

Doctoral thesis

Linking hydrological process patterns in a small catchment

submitted in satisfaction of the requirements for the degree of
Doctor of Science in Civil Engineering
of the Vienna University of Technology, Faculty of Civil Engineering

as part of the
Vienna Doctoral Programme on Water Resource Systems

by

Borbála Széles

Student registration number: 1332010

Taubstummengasse 11, 1040 Vienna, Austria

under the supervision of

Assoc. Prof. Juraj Parajka

Institute of Hydraulic Engineering and Water Resources Management

Vienna University of Technology

Vienna, 2020

.....
Borbála Széles

Dissertation

Zusammenhänge zwischen hydrologischen Prozessen in einem kleinen Einzugsgebiet

ausgeführt zum Zwecke der Erlangung des akademischen Grades einer Doktorin
der technischen Wissenschaft

eingereicht an der Technischen Universität Wien Fakultät für Bauingenieurwesen
VON

Borbála Széles

Matrikelnummer: 1332010

Taubstummengasse 11, 1040 Wien, Österreich

unter der Leitung von

Assoc. Prof. Juraj Parajka

Institute für Wasserbau und Ingenieurhydrologie

Technische Universität Wien

Wien, 2020

.....
Borbála Széles

Examiners

Examiners

Assoc. Prof. Juraj Parajka

Institute of Hydraulic Engineering and Water Resources Management, Department of Civil Engineering, Vienna University of Technology
Karlsplatz 13/222/2, 1040 Vienna, Austria

Ao. Univ. Prof. Dipl.-Ing. Dr. techn. Matthias Zessner

Research Unit for Water Quality Management, Institute for Water Quality and Resource Management, Department of Civil Engineering, Vienna University of Technology
Karlsplatz 13/226, 1040 Vienna, Austria

Prof. Dr. Zoltán Gribovszki

Institute of Geomatics and Civil Engineering, Faculty of Forestry, University of Sopron
Bajcsy-Zsilinszky u. 4, Sopron, 9400, Hungary

Gutachter

Gutachter

Assoc. Prof. Juraj Parajka

Institut für Wasserbau und Ingenieurhydrologie, Fakultät für Bauingenieurwesen, Technische Universität Wien

Karlsplatz 13/222/2, 1040 Wien, Österreich

Ao. Univ. Prof. Dipl.-Ing. Dr. techn. Matthias Zessner

Forschungsbereich Wassergütwirtschaft, Institut für Wassergüte und Ressourcenmanagement, Fakultät für Bauingenieurwesen, Technische Universität Wien

Karlsplatz 13/226, 1040 Wien, Österreich

Prof. Dr. Zoltán Gribovszki

Institut für Geomatik und Bauingenieurwesen, Fakultät für Forstwirtschaft, Universität Sopron
Bajcsy-Zsilinszky u. 4, Sopron, 9400, Ungarn

Abstract

Investigating hydrological process patterns in small catchments provides a way to better understand the rainfall-runoff transformation, which is useful not only for the wider scientific community, but also for practitioners working in water resources management, risk and operational forecasting in order to improve hydrological predictions.

The objective of this thesis is to understand and describe the links between observed hydrological process patterns of various variables in a small experimental catchment, the Hydrological Open Air Laboratory (HOAL) by using comparative data analyses and integrated hydrological modelling. The main aim is to advance the understanding of dynamic catchment behavior and to investigate methods of linking field observations with hydrological modelling of different water storages and their interactions. The main research question is to explore the value of extensive field observations for improving hydrological models with and without runoff observations.

The thesis is organized into five chapters. Chapter 1 presents the main objectives and introduces the hydrological modelling context for the following chapters. Chapter 2 examines the interaction between streamflow, the riparian zone, and the crop fields during rainless periods. In this chapter, the spatial and temporal variabilities of daily streamflow fluctuations are explored and a solar radiation driven model is proposed to reproduce these fluctuations and the time lags between the radiative forcing and diel streamflow fluctuations. The results indicate that vegetation cover and runoff generation mechanisms determine the magnitudes of the streamflow fluctuations. Most of the volume associated with the diel streamflow fluctuations at the catchment outlet is explained by transpiration from the riparian forest along the main stream. Results of the model simulations show that the time lags have a strong seasonality increasing from spring to summer and decreasing from summer to autumn. These results suggest that the diel variations in streamflow are mostly controlled by the area close to the stream while the seasonal variations in streamflow are controlled by the entire catchment area.

Chapter 3 investigates the value of different field observations for calibrating a rainfall runoff model for small basins. An HBV type, spatially lumped hydrological model is calibrated using a new step-by-step parametrization approach, where the three modules of the model are calibrated using different data types besides runoff. Results indicate that by using the new approach, the overall process consistency improves, compared to model simulations when only runoff is used for model calibration. Soil moisture and evapotranspiration observations are found to have the largest influence on simulated runoff, while the parameterization of the snow and runoff generation modules have a smaller influence in case of this catchment. Chapter 4 investigates the prediction of runoff using proxy data in a small ungauged catchment. Similarly to Chapter 3, the three modules of a conceptual hydrological model are calibrated step-by-step using all the available field observations except runoff. The results indicate that by using snow and soil moisture information for model calibration, the runoff model performance is comparable to the scenario when the model is calibrated using only runoff data and the simulation of state variables such as snow cover and soil moisture also improves. The thesis concludes with a summary of the main findings and conclusions.

This thesis has advanced the understanding of dynamic catchment behavior and the interconnection between different storages of water within the catchment. This work has developed new techniques for estimating evapotranspiration rates during rainless periods and for

Abstract

parameterizing conceptual hydrological models using field observations besides and without runoff.

By using additional observations of water fluxes and storages, we are able to better understand and estimate the water balance components in gauged and ungauged small agricultural catchments. By using streamflow observations related to different runoff generation mechanisms, net radiation, evapotranspiration (measured by eddy covariance systems), and groundwater measurements, we now have a better understanding of how the catchment behaves during rainless periods. For example, the crop fields further away from the stream tend to disconnect from the stream and therefore evapotranspiration seen by the streamflow fluctuations is not influenced by the crop fields at the daily scale. By using additional observations of the hydrological cycle besides runoff, such as the precipitation phase, snow, soil moisture, actual evapotranspiration, overland flow and storage change in the saturated zone, we are able to improve the simulations of state variables and runoff, and therefore the model consistency. This step is useful for reproducing reality more accurately which is the general goal of every modelling study.

Kurzfassung

Die vorliegende Untersuchung hydrologischer Prozesse in kleinen Einzugsgebieten ist ein Beitrag, der sich zur Aufgabe gemacht hat, Niederschlag-Abfluss Prozesse besser zu verstehen. Die Ergebnisse sind nicht nur aus wissenschaftlicher Perspektive hilfreich, sondern auch für die wasserwirtschaftliche Praxis um das Hochwasser- und Dürrierisiko genauer beurteilen zu können, und operationelle Vorhersagen des Abflusses zu verbessern.

Das Ziel dieser Dissertation ist es, die Zusammenhänge zwischen beobachteten hydrologischen Prozessen in einem kleinen Einzugsgebiet, dem Hydrological Open Air Laboratory (HOAL), mit Hilfe von vergleichenden Datenanalysen und hydrologischer Modellierung zu beschreiben und zu verstehen. Mit Hilfe dieser Untersuchungen sollen die Kenntnisse über die Niederschlag-Abfluss Prozesse erweitert und Wege aufgezeigt werden, Feldbeobachtungen mit Modellsimulationen der Wasserspeicherung in unterschiedlichen Bereichen eines Gebietes zu verknüpfen. Die wissenschaftliche Fragestellung behandelt den Mehrwert von umfangreichen Feldbeobachtungen für die Güte der Modellkalibrierung mit und ohne Nutzung von Abflussdaten.

Die Dissertation ist in fünf Kapitel gegliedert. Kapitel 1 beschreibt das Ziel der Arbeit, den Kontext und beinhaltet eine kurze Einführung in die Niederschlag-Abfluss Modellierung. Kapitel 2 untersucht die Interaktion zwischen Abfluss, der Uferrandzone und den Ackerflächen während niederschlagsfreien Perioden. Die raum-zeitliche Variabilität der täglichen Abflussfluktuationen wird untersucht und ein Modell wird vorgestellt, das durch Sonneneinstrahlung angetrieben wird, um das Verdunstungsvolumen und die Verzögerungszeiten zwischen Verdunstung und Strahlungsantrieb zu bestimmen. Es wird gezeigt, dass die raum-zeitliche Variabilität der Abflussfluktuationen durch die Vegetation und die Abflussprozesse erklären werden kann und der größte Teil der täglichen Abflussfluktuationen beim Auslass des Einzugsgebiets durch die Verdunstung der Uferrandzone bestimmt wird. Die Ergebnisse der Modellsimulationen zeigen, dass die Verzögerungszeiten zwischen Sonneneinstrahlung und deren Auswirkung auf den Abfluss eine starke Saisonalität aufweisen, wobei die Verzögerungszeiten die sich vom Frühjahr bis zum Sommer zunehmen, und vom Sommer bis zum Herbst zurückgehen. Diese Ergebnisse deuten darauf hin, dass die täglichen Schwankungen des Abflusses hauptsächlich durch das Gebiet in der Nähe des Baches beeinflusst werden, während die saisonalen Schwankungen des Abflusses durch das gesamte Einzugsgebiet beeinflusst werden.

