altered forests, making them more susceptible and vulnerable to disturbances (Schelhaas et al., 2003; Levia et al., 2020). However,

humans also have the ability to positively influence the water cycle through vegetation, by promoting sustainable agricultural practices and integrated forest management with a simultaneous focus on biodiversity, recreation and timber production. Additionally, the conversion of exotic to native species may also increase biodiversity and the resilience of ecosystems to climate change (Klingen, 2017). However, these processes are slow, implying that current management practices shape the forests of decades and centuries to come in an uncertain future. Increasing our understanding on how to include these changes in hydrological models to reliably quantify their impact is a way forward in the development of strategies to mitigate the adverse effects of climate change.

We quantify the changes in the hydrological response as a result of a changing climate in combination with several land-use scenarios (historical, conversion of broadleaved forests to coniferous plantations and agriculture and vice-versa). These changes are integrated in the root-zone storage capacity as single parameter. However, climate and land-use changes likely affect other aspects of catchment functioning (Seibert and van Meerveld, 2016). For example, changes in the maximum interception storage capacity (Calder et al., 2003) are not explicitly considered in the estimation of the adapted root-zone storage capacities, as the impact was shown to be relatively minor in Bouaziz et al. (2020). Additional effects of soil compaction and artificial drainage on peak flows as a result of future land conversion (Buytaert and Beven, 2009; Seibert and van Meerveld, 2016) are difficult to quantify but may partly be captured in the changed ω-values.

In a first step towards temporally adaptive models, and trading space-for-time for different land-use scenarios, we did not consider any additional vertical movements in the Budyko space due to the effects of increasing CO_2 levels in terms of increased productivity through fertilization, on the one hand, or water use efficiency on the other hand (Keenan et al., 2013; van der Velde et al., 2014; van Der Sleen et al., 2015; Ukkola et al., 2016; Jaramillo et al., 2018; Yang et al., 2019), as these effects remain problematic to quantify in a meaningful way. Neither did we, for the same reason, consider how the relatively high ω-values may be related to intercatchment groundwater losses (Bouaziz et al., 2018). Note that as our analyses should be understood in the context of a sensitivity analyses of the impact of potential additional vertical shifts in the Budyko space as a result of a changing land use (Fig. 6.2), the potential effects on groundwater losses on the results are likely to be minor.

In addition, we performed a limited calibration of the hydrological model to retain an ensemble of plausible solutions and only used a single climate simulation despite the uncertainty in initial and boundary conditions of regional climate models. Our analyses are intended as a proof-of-concept to introduce a top-down methodology to quantify non-stationarity of the root-zone storage capacity parameter through optimal use of projected climate data, rather than a comprehensive climate change impact assessment of the Meuse basin.

6.7 CONCLUSIONS

Understanding non-stationarity of hydrological systems under climate and environmental changes has been recognized as a major challenge in hydrology (Blöschl et al., 2019). Despite our strong awareness of non-stationarity of hydrological parameters, we often lack knowledge to implement system changes in hydrological models. In this proof-of-concept study in the Meuse basin, we propose a top-down approach to introduce a time-dynamic representation of the root-zone storage capacity parameter within process-based hydrological models, using regional climate model simulations. Our approach relies, on the one hand, on a space-for-time exchange of Budyko characteristics of dominant land-use types to estimate the hydrological behavior of potential land-use changes and, on the other hand, on the interplay between the long-term and seasonal water budgets to represent climate-vegetation interactions under climate and land-use change. Despite knowledge gaps on future ecosystem water use, we implement potential system changes in a hydrological model based on our current understanding of hydrological systems. The predicted hydrological response to 2K warming is strongly altered if we consider that vegetation has adapted its root-zone storage capacity to offset the more pronounced hydro-climatic seasonality under 2K global warming compared to a stationary system. The increased vegetation water demand under global warming results on average annually in -7 % less streamflow, +4 % more evaporation and -6 % less recharge for the scenarios assuming non-stationary conditions compared to a stationary system. These differences even lead to a distinct change of sign of median annual streamflow. Our study contributes to the quest for more plausible representations of catchment properties under change and, therefore, more reliable long-term hydrological predictions.

7

SYNTHESIS

Daarna was het zoals het altijd gaat: als je eindelijk hebt bereikt wat je wilde bereiken, is het niet meer wat je wilde bereiken, maar eenvoudig datgene wat je hebt bereikt. Dan is het vanzelfsprekend geworden.

Harry Mulisch - *De ontdekking van de hemel* (1992)

In the previous chapters, several directions are explored to improve and evaluate the representation of internal processes in hydrological models of the Meuse basin. We have evaluated the relevance of specific processes and the use of additional data for reliable model predictions under present-day and projected climate change conditions. In this synthesis chapter, main findings for the Meuse basin are first outlined in terms of hydrological processes and the use of additional data. Then, a synthesis is provided for modelers on how to achieve more reliable hydrological predictions, based on insights gained from this thesis. Next, a specific advice for Rijkswaterstaat is presented to carry on with hydrological modeling of the Meuse basin for operational and policy applications. Finally, ideas for future research and challenges ahead are discussed.

7.1 MAIN FINDINGS FOR THE MEUSE BASIN

7.1.1 HYDROLOGICAL PROCESSES

Despite the shared temperate climate, catchments within the Meuse basin reveal a substantial spatial variability of hydrological processes. Productive aquifers, gentle slopes and more natural broadleaved forests dominate the French part. In contrast, poorly productive aquifers, steeper slopes and coniferous forests, agriculture and urban areas are dominantly represented in the Belgian part. However, this simplified distinction is far too rough and the spatial patterns of hydrological processes are influenced by local heterogeneity of hydrogeological, topographical and land-use properties.

Net intercatchment groundwater losses were identified in several headwater catchments (< 500 km^2) with productive aquifers in both the French and Belgian part of the basin and represented on average 10 % of mean annual precipitation. Some of these underground flow paths may lead to downstream catchments along the same river, while others are diverted to neighboring basins, which explains why net losses could be considerable at the scale of nested headwater catchments and negligible at the scale of larger catchments. Including intercatchment groundwater flows increases high and low flow model performance as well as the plausibility of long-term mean annual actual evaporation estimates (Chapter 3).

Vegetation plays a key role in regulating the hydrological cycle by developing root systems which can buffer a sufficient amount of water to survive dry periods. Also the type of vegetation matters, as coniferous plantations in the Ardennes tend to have a higher water use than the more natural broadleaved forests in the French part of the basin (Chapter 6). The size of the root-zone storage capacity reflects the interactions between water supply through precipitation and transpiration water demand, as controlled by potential evaporation and vegetation type in the catchment. We found relatively large differences in the size of the root-zone storage capacity across catchments within the Meuse (~100-250 mm), suggesting dissimilarities in seasonal climate properties and water demand at the catchment scale.

More importantly, this key parameter of hydrological models can directly be estimated from readily available water balance data for present-day and projected climate conditions. Ecosystems are likely to adapt by increasing their root-zone storage capacity to cope with changing environmental conditions due to an increase of mean global temperatures. This ecosystem adaptation in combination with potential land-use changes strongly alters the monthly patterns of the hydrological response (Chapter 6).

To accommodate the spatial variability of hydrological processes, a flexible process-based model is set-up for 35 nested catchments within the Meuse basin and three hydrological response units, derived from landscape characteristics. With the root-zone storage capacity as time-variant parameter, we are able to parsimoniously predict historical hydrological behavior and potential future behavior under change.

7.1.2 HYDROLOGICAL DATA

To identify, quantify and evaluate hydrological processes, we rely on expert knowledge and the availability of data. For the Meuse basin, the relatively large availability of in-situ observations of precipitation, temperature and streamflow forms the base of the studies in this thesis. In addition, the use of remote-sensing data provides independent information for hypotheses testing on hydrological functioning.

The increased correspondence between modeled actual evaporation and remote sensing estimates, when including intercatchment groundwater flow processes, provides additional evidence for the presence of this flux (Chapter 3). From the combined use of expert knowledge and remote-sensing data of evaporation, vegetation indices, soil moisture and total storage anomalies, we identified the drying-out of the root-zone storage each summer as unlikely and, therefore, as a model deficiency to be improved in a next step of the model comparison of Chapter 5. More generally, advances in model development benefit from testing the consistency between independent remote-sensing data and modeled behavior of hydrological state and flux variables.

However, uncertainty in remote-sensing data may be large and unquantified, and many aspects of the heterogeneous hydrological world remain unknown. The variables really measured through remote sensing often lack direct relevance for hydrological applications. As a result, a same variable, e.g. near-surface soil moisture, estimated indirectly by different sensors and processing techniques displays substantial dissimilarities (Chapter 4). As additional step, this near-surface soil moisture may require to be converted to root-zone soil moisture, through the Soil Water Index, to correspond to the variables of interest represented in hydrological models. To do so in a meaningful way, we have linked the time-scale parameter of the Soil Water Index to the root-zone storage capacity at the catchment scale (Chapter 4). This is an important step to adequately represent the spatial variability of hydrological processes in remote-sensing data.

7.2 SYNTHESIS ON RELIABLE HYDROLOGICAL PREDICTIONS

How to improve the reliability of hydrological predictions is a question with many answers. While there may be no wrong answers, surely there are different personal emphases or even contrasting views on the matter. For some, the key lies in improving calibration strategies (Gharari, 2016); or on the contrary, reducing the need for calibration (Nijzink, 2018). The focus may be to assimilate data or to improve the interpolation of forcing data (van Osnabrugge, 2020). Others choose to distribute model structures for a few landscape classes (de Boer-Euser, 2017; Gao, 2015), as opposed to hyper-resolution approaches with increasing model complexity (Bierkens et al., 2015; Imhoff et al., 2020). For some the crux is a Darwinian understanding of the system as opposed to a Newtonian approach (Harman and Troch, 2014; Savenije and Hrachowitz, 2017). For many, it will involve using (or performing) additional measurements (Winsemius, 2009; van Emmerik, 2017; Jiménez-Rodríguez, 2020). Answers likely combine several of these pathways with the common aim to increase our understanding of the variability of hydrological processes with respect to (changing) climate and catchment characteristics across

spatial and temporal scales. From the previously conducted studies, several steps are proposed as a way forward for improving the reliability of hydrological model predictions.

7.2.1 INTERCATCHMENT GROUNDWATER FLOWS IN THE LONG-TERM WATER BALANCE

Reliable streamflow estimates are often difficult and rely on indirect measurements. Yet, observations of catchment-scale evaporation and intercatchment groundwater flows are even harder or infeasible. Hence, it is unsurprising that the potential presence of intercatchment groundwater flows is often silently omitted from the long-term water balance. In a recent study by Berghuijs et al. (2020), an adapted representation of the Budyko framework is shown for 410 catchments of the US MOPEX dataset (Duan et al., 2006), where the aridty index (E_P/P) is plotted against evaporation (estimated as $E_A = P - Q$) and streamflow (Q), thereby omitting the common denominator P. Interestingly, the aridity index is strongly correlated to streamflow and the evaporative index (E_A/P), but very weakly to evaporation. In my opinion, this is not only related to data uncertainty, but also likely an indication of the potential presence of intercatchment groundwater flows.

In a proof-of-concept study in the Meuse, we show that neglecting intercatchment groundwater losses comes at the cost of overestimating actual evaporation to compensate for this lack (Chapter 3). The Meuse basin has the benefit of being relatively well-gauged and the subsurface is well-documented, which reduces the uncertainty. However, potential "inconsistencies" in the water balance should not blindly be attributed to data uncertainty or intercatchment groundwater flows in relatively ungauged catchments of the world. Instead, the potential presence of intercatchment groundwater flows can be detected through a careful evaluation of uncertainty in (remote sensing) precipitation, evaporation and streamflow within the Budyko framework, hypothesis testing through modeling and expert knowledge. If intercatchment groundwater flows are to be neglected in hydrological modeling studies, the decision should be conscious and clearly communicated.

7.2.2 ROOT-ZONE STORAGE CAPACITY FROM THE SEASONAL WATER BALANCE

Estimating the long-term water balance of a catchment is a sensible first step of any modeling study, as it provides insight in the long-term partitioning of precipitation into streamflow, evaporation and potential intercatchment groundwater flows. A next step is to evaluate the seasonal water balance, as it reflects important ecosystems properties. Vegetation tends to efficiently allocate its resources to balance above- and below-ground growth to simultaneously meet requirements for light, nutrients and water (Hrachowitz et al., 2020). In other words, vegetation efficiently adapts its root-zone storage capacity to supply canopy water demand. If seasonal water supply and demand are out of phase, a larger root-zone storage capacity is required to bridge dry spells than when supply and demand are in phase.

Hence, vegetation strongly regulates the water cycle by storing water when it is available and releasing it to the atmosphere through transpiration during dry periods. The root-zone storage capacity is, therefore, an important parameter of hydrological models, as it controls the long-term and seasonal hydrological partitioning. The water balance method to estimate the catchment-scale root-zone storage capacity is robust and relies on readily available information (Chapter 4 and 6). Moreover, it can be estimated as a time-variant parameter to accommodate the dynamic state of ecosystems as a result of natural and anthropogenic changes (Chapter 6). Thus, it is a key element to improve the reliability of hydrological predictions under change, which has been recognized as one of the major challenges in hydrological science (Blöschl et al., 2019).

7.2.3 SPECIFIC EVENTS AND THE SHORT-TERM WATER BALANCE

The third relevant time scale to consider is the event-based or short-term water balance. As for the other time scales, an expert judgement on dominant hydrological processes is developed by observing and evaluating catchment and climate characteristics, including topography, land use, slope, geology, soils, seasonality and aridity index. Streamflow signatures are also useful to quantify a specific aspect of hydrological functioning, as well as a visual inspection of the hydrograph. Furthermore, remote-sensing data provides additional insights on the magnitudes and dynamics of hydrological processes (Chapter 5). For example, accurate estimates of soil moisture contents in the unsaturated root-zone are required for reliable streamflow predictions, e.g. floods caused by heavy precipitation with high antecedent soil moisture. In Chapter 4, we show that we can generate meaningful estimates of the water contents in the root zone from globally available remote-sensing estimates of near-surface soil moisture, through the established link with catchment-scale root-zone storage capacities. Other relevant aspects of hydrological functioning related to dry periods and snow dynamics in a catchment can also be inferred from remote-sensing data of snow, total storage anomalies, soil moisture and evaporation (Chapter 5).

7.2.4 MODEL DEVELOPMENT AND CALIBRATION

Once the modeler has reached a perceptual understanding of overall catchment behavior from the evaluation of long-term, seasonal and event-based (remote-sensing) data, his/her theories can be translated to hydrological models. These models are simplified representations of the complex reality and heterogeneity of landscapes. The degree to which processes are simplified in models is highly variable and the possibilities are endless. In the resulting plethora of hydrological models, several conceptualization can give a similar aggregated response (i.e. streamflow) despite differences in internal process representations (Chapter 5). This great challenge of hydrology, referred to as equifinality, is unlikely to be solved soon. Even if more data becomes available, the uncertainties remain high and often unknown. If the use of several plausible hypotheses/models in hydrological studies is a way to explicitly address our lack of knowledge and understanding, it is not yet a cure.

Progress in computational resources has enabled us to run models at increasing spatial resolution. Fast run times are herein essential to efficiently allow thorough calibration and to adequately consider the uncertainty of catchment properties and their link to parameter values. Despite the attractiveness of developing calibration-free models, some degree of calibration seems inevitable to ensure a consistency with the available evaluation data and to account for uncertainty. Of course, we should also aim to develop methods to derive (time-dynamic) parameters from the available data, such as for the root-zone storage capacity, to somehow limit the degrees of freedom during calibration (Chapter 6). The calibration process should result in an ensemble of feasible parameters from a multi-objective selection approach.

Alternatively, the spatial variability of hydrological processes may be represented in a semi-distributed way, splitting a large basin in many smaller catchments. In each of these unit catchments, it is then possible to distinguish a small number of hydrological response units with different dominant hydrological processes, based on topography, land use, and/or (hydro)geology (Savenije, 2010; Fenicia et al., 2016; de Boer-Euser, 2017). On the one hand, this may increase the subjectivity of model development, but on the other hand, it promotes critical hydrological thinking and the development of tailor-made models (Savenije, 2009).

7.2.5 EVALUATION AND DATA UNCERTAINTY

The alternative models are then confronted against independent observations in a period outside the calibration period. If available, streamflow is typically used for model evaluation, using hydrological signatures and inspecting specific events of the hydrograph (Euser et al., 2013; de Boer-Euser et al., 2017). However, there is a growing interest to test the consistency between model behavior and remote-sensing data, which are available globally. These data can increase our understanding of processes that are not directly observable, such as the catchment-scale root-zone soil moisture or intercatchment groundwater flows (Chapter 3 and 4).

Although the uncertainty of remote-sensing data may be high and unknown, it can partly be estimated for a specific use when combined with expert knowledge (Chapter 5). Note that streamflow itself also relies on a rating-curve model with associated uncertainty (Westerberg and McMillan, 2015; Westerberg et al., 2016). Nevertheless, the combined use of uncertain streamflow, ground measurements and remote-sensing data is valuable to increase the reliability of model predictions, as knowing something with uncertainty is more valuable than knowing nothing with certainty.

The development of adequate models for a specific purpose is rarely straightforward and requires an iterative process. Insights gained from a first iteration lead to new questions and ideas for model structure improvements. The iterative process may never end, especially if the goal is exploratory science of hypotheses testing and hydrological process understanding. If the purpose is to develop a model to support decision making, a few iterations may be sufficient to achieve an a priori defined set of requirements for model performance.

7.3 ADVICE FOR PRACTITIONERS

The studies presented in this thesis result from a collaboration between Delft University of Technology (TU Delft), Deltares research institute and the Dutch Ministry of Infrastructure and Water Management (Rijkswaterstaat) and contribute to improving hydrological modeling of the Meuse basin for operational and policy applications. If in theory, the choice of a modeling framework is considerable, e.g. SUMMA (Clark et al., 2015), SUPERFLEX (Fenicia et al., 2011; Dal Molin et al., 2020), wflow (de Boer-Euser, 2017; Schellekens et al., 2020), it is strongly narrowed in practice.

Having an open source code, which is able to communicate with other used software and, most importantly, which is well known by the people who develop, maintain and use the models are determining practicalities. Hence, the wflow framework is a logical step for operational and policy applications of Rijkswaterstaat, as it is developed and maintained by Deltares and because it hosts multiple model structures (e.g. wflow_FLEX-Topo and wflow_sbm). Both the wflow_sbm and wflow_FLEX-Topo models have been extensively developed by Deltares and TU Delft in the past years. An important recommendation of this thesis for decision-makers is to explicitly consider model structure uncertainty by selecting multiple models (Chapter 5). Additionally, the comparison of a process-based model (i.e. wflow_FLEX-Topo) with a physically-based model (i.e. wflow_sbm) can foster model improvements by exposing model deficiencies.

The achieved model performance is already high following a brief calibration (Chapter 6) and is expected to further increase in a recommended next iteration. Hence, it is questionable if data-assimilation of remotely-sensed soil moisture data can be of significant added value to improve operational predictions. López López et al. (2016) showed that data-assimilation of soil moisture data only strongly increases model performance when applied in global models forced with global data, as opposed to local models with local data. An improvement could possibly be achieved in the French part of the basin, due to the lower density of precipitation stations. However, it is perhaps more sensible to first study the potential benefit of including more precipitation stations, in an attempt not to further complexify the operational system. On the other hand, it is recommended to study the impact of data-assimilation of streamflow data to improve operational forecasts. In addition, a system which allows to easily retrieve archived operational data and run hindcasts with complete model outputs would facilitate the continuous improvements of hydrological predictions.

For policy applications addressing long-term hydrological projections, an important recommendation of this thesis is to consider the effect of ecosystem adaptation in response to climate change (Chapter 6). This implies to evaluate the impact of making the vegetation parameters (i.e. root-zone storage capacity) time-variant, in response to changes in magnitude and seasonality of hydro-climatic variables. The proposed approach in Chapter 6 also enables to consider changes in land-use management by exchanging space-for-time, whereby current hydro-climatic characteristics of a given location allow to estimate potential future hydrological behavior of another location.

7.4 CHALLENGES AHEAD

7.4.1 ECOSYSTEM ADAPTATION AND HUMAN INTERFERENCE

One grand challenge of hydrology is the prediction of future hydrological behavior under change (Blöschl et al., 2019; Beven, 2019b). Nature is a living whole, not a dead aggregate (Wulf, 2015), meaning that ecosystems are likely to adapt in response to changing environmental conditions (Savenije and Hrachowitz, 2017). Adapting the root-zone storage capacity in hydrological models is a sensible way to (partly) represent non-stationarity of hydrological systems.