Kapitel 3 beschäftigt sich mit dem Mehrwert von Feldbeobachtungen für die Kalibrierung von Abflussmodellen für kleine Einzugsgebiete. Ein konzeptionelles Modell ähnlich zu HBV wird durch eine stufenweise Parametrisierung kalibriert. Die drei Module des Modells werden mit Hilfe von verschiedenen Feldbeobachtungen, die über die Beobachtung des Abfluss hinaus gehen, kalibriert. Die Ergebnisse zeigen, dass die gesamte Prozesskonsistenz mit dieser neuen Herangehensweise, im Vergleich zu einer reinen Abflusskalibrierung, verbessert wird. Beobachtungen von Bodenfeuchte und Verdunstung haben den größten Einfluss auf den simulierten Abfluss in diesem Einzugsgebiet. Die Parametrisierung der Module von Schnee und Abflussbildung hat einen geringeren Einfluss.

In Kapitel 4 wird der Nutzen von Proxy Daten für die Kalibrierung eines konzeptuellen Modells für ein kleines, unbeobachtetes Einzugsgebiet untersucht. Ähnlich zu Kapitel 3 werden die drei Module des Modells, unter Berücksichtigung aller Feldbeobachtungen mit Ausnahme des Abflusses, schrittweise kalibriert. Es konnte gezeigt werden, dass es möglich ist, nur mit Hilfe von Schnee- und Bodenfeuchtemessungen den Abfluss so gut wie mit reiner Abflusskalibrierung zu

simulieren. Zusätzlich kann durch diese Vorgangsweise die Simulation von Zustandsvariablen verbessert werden. Die Dissertation schließt mit einer Zusammenfassung der Ergebnisse und Schlussfolgerungen.

Diese Dissertation erweitert die Kenntnisse über die Niederschlag-Abflussprozesse und die Zusammenhänge zwischen verschiedenen Komponenten der Wasserspeicherung innerhalb eines Einzugsgebiets. Es werden neue Techniken entwickelt, um das Verdunstungsvolumen während trockener Perioden zu bestimmen und um konzeptuelle hydrologische Modelle mit Hilfe von Feldbeobachtungen mit und ohne Berücksichtigung von Abflussmessungen zu parametrisieren.

Durch die zusätzliche Messung von verschiedenen Elementen des Wasserkreislaufs erreichen wir ein besseres Verständnis der Wasserbilanz in kleinen, landwirtschaftlichen Einzugsgebieten, mit und ohne Abflussbeobachtungen. Dadurch, dass wir die Abflussmessungen mit dem Verständnis verschiedener Abflussprozessen, der Nettostrahlung, der mittels Eddy-Kovarianz gemessenen Verdunstung und Grundwasserspiegelstände verknüpfen konnten, verstehen wir nun besser, wie sich das Einzugsgebiet während Trockenperioden verhält. Wir verstehen, dass die Verdunstung auf den Ackerflächen in größerer Entfernung des Baches wenig Einfluss auf den Abfluss auf der Tagesskala haben, da diese Flächen vom Abfluss im Bach hydrologisch entkoppelt sind. Mit Hilfe von zum Abfluss ergänzenden Feldbeobachtungen, wie zum Beispiel Niederschlagstyp, Schneelage, Bodenfeuchte, Verdunstung, Oberflächenabfluss, Veränderungen in der Speicherung der gesättigten Zone, kann die Simulation des Abflusses, der hydrologischen Zustandsvariablen und damit auch die Modellkonsistenz verbessert werden. Diese Verbesserungen sind ein wichtiger Schritt für die genauere Simulation der Realität durch das Modell - dem Hauptziel jeder Modellierung.

Funding information

Funding information

This work has been conducted as part of the following research project:

Vienna Doctoral Programme on Water Resource Systems (DK W1219-N28) of the Austrian Science Funds (FWF).

Acknowledgements

Acknowledgements

I would like to express my gratitude to Professor Juraj Parajka and Professor Günter Blöschl for the supervision of my PhD thesis and the relentless support. I am grateful for the constructive suggestions and ideas so I could continuously improve my research.

I would like to thank all my colleagues from the Institute of Hydraulic Engineering and Water Resources Management at the Vienna University of Technology for their support during the last couple of years.

I would also like to thank my colleagues in Petzenkirchen, who helped me with the field work and gave me constructive feedback on my work.

I would also like to thank the students of the Vienna Doctoral Programme on Water Resources Systems for helping my work and stay in Vienna.

I wish to thank my family and friends for always being there and supporting me.

Content

Examiners 3

Gutachter..... 4

Abstract 5

Kurzfassung 7

Funding information 9

Acknowledgements..... 10

Content..... 11

1. Introduction..... 14

2. Separation of scales in transpiration effects on low flows – A spatial analysis in the Hydrological Open Air Laboratory..... 16

 2.1. General 16

 2.2. Key Points 16

 2.3. Abstract 16

 2.4. Introduction 17

 2.5. Study area and data 18

 2.5.1. Study area..... 18

 2.5.2. Data..... 23

 2.6. Methodology 25

 2.6.1. Identification of periods with streamflow fluctuations..... 25

 2.6.2. Modeling of streamflow fluctuations..... 25

 2.7. Results 27

 2.7.1. Amplitude of the observed diel signals..... 27

 2.7.2. Timing of the observed diel signals 32

 2.7.3. Model performance 33

 2.7.4. Process controls on amplitude of diel streamflow variation..... 34

 2.7.5. Process controls on lag times of diel streamflow variation 36

 2.8. Discussion 38

 2.8.1. Spatio-temporal patterns of streamflow fluctuations 38

 2.8.2. Simplified process representation 39

 2.8.3. Estimated evapotranspiration volumes and rates 40

 2.8.4. Separation of scales in time implies a separation of scales in space 42

 2.9. Conclusions 45

 2.10. Acknowledgements 46

3.	The added value of different data types for calibrating and testing a hydrological model in a small catchment	47
3.1.	General	47
3.2.	Key points	47
3.3.	Abstract	47
3.4.	Introduction	48
3.5.	Study area and data	50
3.5.1.	Study area.....	50
3.5.2.	Data.....	51
3.6.	Methodology	53
3.6.1.	Hydrological model	53
3.6.2.	Model calibration with only runoff data	58
3.6.3.	Model calibration with runoff and additional data	59
3.6.3.1.	Snow parameters.....	59
3.6.3.2.	Soil moisture and evapotranspiration parameters.....	61
3.6.3.3.	Runoff generation parameters	63
3.6.3.4.	Runoff simulation	65
3.7.	Results	65
3.7.1.	Model calibration with only runoff data	65
3.7.2.	Model calibration with runoff and additional data	66
3.7.2.1.	Snow simulation	66
3.7.2.2.	Soil moisture and evapotranspiration simulation	68
3.7.2.3.	Runoff generation and routing simulation.....	73
3.7.2.4.	Runoff simulation	75
3.8.	Discussion	80
3.9.	Conclusions	82
3.10.	Acknowledgements	83
4.	Stepwise prediction of runoff using proxy data in a small agricultural catchment	84
4.1.	General	84
4.2.	Key points	84
4.3.	Abstract	84
4.4.	Introduction	84
4.5.	Study area and data	86
4.5.1.	Study area.....	86

Content

4.5.2. Data	86
4.6. Methodology	86
4.6.1. Hydrological model	86
4.6.2. Model calibration without runoff data	87
4.6.2.1. Calibration of snow module	87
4.6.2.2. Calibration of the soil moisture accounting module.....	88
4.6.2.3. Calibration of the runoff generation module	89
4.6.3. Model calibration with runoff data	89
4.7. Results	91
4.7.1. Snow module simulations	91
4.7.2. Soil moisture accounting module simulations	91
4.7.3. Runoff generation module simulations	94
4.7.4. Runoff simulations.....	95
4.8. Discussion	97
4.9. Conclusions	99
4.10. Acknowledgements	99
5. Summary and conclusions	100
References.....	103
Appendix A1	117
Appendix A2.....	117
Appendix A3.....	118
Appendix A4.....	119
Appendix A5.....	119
Appendix A6.....	120
Appendix A7.....	120
Appendix A8.....	120
Appendix A9.....	120
Appendix A10.....	121
Appendix A11.....	122

Introduction

1. Introduction

The changing human systems induce changes in the hydrological systems and conditions. The elements of the hydrological cycle show increasingly varying spatial and temporal patterns. Therefore, the new Scientific Decade 2013-2022 of the International Association of Hydrological Sciences (IAHS), *Panta Rhei* (Everything Flows) has been dedicated to understand changes in hydrology and society to develop more accurate predictions. The main target of *Panta Rhei* is to understand the hydrological systems based on innovative research ideas, interdisciplinary collaboration, cutting-edge monitoring network and maintenance of the existing traditional continuous measurement networks (Montanari et al., 2013).

Hydrological model simulations provide a way to understand the rainfall response of catchments with different characteristics, the overall catchment behavior and its seasonal variability. Furthermore, the results of hydrological modelling can be applied by different disciplines, for instance by hydrodynamics and water quality. However, given the fact that catchments are complex systems characterized by spatially heterogeneous properties, their simplified representations, i.e. the hydrological models are inevitably poorly defined systems. In order to capture the heterogeneities and nonlinear behavior of rainfall-runoff processes, the complexity of hydrological models have been increasing. Hydrological processes, their spatial and temporal patterns are linked on different levels of complexity by different modelling approaches (Beven, 2001b; Singh, 1988).

In the literature, there is a description and analysis of a large number of different hydrological models (Moradkhani & Sorooshian, 2009). For example, NAM (Nedbor Afstromnings Model) was developed to estimate the interaction between groundwater and surface water (Nielsen & Hansen, 1973). The TOPMODEL (TOPography based hydrological MODEL) approach assumes that runoff is determined by topography and parts of the catchment with the same topographic index behave in a hydrologically similar way (Beven & Kirkby, 1979; Beven et al., 1984). HCYMODEL (Hydrological Cycle Model) was designed for forested mountain catchments, where the separated river and hillslope systems are represented by five storage tanks (Fukushima, 1988). GR4J (modèle du Génie Rural à 4 paramètres Journalier) is a four-parameter lumped hydrological model using one production and one routing store, where the flow routing modelled by two unit hydrographs is influenced by groundwater exchange (Perrin et al., 2003). There exist more complex, integrated models providing a variety of modules to be combined, such as the lumped HEC-HMS (Hydrologic Engineering Center - Hydrologic Modeling System) (Scharffenberg, 2001) for dendritic watershed systems or the physically based, spatially distributed MIKE SHE (MIKE Système Hydrologique Européen) (DHI, 1998). In Austria, a concept of HBV-type-of hydrological model (Hydrologiska Byråns Vattenbalansavdelning model) (Bergström, 1976; Bergström, 1995; Bergström & Linström, 2015; Blöschl et al., 2008; Göttinger & Bardossy, 2007; Lindström et al., 1997) has been successfully applied for regional hydrological modelling, but also for operational river flood forecasting (Nester et al., 2016). The general advantages of using the conceptual HBV-type-of hydrological model for operational and scientific purposes are the relative small number of parameters, the continuous representation of the main water storages and fluxes, the high computational speed and open access source code. However, it is still not well understood, how to link the model parameters to different field observations.

As hydrological models evolved parallel with the advances in computer science, remote sensing and measurement techniques, new challenges have arisen (Beven, 2001a; 2007), such as

Introduction

overparameterisation, equifinality, intercorrelation between model parameters, uncertainty (model structure, parameters, boundary conditions), uniqueness and non-uniqueness, parameter nonidentifiability, nontransferability of parameters across calibration scales and locations and demanding computational time. Furthermore, “there are scales and physical processes that cannot be represented by a numerical model regardless of the resolution” (Stensrud, 2007). Hydrological modelling hinges on the existence and availability of long-term observations with a spatial and temporal resolution representative enough for the specific modelling purpose. Generally, measurements of different hydrological states and fluxes are not available on very fine spatial and/or temporal resolution. The point scale measurements are usually difficult to be upscaled to the catchment scale. Different measurement techniques have different footprints (for example evapotranspiration estimation based on water balance or by eddy covariance technique), which makes challenging their comparison and assimilation with hydrological models.

Experimental catchments provide an excellent opportunity to study the rainfall-runoff generation processes in very fine spatial and temporal resolution. There is only a limited number of well-equipped open air laboratories in the world and one of them is the Hydrologic Open Air Laboratory (HOAL) (Blöschl et al., 2016). The advantage of such experimental sites is that they allow to combine and supplement traditional observation techniques and methods with more dedicated monitoring of meteorological and hydrological processes and hence to investigate more the patterns and links between runoff generation mechanisms and their upscaling to larger catchment scales. Dedicated monitoring and state-of-the-art sensing techniques (such as Eddy covariance and geophysical measurements, satellite products, in situ soil moisture sensors, etc.) provide additional observations of patterns and processes at a very fine resolution. This allows us to improve the understanding of the dynamic behavior and non-linear relationship between rainfall and runoff and the interconnection between different storages of water within a catchment. The lessons learnt from the hypothesis testing in experimental catchments is useful not only for the wider scientific community, but also for practitioners working in water resources management, risk and operational forecasting.

The objective of the thesis is to better understand and describe the links between the observed hydrological process patterns in a small experimental catchment, the Hydrological Open Air Laboratory, using comparative data analyses and integrated hydrological modelling framework. The main aim of this thesis is to advance the knowledge of the dynamic catchment behavior and the interconnection between different storages of water within the catchment. The research part of the thesis is organized into three chapters (Chapter 2-4). The main objective of Chapter 2 is to understand the catchment behavior during longer rainless periods, more specifically it is explored which part of the catchment induces the diel streamflow fluctuations, i.e. the daily fluctuations of the streamflow rates, and how the spatial differences in land use and runoff generation mechanisms influence these fluctuations. The research question of Chapter 3 is how to parameterize a lumped, conceptual hydrological model for a small agricultural catchment using a variety of proxy data besides runoff. In this chapter it is explored how to link field observations with hydrological model simulations. And the science question of Chapter 4 is how accurately we can predict runoff in a small ungauged catchment using all the available information except runoff. Finally, Chapter 5 concludes the thesis with a summary of key findings and discusses potential future extensions of the research work and implications for hydrology and water resources management.

2. Separation of scales in transpiration effects on low flows – A spatial analysis in the Hydrological Open Air Laboratory

2.1. General

The goal of this chapter was to better understand the catchment behavior during rainless, lowflow periods.

This chapter investigated the interaction between the stream, the riparian zone and the rest of a small agricultural catchment in low flow, rainless periods. The observed streamflow rates featuring different runoff generation mechanisms, such as wetland and tile drain runoff, were linked with estimated evapotranspiration patterns.

The present chapter corresponds to the following scientific publication in its original form:

Széles, B., Broer, M., Parajka, J., Hogan, P., Eder, A., Strauss, P., & Blöschl, G. (2018). Separation of scales in transpiration effects on low flows – A spatial analysis in the Hydrological Open Air Laboratory. *Water Resources Research*, 54, 6168-6188. doi: 10.1029/2017WR022037

2.2. Key Points

1. The amplitude of the diel streamflow fluctuations from wetlands was one magnitude larger than those from tile drains and springs.
2. Three quarters of the volume associated with diel streamflow fluctuations was due to riparian transpiration along the main stream.
3. Lag times between radiative forcing and *ET* increased from 3 to 11 hours from spring to autumn.

2.3. Abstract

The objective of this study was to understand whether spatial differences in runoff generation mechanisms affect the magnitudes of diel streamflow fluctuations during low flow periods and which part of the catchment induces the diel streamflow signal. The spatio-temporal variability of the streamflow fluctuations observed at twelve locations in the 66 ha Hydrological Open Air Laboratory experimental catchment in Austria was explained by differences in the vegetation cover and runoff generation mechanisms. Almost a quarter of the volume associated with diel streamflow fluctuations at the catchment outlet was explained by transpiration from vegetation along the tributaries; more than three quarters was due to transpiration by the riparian forest along the main stream. The lag times between radiative forcing and evapotranspiration estimated by a solar radiation driven model increased from 3 to 11 hours from spring to autumn. The recession time scales increased from 21 days in spring to 54 days in autumn. Observations and model simulations suggest that a separation of scales in transpiration effects on low flows exist both in time and space, i.e. the diel streamflow fluctuations are induced by transpiration from the riparian vegetation, while most of the catchment evapotranspiration, such as evapotranspiration from the crop fields further away from the stream, do not influence the diel signal in streamflow.

2.4. Introduction

Evaporation and transpiration in mid latitude humid catchments affect streamflow at two main time scales. At the seasonal time scale, the energy input is at a maximum in summer, therefore evapotranspiration is also at a maximum. This depletes soil moisture in summer below the annual mean, which affects runoff generation during storms. Soil moisture depletion also tends to reduce groundwater recharge and hence discharge to the streams. Streamflow during recession periods is the net result of the interplay of the hydraulic aquifer characteristics and evapotranspiration within the catchment. Several studies observed faster streamflow recessions in summer than during the rest of the year due to summer evapotranspiration (Federer, 1973; Shaw and Riha, 2012; Szilágyi et al., 2007). At the daily time scale, there are similar fluctuations in the energy input between day and night, leading to diel fluctuations in evaporation and transpiration, which again, affect soil moisture and subsurface flow. In small headwater catchments these diel fluctuations of transpiration usually imprint a diel signal on the streamflow during low flow periods (Gribovszki et al., 2010b).