A time-dynamic root-zone storage capacity has been implemented in studies describing sudden changes, i.e. deforestation (Nijzink et al., 2016a; Hrachowitz et al., 2020), and projected long-term changes in climatic conditions due to global warming (Chapter 6). However, there is more to learn from past gradual changes in the root-zone storage capacity over the last decades. The study of Merz et al. (2011) shows almost a doubling of this parameter over the last 40 years using data from over 250 Austrian catchments. This is intriguing and deserves additional analyses with the water balance method. Also for the Meuse basin, where long-time series are available, it is worthwhile to analyze the changes which have already occurred until now, to increase our understanding of potential future changes.

The Budyko framework has been very valuable to study intercatchment groundwater flows (Chapter 3) and to predict the long-term hydrological partitioning under climate and land-use change (Chapter 6). However, our limited deeper understanding of the framework hinders us from accurately predicting movements in the Budyko space in relation to natural and anthropogenic changes in vegetation and climate. In addition, we lack a quantitative understanding of the contrasting effects of increased CO_2 levels on both increased productivity through fertilization and water use efficiency. Furthermore, it is highly uncertain how present-day vegetation may adapt its overall behavior and water use to changes in water availability and environmental conditions. Finding a region with a current climate that resembles the future climate of a region elsewhere may be a way forward to predict future hydrological behavior. These fascinating topics require further research.

The central role of vegetation in regulating the water cycle has repetitively been emphasized in this thesis. However, in many parts of the world, humans are the architects that have reshaped nature to meet food, energy and water demands. The growing shift from diverse and natural vegetation to monocultures has had adverse effects on the water cycle and the resilience of ecosystems to natural threats (Levia et al., 2020). However, humans also have the ability to reverse the past, enhance biodiversity, improve agricultural and forestry practices and positively influence the hydrological cycle through vegetation. To optimally mitigate the adverse effects of climate change, future research is required on how to meaningfully implement these changes in (hydrological) models to reliably quantify their impact.

7.4.2 TO MEASURE, OR NOT TO KNOW?

In hydrological science, our knowledge is often limited by a lack of data at the required spatial and temporal resolution. Many aspects of the hydrological cycle remain unknown and are hardly observable. Flow paths in the heterogeneous subsurface are only visible in exceptional cases. The partitioning of terrestrial evaporation into interception, transpiration and soil evaporation is highly uncertain (Coenders-Gerrits et al., 2014). Yet, efficiently measuring these different fluxes at the catchment scale is so far infeasible, despite promising advances with accelerometers and distributed temperature sensing (van Emmerik et al., 2017; Schilperoort et al., 2018). While there is no doubt that our hydrological process understanding benefits from additional innovative measurements, it is less clear how these should be organized to be coherent and extrapolatable across temporal and spatial scales.

Besides innovations in measuring techniques to improve hydrological data and theories, there is also more to gain from what is currently available. In addition to the petabytes of remote-sensing data openly available, dozens of new studies are published every week, not only in the field of hydrology, but also in relevant related fields of e.g. ecology, agriculture, forestry, meteorology, geology and remote sensing. Innovative overarching theories could be within reach by integrating and linking the knowledge gained from these studies. The crux is to creatively filter and ingest this huge amount of information to advance science and make these progresses also within reach of engineers and practitioners. Undoubtedly, uncertainty is here to stay, but the mysteries of the natural world are also the beauty of hydrological science.

A

MODEL EQUATIONS AND PARAMETERS

The equations and parameter of the models used in Chapters 3, 4 and 6 are provided in this Appendix.

A.1 MODEL EQUATIONS

The symbols for fluxes, storages and parameters are given in Tables A.1 to A.3. The equations of the four alternative process-based flexible model structures used in Chapter 3 (Fig. 3.4) are shown in Tables A.4 and A.5. From these models, the preferential model structure is also used in Chapter 4 (Fig. 4.2).

In Chapter 6, a flexible model structure with three hydrological response units for wetland, hillslope and plateau is used instead (Fig. 6.3). Model equations are provided in Tables A.6 and A.7.

Table A.1: Definitions of the symbols used to denote the different stores in the models of Chapter 3, 4 and 6.

Storage (mm)	Definition
S_W	Snow storage
S_I	Interception storage
S_R	Root-zone storage
S_F	Fast runoff storage
S_S	Slow runoff storage

Table A.2: Definitions of the symbols used to denote the different fluxes in the flexible model structures of Chapter 3, 4 and 6.

Fluxes (mm hr^{-1})	Definition
P	Precipitation
P_R	Rainfall
P_S	Snowfall
P_M	Snow melt
E_P	Potential evaporation
E_W	Evaporation from snow storage
E_I	Evaporation from interception
E_R	Evaporation from the root-zone storage
P_E	Effective precipitation
R_R	Outflow from the root-zone storage
R_{RS}	Recharge to the slow storage
R_{RF}	Recharge to the fast storage
R_P	Percolation
R_C	Capillary rise
Q_F	Fast runoff
Q_S	Slow runoff
Q_{IGF}	Net intercatchment groundwater flows
Q_{River} (or Q)	Streamflow which ends up in the river

Table A.3: Model parameters used in Chapter 3 and 4. For wflow_FLEX-Topo see Table A.12.

Parameter	unit	Definition
I_{max}	mm	Maximum interception capacity
$S_{R,max}$	mm	Root-zone storage capacity
β	-	Shape parameter of storage capacity distribution
L_P	-	Reduction parameter for potential evaporation
K_F	h	Characteristic time scale of the fast recession
K_S	h	Characteristic time scale of the slow recession
T_F	h	Time lag
D	-	Fraction to slow storage
P_{max}	mm h^{-1}	Maximum percolation rate
α	-	Non-linear coefficient of the fast storage
C_{IGF}	mm h^{-1}	Constant net intercatchment groundwater flow (IGF_{net})
μ	mm h^{-1}	Threshold of the recharge above which IGF_{net} occurs
P_{erc}	-	Fraction of the recharge to IGF_{net}
K_{IGF}	h	Characteristic time scale of the IGF_{net}
K_{River}	h	Characteristic time scale of the river flow (overflow model)
$S_{S,max}$	mm	Maximum capacity of underground storage
D_{IGF}	-	Fraction to IGF_{net} storage

Table A.4: Water balance equations for the zero, constant, preferential and overflow intercatchment groundwater flows models (Chapter 3). The ✓ indicates for which model(s) the water balance equations apply. In Chapter 4, the preferential model is used.

Water balance equation	Zero	Constant	Preferential	Overflow
$\frac{dS_I}{dt} = P - E_I - P_E$	✓	✓	✓	✓
$\frac{dS_R}{dt} = P_E - E_R - R_P - R_{RS} - R_{RF}$	✓	✓	✓	
$\frac{dS_R}{dt} = P_E - E_R - R_{RS}$				✓
$\frac{dS_F}{dt} = R_{RF} - Q_F$	✓	✓	✓	
$\frac{dS_F}{dt} = R_{SF} - Q_{River}$				✓
$\frac{dS_S}{dt} = R_{RS} + R_P - Q_S$	✓		✓	
$\frac{dS_S}{dt} = R_{RS} + R_P - Q_S - Q_{IGF}$		✓		
$\frac{dS_S}{dt} = R_{RS} - R_{SF} - Q_{IGF}$				✓
$Q_{River} = Q_S + Q_F$	✓	✓	✓	
$Q_{Tot} = Q_{River} + Q_{IGF}$	✓	✓	✓	✓

A

Table A.5: Constitutive functions for the zero, constant, preferential and overflow intercatchment groundwater flows models (Chapter 3). The ✓ indicates for which model(s) the constitutive functions apply. The following values are fixed for the smoothing parameters $m_1 = m_2 = m_4 = 0.005$ and $m_3 = 0.05$ (σ value of the error function). In Chapter 4, the preferential model is used.

Constitutive functions	Zero	Constant	Preferential	Overflow
$\overline{S_I} = S_I / I_{max}$	✓	✓	✓	✓
$\overline{S_R} = S_R / S_{R,max}$	✓	✓	✓	✓
$\overline{S_S} = S_S / S_{S,max}$				✓
$E_I = E_P \cdot \frac{\overline{S_I} \cdot (1+m_1)}{\overline{S_I}+m_1}$	✓	✓	✓	✓
$P_E = P \cdot (1 - \frac{(1-\overline{S_I})(1+m_2)}{1-\overline{S_I}+m_2})$	✓	✓	✓	✓
$E_R = (E_P - E_I) \cdot \frac{\overline{S_R}}{L_P}$	✓	✓	✓	✓
$R_R = R_{RS} + R_{RF}$	✓	✓	✓	
$R_R = P \cdot \overline{S_R}^{\beta}$	✓	✓	✓	
$R_{RS} = R_R \cdot D$	✓	✓	✓	
$R_{RF} = R_R \cdot (1 - D)$	✓	✓	✓	
$R_P = P_{max} \cdot \overline{S_R}$	✓	✓	✓	
$Q_F = K_F^{-1} \cdot S_F^{\alpha}$	✓	✓	✓	
$Q_S = K_S^{-1} \cdot S_S$	✓	✓	✓	
$Q_{IGF} = C_{IGF}$		✓		
$Q_{IGF} = \text{erf}(R_{RS}, \mu, m_3) \cdot P_{erc} \cdot R_{RS}$			✓	
$Q_{IGF} = K_{IGF}^{-1} \cdot S_S$				✓
$R_{RS} = P \cdot \overline{S_R}^{\beta}$				✓
$R_{SF} = R_{RS} \cdot \frac{\overline{S_S} \cdot (1+m_4)}{\overline{S_S}+m_4}$				✓
$Q_{River} = K_{River}^{-1} \cdot S_F$				✓

Table A.6: Water balance equations for each class of the wflow_FLEX-Topo model used in Chapter 6. The three classes share a common groundwater storage S_S. $Q_{F,P}$, $Q_{F,H}$ and $Q_{F,W}$ denote the fast runoff of the classes plateau, hillslope and wetland, respectively.

Water balance equation	Plateau	Hillslope	Wetland
$\frac{dS_W}{dt} = P_S - E_W - P_M$	✓	✓	✓
$\frac{dS_I}{dt} = P_R - E_I - P_E$	✓	✓	✓
$\frac{dS_R}{dt} = P_E + P_M - E_R - R_{RS} - R_{RF} - R_P$	✓		
$\frac{dS_R}{dt} = P_E + P_M - E_R - R_{RS} - R_{RF}$		✓	
$\frac{dS_R}{dt} = P_E + P_M - E_R - R_{RS} - R_{RF} + R_C$			✓
$\frac{dS_F}{dt} = R_{RF} - Q_F$	✓	✓	✓
$\frac{dS_S}{dt} = R_{RS} + R_P - Q_S$	✓		
$\frac{dS_S}{dt} = R_{RS} - Q_S$		✓	
$\frac{dS_S}{dt} = -R_C - Q_S$			✓
$Q = Q_S + Q_{F,P} + Q_{F,H} + Q_{F,W}$			

A

Table A.7: Constitutive functions of the wflow_FLEX-Topo model used in Chapter 6. T denotes temperature. The groundwater storage is shared between all classes. Symbols for the parameters are detailed in Table A.12.

Constitutive functions	Plateau	Hillslope	Wetland
Snow			
$P_S = \begin{cases} P, & \text{if } T < T_T \\ 0, & \text{if } T \geq T_T \end{cases}$	✓	✓	✓
$E_W = \min(E_P, S_W/dt)$	✓	✓	✓
$P_M = \begin{cases} 0, & \text{if } T < T_T \\ \min(F_M \cdot (T - T_M), S_W/dt), & \text{if } T \geq T_T \end{cases}$	✓	✓	✓
Interception			
$\overline{S_I} = S_I / I_{max}$	✓	✓	✓
$P_R = \begin{cases} 0, & \text{if } T < T_T \\ P, & \text{if } T \geq T_T \end{cases}$	✓	✓	✓
$P_E = \max(0, (S_I - I_{max})/dt)$	✓	✓	✓
$E_I = \min(E_P - E_W, (S_I - I_{max})/dt)$	✓	✓	✓
Root-zone			
$\overline{S_R} = S_R / S_{R,max}$	✓	✓	✓
$R_R = R_{RS} + R_{RF}$	✓	✓	
$E_R = \min((E_P - E_I) \cdot \min(\overline{S_R}/L_P, 1), S_U/dt)$	✓	✓	✓
$R_R = (P_E + P_M) \cdot (1 - (1 - \overline{S_R})^\beta)$	✓	✓	✓
$R_P = R_{P,max} \cdot \overline{S_R}$	✓		
$R_C = R_{C,max} \cdot (1 - \overline{S_R})$			✓
Fast storage			
$R_{RF} = R_R \cdot (1 - D)$	✓	✓	
$R_{RF} = R_R$			✓
$Q_F = K_F^{-1} \cdot S_F^\alpha$	✓	✓	
$Q_F = K_F^{-1} \cdot S_F$			✓
Slow storage			
$R_{RS} = R_R \cdot D$	✓	✓	
$Q_S = K_S^{-1} \cdot S_S$	✓	✓	✓

A.2 PRIOR AND POSTERIOR PARAMETER DISTRIBUTIONS

Prior parameter distributions for the models used in Chapter 3 and 4 are shown in Table A.8. Posterior parameter ranges for the zero, constant and preferential models of Chapter 3 are given in Table A.9. For the Aroffe catchment, the posterior parameter distributions of the overflow model are shown in Table A.10. Posterior parameter ranges for the preferential model used in Chapter 4 are given in Table A.11. Prior and posterior parameter distributions for the wflow_FLEX-Topo flexible model with three hydrological response units used in Chapter 6 are given in Table A.12.

Table A.8: Model parameters, units and prior range (*MRC denotes the value determined with a master recession curve ± 10 days). The ✓ indicates for which model(s) the parameters apply (Chapter 3). In Chapter 4, the preferential model is used with the same prior distributions.

Parameter	unit	Range	Zero	Constant	Preferential	Overflow
I_{max}	mm	1 - 3	✓	✓	✓	✓
$S_{R,max}$	mm	50 - 350	✓	✓	✓	✓
β	-	1 - 5	✓	✓	✓	✓
L_P	-	0 - 1	✓	✓	✓	✓
K_F	h	2 - 960	✓	✓	✓	
K_S	h	MRC*	✓	✓	✓	
T_F	h	1 - 20	✓	✓	✓	
D	-	0 - 1	✓	✓	✓	
P_{max}	mm h^{-1}	0 - 0.05	✓	✓	✓	
α	-	1 - 2	✓	✓	✓	
C_{IGF}	mm h^{-1}	-0.01 - 0.02		✓		
μ	mm h^{-1}	0.005 - 0.9			✓	
P_{erc}	-	-0.5 - 1			✓	
K_{IGF}	h	5 - 600				✓
K_{River}	h	5 - 600				✓
$S_{S,max}$	mm	1 - 60				✓
D_{IGF}	-	0.5 - 1				✓

Table A.9: Posterior parameter range (5-95 percentiles) for the zero, constant and preferential models for a selection of parameters (Chapter 3)

Parameter	$S_{R,max}$	L_P	K_F	D	P_{max}	α	C_{IGF}	P_{erc}
Unit	mm	-	h	-	mm h^{-1}		mm h^{-1}	-
Prior	50 - 350	0 - 1	2 - 960	0 - 1	0 - 0.05	1 - 2	-0.01 - 0.02	-0.5 - 1
Sainte-Marie - Zero	224 - 324	0.1 - 0.5	108 - 908	0.02 - 0.16	0.007 - 0.020	1.5 - 2.0		
Sainte-Marie - Constant	220 - 309	0.5 - 0.9	83 - 747	0.00 - 0.15	0.016 - 0.038	1.3 - 2.0	0.004 - 0.016	
Sainte-Marie - Pref.	64 - 273	0.0 - 0.8	68 - 542	0.17 - 0.42	0.003 - 0.020	1.1 - 1.9		0.7 - 1.0
Straimont - Zero	162 - 226	0.1 - 0.5	316 - 904	0.07 - 0.18	0.001 - 0.013	1.2 - 1.6		
Straimont - Constant	102 - 292	0.0 - 0.9	234 - 934	0.06 - 0.20	0.001 - 0.020	1.1 - 1.5	-0.003 - 0.006	
Straimont - Pref.	152 - 291	0.1 - 0.7	258 - 886	0.03 - 0.16	0.001 - 0.010	1.1 - 1.5		-0.3 - 0.6
Tintigny - Zero	144 - 318	0.3 - 0.8	177 - 931	0.02 - 0.12	0.002 - 0.012	1.2 - 1.6		
Tintigny - Constant	125 - 248	0.3 - 0.8	151 - 896	0.02 - 0.13	0.002 - 0.028	1.2 - 1.7	-0.004 - 0.008	
Tintigny - Pref.	152 - 303	0.4 - 0.8	108 - 876	0.03 - 0.20	0.001 - 0.011	1.1 - 1.6		-0.4 - 0.9
Chiny - Zero	166 - 283	0.2 - 0.8	203 - 948	0.01 - 0.14	0.003 - 0.016	1.2 - 1.6		
Chiny - Constant	140 - 314	0.0 - 0.8	182 - 901	0.01 - 0.13	0.003 - 0.025	1.1 - 1.6	-0.002 - 0.008	
Chiny - Pref.	111 - 268	0.2 - 0.7	122 - 865	0.03 - 0.18	0.004 - 0.016	1.1 - 1.6		-0.4 - 0.9
Membre-Pont - Zero	114 - 232	0.1 - 0.6	307 - 884	0.01 - 0.16	0.004 - 0.022	1.2 - 1.5		
Membre-Pont - Constant	107 - 237	0.0 - 0.7	258 - 922	0.01 - 0.14	0.002 - 0.031	1.2 - 1.5	-0.005 - 0.003	
Membre-Pont - Pref.	129 - 260	0.1 - 0.7	195 - 919	0.02 - 0.15	0.004 - 0.020	1.1 - 1.5		-0.4 - 0.9
Huccorgne Pref.	146 - 316	0.3 - 0.8	86 - 837	0.07 - 0.52	0.006 - 0.015	1.3 - 1.8		-0.5 - 1.0
Yvoir Pref.	110 - 250	0.4 - 0.9	179 - 908	0.39 - 0.65	0.009 - 0.024	1.4 - 1.9		0.7 - 0.9
Sormonne Pref.	119 - 299	0.0 - 0.8	70 - 818	0.41 - 0.58	0.000 - 0.017	1.3 - 2.0		0.6 - 1.0
Crusnes Pref.	112 - 295	0.2 - 0.7	271 - 782	0.36 - 0.58	0.001 - 0.017	1.0 - 1.4		0.4 - 0.9

Table A.10: Posterior parameter range (5-95 percentiles) for the overflow model used in the Aroffe catchment at Vannes-le-Châtel (Chapter 3).