Even though there are two distinct time scales of variability in evaporation and transpiration, diel and seasonal, a relationship between the two time scales would be expected. Only a few studies examined the seasonality in the diel fluctuations and how the diel transpiration was related to the seasonal transpiration that determines the catchment water balance. Bond et al. (2002) and Wondzell et al. (2007, 2010) showed that the area contributing to streamflow fluctuations decreased as the catchment gradually dried out. This dynamic was explained by the weakening of the coupling between the vegetation and stream during summer as the groundwater levels dropped. The time lags between transpiration and streamflow also varies seasonally due to changes in the flow paths (Barnard et al., 2010; Cadol et al., 2012; Deutscher et al., 2016; Fonley et al., 2016; Graham et al., 2013; Gribovszki et al., 2008; Gribovszki et al., 2010a; Gribovszki et al., 2011; Kirchner, 2009; Szeftel, 2010; Szilágyi et al., 2008; Wondzell et al., 2007, 2010; Yue et al., 2016).

One of the main questions to investigate diel streamflow fluctuations during low flows in the past was to determine where the streamflow fluctuations originate. Numerous studies have confirmed that these fluctuations were induced by transpiration from the riparian and near river vegetation (e.g. Meyboom, 1965). Experiments where the riparian forest was removed showed that the streamflow fluctuations stopped (e.g. Dunford and Fletcher, 1947; Lawrence, 1990; O'Loughlin et al., 1982), however, when the hillslope vegetation was removed instead, the fluctuations persisted but in a modified way (Bren, 1997). Another question was related to the main flow paths in the subsurface between the vegetation and the stream. The integrated effect of site conditions and the species assemblage may determine the sources of root water uptake (Snyder and Williams, 2000). Certain trees may extract water directly from the groundwater (Barbeta and Peñuelas, 2017; Gribovszki et al., 2010b; Snyder and Williams, 2000), but the amplitude of the water table fluctuations depends on the distance from the stream and the vegetation type. For example, Reigner (1966) showed that the groundwater fluctuations were significant only up to a 2 meter distance from the stream. Irrigation experiments of Barnard et al. (2010) in Oregon suggested, however, that during higher soil moisture conditions after irrigation, hillslope vegetation located further from the stream could contribute to the diel streamflow fluctuations. Yue et al. (2016) observed diel groundwater table fluctuations along a part of a river that was covered by woody species and wet slough, however, the middle section of the river with shallower-rooted grasses did not exhibit water table fluctuations. When the depth to the water table exceeded a certain threshold during a

long recession period, the amplitudes of the diel water table fluctuations observed in an area covered by wet slough decreased.

There are no robust methods for measuring the evapotranspiration rates in riparian forests with mixed vegetation types (Drexler et al., 2004; Goodrich et al., 2000; Landon et al., 2009; Leenhouts et al., 2006; Loheide et al., 2005). For many of the narrow riparian corridors, the fetch requirement of the eddy covariance method often exceeds the width of the riparian forest (Goodrich et al., 2000). Upscaling sap flux measurements from tree to stand level can also be problematic (Cermák et al., 2004; Oishi et al., 2008; Schaeffer et al., 2000). The empirical approaches based on crop coefficients and vegetation index-based crop coefficients (Glenn et al., 2011; Nagler et al., 2005) are not widely used for riparian ecosystems due to the heterogeneous species composition. Therefore, estimating the evapotranspiration rates for the riparian zone based on measurements of diel fluctuations of soil moisture (e.g. the calibration free methods from Gribovszki, 2014; Gribovszki, 2018; Nachabe et al., 2005) or water levels (e.g. Bond et al., 2002; Cadol et al., 2012; Cernohous and Šach, 2008; Dvorakova et al., 2014; Gribovszki et al., 2008; Gribovszki et al., 2010a; Kirchner, 2009; Loheide et al., 2005; White, 1932; Wondzell et al., 2007, 2010) can be a valuable alternative. However, estimating riparian zone evapotranspiration in terms of volumes might be uncertain at some study sites where estimated evapotranspiration rates are only representative of a fraction of the riparian zone (Butler et al., 2007; Loheide et al., 2005; White, 1932).

Previous studies examined diel streamflow fluctuations and their temporal changes in single or nested catchments with spatially rather uniform runoff generation mechanisms (e.g. Bond et al., 2002; Szeftel, 2010). However, it is not clear, how spatial differences in the runoff generation mechanisms affect the flow paths and total evapotranspiration. The main objective of our study therefore was to evaluate spatial and temporal patterns of the diel streamflow fluctuations in relation to the seasonal cycle of evapotranspiration in a small experimental catchment with different runoff generation mechanisms and mixed land cover types. The study was performed in the Hydrological Open Air Laboratory (HOAL), a 66 ha Austrian experimental catchment, where both the main stream and the tributaries draining nearly all the surface contributing area to the stream are gauged, allowing the main stream and tributary contributions to be separated. The goal of our study was to determine where the diel low flow fluctuations originate, i.e. how much the streamflow fluctuations observed at the tributaries with different characteristics and runoff generation mechanisms (such as wetlands, springs and tile drainage systems) contribute to the streamflow fluctuations observed at the catchment outlet. Furthermore, we aimed to understand which part of the catchment induces the streamflow fluctuations on the diel time scale, and which part influences the evapotranspiration on the seasonal time scale. In order to explore the causal link between the process drivers and the streamflow signal, we used a solar radiation driven model. The estimated model parameters assisted in generalizing the response time scales associated with the diel fluctuations and the seasonal recession of streamflow.

2.5. Study area and data

2.5.1. Study area

The study was conducted in a small experimental catchment, the Hydrological Open Air Laboratory (HOAL) in Petzenkirchen, located in the western part of Lower Austria (Figure 1a), approximately 100 km west of Vienna (48° 9' N, 15° 9' E). The drainage area of the catchment is

Separation of scales in transpiration effects on low flows – A spatial analysis in the Hydrological Open Air Laboratory

66 ha at the catchment outlet (MW). The natural surface water outlet of the catchment is the Seitengraben stream. The elevation of the catchment ranges from 268 to 323 m above sea level, with a mean slope of 8%. The stream is approximately 620 m long and has a medium slope of 2.4% (Blöschl et al., 2016; Eder et al., 2010, 2014).

The climate is humid. Mean annual (2002-2015) air temperature, precipitation, runoff and evapotranspiration are 9.6 °C, 784 mm/yr, 178 mm/yr and 606 mm/yr, respectively. Seasonal maxima of air temperature, rainfall amount and intensity occur in the summer (Figure 3 in Blöschl et al., 2016). Mean monthly runoff tends to peak in winter or early spring. The seasonal variability in evapotranspiration plus storage change estimated from the water balance is presented in Figure 2.

The geology of the HOAL consists of Tertiary fine sediments and fractured siltstone of the Molasse zone. The dominant soil types are Cambisols (57%), Kolluvisol (16%) and Planosols (21%) with moderate to low permeability. Gleysols (6%) occur close to the stream (Blöschl et al., 2016; FAO, ISRIC and ISSS, 1998).

The catchment is dominated by agricultural land use. At the time of the study 87% of the catchment area was arable land, the rest was forested, paved or used as pasture. The crops were mainly winter wheat, maize, winter barley and winter oilseed rape. Winter wheat and winter barley resume growing in March, reaching their full height in June and are harvested typically in July. Maize is planted mid/end of April and harvested in late September. Oil seed rape develops from March to June and is harvested in July. Most of the forested area is located close to the stream (Figure 1). The most common trees in the riparian zone are different species of willow, poplar, ash, field maple and black alder. The main shrub species in the riparian zone are cornel, elder, hazelnut, honeysuckle and guelder rose. Based on the heterogeneity of the vegetation, i.e. the dominant tree types, the riparian zone can be divided into three parts. The upstream area is dominated by field maple and black alder, the middle section by poplar, the downstream part by ash (Appendix A4). The Leaf Area Index (LAI) is close to zero in winter, when the fields are bare and the trees shed their leaves. The maximum of LAI is reached in different months in the vegetation period depending on the crop type. On the maize fields the LAI reaches LAI=6 in July/August, on the winter wheat and barley fields the maximum LAI of around 3 is reached at the end of spring. In the riparian forest the maximum LAI of around 4 is reached at the beginning of summer (European Space Agency – ESA, 2020).

Separation of scales in transpiration effects on low flows – A spatial analysis in the Hydrological Open Air Laboratory

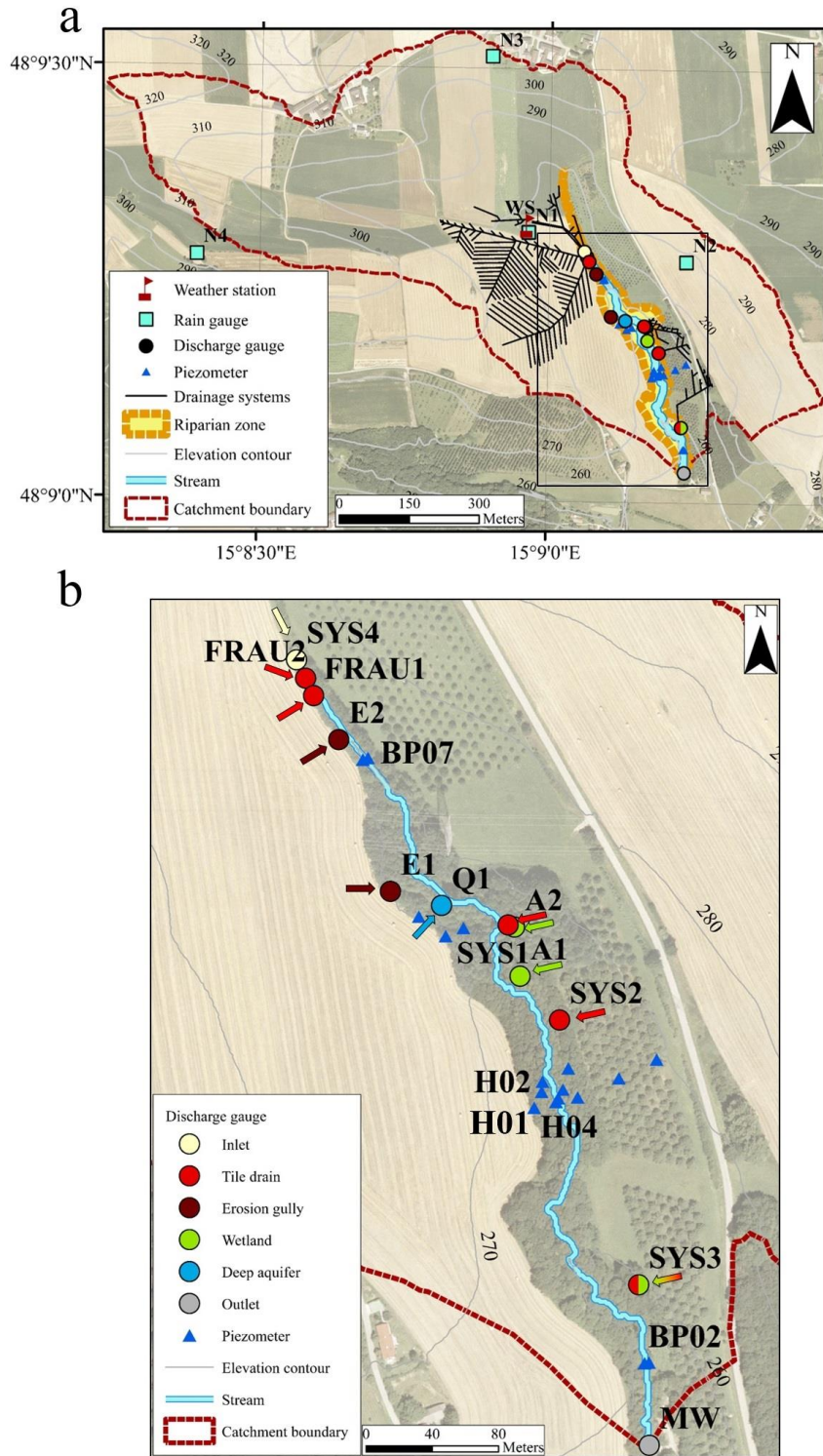


Figure 1. Hydrological Open Air Laboratory (HOAL) in Petzenkirchen, Lower Austria (panel a). Streamflow is monitored at the main outlet (MW) and at 11 tributaries (detail, panel b) that have different dominant runoff generation mechanisms.

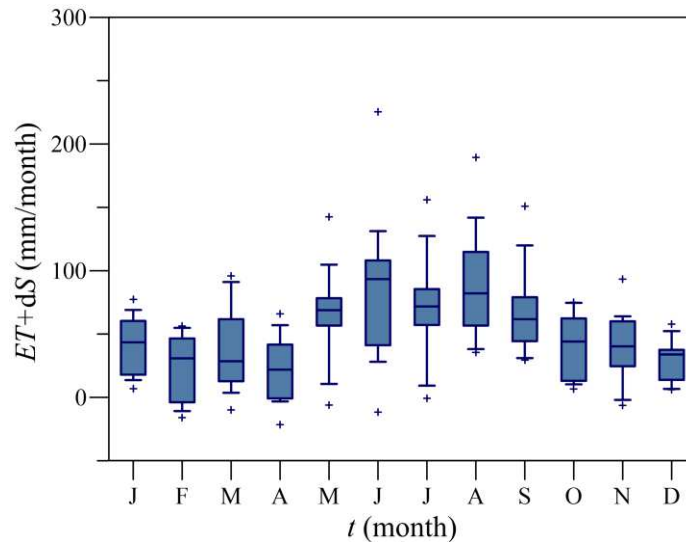


Figure 2. Monthly evapotranspiration (ET) and storage change (dS) estimated from the water balance in the HOAL for the period 2002-2015.

One of the main features of the HOAL is the wide range of observed runoff generation mechanisms, such as infiltration excess overland flow, re-infiltration of overland flow, saturation excess runoff from wetlands, tile drainage flow, shallow aquifer seepage flow and groundwater discharge from springs (Blöschl et al., 2016). Due to the thin soil layers with low permeability and the agricultural land use, subsurface tile drainage systems were installed in the 1950s to eliminate waterlogging (Figure 1a). These systems contribute to the stream at five points: Sys1 and Sys2 flow throughout the whole year, while Frau1, Frau2 and Sys3 are ephemeral. Sys3 behaves as a combination of a tile drain and a wetland during low flow conditions as it also collects the water from the neighboring highly saturated parts of the catchment. Exner-Kittridge et al. (2016) found that, chemically and dynamically, the water at Sys1 tile drain originates from the deep aquifer and not the shallow aquifer, therefore Sys1 behaves like a spring rather than a drainage system. The upper 25% of the stream length was piped in the 1950s to expand the agricultural area. The concrete pipe enters the main stream at inlet Sys4. The dynamics of Sys4 are similar to the perennial drains. A perennial spring, Q1 originates from a fractured siltstone aquifer and directly enters the main stream. Two wetlands (A1 and A2) are fed by springs which seep into the stream through rivulets in the south-eastern part of the catchment. Due to their high saturation, the runoff response of the wetlands is fast. During major storm events, saturation overland flow laterally enters the stream at E1 and E2 erosion gullies, and potentially, a few other points, depending on the event magnitude, soil moisture state and the direction of the ploughing. As both the main stream and the tributaries draining nearly all the surface contributing area to the stream are gauged (12 gauging stations in total), the main stream and tributary contributions can be separated.

The estimated drainage area, the percentage of the total drainage area occupied by the riparian forest, the forest cover, the mean streamflow and the mean low flow episode streamflow (according to sections 3.1 and 4.1) are summarized in Table 1 for each gauge in the HOAL. Comparing the mean streamflow at MW catchment outlet with the sum of the tributaries for the time period 2013-2015, the tributaries contribute approximately 56% to the total runoff at MW outlet. The contribution of the tributaries to the main stream is higher in the low flow periods (63%). This

Separation of scales in transpiration effects on low flows – A spatial analysis in the Hydrological Open Air Laboratory

means that about 40% of the streamflow observed at MW enters the stream laterally in a diffuse way, mostly through the subsurface.

Table 1. Estimated drainage area, proportion of the catchment drainage area occupied by riparian forest, forest cover, mean streamflow and mean low flow episode streamflow for 2013-2015 (for the MW catchment outlet streamflow data for 2002-2015 are in brackets) for each gauge in the HOAL (Figure 1 shows the location of the gauges).

Gauge	Runoff generation mechanism	Estimated drainage area (ha)	Riparian zone (%)	Forest cover (ha)	Mean streamflow (l/s)	Mean low flow episode streamflow (l/s)
MW	Outlet	65.8	3.5	6.32	4.23 (3.94)	3.12 (2.82)
Sys4	Inlet pipe	37.4	0.2	1.73	0.74	0.62
Frau1	Tile drain	3.1	0.1	0.00	0.03	0.00
Frau2	Tile drain	4.8	0.3	0.01	0.15	0.09
Sys1	Tile drain (deep aquifer)	6.5	5.0	0.77	0.43	0.41
Sys2	Tile drain	2.4	0.7	0.45	0.18	0.16
Sys3	Tile drain/Wetland	4.3	0.1	0.61	0.09	0.06
A1	Wetland	2.1	1.6	0.25	0.09	0.09
A2	Wetland	1.1	7.0	0.17	0.09	0.06
Q1	Deep aquifer	2.0	0.8	0.02	0.46	0.46
E1	Erosion gully	0.8	0.9	0.01	0.01	0.00
E2	Erosion gully	1.0	0.2	0.00	0.09	0.00

2.5.2. Data

Streamflow has been monitored at MW outlet of the catchment by a calibrated H-flume with a pressure transducer since 2001 with one-minute temporal resolution (Figure 1). Streamflow measurements at the tributaries started in 2011. At Sys4 inlet, at Frau1, Frau2 and Sys2 tile drainage systems, Sys1 tile drain (deep aquifer), Sys3 tile drain/wetland, at the wetlands (A1 and A2) and at the erosion gullies (E1 and E2) H-flumes and pressure transducers are used to monitor the flow; at Q1 spring a V-notch weir and a pressure transducer are used to measure the flow. Details on the sensors are given in Blöschl et al. (2016).