Parameter	I_{max}	$S_{R,max}$	β	L_P	K_{River}	K_{IGF}	$S_{S,max}$	d_{IGF}
Unit	mm	mm	-	-	h	h	mm	-
Prior	1 - 3	50 - 350	1 - 5	0 - 1	5 - 600	5 - 600	1 - 60	0.5 - 1
Aroffe - Overflow	1.0 - 2.9	88 - 210	2.0 - 4.9	0.2 - 0.8	87 - 226	212.7 - 594.1	30.9 - 52.0	0.75 - 0.83

Table A.11: Posterior parameter range (5-95 percentiles) for a selection of parameters for the model runs of Chapter 4, *MRC denotes the value determined with a master recession curve ± 10 days

Parameter	I_{max}	$S_{R,max}$	β	L_P	D	K_F	K_S	P_{max}	α	μ	P_{erc}
Unit	mm	mm	-	-	-	h	h	mm h^{-1}	-	mm h^{-1}	-
Prior	1 - 3	50 - 350	1 - 5	0 - 1	0 - 1	2 - 960	MRC*	0 - 0.05	1 - 2	0.005 - 0.9	-0.5 - 1
(1) Straimont	1.2-2.9	160-282	3.5-4.8	0.03-0.48	0.01-0.17	336.06-890.69	1272-1713	0.001-0.013	1.2-1.5	0.09-0.76	-0.04-0.81
(2) Daverdisse	1.3-2.9	82-267	2.2-4.5	0.03-0.56	0.11-0.37	217.00-919.64	1554-1995	0.001-0.017	1.1-1.7	0.02-0.41	0.36-0.91
(3) Jemelle	1.1-2.6	139-207	3.1-4.8	0.08-0.58	0.11-0.29	168.53-904.74	1651-2013	0.005-0.016	1.1-1.7	0.18-0.77	-0.43-0.81
(4) Mabompre	1.2-3.0	147-277	2.8-4.8	0.19-0.81	0.01-0.17	183.66-894.08	1550-1959	0.001-0.009	1.0-1.4	0.04-0.85	-0.32-0.74
(5) Ortho	1.1-2.9	119-200	2.6-5.0	0.22-0.64	0.04-0.23	192.87-936.27	1627-2054	0.002-0.013	1.0-1.5	0.04-0.82	-0.44-0.93
(6) Treignes	1.1-2.8	126-231	2.3-4.9	0.07-0.48	0.06-0.33	147.23-807.97	1614-2009	0.001-0.013	1.3-1.9	0.05-0.58	0.47-0.96
(7) Sainte-Marie	1.4-2.9	80-249	2.0-4.8	0.16-0.81	0.27-0.46	76.33-715.85	1901-2290	0.001-0.015	1.3-2.0	0.05-0.37	0.65-0.94
(8) Longlaville	1.1-2.9	144-268	2.3-4.6	0.37-0.87	0.05-0.26	86.18-485.06	1572-1949	0.009-0.025	1.6-2.0	0.16-0.82	-0.50-0.93
(9) Wiheries	1.1-2.5	121-191	2.9-4.9	0.08-0.70	0.02-0.20	224.57-813.54	1588-2045	0.003-0.015	1.5-2.0	0.28-0.84	-0.15-0.94
(10) Yvoir	1.2-2.8	106-288	1.9-4.5	0.16-0.93	0.50-0.70	190.35-859.00	2758-3187	0.008-0.022	1.5-2.0	0.06-0.47	0.57-0.97
(11) Warnant	1.1-2.7	140-251	2.1-4.7	0.25-0.64	0.29-0.64	188.22-859.56	2648-3070	0.010-0.030	1.6-2.0	0.04-0.70	0.22-0.88
(12) Hastiere	1.1-2.8	119-176	2.6-5.0	0.14-0.64	0.01-0.20	136.48-738.16	1544-1936	0.003-0.015	1.4-2.0	0.20-0.79	-0.22-0.66
(13) Soulosse	1.1-2.9	51-113	1.5-4.7	0.13-0.96	0.02-0.18	187.97-920.87	1326-1728	0.001-0.014	1.5-1.9	0.06-0.82	-0.36-0.90
(14) Circourt	1.3-2.8	53-83	1.9-4.8	0.31-0.85	0.01-0.14	65.95-753.43	1041-1398	0.000-0.007	1.2-1.9	0.02-0.85	-0.20-0.72
(15) Goncourt	1.2-2.8	54-91	2.4-4.9	0.08-0.68	0.02-0.17	42.03-891.56	952-1282	0.000-0.005	1.1-1.9	0.07-0.82	-0.17-0.82
(16) Huccorgne	1.1-2.9	131-314	1.5-4.0	0.29-0.83	0.17-0.61	155.33-889.04	1869-2268	0.002-0.010	1.2-1.9	0.04-0.30	0.54-0.97

Table A.12: Calibrated model parameters, units and prior range for the wflow_FLEX-Topo model used in Chapter 6 (*MRC denotes the value determined with a master recession curve ± 30 %).

Parameter	unit	Description	Prior range	Plateau	Hillslope	Wetland
T_T	°C	Threshold temp. snow and rain	0.7 - 1.9	0.7 - 1.7	0.7 - 1.7	0.7 - 1.7
T_M	°C	Threshold temp. snow melt	0.7 - 2.3	0.8 - 2.2	0.8 - 2.2	0.8 - 2.2
F_M	mm d^{-1} $°C^{-1}$	Degree day factor	2.0 - 5.0	2.3 - 5.0	2.3 - 5.0	2.3 - 5.0
I_{max}	mm	Max. interception capacity	0.5 - 4.0	0.5 - 3.0	0.9 - 4.0	0.5 - 3.0
β	-	Shape parameter	0.2 - 0.4	0.2 - 0.4	0.2 - 0.4	0.2 - 0.4
L_P	-	Evap. reduction coefficient	0.1 - 0.6	0.1 - 0.6	0.1 - 0.4	0.1 - 0.6
$R_{C,max,W}$	mm d^{-1}	Max. capillary rise	0.1 - 0.5			0.1 - 0.5
$R_{P,max,P}$	mm d^{-1}	Max. percolation	0.05 - 0.72	0.05 - 0.72		
α	-	Non-linear coefficient	1 - 1.8	1.0 - 1.8	1.0 - 1.4	
K_F	d	Fast recession time scale	10 - 100	10 - 100	10 - 100	10 - 100
D	-	Fraction to slow storage	0.04 - 1	0.05 - 1	0.05 - 1	
K_S	d	Slow recession time scale	MRC*			

B

SOIL WATER INDEX (SWI)

The SWI aims to represent time series of soil moisture in deeper layers from past near-surface soil moisture estimates. A two-layer water balance model is defined with the upper layer representing the near-surface soil moisture and a second layer representing a deeper reservoir with length L which is only in contact with the surface layer. The following equation is used to link the water content in the surface and deeper reservoir (Wagner et al., 1999).

$$L \cdot \frac{\mathrm{d}W(t)}{\mathrm{d}t} = C \cdot (W_s(t) - W(t)) \tag{B.1}$$

With W the moisture content in the lower reservoir ($m^3\ m^{-3}$ or percentage saturation), L the depth of the reservoir [m], W_s the surface soil moisture content ($m^3\ m^{-3}$ or percentage saturation) and C represents a pseudo-diffusivity coefficient dependent on soil moisture and is assumed constant [m d^{-1}]. The equation is solved by assuming a constant parameter T [d]:

$$T = \frac{L}{C} \tag{B.2}$$

With increasing soil depth or decreasing pseudo-diffusivity constant, the value of T increases.

The solution of the differential Equation B.1 is:

$$W(t) = \frac{1}{T} \cdot \int_{-\infty}^{t} W_s(\tau) \cdot \exp(-\frac{t-\tau}{T}) \mathrm{d}\tau \tag{B.3}$$

and:

$$T = \int_{-\infty}^{t} \exp(-\frac{t-\tau}{T}) \mathrm{d}\tau \tag{B.4}$$

where τ represents the time of the previous measurement.

As satellite-based estimates of near-surface soil moisture are available at irregular intervals, the continuous formulation of Equation B.3 is replaced by a discrete equation given in Equation B.5.

$$SWI(t_n) = \frac{\sum_i^n SSM(t_i) \cdot e^{-\frac{t_n - t_i}{T}}}{\sum_i^n e^{-\frac{t_n - t_i}{T}}} \tag{B.5}$$

Where SWI, the Soil Water Index at time t_n, has replaced the continuous W and SSM, the near-surface soil moisture estimate at time t_i, has replaced the continuous W_s (all in $m^3\ m^{-3}$ or percentage saturation).

Stroud (1999) and Albergel et al. (2008) reformulated this exponential filter to a recursive formulation to handle irregular data more easily than the original exponential filter:

$$SWI(t_n) = SWI(t_{n-1}) + K_n \cdot (SSM(t_n) - SWI(t_{n-1})) \tag{B.6}$$

The gain K_n [-] at time t_n is given by the following recursive formula:

$$K_n = \frac{K_{n-1}}{K_{n-1} + e^{-\frac{(t_n - t_{n-1})}{T}}} \tag{B.7}$$

With initial values $SWI(t_0) = SSM(t_0)$ and $K_0 = 1$, and where t_n and t_{n-1} are the observation times of the current and previous SSM observation in Julian days. The gain K_n ranges between [0,1]. When many observations during the characteristic time length T are available, the gain will be small, meaning that the prior value will only be changed slightly towards the new observation. On the other hand, when no data has been received in quite some time relative to T, the gain will be large implying that the new estimate of SWI will converge towards the value of the new observation.

C

WATER BALANCE METHOD TO ESTIMATE THE ROOT-ZONE STORAGE CAPACITY

There is increasing evidence that catchment-scale root-zone water storage capacities $S_{R,max}$ can be robustly and readily estimated following a simple water balance approach based on daily precipitation and potential evaporation in combination with long-term runoff ratios (Gao et al., 2014; Wang-Erlandsson et al., 2016; Nijzink et al., 2016a; de Boer-Euser et al., 2016). The underlying assumption implies that natural vegetation adapts its storage to overcome dry spells with certain return periods (Milly, 1994; Kleidon and Heimann, 1998; Donohue et al., 2012; Gentine et al., 2012).

The long-term water balance of a catchment used here (Equation C.1), with storage (S), long-term mean annual precipitation (P), interception evaporation (E_I), plant transpiration (E_R), runoff (Q_{River}) and intercatchment groundwater flow losses (Q_{IGF}) provided in mm yr^{-1}, accounts for deep groundwater losses as they were shown to be significant in several catchments of the Meuse (Bouaziz et al., 2018). The long-term water balance considered here therefore reads as:

$$\frac{dS}{dt} \approx 0 \approx P - E_I - E_R - Q_{River} - Q_{IGF} \tag{C.1}$$

1. At first interception evaporation is estimated and used to quantify the effective precipitation P_E that reaches the soil ($P_E = P - E_I - \frac{dS_I}{dt}$). Here, E_I is a function of interception capacity (I_{max}; see Appendix A for the detailed equations). Due to the lack of more detailed information, we quantified the effect of different interception storage capacities, i.e. I_{max} = 0.5, 1.0, 2.0 and 3.0 mm, in a sensitivity analysis.

2. Long-term mean annual deep groundwater losses are expected in catchments where the difference between long-term mean annual precipitation and runoff approximates or exceeds the energy available for evaporation (Andréassian et al., 2012; Bouaziz et al., 2018). For each catchment, we check if the difference between mean annual precipitation and runoff exceeds potential evaporation. We attribute this gap in the water balance to deep groundwater losses as they were found to be significant in several catchments of the Meuse (Bouaziz et al., 2018). We estimate mean annual deep groundwater losses using the Turc-Mezentsev/Budyko (Turc, 1954; Mezentsev, 1955; Budyko, 1961) framework. We assume that actual evaporation ($E_A = P - Q_{River} = E_I + E_R$) should not deviate more than 5% from the long-term mean annual evaporation determined with the Turc-Mezentsev/Budyko (Turc, 1954; Mezentsev, 1955; Budyko, 1961) framework ($E_{A,TM}$).

$$Q_{IGF} = \begin{cases} E_A - (1.05 \cdot E_{A,TM}), & \text{if } E_A = P - Q_{River} > E_P. \\ \\ 0, & \text{if } E_A = P - Q_{River} < E_P. \end{cases} \tag{C.2}$$

3. We then determine mean annual transpiration E_R from long-term mean annual effective precipitation P_E, runoff Q_{River} and where applicable groundwater losses Q_{IGF} (Equation C.3).

$$E_R = \begin{cases} P_E - Q_{River} - Q_{IGF}, & \text{if } Q_{IGF} > 0. \\ \\ P_E - Q_{River}, & \text{if } Q_{IGF} = 0. \end{cases} \tag{C.3}$$

4. To maintain the seasonal signal, mean transpiration is then scaled to the daily signal of potential evaporation minus interception evaporation, by using the ratio of long-term mean annual transpiration over long-term mean annual potential evaporation minus long-term mean annual interception evaporation (Equation C.4).

$$E_R(t) = (E_P(t) - E_I(t)) \cdot \frac{E_R}{(E_P - E_I)} \tag{C.4}$$

5. As shown by Bouaziz et al. (2018), groundwater losses are expected to occur largely during the wet period when recharge exceeds a certain threshold. In the absence of clear seasonality in the annual precipitation signals, we accounted for this seasonality in groundwater losses by estimating monthly groundwater losses based on the inverse pattern of mean monthly potential evaporation. As mean monthly groundwater losses mostly occur in the wet period, they only slightly affect the maximum annual deficit.

6. An initial storage deficit of zero is assumed at the end of the wet period, in April, and the cumulative deficit is then calculated from daily effective precipitation, transpiration and groundwater losses assuming an "infinite-reservoir". The cumulative deficit increases during summer until it becomes zero again when all excess water is assumed to drain away through groundwater recharge or directly to the river as it cannot be stored. The annual maximum cumulative deficit ($S_{\mathrm{R,def}}$) between the time where the deficit equals zero (T_0) until the time where the total deficit returns to zero (T_1) is calculated following Equation C.5.

$$S_{\mathrm{R,def}} = \min \int_{T_0}^{T_1} P_{\mathrm{E}} - (E_{\mathrm{R}} - Q_{\mathrm{IGF}}) \mathrm{d}t \tag{C.5}$$

   The maximum storage deficit decreases as the maximum interception capacity increases as more water is intercepted and less goes to transpiration. However, the magnitude of the fluctuations due to interception is minor relative to the magnitude of the storage deficits.

7. The annual maxima of the storage deficit $S_{\mathrm{R,def}}$ are fitted to the extreme value distribution of Gumbel and, following Gao et al. (2014) and Nijzink et al. (2016a), the 20 years drought return period was used to estimate the water-balance-derived root-zone moisture capacity $S_{\mathrm{R,max}}$.

REFERENCES

Aalbers, E., van Meijgaard, E., Lenderink, G., de Vries H., and van den Hurk, B.: The 2018 European drought under future climate conditions, in prep., 2021.

Aalbers, E. E., Lenderink, G., van Meijgaard, E., and van den Hurk, B. J.: Local-scale changes in mean and heavy precipitation in Western Europe, climate change or internal variability?, Climate Dynamics, 50, 4745–4766, doi:10.1007/s00382-017-3901-9, URL http://dx.doi.org/10.1007/s00382-017-3901-9, 2018.

Addor, N. and Melsen, L. A.: Legacy, Rather Than Adequacy, Drives the Selection of Hydrological Models, Water Resources Research, 55, 378–390, doi:10.1029/2018WR022958, 2019.

Adnan, M., Merwade, V., and Yu, Z.: Multi-objective calibration of a hydrologic model using spatially distributed remotely sensed / in-situ soil moisture, Journal of Hydrology, 536, 192–207, doi:10.1016/j.jhydrol.2016.02.037, 2016.

Ajami, H., Troch, P. A., Maddock, T., Meixner, T., and Eastoe, C.: Quantifying mountain block recharge by means of catchment-scale storage-discharge relationships, Water Resources Research, 47, 1–14, doi:10.1029/2010WR009598, 2011.

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J. C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., and Martin, E.: From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations, Hydrology and Earth System Sciences, 12, 1323–1337, doi:10.5194/hess-12-1323-2008, 2008.

Albergel, C., de Rosnay, P., Balsamo, G., Isaksen, L., and Muñoz-Sabater, J.: Soil Moisture Analyses at ECMWF: Evaluation Using Global Ground-Based In Situ Observations, Journal of Hydrometeorology, 13, 1442–1460, doi:10.1175/jhm-d-11-0107.1, 2012.

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., and Siebert, S.: Development and testing of the WaterGAP 2 global model of water use and availability, Hydrological Sciences Journal, 48, 317–338, doi:10.1623/hysj.48.3.317.45290, 2003.

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D. D., Hogg, E. H., Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J. H., Allard, G., Running, S. W., Semerci, A., and Cobb, N.: A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests, Forest Ecology and Management, 259, 660–684, doi:10.1016/j.foreco.2009.09.001, 2010.

Ameli, A. A., Beven, K., Erlandsson, M., Creed, I. F., McDonnell, J. J., and Bishop, K.: Primary weathering rates, water transit times, and concentration-discharge relations: A theoretical analysis for the critical zone, Water resources research, 53, 942–960, 2017.

Ameli, A. A., Gabrielli, C., Morgenstern, U., and McDonnell, J. J.: Groundwater subsidy from headwaters to their parent water watershed: A combined field-modeling approach, Water Resources Research, 2018.

Andréassian, V., Parent, E., and Michel, C.: A distribution-free test to detect gradual changes in watershed behavior, Water Resources Research, 39, 1–11, doi:10.1029/2003WR002081, 2003.

Andréassian, V., Le Moine, N., Perrin, C., Ramos, M. H., Oudin, L., Mathevet, T., Lerat, J., and Berthet, L.: All that glitters is not gold: The case of calibrating hydrological models, Hydrological Processes, 26, 2206–2210, doi:10.1002/hyp.9264, 2012.

Archfield, S. A., Clark, M., Arheimer, B., Hay, L. E., McMillan, H., Kiang, J. E., Seibert, J., Hakala, K., Bock, A., Wagener, T., Farmer, W. H., Andréassian, V., Attinger, S., Viglione, A., Knight, R., Markstrom, S., and Over, T.: Accelerating advances in continental domain hydrologic modeling, Water Resources Research, 51, 10 078–10 091, doi:10.1002/2015WR017498, 2015.

Ashagrie, A. G., De Laat, P. J. M., De Wit, M. J. M., Tu, M., and Uhlenbrook, S.: Detecting the influence of land use changes on Floods in the Meuse River Basin? the predictive power of a ninety-year rainfall-runoff relation, Hydrology and Earth System Sciences Discussions, 3, 529–559, 2006.

Attema, J. J., Loriaux, J. M., and Lenderink, G.: Extreme precipitation response to climate perturbations in an atmospheric mesoscale model, Environmental Research Letters, 9, doi:10.1088/1748-9326/9/1/014003, 2014.

Ault, T. W., Czajkowski, K. P., Benko, T., Coss, J., Struble, J., Spongberg, A., Templin, M., and Gross, C.: Validation of the MODIS snow product and cloud mask using student and NWS cooperative station observations in the Lower Great Lakes Region, Remote Sensing of Environment, 105, 341–353, doi:10.1016/j.rse.2006.07.004, 2006.

Ballabio, C., Panagos, P., and Monatanarella, L.: Mapping topsoil physical properties at European scale using the LUCAS database, Geoderma, 261, 110–123, doi:10.1016/j.geoderma.2015.07.006, URL http://dx.doi.org/10.1016/j.geoderma.2015.07.006, 2016.

Balsamo, G., Viterbo, P., Beijaars, A., van den Hurk, B., Hirschi, M., Betts, A. K., and Scipal, K.: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the integrated forecast system, Journal of Hydrometeorology, 10, 623–643, doi:10.1175/2008JHM1068.1, 2009.

Banque Hydro: Banque Hydro, URL http://hydro.eaufrance.fr/, 2018.

Bastin, J. F., Clark, E., Elliott, T., Hart, S., Van Den Hoogen, J., Hordijk, I., Ma, H., Majumder, S., Manoli, G., Maschler, J., Mo, L., Routh, D., Yu, K., Zohner, C. M., and Crowther, T. W.: Correction: Understanding climate change from a global analysis of city analogues (PLoS ONE (2019) 14:7 (e0217592) DOI: 10.1371/journal.pone.0217592), PLoS ONE, 14, 1–13, doi:10.1371/journal.pone.0224120, 2019.

Bauer-Marschallinger, B.: Copernicus Global Land Operations "Vegetation and Energy" "CGLOPS-1" Validation Report Soil Water Index Collection 1km , URL https://land.copernicus.eu/global/sites/cgls.vito.be/files/products/CGLOPS1{_}QAR{_}SWI1km-V1{_}I1.11.pdf, 2020.

Bauer-Marschallinger, B., Paulik, C., Hochstöger, S., Mistelbauer, T., Modanesi, S., Ciabatta, L., Massari, C., Brocca, L., and Wagner, W.: Soil moisture from fusion of scatterometer and SAR: Closing the scale gap with temporal filtering, Remote Sensing, 10, 1–26, doi:10.3390/rs10071030, 2018.

Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., Modanesi, S., Massari, C., Ciabatta, L., Brocca, L., and Wagner, W.: Toward Global Soil Moisture Monitoring with Sentinel-1: Harnessing Assets and Overcoming Obstacles, IEEE Transactions on Geoscience and Remote Sensing, 57, 520–539, doi:10.1109/TGRS.2018.2858004, 2019.

Beck, H. E., de Jeu, R. A., Bruijnzeel, L. A., Schellekens, J., and van Dijk, A. I.: Improving Curve Number Based Storm Runoff Estimates Using Soil Moisture Proxies, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2, 250–259, doi:10.1109/JSTARS.2009.2031227, 2009.

Beck, H. E., Van Dijk, A. I., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and De Roo, A.: MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data, Hydrology and Earth System Sciences, 21, 589–615, doi:10.5194/hess-21-589-2017, 2017.

Bennett, K. E., Cherry, J. E., Balk, B., and Lindsey, S.: Using MODIS estimates of fractional snow cover area to improve streamflow forecasts in interior Alaska, Hydrology and Earth System Sciences, 23, 2439–2459, doi:10.5194/hess-23-2439-2019, 2019.

Berger, H.: Flow forecasting for the River Meuse, Ph.D. thesis, Delft University of Technology, The Netherlands, 1992.

Berghuijs, W. R., Woods, R. A., and Hrachowitz, M.: A precipitation shift from snow towards rain leads to a decrease in streamflow, Nature Climate Change, 4, 583, 2014.

Berghuijs, W. R., Larsen, J. R., van Emmerik, T. H., and Woods, R. A.: A Global Assessment of Runoff Sensitivity to Changes in Precipitation, Potential Evaporation, and Other Factors, Water Resources Research, 53, 8475–8486, doi:10.1002/2017WR021593, 2017.