Two time periods were selected for the analysis because of the differences in the record lengths between the main outlet and the tributaries. While the streamflow fluctuations at the main outlet were analyzed for the period 2002-2015, the fluctuations at the tributaries were analyzed for the period 2013-2015. Typical examples of streamflow fluctuations at the main outlet (MW) during a low flow period in summer and in autumn 2006 are presented in Figure 3. The relatively long rainless period during June 2006 resulted in a recession of the streamflow, gradually decreasing from 5 to 3 l/s over twenty-one days. Due to evapotranspiration, diel streamflow fluctuations were superimposed on the recession curve with an average amplitude of 0.9 l/s (Figure 3a). In the middle of October 2006, the baseflow was around 1.7 l/s and the amplitude of the diel streamflow fluctuations was around 0.4 l/s. Following a storm event (28-31 October), the air temperature dropped below freezing point and the diel streamflow fluctuations stopped. Even though later the air temperature increased above freezing point, the diel streamflow fluctuations did not resume during the rest of the year (Figure 3b).

Nineteen piezometers located in the riparian forest close to the stream have monitored the groundwater level at a resolution of five minutes since 2013. Consistent with the streamflow records, groundwater level fluctuations were analyzed for the period 2013-2015. Five piezometers at different distances from the stream were selected for the analyses. The average depth to the groundwater table between 2013-15 was larger further away from the stream: 0.3 m at BP02 (0.3 m from the stream), 0.4 m at BP07 and H04 (1.6 and 1.4 m from the stream respectively), 2.7 m at H02 (7.4 m from the stream) and 4.3 m at H01 (14.7 m from the stream).

Rainfall has been measured with high temporal resolution (one minute) since 2002 by a weighing rain gauge situated 200 m from the catchment outlet. In 2012 four additional weighing rain gauges were installed in the HOAL (Figure 1).

Since 1986 air temperature has been recorded at 7, 14 and 19h by a thermometer and cumulative daily solar radiation has been measured with a pyranometer located about 500 m from the catchment outlet. Since October 2012 air temperature, incoming and outgoing solar and long wave radiation have been measured at the HOAL weather station at one minute temporal resolution. Three eddy covariance stations have measured crop evapotranspiration in the HOAL since August 2012.

Separation of scales in transpiration effects on low flows – A spatial analysis in the Hydrological Open Air Laboratory

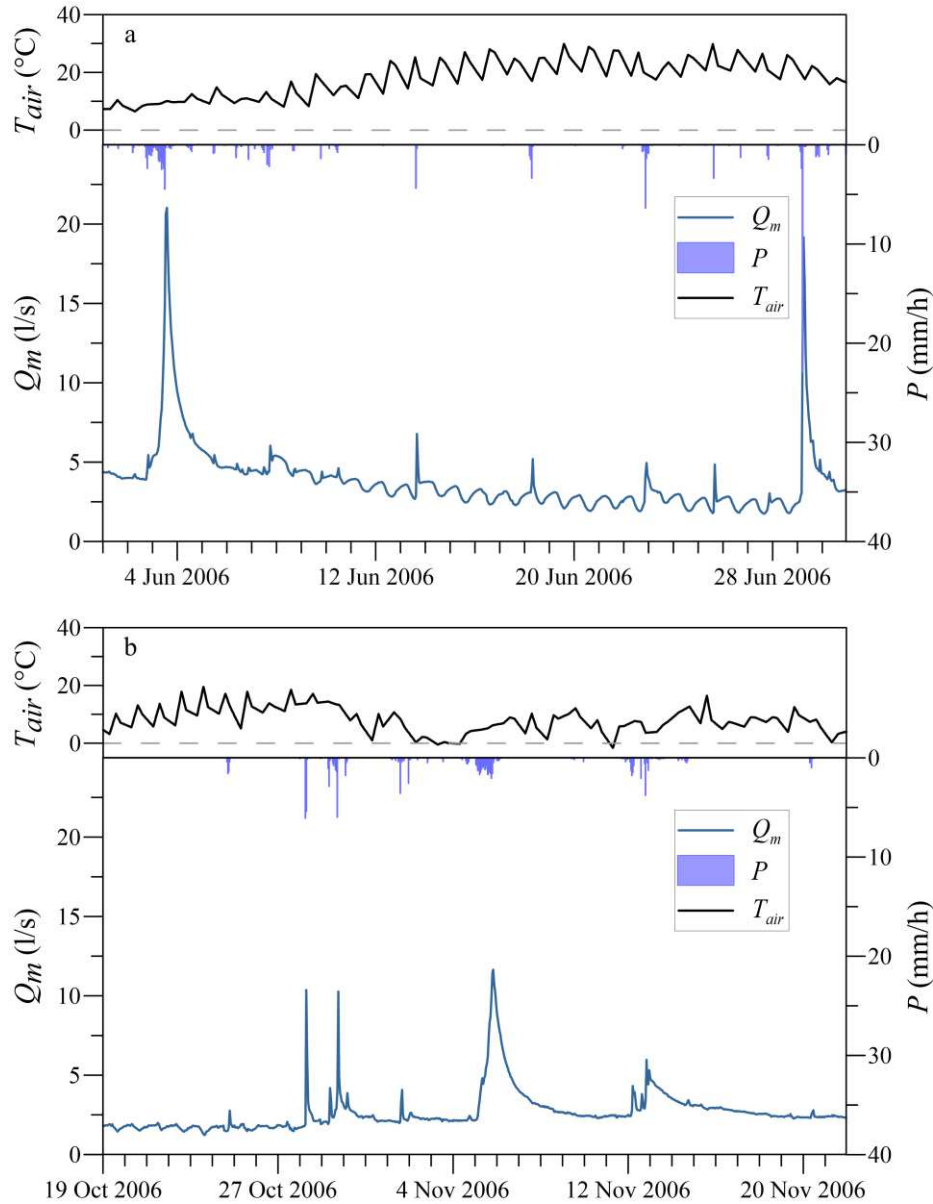


Figure 3. Example of air temperature (T_{air}) interpolated between 7, 14 and 19h, rainfall (P) and streamflow (Q_m) fluctuations at the main outlet (MW) of the HOAL in (panel a) June 2006 and (panel b) October-November 2006.

Streamflow, groundwater level, rainfall and solar radiation data were aggregated to hourly values for the analysis. Test simulations indicated that the difference between hourly and shorter time steps was negligible during low flows. For the time period 2002-2012, hourly solar radiation on a horizontal surface was estimated from measured cumulative daily solar radiation as a function of the day of the year, the solar time and the latitude (Duffie and Beckman, 2013).

2.6. Methodology

2.6.1. Identification of periods with streamflow fluctuations

In order to analyze the diel streamflow fluctuations, periods during streamflow recession with diel streamflow and groundwater level signals were identified for the main outlet, for all tributaries and for two piezometers (one on the left, the other on the right side of the stream) based on four criteria:

- (i) the day and the night preceding the day had no precipitation,
- (ii) the difference between the daily minimum and maximum streamflow or groundwater level exceeded a limit (see Table A1.1 in Appendix A1),
- (iii) the daily minimum streamflow or groundwater level was observed between 8 am and 7 pm,
- (iv) the daily maximum streamflow or groundwater level was observed either before 3 pm or after 9 pm.

For the selected episodes, the amplitude was estimated as the difference between the daily minimum and maximum streamflow or groundwater level (see Appendix A2). If an episode was identified at the main outlet, the streamflow records of the tributaries and the groundwater levels of the two piezometers were checked with regards to the presence of fluctuations. If one or more of the criteria above were not matched at the tributaries or at the piezometers, the episode was considered to have zero amplitude, otherwise the amplitudes were evaluated in a similar way as for the main outlet. The mean streamflow of the episodes was calculated for the same time periods for all streamflow gauges and became zero only if a tributary dried out.

2.6.2. Modeling of streamflow fluctuations

The study proceeded along two approaches of analyzing streamflow fluctuations, i.e. direct analyses and model simulations. The spatio-temporal patterns and the seasonal variability of the low flow fluctuations were described through direct analyses of the observations. The model simulations aimed at estimating the time lags of the streamflow response relative to its forcing. A new modelling approach based on a simple impulse response model was used assuming that

- (i) the amplitude of the diel streamflow fluctuations is proportional to incoming shortwave solar radiation, i.e., solar radiation can be used as a proxy for transpiration (e.g. Renner et al., 2016),
- (ii) the temporal pattern of the diel streamflow fluctuations can be modeled by an exponential response function to solar radiation,
- (iii) the main recession trend during the low flow period is exponential.

The model has three free parameters (f , λ and α). The parameter f expresses the proportion of the maximum available energy in the entire catchment or subcatchment, which influences the amplitude of the diel streamflow fluctuations, assuming that the energy consumed by evapotranspiration is equal to incoming shortwave solar radiation. If this assumption is true, f is equal to one, if the entire maximum catchment energy contributes to diel streamflow variations. The parameter λ represents the time lag between incoming shortwave solar radiation and the diel streamflow fluctuations. The parameter α is the recession time scale. If α is significantly larger

than λ , i.e. there is a magnitude difference between the two time scales, then a separation of scales exists in the time domain.