Berghuijs, W. R., Gnann, S. J., and Woods, R. A.: Unanswered questions on the Budyko framework, Hydrological Processes, pp. 1–5, doi:10.1002/hyp.13958, 2020.

Bergström, S.: The HBV model - its structure and applications, vol. 4, Swedish Meteorological and Hydrological Institute, 1992.

Beven, K.: Changing ideas in hydrology - The case of physically-based models, Journal of Hydrology, 105, 157–172, doi:10.1016/0022-1694(89)90101-7, 1989.

Beven, K.: How far can we go in distributed hydrological modelling?, Hydrology and Earth System Sciences, 5, 1–12, doi:10.5194/hess-5-1-2001, 2001.

Beven, K.: A manifesto for the equifinality thesis, Journal of Hydrology, 320, 18–36, doi:10.1016/j.jhydrol.2005.07.007, 2006.

Beven, K.: Towards a methodology for testing models as hypotheses in the inexact sciences, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 475, doi:10.1098/rspa.2018.0862, 2019a.

Beven, K.: How to make advances in hydrological modelling, Hydrology Research, 50, 1481–1494, doi:10.2166/nh.2019.134, 2019b.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrological Processes, 6, 279–298, doi:10.1002/hyp.3360060305, 1992.

Beven, K., Cloke, H., Pappenberger, F., Lamb, R., and Hunter, N.: Hyperresolution information and hyperresolution ignorance in modelling the hydrology of the land surface, Science China Earth Sciences, 58, 25–35, doi:10.1007/s11430-014-5003-4, 2015.

Beven, K. J.: Uniqueness of place and process representations in hydrological modelling, Hydrology and Earth System Sciences, 4, 203–213, doi:10.5194/hess-4-203-2000, 2000.

Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, Hydrological Sciences Bulletin, 24, 43–69, doi:10.1080/02626667909491834, 1979.

Bierkens, M. F., Bell, V. A., Burek, P., Chaney, N., Condon, L. E., David, C. H., de Roo, A., Döll, P., Drost, N., Famiglietti, J. S., Flörke, M., Gochis, D. J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R. M., Reager, J. T., Samaniego, L., Sudicky, E., Sutanudjaja, E. H., van de Giesen, N., Winsemius, H., and Wood, E. F.: Hyper-resolution global hydrological modelling: What is next?: "Everywhere and locally relevant" M. F. P. Bierkens et al. Invited Commentary, Hydrological Processes, 29, 310–320, doi:10.1002/hyp.10391, 2015.

Blazkova, S., Beven, K. J., and Kulasova, A.: On constraining TOPMODEL hydrograph simulations using partial saturated area information, Hydrological Processes, 16, 441–458, doi:10.1002/hyp.331, 2002.

Blöschl, G. and Montanari, A.: Climate change impacts-throwing the dice?, Hydrological Processes, 24, 374–381, doi:10.1002/hyp.7574, URL http://jamsb.austms.org.au/courses/CSC2408/semester3/resources/ldp/abs-guide.pdf, 2010.

Blöschl, G. and Zehe, E.: On hydrological predictability, Hydrological Processes, 19, 3923–3929, doi:10.1002/hyp.6075, 2005.

Blöschl, G., Bierkens, M. F. P., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., Kirchner, J. W., McDonnell, J. J., Savenije, H. H. G., Sivapalan, M., Stumpp, C., Toth, E., Volpi, E., Carr, G., Lupton, C., Salinas, J., Széles, B., Viglione, A., Aksoy, H., Allen, S. T., Amin, A., Andréassian, V., Arheimer, B., Aryal, S. K., Baker, V., Bardsley, E., Barendrecht, M. H., Bartosova, A., Batelaan, O., Berghuijs, W. R., Beven, K., Blume, T., Bogaard, T., de Amorim, P. B., Böttcher, M. E., Boulet, G., Breinl, K., Brilly, M., Brocca, L., Buytaert, W., Castellarin, A., Castelletti, A., Chen, X., Chen, Y., Chen, Y., Chifflard, P., Claps, P., Clark, M. P., Collins, A. L., Croke, B., Dathe, A., David, P. C., de Barros, F. P. J., de Rooij, G., Baldassarre, G. D., Driscoll, J. M., Duethmann, D., Dwivedi, R., Eris, E., Farmer, W. H., Feiccabrino, J., Ferguson, G., Ferrari, E., Ferraris, S., Fersch, B., Finger, D., Foglia, L., Fowler, K., Gartsman, B., Gascoin, S., Gaume, E., Gelfan, A., Geris, J., Gharari, S., Gleeson, T., Glendell, M., Bevacqua, A. G., González-Dugo, M. P., Grimaldi, S., Gupta, A. B., Guse, B., Han, D., Hannah, D., Harpold, A., Haun, S., Heal, K., Helfricht, K., Herrnegger, M., Hipsey, M., Hlaváčiková, H., Hohmann, C., Holko, L., Hopkinson, C., Hrachowitz, M., Illangasekare, T. H., Inam, A., Innocente, C., Istanbulluoglu, E., Jarihani, B., Kalantari, Z., Kalvans, A., Khanal, S., Khatami, S., Kiesel, J., Kirkby, M., Knoben, W., Kochanek, K., Kohnová, S., Kolechkina, A., Krause, S., Kreamer, D., Kreibich, H., Kunstmann, H., Lange, H., Liberato, M. L. R., Lindquist, E., Link, T., Liu, J., Loucks, D. P., Luce, C., Mahé, G., Makarieva, O., Malard, J., Mashtayeva, S., Maskey, S., Mas-Pla, J., Mavrova-Guirguinova, M., Mazzoleni, M., Mernild, S., Misstear, B. D., Montanari, A., Müller-Thomy, H., Nabizadeh, A., Nardi, F., Neale, C., Nesterova, N., Nurtaev, B., Odongo, V. O., Panda, S., Pande, S., Pang, Z., Papacharalampous, G., Perrin, C., Pfister, L., Pimentel, R., Polo, M. J., Post, D., Sierra, C. P., Ramos, M.-H., Renner, M., Reynolds, J. E., Ridolfi, E., Rigon, R., Riva, M., Robertson, D. E., Rosso, R., Roy, T., Sá, J. H. M., Salvadori, G., Sandells, M., Schaefli, B., Schumann, A., Scolobig, A., Seibert, J., Servat, E., Shafiei, M., Sharma, A., Sidibe, M., Sidle, R. C., Skaugen, T., Smith, H., Spiessl, S. M., Stein, L., Steinsland, I., Strasser, U., Su, B., Szolgay, J., Tarboton, D., Tauro, F., Thirel, G., Tian, F., Tong, R., Tussupova, K., Tyralis, H., Uijlenhoet, R., van Beek, R., van der Ent, R. J., van der Ploeg, M., Loon, A. F. V., van Meerveld, I., van Nooijen, R., van Oel, P. R., Vidal, J.-P., von Freyberg, J., Vorogushyn, S., Wachniew, P., Wade, A. J., Ward, P., Westerberg, I. K., White, C., Wood, E. F., Woods, R., Xu, Z., Yilmaz, K. K., and Zhang, Y.: Twenty-three unsolved problems in hydrology (UPH) – a community perspective, Hydrological Sciences Journal, 64, 1141–1158, doi:10.1080/02626667.2019.1620507, URL https://doi.org/10.1080/02626667.2019.1620507, 2019.

Bogena, H., Herbst, M., Huisman, J., Rosenbaum, U., Weuthen, A., and Vereecken, H.: Potential of Wireless Sensor Networks for Measuring Soil Water Content Variability, Vadose Zone Journal, 9, 1002, doi:10.2136/vzj2009.0173, 2010.

Bonin, J. and Chambers, D.: Uncertainty estimates of a GRACE inversion modelling technique over greenland using a simulation, Geophysical Journal International, 194, 212–229, doi:10.1093/gji/ggt091, 2013.

Booij, M. J.: Impact of climate change on river flooding assessed with dif-

ferent spatial model resolutions, Journal of Hydrology, 303, 176–198, doi:10.1016/j.jhydrol.2004.07.013, 2005.

Bouaziz, L., Weerts, A., Schellekens, J., Sprokkereef, E., Stam, J., Savenije, H., and Hrachowitz, M.: Redressing the balance: Quantifying net intercatchment groundwater flows, Hydrology and Earth System Sciences, 22, 6415–6434, doi:10.5194/hess-22-6415-2018, 2018.

Bouaziz, L. J., Steele-Dunne, S. C., Schellekens, J., Weerts, A. H., Stam, J., Sprokkereef, E., Winsemius, H. H., Savenije, H. H., and Hrachowitz, M.: Improved understanding of the link between catchment-scale vegetation accessible storage and satellite-derived Soil Water Index, Water Resources Research, doi:10.1029/2019WR026365, 2020.

Bouaziz, L. J. E., Fenicia, F., Thirel, G., de Boer-Euser, T., Buitink, J., Brauer, C. C., De Niel, J., Dewals, B. J., Drogue, G., Grelier, B., Melsen, L. A., Moustakas, S., Nossent, J., Pereira, F., Sprokkereef, E., Stam, J., Weerts, A. H., Willems, P., Savenije, H. H. G., and Hrachowitz, M.: Behind the scenes of streamflow model performance, Hydrology and Earth System Sciences, 25, 1069–1095, doi:10.5194/hess-25-1069-2021, URL https://hess.copernicus.org/articles/25/1069/2021/, 2021.

Brauer, C. C., Teuling, A. J., F. Torfs, P. J., and Uijlenhoet, R.: The Wageningen Lowland Runoff Simulator (WALRUS): A lumped rainfall-runoff model for catchments with shallow groundwater, Geoscientific Model Development, 7, 2313–2332, doi:10.5194/gmd-7-2313-2014, 2014a.

Brauer, C. C., Torfs, P. J., Teuling, A. J., and Uijlenhoet, R.: The Wageningen Lowland Runoff Simulator (WALRUS): Application to the Hupsel Brook catchment and the Cabauw polder, Hydrology and Earth System Sciences, 18, 4007–4028, doi:10.5194/hess-18-4007-2014, 2014b.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., and Hasenauer, S.: ASCAT soil wetness index validation through in situ and modeled soil moisture data in central Italy, Remote Sensing of Environment, 114, 2745–2755, doi:10.1016/j.rse.2010.06.009, 2010a.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., and Hasenauer, S.: Improving runoff prediction through the assimilation of the ASCAT soil moisture product, Hydrology and Earth System Sciences, 14, 1881–1893, doi:10.5194/hess-14-1881-2010, 2010b.

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., Dorigo, W., Matgen, P., Martínez-Fernández, J., Llorens, P., Latron, J., Martin, C., and Bittelli, M.: Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe, Remote Sensing of Environment, 115, 3390–3408, doi:10.1016/j.rse.2011.08.003, URL http://dx.doi.org/10.1016/j.rse.2011.08.003, 2011.

Brocca, L., Crow, W. T., Ciabatta, L., Massari, C., De Rosnay, P., Enenkel, M., Hahn, S., Amarnath, G., Camici, S., Tarpanelli, A., and Wagner, W.: A Review of the Applications of ASCAT Soil Moisture Products, IEEE Journal of Selected Topics in Applied Earth

Observations and Remote Sensing, 10, 2285–2306, doi:10.1109/JSTARS.2017.2651140, 2017.

Brogli, R., Kröner, N., Sørland, S. L., Lüthi, D., and Schär, C.: The role of hadley circulation and lapse-rate changes for the future European summer climate, Journal of Climate, 32, 385–404, doi:10.1175/JCLI-D-18-0431.1, 2019.

Brown, A. E., Zhang, L., McMahon, T. A., Western, A. W., and Vertessy, R. A.: A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation, Journal of Hydrology, 310, 28–61, doi:10.1016/j.jhydrol.2004.12.010, 2005.

Brunner, M. I., Farinotti, D., Zekollari, H., Huss, M., and Zappa, M.: Future shifts in extreme flow regimes in Alpine regions, Hydrology and Earth System Sciences Discussions, pp. 1–26, doi:10.5194/hess-2019-144, 2019.

Budyko, M. I.: The heat balance of the earth's surface, Soviet Geography, 2, 3–13, 1961.

Buitink, J., Melsen, L. A., Kirchner, J. W., and Teuling, A. J.: A distributed simple dynamical systems approach (dS2 v1. 0) for computationally efficient hydrological modelling at high spatio-temporal resolution, Geoscientific Model Development, 13, 6093–6110, doi:https://doi.org/10.5194/gmd-13-6093-2020, 2020.

Burt, T. P., McDonnell, J. J.: Whither Field Hydrology?, Water Res. Research, 51, 5919–5928, doi:10.1002/2014WR016839, 2015.

Buytaert, W. and Beven, K.: Regionalization as a learning process, Water Resources Research, 45, 1–13, doi:10.1029/2008WR007359, 2009.

Calder, I. R., Reid, I., Nisbet, T. R., and Green, J. C.: Impact of lowland forests in England on water resources: Application of the Hydrological Land Use Change (HYLUC) model, Water Resources Research, 39, 1–10, doi:10.1029/2003WR002042, 2003.

Cateau, E., Larrieu, L., Vallauri, D., Savoie, J. M., Touroult, J., and Brustel, H.: Ancienneté et maturité: Deux qualités complémentaires d'un écosystème forestier, Comptes Rendus - Biologies, 338, 58–73, doi:10.1016/j.crvi.2014.10.004, 2015.

Ceballos, A., Scipal, K., Wagner, W., and Martínez-Fernández, J.: Validation of ERS scatterometer-derived soil moisture data in the central part of the Duero Basin, Spain, Hydrological Processes, 19, 1549–1566, doi:10.1002/hyp.5585, 2005.

Ceola, S., Arheimer, B., Baratti, E., Blöschl, G., Capell, R., Castellarin, A., Freer, J., Han, D., Hrachowitz, M., Hundecha, Y., Hutton, C., Lindström, G., Montanari, A., Nijzink, R., Parajka, J., Toth, E., Viglione, A., and Wagener, T.: Virtual laboratories: New opportunities for collaborative water science, Hydrology and Earth System Sciences, 19, 2101–2117, doi:10.5194/hess-19-2101-2015, 2015.

Chiew, F. H. and McMahon, T. A.: Estimating groundwater recharge using a surface watershed modelling approach, Journal of Hydrology, 114, 285–304, doi:10.1016/0022-1694(90)90062-3, 1990.

Cisneros Vaca, C., Van Der Tol, C., and Ghimire, C. P.: The influence of long-term changes in canopy structure on rainfall interception loss: A case study in Speulderbos, the Netherlands, Hydrology and Earth System Sciences, 22, 3701–3719, doi:10.5194/hess-22-3701-2018, 2018.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T., and Hay, L. E.: Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, Water Resources Research, 44, 1–14, doi:10.1029/2007wr006735, 2008.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. a., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D., Arnold, J. R., Gochis, D. J., and Rasmussen, R. M.: A unified approach for process-based hydrologicmodeling: 1. Modeling concept, Water Resources Research, 51, 1–17, doi:10.1002/2015WR017200.A, 2015.

Clark, M. P., Schaefli, B., Schymanski, S. J., Samaniego, L., Luce, C. H., Jackson, B. M., Freer, J. E., Arnold, J. R., Moore, R. D., Istanbulluoglu, E., and Ceola, S.: Improving the theoretical underpinnings of process-based hydrologic models, Water Resources Research, 52, 2350–2365, doi:10.1002/2015WR017910, 2016.

Coenders-Gerrits, A. M., Van Der Ent, R. J., Bogaard, T. A., Wang-Erlandsson, L., Hrachowitz, M., and Savenije, H. H.: Uncertainties in transpiration estimates, Nature, 506, 2013–2015, doi:10.1038/nature12925, 2014.

Copernicus Global Land Service: Soil Water Index, 2019.

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J., and Jones, P. D.: An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets, Journal of Geophysical Research: Atmospheres, 123, 9391–9409, doi:10.1029/2017JD028200, 2018.

Coron, L.: Les modèles hydrologiques conceptuels sont-ils robustes face à un climat en évolution ?, URL https://webgr.irstea.fr/wp-content/uploads/2012/11/these{_}Coron.pdf, 2013.

Coron, L., Andréassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M., and Hendrickx, F.: Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments, Water Resources Research, 48, 1–17, doi:10.1029/2011WR011721, 2012.

Coron, L., Thirel, G., Delaigue, O., Perrin, C., and Andréassian, V.: The suite of lumped GR hydrological models in an R package, Environmental Modelling and Software, 94, 166–171, doi:10.1016/j.envsoft.2017.05.002, 2017.

Coron, L., Perrin, C., Delaigue, O., Thirel, G., and Michel, C.: airGR: Suite of GR Hydrological Models for Precipitation-Runoff Modelling, R package version 1.0.10.11, doi:10.15454/EX11NA, URL https://cran.r-project.org/package=airGR., 2019.

Crow, W. T., Bindlish, R., and Jackson, T. J.: The added value of spaceborne passive microwave soil moisture retrievals for forecasting rainfall-runoff partitioning, Geophysical Research Letters, 32, 1–5, doi:10.1029/2005GL023543, 2005.

Dal Molin, M., Kavetski, D., and Fenicia, F.: SuperflexPy 1.2.0: an open source Python framework for building, testing and improving conceptual hydrological models, Geoscientific Model Development Discussions, 2020, 1–39, doi:10.5194/gmd-2020-409, URL https://gmd.copernicus.org/preprints/gmd-2020-409/, 2020.

Dal Molin, M., Schirmer, M., Zappa, M., and Fenicia, F.: Understanding dominant controls on streamflow spatial variability to set up a semi-distributed hydrological model: the case study of the Thur catchment, Hydrology and Earth System Sciences, 24, 1319–1345, doi:10.5194/hess-24-1319-2020, URL https://www.hydrol-earth-syst-sci.net/24/1319/2020/, 2020.

de Boer-Euser, T.: Added value of distribution in rainfall-runoff models for the Meuse basin, Ph.D. thesis, Delft University of Technology, The Netherlands, 2017.

de Boer-Euser, T., McMillan, H. K., Hrachowitz, M., Winsemius, H. C., and Savenije, H. H.: Influence of soil and climate on root zone storage capacity, Water Resources Research, doi:10.1002/2015WR018115, 2016.

de Boer-Euser, T., Bouaziz, L., de Niel, J., Brauer, C., Dewals, B., Drogue, G., Fenicia, F., Grelier, B., Nossent, J., Pereira, F., Savenije, H., Thirel, G., and Willems, P.: Looking beyond general metrics for model comparison; Lessons from an international model intercomparison study, Hydrology and Earth System Sciences, 21, 423–440, doi:10.5194/hess-21-423-2017, 2017.

de Lange, R., Beck, R., van de Giesen, N., Friesen, J., de Wit, A., and Wagner, W.: Scatterometer-derived soil moisture calibrated for soil texture with a one-dimensional water-flow model, IEEE Transactions on Geoscience and Remote Sensing, 46, 4041–4049, doi:10.1109/TGRS.2008.2000796, 2008.

de Niel, J., van Uytven, E., and Willems, P.: Uncertainty Analysis of Climate Change Impact on River Flow Extremes Based on a Large Multi-Model Ensemble, Water Resources Management, 33, 4319–4333, doi:10.1007/s11269-019-02370-0, 2019.

de Wit, M.: Van regen tot {M}aas, chap. De ondergr, Uitgeverij Veen Magazines, 2008a.

de Wit, M.: Van regen tot {M}aas, chap. De ondergr, p. 54, Uitgeverij Veen Magazines, Diemen, 2008b.

de Wit, M. J., van den Hurk, B., Warmerdam, P. M., Torfs, P. J., Roulin, E., and Van Deursen, W. P.: Impact of climate change on low-flows in the river Meuse, Climatic Change, 82, 351–372, doi:10.1007/s10584-006-9195-2, 2007.

de Wit, M. J. M., Warmerdam, P., Torfs, P., Uijlenhoet, R., Roulin, E., Cheymol, A., Van Deursen, W., Van Walsum, P., Ververs, M., Kwadijk, J., and Others: Effect of climate change on the hydrology of the river Meuse, Wageningen UR, 2001.

Dembélé, M., Hrachowitz, M., Savenije, H. H., Mariéthoz, G., and Schaefli, B.: Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets, Water Resources Research, 56, 0–3, doi:10.1029/2019WR026085, 2020.

Demirel, M. C., Mai, J., Mendiguren, G., Koch, J., Samaniego, L., and Stisen, S.: Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model, Hydrology and Earth System Sciences, 22, 1299–1315, doi:10.5194/hess-22-1299-2018, 2018.

Didan, K.: MODIS/Terra Vegetation Indices 16-Day L3 Global 500m SIN Grid V006. [MOD13A1], doi:https://doi.org/10.5067/MODIS/MYD10A1.006., 2015a.

Didan, K.: MODIS/Aqua Vegetation Indices 16-Day L3 Global 500m SIN Grid V006. [MYD13A1], doi:https://doi.org/10.5067/MODIS/MYD10A1.006., 2015b.

Donohue, R. J., Roderick, M. L., and McVicar, T. R.: Roots, storms and soil pores: Incorporating key ecohydrological processes into Budyko's hydrological model, Journal of Hydrology, 436-437, 35–50, doi:10.1016/j.jhydrol.2012.02.033, URL http://dx.doi.org/10.1016/j.jhydrol.2012.02.033, 2012.

Dooge, J. C.: Searching for simplicity in hydrology, Surveys in Geophysics, 18, 511–534, doi:10.1023/A:1006557801884, 1997.

Duan, Q., Schaake, J., Andréassian, V., Franks, S., Goteti, G., Gupta, H. V., Gusev, Y. M., Habets, F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O. N., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T., and Wood, E. F.: Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, Journal of Hydrology, 320, 3–17, doi:10.1016/j.jhydrol.2005.07.031, 2006.

Duethmann, D., Blöschl, G., and Parajka, J.: Why does a conceptual hydrological model fail to correctly predict discharge changes in response to climate change?, Hydrology and Earth System Sciences, 24, 3493–3511, doi:10.5194/hess-24-3493-2020, 2020.

Eagleson, P. S. and Tellers, T. E.: Ecological optimality in water-limited natural soil-vegetation systems: 2. Tests and applications, Water Resources Research, 18, 341–354, doi:10.1029/WR018i002p00341, 1982.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions "should we apply bias correction to global and regional climate model data?", Hydrology and Earth System Sciences, 16, 3391–3404, doi:10.5194/hess-16-3391-2012, 2012.

Eilander, D., Verseveld, W. V., Yamazaki, D., Weerts, A., Winsemius, H. C., and Ward, P. J.: A hydrography upscaling method for scale parametrization of distributed hydrological models, Hydrol. Earth Syst. Sci. Discuss., pp. 1–34, 2020.

Entekhabi, D., Nakamura, H., and Njoku, E. G.: Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely

sensed observations, IEEE Transactions on Geoscience and Remote Sensing, 32, 438–448, doi:10.1109/36.295058, 1994.

Entekhabi, D., Das, N., Njoku, E. G., T., J. J., and J., S.: SMAP L3 Radar/Radiometer Global Daily 9 km EASE-Grid Soil Moisture, Version 3. [L3_SM_P_E], doi:10.5067/7KKNQ5UURM2W, 2016.

European Environment Agency: Corine land cover, 2000.

European Environment Agency: Corine Land Cover (CLC) 2018, Version 2020_20u1, URL https://land.copernicus.eu/pan-european/corine-land-cover/, 2018.

Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., and Savenije, H. H. G.: A framework to assess the realism of model structures using hydrological signatures, Hydrology and Earth System Sciences, 17, 1893–1912, 2013.

Euser, T., Hrachowitz, M., Winsemius, H. C., and Savenije, H. H. G.: The effect of forcing and landscape distribution on performance and consistency of model structures, Hydrol. Process., 29, 3727–3743, doi:DOI: 10.1002/hyp.10445, 2015.

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., and Otero-Casal, C.: Hydrologic regulation of plant rooting depth, Proceedings of the National Academy of Sciences, p. 201712381, doi:10.1073/pnas.1712381114, URL http://www.pnas.org/lookup/doi/10.1073/pnas.1712381114, 2017.

FAO: AQUASTAT website. Food and Agriculture Organization of the United Nations (FAO), 2016.

Feddes, R. A., Kowalik, P. J., and Zaradny, H.: Water uptake by plant roots, Simulation of field water use and crop yield, pp. 16–30, 1978.

Fenicia, F., Savenije, H. H., Matgen, P., and Pfister, L.: Is the groundwater reservoir linear? Learning from data in hydrological modelling, Hydrology and Earth System Sciences, 10, 139–150, doi:10.5194/hess-10-139-2006, 2006.

Fenicia, F., McDonnell, J. J., and Savenije, H. H. G.: Learning from model improvement: On the contribution of complementary data to process understanding, Water Resources Research, 44, 1–13, doi:10.1029/2007WR006386, 2008.

Fenicia, F., Savenije, H. H., and Avdeeva, Y.: Anomaly in the rainfall-runoff behaviour of the Meuse catchment. Climate, land-use, or land-use management?, Hydrology and Earth System Sciences, 13, 1727–1737, doi:10.5194/hess-13-1727-2009, 2009.

Fenicia, F., Kavetski, D., and Savenije, H. H. G.: Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development, Water Resources Research, 47, 1–13, doi:10.1029/2010WR010174, 2011.

Fenicia, F., Kavetski, D., Savenije, H. H., Clark, M. P., Schoups, G., Pfister, L., and Freer, J.: Catchment properties, function, and conceptual model representation: Is there a correspondence?, Hydrological Processes, 28, 2451–2467, doi:10.1002/hyp.9726, 2014a.

Fenicia, F., Kavetski, D., Savenije, H. H., Clark, M. P., Schoups, G., Pfister, L., and Freer, J.: Catchment properties, function, and conceptual model representation: Is there a correspondence?, Hydrological Processes, 28, 2451–2467, doi:10.1002/hyp.9726, 2014b.

Fenicia, F., Kavetski, D., Savenije, H. H., and Pfister, L.: From spatially variable streamflow to distributed hydrological models: Analysis of key modeling decisions, Water Resources Research, doi:10.1002/2015WR017398, 2016.

Ficchì, A., Perrin, C., and Andréassian, V.: Hydrological modelling at multiple sub-daily time steps: Model improvement via flux-matching, Journal of Hydrology, 575, 1308–1327, doi:10.1016/j.jhydrol.2019.05.084, 2019.

Fister, V.: Dynamique des écoulements dans les aquifères calcaires de bas plateaux : de l'identification à la quantification des types de circulation. Exemple des formations triasiques et jurassiques dans le Nord-Est de la France., Ph.D. thesis, Université de Lorraine, Lorraine, France, 2012.

Fitzpatrick, M. C. and Dunn, R. R.: Contemporary climatic analogs for 540 North American urban areas in the late 21st century, Nature Communications, 10, 1–7, doi:10.1038/s41467-019-08540-3, URL http://dx.doi.org/10.1038/s41467-019-08540-3, 2019.

Ford, T. W., Harris, E., and Quiring, S. M.: Estimating root zone soil moisture using near-surface observations from SMOS, Hydrology and Earth System Sciences, 18, 139–154, doi:10.5194/hess-18-139-2014, 2014.

Fovet, O., Ruiz, L., Hrachowitz, M., Faucheux, M., and Gascuel-Odoux, C.: Hydrological hysteresis and its value for assessing process consistency in catchment conceptual models, Hydrology and Earth System Sciences, 19, 105–123, doi:10.5194/hess-19-105-2015, 2015.

Franks, S. W., Gineste, P., Beven, K. J., and Merot, P.: On constraining the predictions of a distributed model: The incorporation of fuzzy estimates of saturated areas into the calibration process, Water Resources Research, 34, 787–797, doi:10.1029/97WR03041, 1998.

Frisbee, M. D., Phillips, F. M., Campbell, A. R., Liu, F., and Sanchez, S. A.: Streamflow generation in a large, alpine watershed in the southern Rocky Mountains of Colorado: Is streamflow generation simply the aggregation of hillslope runoff responses?, Water Resources Research, 47, 2011.

Frisbee, M. D., Phillips, F. M., Weissmann, G. S., Brooks, P. D., Wilson, J. L., Campbell, A. R., and Liu, F.: Unraveling the mysteries of the large watershed black box: Implications for the streamflow response to climate and landscape perturbations, Geophysical Research Letters, 39, 1–6, doi:10.1029/2011GL050416, 2012.

Frisbee, M. D., Tysor, E. H., Stewart-Maddox, N. S., Tsinnajinnie, L. M., Wilson, J. L., Granger, D. E., and Newman, B. D.: Is there a geomorphic expression of interbasin groundwater flow in watersheds? Interactions between interbasin groundwater flow,

springs, streams, and geomorphology, Geophysical Research Letters, 43, 1158–1165, doi:10.1002/2015GL067082, 2016.

Fu, B.: On the calculation of the evaporation from land surface [in Chinese], Scientia Atmospherica Sinica, 5, 23–31, 1981.

Gao, C., Booij, M. J., and Xu, Y.-P.: Assessment of extreme flows and uncertainty under climate change: Disentangling the uncertainty contribution of representative concentration pathways, global climate models and internal climate variability, Hydrology and Earth System Sciences, 24, 3251–3269, doi:10.5194/hess-24-3251-2020, 2020.

Gao, H.: Landscape-based hydrological modelling; understanding the influence of climate, topography, and vegetation on catchment hydrology, Ph.D. thesis, Delft University of Technology, 2015.

Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongsitanon, N., and Savenije, H. H. G.: Climate controls how ecosystems size the root zone storage capacity at catchment scale, Geophysical Research Letters, 41, 7916–7923, doi:10.1002/2014GL061668, 2014.

Gao, H., Ding, Y., Zhao, Q., Hrachowitz, M., and Savenije, H. H.: The importance of aspect for modelling the hydrological response in a glacier catchment in Central Asia, Hydrological Processes, 31, 2842–2859, doi:10.1002/hyp.11224, 2017.

Gao, J., Holden, J., and Kirkby, M.: A distributed TOPMODEL for modelling impacts of land-cover change on river flow in upland peatland catchments, Hydrological Processes, 29, 2867–2879, doi:10.1002/hyp.10408, 2015.

Gash, J. H., Wright, I. R., and Lloyd, C. R.: Comparative estimates of interception loss from three coniferous forests in Great Britain, Journal of Hydrology, 48, 89–105, doi:10.1016/0022-1694(80)90068-2, 1980.

Genereux, D. P. and Jordan, M.: Interbasin groundwater flow and groundwater interaction with surface water in a lowland rainforest, Costa Rica: A review, Journal of Hydrology, 320, 385–399, doi:10.1016/j.jhydrol.2005.07.023, 2006.

Genereux, D. P., Wood, S. J., and Pringle, C. M.: Chemical tracing of interbasin groundwater transfer in the lowland rainforest of Costa Rica, Journal of Hydrology, 258, 163–178, 2002.

Genereux, D. P., Jordan, M. T., and Carbonell, D.: A paired-watershed budget study to quantify interbasin groundwater flow in a lowland rain forest, Costa Rica, Water resources research, 41, 2005.

Gentine, P., D'Odorico, P., Lintner, B. R., Sivandran, G., and Salvucci, G.: Interdependence of climate, soil, and vegetation as constrained by the Budyko curve, Geophysical Research Letters, 39, 2–7, doi:10.1029/2012GL053492, 2012.

Gerrits, A. M., Savenije, H. H., Veling, E. J., and Pfister, L.: Analytical derivation of the Budyko curve based on rainfall characteristics and a simple evaporation model, Water Resources Research, 45, 1–15, doi:10.1029/2008WR007308, 2009.

Gevaert, A. I., Renzullo, L. J., Van Dijk, A. I., Van Der Woerd, H. J., Weerts, A. H., and De Jeu, R. A.: Joint assimilation of soil moisture retrieved from multiple passive microwave frequencies increases robustness of soil moisture state estimation, Hydrology and Earth System Sciences, 22, 4605–4619, doi:10.5194/hess-22-4605-2018, 2018.

Gharari, S.: On the role of model structure in hydrological modeling; understanding models, Ph.D. thesis, Delft University of Technology, 2016.

Gharari, S., Hrachowitz, M., Fenicia, F., and Savenije, H. H. G.: Hydrological landscape classification: Investigating the performance of HAND based landscape classifications in a central European meso-scale catchment, Hydrology and Earth System Sciences, 15, 3275–3291, doi:10.5194/hess-15-3275-2011, 2011.

Gharari, S., Hrachowitz, M., Fenicia, F., and Savenije, H. H.: An approach to identify time consistent model parameters: Sub-period calibration, Hydrology and Earth System Sciences, 17, 149–161, doi:10.5194/hess-17-149-2013, 2013.

Gharari, S., Hrachowitz, M., Fenicia, F., Gao, H., and Savenije, H. H. G.: Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration, Hydrology and Earth System Sciences, 18, 4839–4859, doi:10.5194/hess-18-4839-2014, 2014.

Gleeson, T. and Manning, A. H.: Regional groundwater flow in mountainous terrain: Three-dimensional simulations of topographic and hydrogeologic controls, Water Resources Research, 44, 1–16, doi:10.1029/2008WR006848, 2008.

Gleeson, T., Wang-Erlandsson, L., Porkka, M., Zipper, S. C., Jaramillo, F., Gerten, D., Fetzer, I., Cornell, S. E., Piemontese, L., Gordon, L. J., Rockström, J., Oki, T., Sivapalan, M., Wada, Y., Brauman, K. A., Flörke, M., Bierkens, M. F., Lehner, B., Keys, P., Kummu, M., Wagener, T., Dadson, S., Troy, T. J., Steffen, W., Falkenmark, M., and Famiglietti, J. S.: Illuminating water cycle modifications and Earth system resilience in the Anthropocene, Water Resources Research, 56, 1–24, doi:10.1029/2019WR024957, 2020.

Goswami, M. and O'Connor, K. M.: A "monster" that made the SMAR conceptual model "right for the wrong reasons", Hydrological Sciences Journal, 55, 913–927, doi:10.1080/02626667.2010.505170, 2010.

Goswami, M., O'Connor, K. M., and Bhattarai, K. P.: Development of regionalisation procedures using a multi-model approach for flow simulation in an ungauged catchment, Journal of Hydrology, 333, 517–531, doi:10.1016/j.jhydrol.2006.09.018, 2007.

Güntner, A., Uhlenbrook, S., Seibert, J., and Leibundgut, C.: Multi-criterial validation of TOPMODEL in a mountainous catchment, Hydrological Processes, 13, 1603–1620, doi:10.1002/(SICI)1099-1085(19990815)13:11<1603::AID-HYP830>3.0.CO;2-K, 1999.

Gupta, H. V., Wagener, T., and Liu, Y.: Reconciling theory with observations: Elements of a diagnostic approach to model evaluation, Hydrological Processes, 22, 3802–3813, doi:10.1002/hyp.6989, 2008.

Gupta, H. V., Clark, M. P., Vrugt, J. A., Abramowitz, G., and Ye, M.: Towards a comprehensive assessment of model structural adequacy, Water Resources Research, 48, 1–16, doi:10.1029/2011WR011044, 2012.

Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., and Andréassian, V.: Large-sample hydrology: a need to balance depth with breadth, Hydrol. Earth Syst. Sci., 18, 463–477, doi:10.5194/hess-18-463-2014, 2014.

Guswa, A. J.: The influence of climate on root depth: A carbon cost-benefit analysis, Water Resources Research, 44, 1–11, doi:10.1029/2007WR006384, 2008.

Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G. P., and Yeh, P.: Multimodel estimate of the global terrestrial water balance: Setup and first results, Journal of Hydrometeorology, 12, 869–884, doi:10.1175/2011JHM1324.1, 2011.

Hagemann, S., Chen, C., Clark, D. B., Folwell, S., Gosling, S. N., Haddeland, I., Hanasaki, N., Heinke, J., Ludwig, F., Voss, F., and Wiltshire, A. J.: Climate change impact on available water resources obtained using multiple global climate and hydrology models, Earth System Dynamics, 4, 129–144, doi:10.5194/esd-4-129-2013, 2013.

Hakala, K., Addor, N., Gobbe, T., Ruffieux, J., and Seibert, J.: Risks and opportunities for a Swiss hydropower company in a changing climate, Hydrology and Earth System Sciences Discussions, pp. 1–35, doi:10.5194/hess-2019-475, 2019.

Hall, D. K. and Riggs, G. A.: Accuracy assessment of the MODIS snow products, Hydrological Processes, 21, 1534–1547, doi:10.1002/hyp.6715, 2007.

Hall, D. K. and Riggs, G. A.: MODIS/Terra Snow Cover Daily L3 Global 500m SIN Grid, Version 6. [MOD10A1], doi:https://doi.org/10.5067/MODIS/MOD10A1.006, 2016a.

Hall, D. K. and Riggs, G. A.: MODIS/Aqua Snow Cover Daily L3 Global 500m SIN Grid, Version 6. [MYD10A1], doi:https://doi.org/10.5067/MODIS/MYD10A1.006., 2016b.

Hargreaves, G. H. and Samani, Z. A.: Reference Crop Evapotranspiration from Temperature, Applied Engineering in Agriculture, 1, 96–99, doi:10.13031/2013.26773, 1985.

Harman, C. and Troch, P. A.: What makes Darwinian hydrology "darwinian"? Asking a different kind of question about landscapes, Hydrology and Earth System Sciences, 18, 417–433, doi:10.5194/hess-18-417-2014, 2014.

Hartmann, A., Goldscheider, N., Wagener, T., Lange, J., and Weiler, M.: Karst water resources in a changing world: Review of hydrological modeling approaches, Reviews of Geophysics, 52, 218–242, 2014.

Haylock, M. R., Hofstra, N., Klein Tank, A. M., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006, Journal of Geophysical Research Atmospheres, 113, doi:10.1029/2008JD010201, 2008.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, doi:10.1002/qj.3803, 2020.

Holländer, H. M., Blume, T., Bormann, H., Buytaert, W., Chirico, G. B., Exbrayat, J. F., Gustafsson, D., Hölzel, H., Kraft, P., Stamm, C., Stoll, S., Blöschl, G., and Flühler, H.: Comparative predictions of discharge from an artificial catchment (Chicken Creek) using sparse data, Hydrology and Earth System Sciences, 13, 2069–2094, doi:10.5194/hess-13-2069-2009, 2009.

Hooghart, J. C. and Lablans, W. N.: Van Penman naar Makkink: een nieuwe berekeningswijze voor de klimatologische verdampingsgetallen, De Bilt, Royal Netherlands Meteorological Institute (KNMI), De Bilt, the Netherlands, 1988.

Hrachowitz, M. and Clark, M. P.: HESS Opinions: The complementary merits of competing modelling philosophies in hydrology, Hydrology and Earth System Sciences, 21, 3953–3973, doi:10.5194/hess-21-3953-2017, 2017.

Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H. H., and Gascuel-Odoux, C.: Process consistency in models: The importance of system signatures, expert knowledge, and process complexity, Water Resources Research, 50, 7445–7469, doi:10.1002/2014WR015484, 2014.

Hrachowitz, M., Stockinger, M., Coenders-Gerrits, M., van der Ent, R., Bogena, H., Lücke, A., and Stumpp, C.: Deforestation reduces the vegetation-accessible water storage in the unsaturated soil and affects catchment travel time distributions and young water fractions, Hydrology and Earth System Sciences, i, 1–43, doi:10.5194/hess-2020-293, URL https://doi.org/10.5194/hess-2020-293?utm{_}source=researcher{_}app{&}utm{_}medium=referral{&}utm{_}campaign=RESR{_}MRKT{_}Researcher{_}inbound, 2020.

Hulsman, P., Winsemius, H., Michailovsky, C., Savenije, H., and Hrachowitz, M.: Using altimetry observations combined with GRACE to select parameter sets of a hydrological model in data scarce regions, Hydrology and Earth System Sciences, pp. 1–35, doi:10.5194/hess-2019-346, 2019.

Hulsman, P., Hrachowitz, M., and Savenije, H. H.: Improving the representation of long-term storage variations with conceptual hydrological models in data-scarce regions, Water Resources Research, 57, e2020WR028 837, 2021.

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., Fisk, J. P., Hibbard, K., Houghton, R. A., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T., Klein Goldewijk, K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., van Vuuren, D. P., and Wang, Y. P.: Harmonization of land-use scenarios for the period 1500-2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, Climatic Change, 109, 117–161, doi:10.1007/s10584-011-0153-2, 2011.

Imhoff, R. O., van Verseveld, W. J., van Osnabrugge, B., and Weerts, A. H.: Scaling point-scale (pedo)transfer functions to seamless large-domain parameter estimates for high-resolution distributed hydrologic modeling: An example for the Rhine river, Water Resources Research, n/a, e2019WR026 807, doi:10.1029/2019WR026807, URL https://doi.org/10.1029/2019WR026807, 2020.

Institut National de l'Information Géographique et Forestière: La base de données Forêt version 2.0, URL https://inventaire-forestier.ign.fr/spip.php?rubrique227, 2019.

Jacquemin, F., Kervyn, T., Branquart, E., Delahaye, L., Dufrêne, M., and Claessens, H.: Les forêts anciennes en Wallonie. 1ère partie : concepts généraux, Forêt Wallone, 131, 2014.

Jakeman, A. J. and Hornberger, G. M.: How much complexity is warranted in a rainfall-runoff model?, Water Resources Research, 29, 2637–2649, doi:10.1029/93WR00877, 1993.

Jaramillo, F. and Destouni, G.: Developing water change spectra and distinguishing change drivers worldwide, Geophysical Research Letters, 41, 8377–8386, doi:10.1002/2014GL061848, 2014.