Based on assumption (i) evapotranspiration ET is estimated from the incoming shortwave solar radiation according to (1)

$$ET = f \cdot S \cdot \frac{G}{\rho \Delta H_{vap}} \quad (1)$$

where ET is evapotranspiration in (L^3T^{-1}), f is the dimensionless amplitude factor, S (L^2) is the drainage area of the entire catchment or subcatchment, G (MT^{-3}) is incoming shortwave solar radiation, ρ (ML^{-3}) is the density of water, ΔH_{vap} (L^2T^{-2}) is the latent heat of vaporization of water. This approach assumes no plant regulation of evapotranspiration. Based on assumption (ii), the evapotranspiration pattern is convoluted with an exponential response function to obtain the evapotranspiration signal in the hydrograph Q_d (L^3T^{-1}) according to (2)

$$Q_d(t) = \int_0^t ET(\tau)u(t - \tau)d\tau \quad (2)$$

where τ (T) is the integration variable, and u_t (T^{-1}) is the response function according to (3)

$$u_t = \frac{1}{\lambda} e^{-\frac{t}{\lambda}} \quad (3)$$

where λ (T) is the time lag and t (T) is time. Based on assumption (iii), the recession curve Q_r (L^3T^{-1}) is expressed as an exponential function (4)

$$Q_r = Q_0 e^{-\frac{t}{\alpha}} \quad (4)$$

where α (T) is the recession time scale and Q_0 (L^3T^{-1}) is set to the maximum measured discharge during the first day of the time period analyzed. Streamflow Q (L^3T^{-1}) is calculated according to (5), subtracting the evapotranspiration signal from the recession curve

$$Q = Q_r - Q_d \quad (5)$$

It is important to note that the model is not a predictive model. It was fitted to the streamflow data using a multiple objective calibration approach (combining the root mean square error, the amplitude error and the error of timing, see Appendix A3) for the purpose of interpreting the streamflow fluctuations and low flow recessions at different times of the year. As the model was fitted to each recession period independently, the result was one parameter set (f , λ and α) for each episode.

The modelled evapotranspiration volumes were compared with upscaled, literature based riparian evapotranspiration volumes (see Appendix A4 and Appendix A5). The groundwater evapotranspiration, i.e. the evapotranspiration calculated from the diel fluctuations of the shallow groundwater levels, was estimated by the White method (White, 1932) and the empirical method of Gribovszki et al. (2008) (see Appendix A6).

2.7. Results

2.7.1. Amplitude of the observed diel signals

The number of episodes with diel variations at the main outlet was 549 and 138 for the two time periods (Table 2). Depending on the runoff generation mechanism, the streamflow data of some tributaries featured diel signals simultaneously with the catchment outlet (approximately 60% of the episodes for the wetlands), while other tributaries were dry during long rainless periods (for instance the erosion gullies E1 and E2). While all periods were used for the direct analyses, periods shorter than three days were excluded from the model simulations (Table 2). The mean lengths of the episodes for the direct analyses and model simulations were about 60 and 110 hours, respectively.

Table 2. Number of episodes and days included in the direct analyses and the model simulations and mean episode lengths. Episodes shorter than three days were excluded from the model simulations.

Time period	Gauges	Direct analyses		
		Number of episodes	Total number of days in episodes	Mean episode length (hrs)
2002-2015	MW	549	1364	57
2013-2015	MW and tributaries	138	344	56
Time period	Gauges	Model simulations		
		Number of episodes	Total number of days in episodes	Mean episode length (hrs)
2002-2015	MW	185	832	108
2013-2015	MW and tributaries	42	197	113

The monthly average amplitudes of the streamflow diel signals at the MW catchment outlet and diel fluctuations of the groundwater levels at two piezometers (on the left bank, 1.6 m from the stream at BP07 and on the right bank, 1.4 m away from the stream at H04) show a clear seasonal pattern (Figure 4a and c). The amplitudes of the diel streamflow and groundwater level fluctuations increased from spring to summer. During the summer season (May to September) the amplitude of the streamflow fluctuations was usually larger than 0.4 l/s at MW outlet, while during the winter season (November to March) the amplitudes approached zero. The amplitude is controlled by both the energy input, which peaks in the summer, and the efficiency of the plants as they allocate the available energy to transpiration. The very small values of the amplitudes in the late autumn and

Separation of scales in transpiration effects on low flows – A spatial analysis in the Hydrological Open Air Laboratory

winter months suggest that the efficiency is the lowest in the winter season. This is consistent with observations (e.g. Figure 3b) that the first late autumn frost terminates the diel streamflow fluctuations.

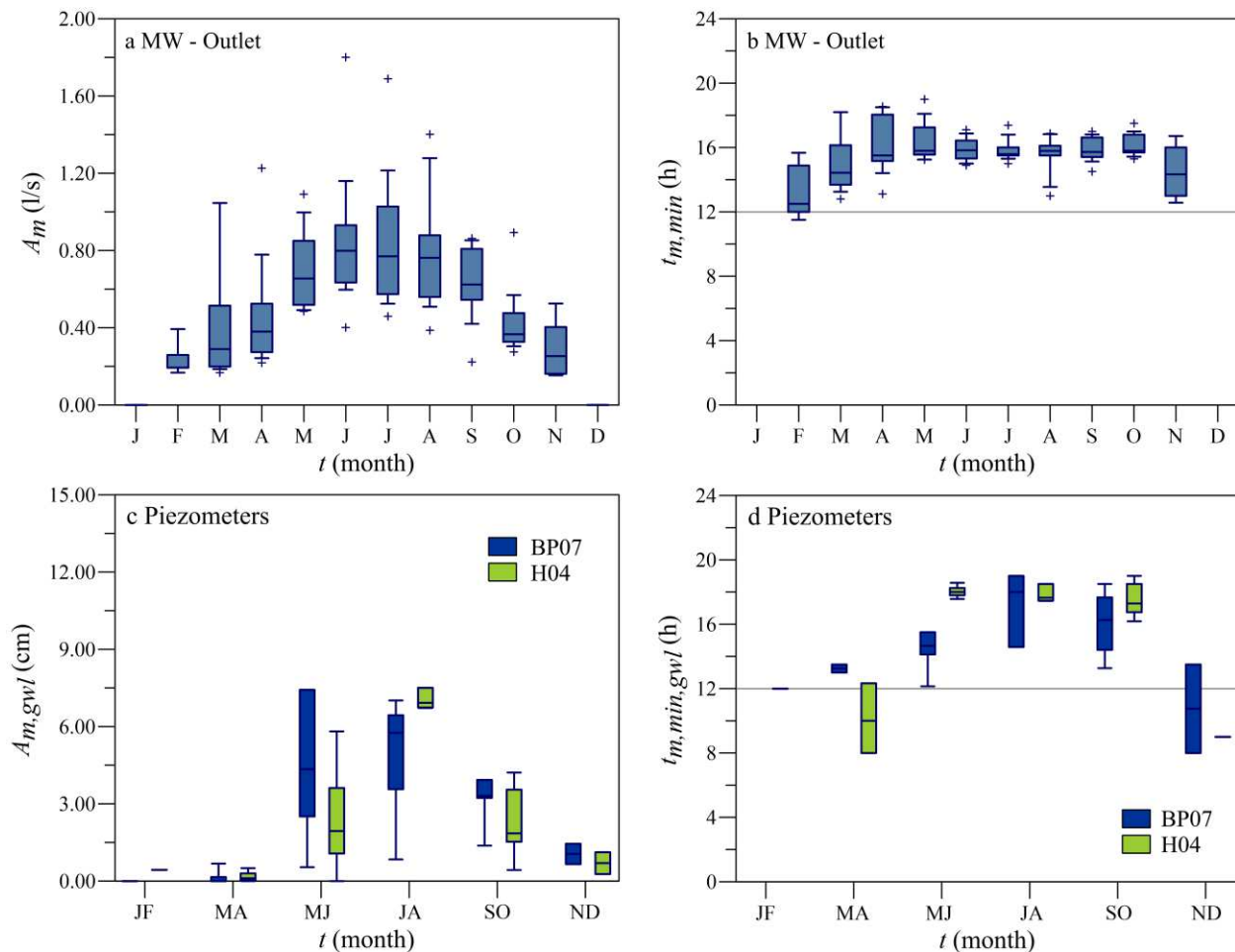


Figure 4. Seasonal variability of the monthly average measured amplitudes A_m of the streamflow diel signals (panel a) and the time of the day when the minimum measured discharge occurs $t_{m,min}$ (panel b) at MW catchment outlet based on observations during 2002-2015; seasonal variability of the monthly average measured amplitudes $A_{m,gwl}$ of the groundwater level diel signals (panel c) and the time of the day when the minimum measured groundwater level occurs $t_{m,min,gwl}$ (panel d) at BP07 and H04 piezometers based on observations during 2013-2015. For the piezometers the results have been lumped into bimonthly bins because of the shorter observation period. Direct analyses, see Appendix A2.

The magnitudes of the streamflow fluctuations varied between the locations (Figure 5 and Figure 6). During a five-day dry period in August 2013 diel signals were observed for most of the tributaries and piezometers in the HOAL, although with very different magnitudes (Figure 5). The differences between the tributaries are highlighted in Figure 6 which shows the relative amplitudes, i.e. the mean measured amplitudes as a function of the mean streamflow of the episodes at the nine gauges of the catchment that usually have non-zero runoff during rainless periods. The other three gauges, the ephemeral Frau1 tile drain and the erosion gullies (E1 and E2) are not included in Figure 6 because they are always dry during rainless periods.