Jaramillo, F., Cory, N., Arheimer, B., Laudon, H., Van Der Velde, Y., Hasper, T. B., Teutschbein, C., and Uddling, J.: Dominant effect of increasing forest biomass on evapotranspiration: Interpretations of movement in Budyko space, Hydrology and Earth System Sciences, 22, 567–580, doi:10.5194/hess-22-567-2018, 2018.

Jiménez-Rodríguez, C. D.: Evaporation partitioning of forest stands; the role of forest structure, Ph.D. thesis, Delft University of Technology, 2020.

Keenan, T. F., Hollinger, D. Y., Bohrer, G., Dragoni, D., Munger, J. W., Schmid, H. P., and Richardson, A. D.: Increase in forest water-use efficiency as atmospheric carbon dioxide concentrations rise, Nature, 499, 324–327, doi:10.1038/nature12291, 2013.

Kervyn, T., Scohy, J.-P., Marchal, D., Collette, O., Hardy, B., Delahaye, L., Wibail, L., Jacquemin, F., Dufrêne, M., and Claessens, H.: La gestion patrimoniale des forêts anciennes de Wallonie (Belgique), Forêt Nature, 148, doi:10.4267/2042/67878, 2018.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resources Research, 42, 1–5, doi:10.1029/2005WR004362, 2006.

Kleidon, A.: Global datasets and rooting zone depth inferred from inverse methods, Journal of Climate, 17, 2714–2722, doi:10.1175/1520-0442(2004)017<2714:GDORZD>2.0.CO;2, 2004.

Kleidon, A. and Heimann, M.: A method of determining rooting depth from a terrestrial biosphere model and its impacts on the global water and carbon cycle, Glob. Chang. Biol., 4, 275–286, doi:10.1046/j.1365-2486.1998.00152.x, 1998.

Klemeš, V.: Operational testing of hydrological simulation models, Hydrological Sciences Journal, 31, 13–24, doi:10.1080/02626668609491024, 1986.

Klingen, S.: Twaalf boslessen, Klingen Bomen, Doorn, 2017.

Knoben, W. J. M., Freer, J. E., Peel, M. C., Fowler, K. J. A., and Woods, R. A.: A brief analysis of conceptual model structure uncertainty using 36 models and 559 catchments, Water Resources Research, doi:10.1029/2019WR025975, URL https://doi.org/10.1029/2019WR025975, 2020.

Koch, J., Cornelissen, T., Fang, Z., Bogena, H., Diekkrüger, B., Kollet, S., and Stisen, S.: Inter-comparison of three distributed hydrological models with respect to seasonal variability of soil moisture patterns at a small forested catchment, Journal of Hydrology, 533, 234–249, doi:10.1016/j.jhydrol.2015.12.002, 2016.

Koch, J., Demirel, M. C., and Stisen, S.: The SPAtial EFficiency metric (SPAEF): Multiple-component evaluation of spatial patterns for optimization of hydrological models, Geoscientific Model Development, 11, 1873–1886, doi:10.5194/gmd-11-1873-2018, 2018.

Kovats, R. S., Valentini, R., Bouwer, L. M., Georgopoulou, E., Jacob, D., Martin, E., Rounsevell, M., and Soussana, J. F.: Europe, Climate Change 2014: Impacts, Adaptation and Vulnerability: Part B: Regional Aspects: Working Group II Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 1267–1326, doi:10.1017/CBO9781107415386.003, 2014.

Kunnath-Poovakka, A., Ryu, D., Renzullo, L. J., and George, B.: The efficacy of calibrating hydrologic model using remotely sensed evapotranspiration and soil moisture for streamflow prediction, Journal of Hydrology, 535, 509–524, doi:10.1016/j.jhydrol.2016.02.018, 2016.

Laiolo, P., Gabellani, S., Campo, L., Silvestro, F., Delogu, F., Rudari, R., Pulvirenti, L., Boni, G., Fascetti, F., Pierdicca, N., Crapolicchio, R., Hasenauer, S., and Puca, S.: Impact of different satellite soil moisture products on the predictions of a continuous distributed hydrological model, International Journal of Applied Earth Observation and Geoinformation, 48, 131–145, doi:10.1016/j.jag.2015.06.002, URL http://dx.doi.org/10.1016/j.jag.2015.06.002, 2016.

Lamb, R., Beven, K., and Myrabø, S.: Use of spatially distributed water table observations to constrain uncertainty in a rainfall-runoff model, Advances in Water Resources, 22, 305–317, doi:10.1016/S0309-1708(98)00020-7, 1998.

Landerer, F. W. and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resources Research, 48, 1–11, doi:10.1029/2011WR011453, 2012.

Lang, C., Freyermuth, A., Gille, E., and François, D.: Le dispositif {PRESAGES} ({PREvisions} et {Simulations} pour l'{Annonce} et la {Gestion} des {Etiages} {Sévères}) : des outils pour évaluer et prévoir les étiages, Géocarrefour, 81, 15–24, doi:10.4000/geocarrefour.1715, 2006.

Latte, N., Lebourgeois, F., Kint, V., Drouet, T., and Claessens, H.: Le hêtre face au changement climatique: Le cas de la Belgique, Revue Forestière Française, 69, 205–218, doi:10.4267/2042/65336, 2017.

Le Moine, N.: Le bassin versant de surface vu par le souterrain : une voie d'amelioration des performances et du realisme des modeles pluie-debit ?, Ph.D. thesis, CEMAGREF, (UR HBAN, Antony), France, URL http://www.cemagref.fr/webgr/Download/Rapports{_}et{_}theses/2008-LE{_}MOINE-THESE.pdf, 2008.

Le Moine, N., Andréassian, V., Perrin, C., and Michel, C.: How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments, Water Resources Research, 43, 1–11, doi:10.1029/2006WR005608, 2007.

Le Moine, N., Andréassian, V., and Mathevet, T.: Confronting surface-and groundwater balances on the La Rochefoucauld-Touvre karstic system (Charente, France), Water Resources Research, 44, 2008.

Lebecherel, L., Andréassian, V., and Perrin, C.: On regionalizing the Turc-Mezentsev water balance formula, Water Resources Research, 49, 7508–7517, doi:10.1002/2013WR013575, 2013.

Lebourgeois, F. and Mérian, P.: La sensibilité au climat des arbres forestiers a-t-elle changé au cours du XXe siècle?, Revue Forestiere Francaise, 63, 17–32, doi:10.4267/2042/43091, 2011.

Leroux, D. J., Pellarin, T., Vischel, T., Cohard, J. M., Gascon, T., Gibon, F., Mialon, A., Galle, S., Peugeot, C., and Seguis, L.: Assimilation of SMOS soil moisture into a distributed hydrological model and impacts on the water cycle variables over the Ouémé catchment in Benin, Hydrology and Earth System Sciences, 20, 2827–2840, doi:10.5194/hess-20-2827-2016, 2016.

Levia, D. F., Creed, I. F., Hannah, D. M., Nanko, K., Boyer, E. W., Carlyle-moses, D. E., Giesen, N. V. D., Grasso, D., Guswa, A. J., Hudson, J. E., Hudson, S. A., Iida, S., Jackson, R. B., Katul, G. G., Kumagai, T., Llorens, P., Ribeiro, F. L., Pataki, D. E., Peters, C. A., Carretero, D. S., and Selker, J. S.: Homogenization of the terrestrial water cycle, 13, 656–658, doi:10.1038/s41561-020-0641-y, 2020.

Li, H.-Y., Sivapalan, M., Tian, F., and Harman, C.: Functional approach to exploring climatic and landscape controls of runoff generation: 1. Behavioral constraints on runoff volume, Water Resources Research, 50, 9300–9322, 2014.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed {HBV}-96 hydrological model, J Hydrol., 201, 272–288, doi:10.1016/S0022-1694(97)00041-3, 1997.

Livneh, B. and Lettenmaier, D. P.: Multi-criteria parameter estimation for the Unified Land Model, Hydrology and Earth System Sciences, 16, 3029–3048, doi:10.5194/hess-16-3029-2012, 2012.

Loizu, J., Massari, C., Álvarez-Mozos, J., Tarpanelli, A., Brocca, L., and Casalí, J.: On the assimilation set-up of ASCAT soil moisture data for improving streamflow catchment simulation, Advances in Water Resources, 111, 86–104, doi:10.1016/j.advwatres.2017.10.034, URL https://doi.org/10.1016/j.advwatres.2017.10.034, 2018.

López López, P., Wanders, N., Schellekens, J., Renzullo, L. J., Sutanudjaja, E. H., and Bierkens, M. F.: Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations, Hydrology and Earth System Sciences, 20, 3059–3076, doi:10.5194/hess-20-3059-2016, 2016.

López López, P., Sutanudjaja, E. H., Schellekens, J., Sterk, G., and Bierkens, M. F.: Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products, Hydrology and Earth System Sciences, 21, 3125–3144, doi:10.5194/hess-21-3125-2017, 2017.

Luo, Y., Yang, Y., Yang, D., and Zhang, S.: Quantifying the impact of vegetation changes on global terrestrial runoff using the Budyko framework, Journal of Hydrology, 590, doi:10.1016/j.jhydrol.2020.125389, 2020.

Magand, C., Ducharne, A., Le Moine, N., and Brigode, P.: Transférabilité des paramètres d'un modèle de surface continentale sous changement climatique dans le bassin versant de la Durance, France, Hydrological Sciences Journal, 60, 1408–1423, doi:10.1080/02626667.2014.993643, URL http://dx.doi.org/10.1080/02626667.2014.993643, 2015.

Mahmood, R. and Hubbard, K. G.: Relationship between soil moisture of near surface and multiple depths of the root zone under heterogeneous land uses and varying hydroclimatic conditions, Hydrological Processes, 21, 3449–3462, doi:10.1002/hyp.6578, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.6578, 2007.

Manfreda, S., Brocca, L., Moramarco, T., Melone, F., and Sheffield, J.: A physically based approach for the estimation of root-zone soil moisture from surface measurements, Hydrology and Earth System Sciences, 18, 1199–1212, doi:10.5194/hess-18-1199-2014, 2014.

Mao, D. and Cherkauer, K. A.: Impacts of land-use change on hydrologic responses in the Great Lakes region, Journal of Hydrology, 374, 71–82, doi:10.1016/j.jhydrol.2009.06.016, URL http://dx.doi.org/10.1016/j.jhydrol.2009.06.016, 2009.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM˜v3: satellite-based land evaporation and \hack{\newline}root-zone soil moisture, Geoscientific Model Development, 10, 1903–1925, doi:10.5194/gmd-10-1903-2017, URL https://www.geosci-model-dev.net/10/1903/2017/, 2017.

Martin, P. and Zany, D.: Géologie de la Lorraine. Les pertes de l'Aroffe, URL http://www4.ac-nancy-metz.fr/base-geol/fiche.php?dossier=167{&}p=3descrip, a.

Martin, P. and Zany, D.: Géologie de la Lorraine. Pertes de la Meuse, URL http://www4.ac-nancy-metz.fr/base-geol/fiche.php?dossier=157{&}p=3descrip, b.

Mathevet, T.: Which rainfall-runoff model at the hourly time-step? {Empirical} development and intercomparison of rainfall runoff model on a large sample of watersheds., Ph.D. thesis, ENGREF University, Paris, France, 2005.

McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., Lucieer, A., Houborg, R., Verhoest, N. E., Franz, T. E., Shi, J., Gao, H., and Wood, E. F.: The future of Earth observation in hydrology, Hydrology and Earth System Sciences, 21, 3879–3914, doi:10.5194/hess-21-3879-2017, 2017.

Melsen, L. A.: Putting hydrological modelling practice to the test, p. 170, 2017.

Melsen, L. A. and Guse, B.: Hydrological Drought Simulations: How Climate and Model Structure Control Parameter Sensitivity, Water Resources Research, 55, 10 527–10 547, doi:10.1029/2019WR025230, 2019.

Melsen, L. A. and Guse, B.: Climate change impacts model parameter sensitivity – implications for calibration strategy and model diagnostic evaluation, Hydrology and Earth System Sciences, 25, 1307–1332, doi:10.5194/hess-25-1307-2021, 2021.

Melsen, L. A., Addor, N., Mizukami, N., Newman, A. J., Torfs, P. J., Clark, M. P., Uijlenhoet, R., and Teuling, A. J.: Mapping (dis)agreement in hydrologic projections, Hydrology and Earth System Sciences, 22, 1775–1791, doi:10.5194/hess-22-1775-2018, 2018.

Melsen, L. A., Teuling, A. J., Torfs, P. J., Zappa, M., Mizukami, N., Mendoza, P. A., Clark, M. P., and Uijlenhoet, R.: Subjective modeling decisions can significantly impact the simulation of flood and drought events, Journal of Hydrology, 568, 1093–1104, doi:10.1016/j.jhydrol.2018.11.046, 2019.

Merz, R., Parajka, J., and Blöschl, G.: Time stability of catchment model parameters: Implications for climate impact analyses, Water Resources Research, 47, 1–17, doi:10.1029/2010WR009505, 2011.

Mezentsev, V.: Back to the computation of total evaporation, Meteorologia i Gidrologia, 5, 24–26, 1955.

Mianabadi, A., Coenders-Gerrits, M., Shirazi, P., Ghahraman, B., and Alizadeh, A.: A global Budyko model to partition evaporation into interception and transpiration, Hydrology and Earth System Sciences Discussions, pp. 1–32, doi:10.5194/hess-2018-638, 2019.

Milly, P. C.: Climate, interseasonal storage of soil water, and the annual water balance, Advances in Water Resources, 17, 19–24, doi:10.1016/0309-1708(94)90020-5, 1994.

Milly, P. C., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R. J.: Climate change: Stationarity is dead: Whither water management?, Science, 319, 573–574, doi:10.1126/science.1151915, 2008.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrology and Earth System Sciences, 15, 453–469, doi:10.5194/hess-15-453-2011, URL http://www.hydrol-earth-syst-sci.net/15/453/2011/, 2011.

Miralles, D. G., Jiménez, C., Jung, M., Michel, D., Ershadi, A., Mccabe, M. F., Hirschi, M., Martens, B., Dolman, A. J., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F., and Fernández-Prieto, D.: The WACMOS-ET project - Part 2: Evaluation of global terrestrial evaporation data sets, Hydrology and Earth System Sciences, 20, 823–842, doi:10.5194/hess-20-823-2016, 2016.

Mostbauer, K., Kaitna, R., Prenner, D., and Hrachowitz, M.: The temporally varying roles of rainfall, snowmelt and soil moisture for debris flow initiation in a snow-dominated system, Hydrology and Earth System Sciences, 22, 3493–3513, doi:10.5194/hess-22-3493-2018, 2018.

Mouelhi, S., Michel, C., Perrin, C., and Andréassian, V.: Stepwise development of a two-parameter monthly water balance model, Journal of Hydrology, 318, 200–214, doi:10.1016/j.jhydrol.2005.06.014, 2006.

Newman, B.: The lazy Meuse, Jenkins, 1949.

Nielsen, S. A. and Hansen, E.: Numerical simulation of the rainfall runoff process on a daily basis, Nord Hydrol, 4, 171–190, 1973.

Nijzink, R., Hutton, C., Pechlivanidis, I., Capell, R., Arheimer, B., Freer, J., Han, D., Wagener, T., McGuire, K., Savenije, H., and Hrachowitz, M.: The evolution of root-zone moisture capacities after deforestation: A step towards hydrological predictions under change?, Hydrology and Earth System Sciences, 20, 4775–4799, doi:10.5194/hess-20-4775-2016, 2016a.

Nijzink, R. C.: Selecting model formulations and parameterizations; Reducing the need for calibration using open data and landscape characteristics, Ph.D. thesis, Delft University of Technology, 2018.

Nijzink, R. C., Samaniego, L., Mai, J., Kumar, R., Thober, S., Zink, M., Schäfer, D., Savenije, H. H., and Hrachowitz, M.: The importance of topography-controlled sub-grid process heterogeneity and semi-quantitative prior constraints in distributed hydrological

models, Hydrology and Earth System Sciences, 20, 1151–1176, doi:10.5194/hess-20-1151-2016, 2016b.

Nijzink, R. C., Almeida, S., Pechlivanidis, I. G., Capell, R., Gustafssons, D., Arheimer, B., Parajka, J., Freer, J., Han, D., Wagener, T., van Nooijen, R. R., Savenije, H. H., and Hrachowitz, M.: Constraining Conceptual Hydrological Models With Multiple Information Sources, Water Resources Research, 54, 8332–8362, doi:10.1029/2017WR021895, 2018.

Orth, R., Staudinger, M., Seneviratne, S. I., Seibert, J., and Zappa, M.: Does model performance improve with complexity ? A case study with three hydrological models, Journal of Hydrology, 523, 147–159, doi:10.1016/j.jhydrol.2015.01.044, 2015.

Ostberg, S., Schaphoff, S., Lucht, W., and Gerten, D.: Three centuries of dual pressure from land use and climate change on the biosphere, Environmental Research Letters, 10, doi:10.1088/1748-9326/10/4/044011, 2015.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2 - Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling, Journal of Hydrology, 303, 290–306, doi:10.1016/j.jhydrol.2004.08.026, 2005.

Owe, M., de Jeu, R., and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, Journal of Geophysical Research: Earth Surface, 113, doi:10.1029/2007JF000769, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JF000769, 2008.

Parajka, J. and Blöschl, G.: Validation of MODIS snow cover images over Austria, Hydrology and Earth System Sciences, 10, 679–689, doi:10.5194/hess-10-679-2006, 2006.

Parajka, J., Naeimi, V., Blöschl, G., and Komma, J.: Matching ERS scatterometer based soil moisture patterns with simulations of a conceptual dual layer hydrologic model over Austria, Hydrology and Earth System Sciences, 13, 259–271, doi:10.5194/hess-13-259-2009, 2009.

Paulik, C., Dorigo, W., Wagner, W., and Kidd, R.: Validation of the ASCAT soil water index using in situ data from the International Soil moisture network, International Journal of Applied Earth Observation and Geoinformation, 30, 1–8, doi:10.1016/j.jag.2014.01.007, 2014.

Peel, M. C. and Blöschl, G.: Hydrological modelling in a changing world, Progress in Physical Geography, 35, 249–261, doi:10.1177/0309133311402550, 2011.

Perrin, C., Michel, C., and Andréassian, V.: Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, Journal of Hydrology, 242, 275–301, doi:10.1016/S0022-1694(00)00393-0, 2001.

Perrin, C., Michel, C., and Andréassian, V.: Improvement of a parsimonious model for streamflow simulation, Journal of Hydrology, 279, 275–289, doi:10.1016/S0022-1694(03)00225-7, 2003.

Pomeroy, J., Fang, X., and Ellis, C.: Sensitivity of snowmelt hydrology in Marmot Creek, Alberta, to forest cover disturbance, Hydrological Processes, 26, 1891–1904, doi:10.1002/hyp.9248, 2012.

Prein, A. F., Rasmussen, R. M., Ikeda, K., Liu, C., Clark, M. P., and Holland, G. J.: The future intensification of hourly precipitation extremes, Nature Climate Change, 7, 48–52, doi:10.1038/nclimate3168, 2017.

Priestley, C. H. B. and Taylor, R. J.: On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters, Monthly Weather Review, 100, 81–92, doi:10.1175/1520-0493(1972)100<0081:otaosh>2.3.co;2, 1972.

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., Fekete, B. M., Franssen, W., Gerten, D., Gosling, S. N., Hagemann, S., Hannah, D. M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y., and Wisser, D.: Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment, Proceedings of the National Academy of Sciences of the United States of America, 111, 3262–3267, doi:10.1073/pnas.1222473110, 2014.

Ragab, R.: Towards a continuous operational system to estimate the root-zone soil moisture from intermittent remotely sensed surface moisture, Journal of Hydrology, 173, 1–25, doi:10.1016/0022-1694(95)02749-F, 1995.

Rakovec, O., Kumar, R., Attinger, S., and Samaniego, L.: Improving the realism of hydrologic model functioning through multivariate parameter estimation, Water Resources Research, 52, 7779–7792, doi:10.1002/2016WR019430, 2016a.

Rakovec, O., Kumar, R., Mai, J., Cuntz, M., Thober, S., Zink, M., Attinger, S., Schãfer, D., Schrön, M., and Samaniego, L.: Multiscale and multivariate evaluation of water fluxes and states over european river Basins, Journal of Hydrometeorology, 17, 287–307, doi:10.1175/JHM-D-15-0054.1, 2016b.

Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A., and Gratzki, A.: A Central European precipitation climatology–Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS), Meteorologische Zeitschrift, 22, 235–256, 2013.

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., and Seo, D. J.: Overall distributed model intercomparison project results, in: Journal of Hydrology, vol. 298, pp. 27–60, doi:10.1016/j.jhydrol.2004.03.031, 2004.

Reggiani, P. and Rientjes, T. H.: Closing horizontal groundwater fluxes with pipe network analysis: An application of the REW approach to an aquifer, Environmental Modelling and Software, 25, 1702–1712, doi:10.1016/j.envsoft.2010.04.019, URL http://dx.doi.org/10.1016/j.envsoft.2010.04.019, 2010.

Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., Tomasella, J., and Waterloo, M. J.: HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia, Remote Sensing of Environment, 112, 3469–3481, doi:http://dx.doi.org/10.1016/j.rse.2008.03.018, URL http://www.sciencedirect.com/science/article/pii/S003442570800120X, 2008.

Reu, B., Zaehle, S., Bohn, K., Pavlick, R., Schmidtlein, S., Williams, J. W., and Kleidon, A.: Future no-analogue vegetation produced by no-analogue combinations of temperature and insolation, Global Ecology and Biogeography, 23, 156–167, doi:10.1111/geb.12110, 2014.

Riboust, P., Thirel, G., Moine, N. L., and Ribstein, P.: Revisiting a simple degree-day model for integrating satellite data: implementation of swe-sca hystereses, Journal of Hydrology and Hydromechanics, 67, 70–81, doi:10.2478/johh-2018-0004, 2019.

Rohat, G., Goyette, S., and Flacke, J.: Characterization of European cities' climate shift – an exploratory study based on climate analogues, International Journal of Climate Change Strategies and Management, 10, 428–452, doi:10.1108/IJCCSM-05-2017-0108, 2018.

Rottler, E., Bronstert, A., Bürger, G., and Rakovec, O.: Projected changes in Rhine River flood seasonality under global warming, Hydrol. Earth Syst. Sci. Discuss., pp. 1–25, 2020.

Royal Meteorological Institute Belgium: Klimaatatlas, gemiddeld aantal dagen met sneeuw available at: https://www.meteo.be/nl/klimaat/klimaatatlas/klimaatkaarten/sneeuw, last access: 2020-03-26, 2015.

Royal Netherlands Meteorological Institute: No Title, 2018.

Sabater, J. M., Jarlan, L., Calvet, J.-C., Bouyssel, F., and De Rosnay, P.: From Near-Surface to Root-Zone Soil Moisture Using Different Assimilation Techniques, Journal of Hydrometeorology, 8, 194–206, doi:10.1175/jhm571.1, 2007.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resources Research, 46, 1–25, doi:10.1029/2008WR007327, 2010.

Samaniego, L., Kumar, R., and Jackisch, C.: Predictions in a data-sparse region using a regionalized grid-based hydrologic model driven by remotely sensed data, Hydrology Research, 42, 338–355, 2011.

Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., Eisner, S., Müller Schmied, H., Sutanudjaja, E. H., Warrach-Sagi, K., and Attinger, S.: Toward seamless hydrologic predictions across scales, Hydrology and Earth System Sciences Discussions, pp. 1–36, doi:10.5194/hess-2017-89, URL https://www.hydrol-earth-syst-sci-discuss.net/hess-2017-89/, 2017.

Savenije, H. H.: HESS opinions: "The art of hydrology", Hydrology and Earth System Sciences, 13, 157–161, doi:10.5194/hess-13-157-2009, 2009.

Savenije, H. H.: HESS opinions "topography driven conceptual modelling (FLEX-Topo)", Hydrology and Earth System Sciences, 14, 2681–2692, doi:10.5194/hess-14-2681-2010, 2010.

Savenije, H. H. G. and Hrachowitz, M.: HESS Opinions Catchments as meta-organisms-a new blueprint for hydrological modelling, Hydrology and Earth System Sciences, 21, 1107, 2017.

Schaller, M. F. and Fan, Y.: River basins as groundwater exporters and importers: Implications for water cycle and climate modeling, Journal of Geophysical Research: Atmospheres, 114, 2009.

Schär, C., Frei, C., Lüthi, D., and Davies, H. C.: Surrogate climate-change scenarios for regional climate models, Geophysical Research Letters, 23, 669–672, doi:10.1029/96GL00265, 1996.

Schattan, P., Zappa, M., Lischke, H., Bernhard, L., Thürig, E., and Diekkrüger, B.: An approach for transient consideration of forest change in hydrological impact studies, IAHS-AISH Proceedings and Reports, 359, 311–319, 2013.

Schelhaas, M. J., Nabuurs, G. J., and Schuck, A.: Natural disturbances in the European forests in the 19th and 20th centuries, Global Change Biology, 9, 1620–1633, doi:10.1046/j.1365-2486.2003.00684.x, 2003.

Schellekens, J., Scatena, F. N., Bruijnzeel, L. A., and Wickel, A. J.: Modelling rainfall interception by a lowland tropical rain forest in northeastern Puerto Rico, Journal of Hydrology, 225, 168–184, doi:10.1016/S0022-1694(99)00157-2, 1999.

Schellekens, J., Verseveld, W. V., Visser, M., Winsemius, H. H., de Boer-Euser, T., Bouaziz, L. J., Thiange, C., de Vries, S., Boisgontier, H., Eilander, D., Tollenaar, D., Weerts, A. H., Baart, F., Hazenberg, P., Lutz, A., ten Velden, C., Jansen, M., and Benedict, I.: openstreams/wflow, URL `https://github.com/openstreams/wflow`, 2020.

Schenk, H. J. and Jackson, R. B.: Rooting depths, lateral root spreads and belowground/above-ground allometries of plants in water-limited ecosystems, Journal of Ecology, 90, 480–494, doi:10.1046/j.1365-2745.2002.00682.x, 2002.

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N. W., Clark, D. B., Dankers, R., Eisner, S., Fekete, B. M., Colón-González, F. J., Gosling, S. N., Kim, H., Liu, X., Masaki, Y., Portmann, F. T., Satoh, Y., Stacke, T., Tang, Q., Wada, Y., Wisser, D., Albrecht, T., Frieler, K., Piontek, F., Warszawski, L., and Kabat, P.: Multimodel assessment of water scarcity under climate change, Proceedings of the National Academy of Sciences of the United States of America, 111, 3245–3250, doi:10.1073/pnas.1222460110, 2014.

Schilperoort, B., Coenders-Gerrits, M., Luxemburg, W., Rodríguez, C. J., Cisneros Vaca, C., and Savenije, H.: Technical note: Using distributed temperature sensing for Bowen ratio evaporation measurements, Hydrology and Earth System Sciences, 22, 819–830, doi:10.5194/hess-22-819-2018, 2018.

Schymanski, S. J., Sivapalan, M., Roderick, M. L., Beringer, J., and Hutley, L. B.: An optimality-based model of the coupled soil moisture and root dynamics, Hydrology and Earth System Sciences, 12, 913–932, doi:10.5194/hess-12-913-2008, URL https://www.hydrol-earth-syst-sci.net/12/913/2008/, 2008.

Seibert, J. and van Meerveld, H. I.: Hydrological change modeling: Challenges and opportunities, Hydrological Processes, 30, 4966–4971, doi:10.1002/hyp.10999, 2016.

Seibert, J., Bishop, K. H., and Nyberg, L.: A test of TOPMODEL'a ability to predict spatially distributed groundwater levels, Hydrological Processes, 11, 1131–1144, doi:10.1002/(sici)1099-1085(199707)11:9<1131::aid-hyp549>3.3.co;2-r, 1997.

Service Public de Wallonie: Direction générale opérationnelle de la Mobilité et des Voies hydrauliques, Département des Etudes et de l'Appui à la Gestion, Direction de la Gestion hydrologique intégrée (Bld du Nord 8-5000 Namur, Belgium, 2018.

Sheffield, J., Wood, E. F., Pan, M., Beck, H., Coccia, G., Serrat-Capdevila, A., and Verbist, K.: Satellite Remote Sensing for Water Resources Management: Potential for Supporting Sustainable Development in Data-Poor Regions, Water Resources Research, 54, 9724–9758, doi:10.1029/2017WR022437, 2018.

Silvestro, F., Gabellani, S., Rudari, R., Delogu, F., Laiolo, P., and Boni, G.: Uncertainty reduction and parameter estimation of a distributed hydrological model with ground and remote-sensing data, Hydrology and Earth System Sciences, 19, 1727–1751, doi:10.5194/hess-19-1727-2015, 2015.

Singh, R., Wagener, T., Van Werkhoven, K., Mann, M. E., and Crane, R.: A trading-space-for-time approach to probabilistic continuous streamflow predictions in a changing climate-accounting for changing watershed behavior, Hydrology and Earth System Sciences, 15, 3591–3603, doi:10.5194/hess-15-3591-2011, 2011.

Smith, M. B., Koren, V., Reed, S., Zhang, Z., Zhang, Y., Moreda, F., Cui, Z., Mizukami, N., Anderson, E. A., and Cosgrove, B. A.: The distributed model intercomparison project – Phase 2 : Motivation and design of the Oklahoma experiments, Journal of Hydrology, 418-419, 3–16, doi:10.1016/j.jhydrol.2011.08.055, 2012a.

Smith, M. B., Koren, V., Zhang, Z., Zhang, Y., Reed, S. M., Cui, Z., Moreda, F., Cosgrove, B. A., Mizukami, N., Anderson, E. A., and Participants, D.: Results of the DMIP 2 Oklahoma experiments, Journal of Hydrology, 418-419, 17–48, doi:10.1016/j.jhydrol.2011.08.056, 2012b.

Speich, M. J., Lischke, H., and Zappa, M.: Testing an optimality-based model of rooting zone water storage capacity in temperate forests, Hydrology and Earth System Sciences, 22, 4097–4124, doi:10.5194/hess-22-4097-2018, 2018.

Speich, M. J., Zappa, M., Scherstjanoi, M., and Lischke, H.: FORests and HYdrology under Climate Change in Switzerland v1.0: A spatially distributed model combining hydrology and forest dynamics, Geoscientific Model Development, 13, 537–564, doi:10.5194/gmd-13-537-2020, 2020.

Stisen, S., Koch, J., Sonnenborg, T. O., Refsgaard, J. C., Bircher, S., Ringgaard, R., and Jensen, K. H.: Moving beyond run-off calibration—Multivariable optimization of a surface–subsurface–atmosphere model, Hydrological Processes, 32, 2654–2668, doi:10.1002/hyp.13177, 2018.

Stroud, P.: A recursive exponential filter for time-sensitive data, Rep. LAUR-99, pp. 1–8, URL http://public.lanl.gov/stroud/ExpFilter/ExpFilter995573.pdf, 1999.

Sutanudjaja, E. H., Van Beek, L. P., De Jong, S. M., Van Geer, F. C., and Bierkens, M. F.: Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data, Water Resources Research, 50, 687–705, doi:10.1002/2013WR013807, 2014.

Swenson, S. and Wahr, J.: Post-processing removal of correlated errors in GRACE data, Geophysical Research Letters, 33, 1–4, doi:10.1029/2005GL025285, 2006.

Swenson, S. C.: GRACE monthly land water mass grids NETCDF RELEASE 5.0. Ver. 5.0. PO.DAAC, CA, USA, doi:http://dx.doi.org/10.5067/TELND-NC005, 2012.

Teuling, A. J., De Badts, E. A., Jansen, F. A., Fuchs, R., Buitink, J., Van Dijke, A. J., and Sterling, S. M.: Climate change, reforestation/afforestation, and urbanization impacts on evapotranspiration and streamflow in Europe, Hydrology and Earth System Sciences, 23, 3631–3652, doi:10.5194/hess-23-3631-2019, 2019.

Thirel, G., Delaigue, O., and Ficchi, A.: Latest developments of the airGR rainfall-runoff modelling R-package: inclusion of an interception store in the hourly model, doi:https://doi.org/10.5194/egusphere-egu2020-15275, 2020.

Tian, S.: Water Resources Research, Journal of the American Water Resources Association, 5, 2–2, doi:10.1111/j.1752-1688.1969.tb04897.x, 2007.

Tóth, J.: A theoretical analysis of groundwater flow in small drainage basins, Journal of geophysical research, 68, 4795–4812, 1963.

Trigo, I. F., Dacamara, C. C., Viterbo, P., Roujean, J. L., Olesen, F., Barroso, C., Camacho-De-Coca, F., Carrer, D., Freitas, S. C., García-Haroj, J., Geiger, B., Gellens-Meulenberghs, F., Ghilain, N., Meliá, J., Pessanha, L., Siljamo, N., and Arboleda, A.: The satellite application facility for land surface analysis, International Journal of Remote Sensing, 32, 2725–2744, doi:10.1080/01431161003743199, 2011.

Troch, P. A., Carrillo, G., Sivapalan, M., Wagener, T., and Sawicz, K.: Climate-vegetation-soil interactions and long-term hydrologic partitioning: Signatures of catchment co-evolution, Hydrology and Earth System Sciences, 17, 2209–2217, doi:10.5194/hess-17-2209-2013, 2013.

Tu, M.: Assessment of the effects of climate variability and land use change on the hydrology of the Meuse River Basin, Ph.D. thesis, Vrije Universiteit Amsterdam, UNESCO-IHE, The Netherlands, 2006.

Turc, L.: Le Bilan d'eau des sols: relations entre les précipitations, l'évaporation et l'écoulement, Annales Agronomiques, 1954.

Ukkola, A. M., Prentice, I. C., Keenan, T. F., Van Dijk, A. I., Viney, N. R., Myneni, R. B., and Bi, J.: Reduced streamflow in water-stressed climates consistent with CO2 effects on vegetation, Nature Climate Change, 6, 75–78, doi:10.1038/nclimate2831, 2016.

Valéry, A., Andréassian, V., and Perrin, C.: 'As simple as possible but not simpler': What is useful in a temperature-based snow-accounting routine? Part 2 - Sensitivity analysis of the Cemaneige snow accounting routine on 380 catchments, Journal of Hydrology, 517, 1176–1187, doi:10.1016/j.jhydrol.2014.04.058, 2014.

van der Schalie, R., Kerr, Y., Wignerond, J. P., Rodriguez-Fernandez, N. J., Al-Yaari, A., and de Jeu, R. A. M.: {G}lobal {SMOS} soil moisture retrievals from the {L}and {P}arameter {R}etrieval {M}odel, {I}nternational {J}ournal of {A}pplied {E}arth {O}bservation and {G}eoinformation, 45, 125–134, doi:10.1016/j.jag.2015.08.005, URL http://www.documentation.ird.fr/hor/{%}7BPAR{%}7D00014174, 2016.

van Der Sleen, P., Groenendijk, P., Vlam, M., Anten, N. P., Boom, A., Bongers, F., Pons, T. L., Terburg, G., and Zuidema, P. A.: No growth stimulation of tropical trees by 150 years of CO2 fertilization but water-use efficiency increased, Nature Geoscience, 8, 24–28, doi:10.1038/ngeo2313, 2015.

van der Velde, Y., Vercauteren, N., Jaramillo, F., Dekker, S. C., Destouni, G., and Lyon, S. W.: Exploring hydroclimatic change disparity via the Budyko framework, Hydrological Processes, 28, 4110–4118, doi:10.1002/hyp.9949, 2014.

van Dijk, A. I. J. M.: Climate and terrain factors explaining streamflow response and recession in Australian catchments, Hydrology and Earth System Sciences, 14, 159–169, doi:10.5194/hess-14-159-2010, 2010.

van Emmerik, T., Mulder, G., Eilander, D., Piet, M., and Savenije, H.: Predicting the ungauged basin: Model validation and realism assessment, Frontiers in Earth Science, 3, 1–11, doi:10.3389/feart.2015.00062, 2015.

van Emmerik, T., Steele-Dunne, S., Hut, R., Gentine, P., Guerin, M., Oliveira, R. S., Wagner, J., Selker, J., and Van De Giesen, N.: Measuring tree properties and responses using low-cost accelerometers, Sensors (Switzerland), 17, 1–17, doi:10.3390/s17051098, 2017.

van Emmerik, T., Popp, A., Solcerova, A., Müller, H., and Hut, R.: Reporting negative results to stimulate experimental hydrology: discussion of "The role of experimental work in hydrological sciences–insights from a community survey", Hydrological Sciences Journal, 63, 1269–1272, doi:10.1080/02626667.2018.1493203, 2018.

van Emmerik, T. H.: Water stress detection using radar, Ph.D. thesis, Delft University of Technology, 2017.

van Meijgaard, E., Ulft, L. H. V., Bosveld, F. C., Lenderink, G., and Siebesma, a. P.: The KNMI regional atmospheric climate model RACMO version 2.1, Technical report; TR - 302, p. 43, 2008.

van Osnabrugge, B.: Interpolate, simulate, assimilate, operational aspects of improving hydrological forecasts in the Rhine basin, Ph.D. thesis, Wageningen University & Research, 2020.

van Osnabrugge, B., Weerts, A. H., and Uijlenhoet, R.: genRE: A Method to Extend Gridded Precipitation Climatology Data Sets in Near Real-Time for Hydrological Forecasting Purposes, Water Resources Research, 53, 9284–9303, doi:10.1002/2017WR021201, 2017.

Vaze, J., Post, D. A., Chiew, F. H., Perraud, J. M., Viney, N. R., and Teng, J.: Climate non-stationarity - Validity of calibrated rainfall-runoff models for use in climate change studies, Journal of Hydrology, 394, 447–457, doi:10.1016/j.jhydrol.2010.09.018, URL http://dx.doi.org/10.1016/j.jhydrol.2010.09.018, 2010.

Velde, Y., Vercauteren, N., Jaramillo, F., Dekker, S. C., Destouni, G., and Lyon, S. W.: Exploring hydroclimatic change disparity via the Budyko framework, Hydrological Processes, 28, 4110–4118, 2014.

Veldkamp, T. I., Zhao, F., Ward, P. J., De Moel, H., Aerts, J. C., Schmied, H. M., Portmann, F. T., Masaki, Y., Pokhrel, Y., Liu, X., Satoh, Y., Gerten, D., Gosling, S. N., Zaherpour, J., and Wada, Y.: Human impact parameterizations in global hydrological models improve estimates of monthly discharges and hydrological extremes: A multi-model validation study, Environmental Research Letters, 13, doi:10.1088/1748-9326/aab96f, 2018.

Vidon, P. G.: Field hydrologists needed: A call for young hydrologists to (re)-focus on field studies, Hydrological Processes, 29, 5478–5480, doi:10.1002/hyp.10614, 2015.

Wagener, T.: Can we model the hydrological impacts of environmental change?, Hydrological Processes, 21, 3233–3236, doi:10.1002/hyp.6873, URL http://jamsb.austms.org.au/courses/CSC2408/semester3/resources/ldp/abs-guide.pdf, 2007.

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V., and Sorooshian, S.: A framework for development and application of hydrological models, Hydrology and Earth System Sciences, 5, 13–26, doi:10.5194/hess-5-13-2001, URL http://www.hydrol-earth-syst-sci.net/5/13/2001/, 2001.

Wagner, W., Lemoine, G., and Rott, H.: A method for estimating soil moisture from ERS Scatterometer and soil data, Remote Sensing of Environment, 70, 191–207, doi:10.1016/S0034-4257(99)00036-X, 1999.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., De Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Blöschl, G., Eitzinger,

J., Steinnocher, K., Zeil, P., and Rubel, F.: The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications, Meteorologische Zeitschrift, 22, 5–33, doi:10.1127/0941-2948/2013/0399, 2013.

Wanders, N. and Wada, Y.: Human and climate impacts on the 21st century hydrological drought, Journal of Hydrology, 526, 208–220, doi:10.1016/j.jhydrol.2014.10.047, URL http://dx.doi.org/10.1016/j.jhydrol.2014.10.047, 2015.

Wanders, N., Karssenberg, D., De Roo, A., De Jong, S. M., and Bierkens, M. F.: The suitability of remotely sensed soil moisture for improving operational flood forecasting, Hydrology and Earth System Sciences, 18, 2343–2357, doi:10.5194/hess-18-2343-2014, 2014.

Wang, R., Gentine, P., Yin, J., Chen, L., Chen, J., and Li, L.: Long-term relative decline in evapotranspiration with increasing runoff on fractional land surfaces, Hydrol. Earth Syst. Sci. Discuss., pp. 1–20, 2020.

Wang, T., Franz, T. E., You, J., Shulski, M. D., and Ray, C.: Evaluating controls of soil properties and climatic conditions on the use of an exponential filter for converting near surface to root zone soil moisture contents, Journal of Hydrology, 548, 683–696, doi:10.1016/j.jhydrol.2017.03.055, 2017.

Wang-Erlandsson, L.: Root for Rain, towards understanding land-use change impacts on the water cycle, Phd thesis, Delft University of Technology, the Netherlands, 2017.

Wang-Erlandsson, L., Bastiaanssen, W. G., Gao, H., Jägermeyr, J., Senay, G. B., Van Dijk, A. I., Guerschman, J. P., Keys, P. W., Gordon, L. J., and Savenije, H. H.: Global root zone storage capacity from satellite-based evaporation, Hydrology and Earth System Sciences, 20, 1459–1481, doi:10.5194/hess-20-1459-2016, 2016.

Welch, L. A. and Allen, D. M.: Consistency of groundwater flow patterns in mountainous topography: Implications for valley bottom water replenishment and for defining groundwater flow boundaries, Water Resources Research, 48, 1–17, doi:10.1029/2011WR010901, 2012.

Werth, S. and Güntner, A.: Calibration analysis for water storage variability of the global hydrological model WGHM, Hydrology and Earth System Sciences, 14, 59–78, doi:10.5194/hess-14-59-2010, 2010.

Westerberg, I. K. and McMillan, H. K.: Uncertainty in hydrological signatures, Hydrology and Earth System Sciences, 19, 3951–3968, doi:10.5194/hess-19-3951-2015, 2015.

Westerberg, I. K., Wagener, T., Coxon, G., McMillan, H. K., Castellarin, A., Montanari, A., and Freer, J.: Uncertainty in hydrological signatures for gauged and ungauged catchments, Water Resources Research, 52, 1847–1865, doi:10.1002/2015WR017635, 2016.

Willems, P.: Parsimonious rainfall-runoff model construction supported by time series processing and validation of hydrological extremes - Part 1: Step-wise model-structure identification and calibration approach, Journal of Hydrology, 510, 578–590, doi:10.1016/j.jhydrol.2014.01.017, 2014.

Willems, P., Mora, D., Vansteenkiste, T., Taye, M. T., and Van Steenbergen, N.: Parsimonious rainfall-runoff model construction supported by time series processing and validation of hydrological extremes - Part 2: Intercomparison of models and calibration approaches, Journal of Hydrology, 510, 591–609, doi:10.1016/j.jhydrol.2014.01.028, 2014.

Winsemius, H. C.: Satellite data as complementary information for hydrological modelling, Phd thesis, Delft University of Technology, the Netherlands, 2009.

Winsemius, H. C., Savenije, H. H., Gerrits, A. M., Zapreeva, E. A., and Klees, R.: Comparison of two model approaches in the Zambezi river basin with regard to model reliability and identifiability, Hydrology and Earth System Sciences, 10, 339–352, doi:10.5194/hess-10-339-2006, 2006.

Wulf, A.: The Invention of Nature: The Adventures of Alexander Von Humboldt, the Lost Hero of Science, Alfred A. Knopf, 2015.

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., and Pavelsky, T. M.: MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset, Water Resources Research, 55, 5053–5073, doi:10.1029/2019WR024873, 2019.

Yang, Y., Roderick, M. L., Zhang, S., McVicar, T. R., and Donohue, R. J.: Hydrologic implications of vegetation response to elevated CO2 in climate projections, Nature Climate Change, 9, 44–48, doi:10.1038/s41558-018-0361-0, URL http://dx.doi.org/10.1038/s41558-018-0361-0, 2019.

Yassin, F., Razavi, S., Wheater, H., Sapriza-Azuri, G., Davison, B., and Pietroniro, A.: Enhanced identification of a hydrologic model using streamflow and satellite water storage data: A multicriteria sensitivity analysis and optimization approach, Hydrological Processes, 31, 3320–3333, doi:10.1002/hyp.11267, 2017.

Zhang, B., Hautier, Y., Tan, X., You, C., Cadotte, M. W., Chu, C., Jiang, L., Sui, X., Ren, T., Han, X., and Chen, S.: Species responses to changing precipitation depend on trait plasticity rather than trait means and intraspecific variation, Functional Ecology, pp. 2622–2633, doi:10.1111/1365-2435.13675, 2020.

Zhang, G. P. and Savenije, H. H. G.: Rainfall-runoff modelling in a catchment with a complex groundwater flow system: application of the Representative Elementary Watershed (REW) approach, Hydrology and Earth System Sciences Discussions, 2, 639–690, doi:10.5194/hessd-2-639-2005, 2005.

Zhang, G. P., Fenicia, F., Rientjes, T. H. M., Reggiani, P., and Savenije, H. H. G.: Modeling runoff generation in the Geer river basin with improved model parameterizations to the REW approach, Physics and Chemistry of the Earth, Parts A/B/C, 30, 285–296, 2005.

Zhang, L., Hickel, K., Dawes, W. R., Chiew, F. H., Western, A. W., and Briggs, P. R.: A rational function approach for estimating mean annual evapotranspiration, Water Resources Research, 40, 1–14, doi:10.1029/2003WR002710, 2004.

Zhong, F., Martens, B., van Dijk, A., Ren, L., Jiang, S., and Miralles, D. G.: Global estimates of rainfall interception loss from satellite observations: recent advances in GLEAM, doi:https://doi.org/10.5194/egusphere-egu2020-13975, 2020.

ACKNOWLEDGEMENTS

This is the end of an incredible journey which has been enjoyable as I was lucky to be surrounded by a great team of supervisors, colleagues, friends and family. To all of you, thanks, dank, merci!

Markus Hrachowitz, I am very thankful for your supervision, you have been enthusiastic, critical, patient and I have learned a lot. You were always there when I needed guidance and you motivated me to always go one step further. Thanks for all your valuable hydrological and non-hydrological advice!

Huub Savenije, jouw unieke manier van college geven in het eerste jaar van de master is beslissend geweest voor mijn toekomst. Mijn fascinatie voor hydrologie is toen aangewakkerd en sindsdien nog veel gegroeid. Ik ben je ontzettend dankbaar voor de mooie kans die je me vier jaar geleden hebt gegeven en je enthousiaste begeleiding.

Albrecht Weerts, ontzettend bedankt voor je begeleiding en de prettige samenwerking. Ik vond je nuchtere blik heel waardevol en je hebt me doeltreffend advies gegeven over het opzetten van een onderzoek.

Jaap Schellekens, het is altijd leuk om met jou samen te werken en ik hoop dat zich weer kansen zullen aandienen in de toekomst. Dank voor al je hulp, goede ideeën en het delen van de bodemvocht data.

Susan Steele-Dunne, it has been really instructive and a great pleasure to work with you. Many thanks for your help and sharp humor.

Ik ben Rijkswaterstaat heel erg dankbaar voor de financiering van mijn onderzoek. Eric Sprokkereef en Jasper Stam, dank voor jullie enthousiasme, vertrouwen en de interessante discussies afgelopen 4 jaar. Ook wil ik graag Hendrik Buiteveld en Rita Lammersen bedanken voor de prettige samenwerking in het GRADE project.

Lucas Janssen, Jaap Kwadijk en Gerard Blom, dank voor het mogelijk maken van deze PhD en de ondersteuning vanuit Deltares. Lucas, dank ook voor je oplossingsgerichtheid, je interesse en betrokkenheid tijdens alle voortgangsbesprekingen. Jan Verkade en Daniel Twigt, dank voor jullie ondersteuning vanuit Strategisch Onderzoek.

I also thank the members of my doctoral committee for their time and positive assessment of my thesis, I am looking forward for fruitful discussions!

As data is key to research, I am grateful to Philippe Dierickx from the Service Public de Wallonie for kindly providing me precipitation and streamflow data. I would also like to thank Bernhard Bauer-Marschallinger for sharing the soil moisture data after his interesting presentation at EGU.

It was a great pleasure to spend time at the University again. Over the years, I have had the opportunity to share my office with great personalities. Tim van Emmerik and Anna Solcerová, thank you for the good times, Family Guy and the most delicious barbecues. Bart Schilperoort, dank dat je op alles een antwoord hebt! Jerom Aerts, dank voor de gezelligheid en fijn om met jou ideeën uit te wisselen over modelleren en satelliet data. Gaby Gruendemann, dank voor alle gezellige havermelk koffies, lunchwandelingen, rubber duckies, recepten en banana breads. Ik ben blij dat we nog steeds zo vaak en zo goed contact hebben, ook nu je in Canada bent. Rolf Hut, dank voor je inspirerende McGyver ideeën tijdens koffiepauzes en het organiseren van de FAIR workshop. Petra Hulsman, dank voor de gedeelde ervaringen en al je goede tips. Ook ben ik de wandelclub dankbaar, Wim Luxemburg, Lydia de Hoog, Tamara Auperlé, Betty Rothfusz, Erik Mostert, Paul Vermunt, Fares Al Hasan, Bas des Tombe, Miriam Coenders, Ruud van der Ent, Fransje van Oorschot en Ties van der Heijden. Thanks to Boran Ekin Aydin, César Dionisio Jimnez-Rodríguez, Juan Pablo Aguilar-López and Jerom Aerts for the Thursday beers. Ruud van der Ent, het was erg leuk en leerzaam om samen met jou de BSc eindwerken van Irene van der Veer, Charlotte van Strien en Joris Dekker te begeleiden, ook dank aan alle drie voor het goede werk! During my research, I was also inspired by the related work of my fellow Delft hydrological modelers Lan Wang-Erlandsson, Remko Nijzink, Shervan Gharari and Hongkai Gao. Martijn Wackers, Ozan Çelik, Nitish Anand, thank you for your feedback during the Writing a dissertation course. Marij Zwart, dank voor je inspirerende ervaringen! Chelsea Kaandorp, dank voor alle gezellige wandelingen in Delft in deze eenzame thuiswerk periode. Vanuit de universiteit van Wageningen wil ik graag ook Joost Buitink en Bart van Osnabrugge bedanken voor de samenwerking en gedachtewisselingen.

Vanuit Deltares was ik goed omringd met slimme en motiverende collega's, afdeling HYD bedankt. Hessel Winsemius en Mark Hegnauer, het is altijd inspirerend om met jullie samen te werken, dank voor jullie zijdelingse betrokkenheid bij mijn onderzoek. Frederiek Sperna Weiland, Ronald Vernimmen, Arjen Haag, Hans Gehrels, Patricia Trambauer, Anke Becker, dank voor de samenwerking in projecten voor mijn Deltares dag in de week. Willem van Verseveld, Martijn Visser en Dirk Eilander, dank voor de wflow, linux, github en hydromt hulp. Nienke Kramer, Corine ten Velden, Carine Wesselius, Eelco Verschelling, Margreet van Marle en Pieter Hazenberg dank voor de gezellige koffies en lunchwandelingen. Jan Verkade, dank voor de tony's en de wijn. Mark Hegnauer en Hélène Boisgontier thanks for the good times in Delft.

The International Meuse Symposium initiated my fascination for the Meuse basin, thank you Bernhard Becker and Benjamin Dewals for organizing this great event each year. To all the co-authors of the hydrological model comparison studies, Fabrizio Fenicia, Guillaume Thirel, Tanja de Boer-Euser, Lieke Melsen, Joost Buitink, Claudia Brauer, Jiri Nossent, Jan de Niel, Sotirios Moustakas, Patrick Willems, Gilles Drogue, Benjamin Grelier, Fernando Pereira, et al. thanks for the great cooperation!

A toute l'équipe d'hydrologie IRSTEA (INRAE) de Paris et en particulier à Guillaume Thirel, Vazken Andréassian et Charles Perrin, merci de votre chaleureux accueil en juin 2019. Ce fût une expérience formidable de mon doctorat et je garde aussi de bons souvenirs des parties de pétanque et de frisbee.

Tanja, het was mij een eer om door te gaan op jouw uitgezette pad, het heeft onze band nog verder versterkt. Fijn om samen te werken aan een vervolg van de Maas studie. Dank voor het meedenken, je morele en hydrologische support bij al mijn vragen!

Emma, het laatste deel van mijn onderzoek was niet mogelijk geweest zonder jou, dank voor je klimaatdata en alle interessante hydrologische discussies! Bovenal wil ik jou bedanken voor al het avontuur samen, de pret, je begrip, troost en geruststelling, onze mooie vriendschap.

Tim, vanaf het moment dat onze vriendschap begon, helemaal aan het begin van onze studies, begon ik me echt thuis te voelen in Nederland. En op dezelfde manier heb je me verwelkomd in 4.84 en wegwijs gemaakt aan het begin van mijn PhD, dank hiervoor!

Franca, dank voor de fijne gesprekken tijdens onze wandelingen, eerst rondom de TU toen we tegelijk bezig waren met onze PhD en nu in het Haagsche bos of in de duinen van Meijendael. Onze vriendschap heeft me zeker houvast gegeven om naar de overkant te varen! Dank voor de gedeelde ervaringen en al je lekkerste gerechten.

Dirk, het was een steun voor mij om afgelopen jaren allebei zowel bij Deltares als aan onze PhD te werken. Dank voor je inspirerende werk, de mooie vakanties, sportieve uitjes en natuurlijk dat we bij je in het wiel kunnen fietsen op de langste en kortste dag, en ook op andere dagen.

Wouter, dank dat je zo mooi laat zien dat waar een wil is, is een weg. Al loopt die weg over alle bergpieken van Zwitserland. Dank voor het goede ontvangst en de mooie herinneringen aan Zürich en de Susten Pass. Op naar nieuwe avonturen in de lage landen.

Marijn en Emma, dank voor de vreugde als we elkaar zien, de fijne herinneringen aan het strandhuis, de picknick bij Reeuwijk en alle andere uitstapjes en etentjes. Dank voor de gedeelde verhalen, jullie adviezen en de inspiratie voor lekkere recepten.

Joost en Roebijn, dank dat jullie ons hebben ingewijd in het fietsen, het werd mijn mentale redding in 2020. Dank voor het boulderen, alle leuke tripjes, de lekkere maaltijden, jullie vlijmscherpe humor, g'woon alles. Die vakantie in Griekenland houden we tegoed.

Gedurende deze pandemie, besef ik des te meer hoe belangrijk etentjes met mijn dierbare vrienden zijn en hoeveel ik die mis. Annabel, Isabelle en Marleen, vanaf dat onze vriendschap begon op dag één van het introductieweekend van Civiel, ben ik jullie dankbaar voor de avondjes uit en de interessante en grappige verhalen die we met elkaar delen. Maurits en Marieke, dank voor de gezellige tijden samen sinds London baby. Gaby & Dave, looking forward for more dinners and bike rides with you in Delft (or Canada!). Franca, Dirk, Tim, Emma, Wouter, nu we allemaal weer in Nederland wonen, kijk ik uit naar meer wine and woods samen. Dimitri & Manon, Zjenja & Britt, Bas & Greetje, dank voor jullie geweldige humor. Erik & Mandy, ik kijk uit naar meer portobello's en goede gesprekken met jullie. Sarah, Noor en Sophie, dank voor onze vriendschap in post-Atlas tijd. Rutger & Meagan, thanks for the delicious and lovely dinners. Bernhard & Charlotte, dank voor de lekkere whisky en barbecue avonden, gezellige vakanties, het fietsen,

het boulderen. Iris & Kjeld, dank voor de ballonnen met helium. Simon & Larissa, it is such a pleasure to have you so close by in Delft for a chat, hang-out in your sunny garden and a ride to Bruges. Eric & Rianne, dank voor de ontspannende biertjes en koffies. Jonathan, dank voor je heerlijke kookkunsten. Rens & Rianne, dank voor de fijne tijd in Zwitserland. Gert-Jan en Noor, ik kijk er naar uit om jullie op te zoeken in Praag. Rrrrrrrrrrrrrrrrrrria, dank voor de lekkerrrrrnijen!

There is no better distraction for the mind than sporting with friends. Meagan thanks for all the fun nights bouldering, always reassuring me that everything would go well. Iris, dank voor je humor, je inspirerende ondernemingen en interesses. Dirk-Jan, dank voor al je goede tips en aanmoedigingen bij het boulderen en mountainbiken. Sondeerfanaten! Jan, Ruud en Emma, wat een memorabele finish met jullie op de Cauberg. De dinsdagavond klimgroep, Job & Annemarie en Melle & Katinka, ik kijk ernaar uit dat we weer goede tijden beleven samen in de hal.

Merci Julie & Geoffrey, Charlotte & Tim, Séverine & Matt, Iris, Ingrid & David, Natsumé, Gilda, Léa & Joe et Elissavet & Albin pour notre précieux amitié, qui ne cesse de grandir depuis plus de 20 ans, malgré la distance qui nous sépare. Nos zooms du dimanche soir et nos beaux voyages me procurent de l'énergie et du bonheur. Merci de m'avoir soutenue et guidée dans mes pensées quand j'avais des doutes pendant ces années de thèse. Coraline, merci pour tes bons conseils et ton aide mathématique. Armelle merci de me faire rêver avec tes photos d'escalade à Fontainebleau. Maya & Julien, merci de me faire rire et des bons moments à Delft et Paris.

Marion & Drewin, Carolien & Daan, dank voor alle gezellige etentjes in Den Haag! Cloë dank voor de bijzondere band die ons bindt sinds altijd. Tami, Freddy, Eyal et Noam, c'est toujours un plaisir de passer du temps ensemble, à Paris ou en Hollande.

Lex, Annemarie, Hein, Maartje, Sofie en Julia, dank dat jullie zo meelevend, betrokken en geïnteresseerd zijn. Dank voor jullie inzichten in onderzoek, geologie en biologie. Lex dank voor de discussies over mijn stellingen. Annemarie, ik ben onder de indruk van je grote kracht de afgelopen maanden en ik ben ontzettend blij dat we mijn promotie met elkaar kunnen vieren!

Papa, nous voila tous les deux des docteurs Bouaziz à présent! Merci de toujours me dire que ce que je fais est tout à fait incroyable et merveilleux ("tu as fait 100 km?!"). Maman, merci pour tous ces bons repas à Rotterdam et de m'ouvrir à de nouveaux mondes. François, merci d'être accueillant, drôle et plein d'idées originales. Marjolein, dank voor jouw artistieke blik en tips voor de kaft. Maxim et Yannick, merci de chasser les fantômes. Ma boule et Greg merci pour toutes les supers vacances chez vous dans les Pyrénées. Zoé, merci de ton joli rire qui me donne le sourire. Quelle chance j'ai d'avoir une famille comme vous!

Tot slot, lieve Pieter, dank voor de mooie fietsvakantie langs de Maas, je creatieve geest, je goede raad, je liefde en al het andere wat het leven met jou zo mooi en bijzonder maakt.

CURRICULUM VITÆ

Laurène Judith Estelle BOUAZIZ

01-09-1988	Born in Les Lilas, France.

EDUCATION

1999–2006	Baccalauréat (International Option in Dutch - OIB) Lycée International de Saint-Germain-en-Laye, France
2006–2009	BSc Civil Engineering Delft University of Technology, the Netherlands
2009–2012	MSc Civil Engineering Delft University of Technology, the Netherlands
2017–2021	PhD Hydrology Delft University of Technology, the Netherlands

EXPERIENCE

2013–present	Hydrologist at Deltares Delft, The Netherlands

LIST OF PUBLICATIONS

Bouaziz, L.J.E., Aalbers, E., Weerts, A., Hegnauer, M., Buiteveld, H., Lammersen, R., Stam, J., Sprokkereef, E., Savenije, H. & Hrachowitz, M. (2021). *The importance of ecosystem adaptation on hydrological model predictions in response to climate change*. Hydrology and Earth System Sciences Discussion [preprint], in review. https://doi.org/10.5194/hess-2021-204

Bouaziz, L.J.E., Fenicia, F., Thirel, G., de Boer-Euser, T., Buitink, J., Brauer, C.C., De Niel J., Dewals, B.J., Drogue, G., Grelier, B., Melsen, L.A., Moustakas, S., Nossent, J., Pereira, F., Sprokkereef, E., Stam J., Weerts A.H., Willems, P., Savenije, H.H.G. & Hrachowitz, M. (2021). *Behind the scenes of streamflow model performance*. Hydrology and Earth System Sciences, 25(2), 1069–1095. https://doi.org/10.5194/hess-25-1069-2021

Bouaziz, L. J. E., Steele-Dunne, S. C., Schellekens, J., Weerts, A. H., Stam, J., Sprokkereef, E., Winsemius, H.C., Savenije, H.H.G & Hrachowitz, M. (2020). *Improved understanding of the link between catchment-scale vegetation accessible storage and satellite-derived Soil Water Index*. Water Resources Research, 56(3), e2019WR026365. https://doi.org/10.1029/2019WR026365

Bouaziz, L., Weerts, A., Schellekens, J., Sprokkereef, E., Stam, J., Savenije, H., & Hrachowitz, M. (2018). *Redressing the balance: Quantifying net intercatchment groundwater flows*. Hydrology and Earth System Sciences, 22(12), 6415-–6434. https://doi.org/10.5194/hess-22-6415-2018

de Boer-Euser, T., **Bouaziz, L.**, de Niel, J., Brauer, C., Dewals, B., Drogue, G., Fenicia, F., Grelier, B., Nossent, J., Pereira, F., Savenije, H., Thirel, G. & Willems, P. (2017). *Looking beyond general metrics for model comparison; Lessons from an international model intercomparison study*. Hydrology and Earth System Sciences, 21(1), 423-–440. https://doi.org/10.5194/hess-21-423-2017