Separation of scales in transpiration effects on low flows – A spatial analysis in the Hydrological Open Air Laboratory

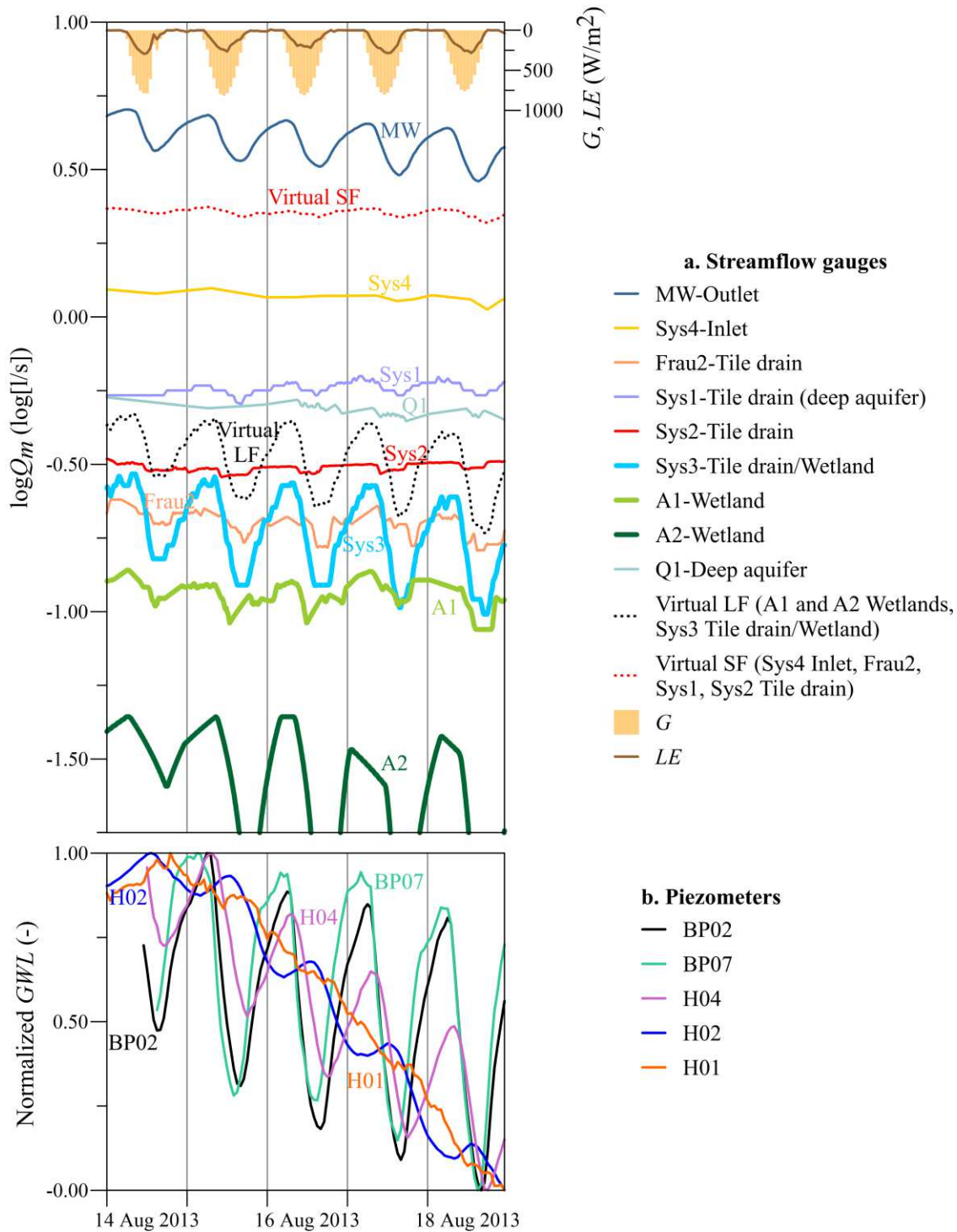


Figure 5. A five-day episode in the HOAL, 14-18 August 2013. (Panel a) Measured runoff is shown as solid lines, two virtual gauges (i.e. combination of gauges, LF with large amplitudes and SF with small amplitudes) are shown as dotted lines. Incoming shortwave radiation G and latent heat of vaporization LE at the weather station are shown at the top. (Panel b) Normalized (between 0 and 1) groundwater levels (GWL) of five piezometers (BP02, H04, BP07, H02, H01, Figure 1).

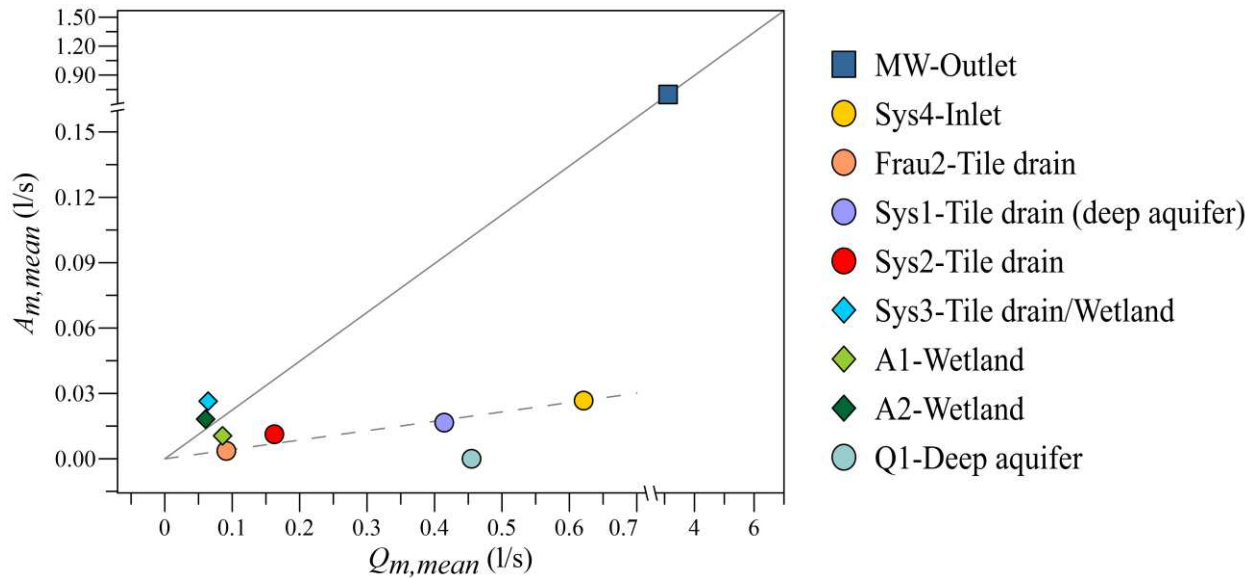


Figure 6. Mean measured streamflow amplitudes ($A_{m,mean}$) as a function of the mean measured streamflow ($Q_{m,mean}$) of the episodes at nine stream gauges. Solid line indicates the relative low flow amplitudes of the MW outlet, dashed line indicates the trend of the relative average low flow amplitudes of the tributaries.

In order to understand from which tributaries the fluctuations detected at MW catchment outlet originate, vegetation cover, dominant soil types, depths to groundwater level and runoff generation mechanisms were explored for the respective subcatchments. Sys3 tile drain/wetland is located on the left side of the stream, where the depth to groundwater level is small. The drainage area of Sys3 tile drain/wetland is dominated by Kolluvisol around the outlet; the soil texture is characterized by a large percentage of silty loam causing low permeability and generally wet soil conditions. The drainage area around the outlet is covered by forest, the dominant tree types are ash, poplar and black alder. The root system of these tree species in moisture retaining and organic rich soil types can reach depths of 2 m (Crow, 2005), therefore they are apparently exceptionally well-connected to the shallow groundwater table. Similarly to Sys3 tile drain/wetland, A1 and A2 wetlands were characterized by diel fluctuations with large magnitudes (Figure 5a and Figure 6). The wetlands are also located on the left side of the stream and due to their high wetness conditions and large riparian forest cover, the trees can be expected to be well-connected to the shallow groundwater table as well.

The relative amplitude was one magnitude smaller at Sys4 inlet and Sys1, Sys2, Frau2 tile drains compared to the wetlands (Figure 5a and Figure 6). One of the reasons for the smaller relative amplitudes is the different vegetation cover (Table 1). The catchment area of Sys4 inlet and Frau2 tile drains are covered mainly by crop, the forest cover is minimal (Table 1). The fluctuations at these tributaries are possibly caused by the narrow riparian forest zone close to the outlets. Furthermore, Frau2 tile drain enters the main stream in an oblique way from the crop fields so is more exposed to riparian vegetation, unlike Frau1 tile drain which enters perpendicularly and is dry in rainless periods. Even though Sys1 tile drain behaves as a spring according to chemical analyses and it has basically constant contribution to the outflow at the catchment outlet even in the driest months, the root zone of the riparian forest may reach the aquifer close to the outlet, causing diel fluctuations of the streamflow. At Sys1 tile drain (deep aquifer) the relative amplitudes

of the fluctuations induced by the riparian vegetation were smaller than the relative amplitudes at the wetlands and they were superimposed on the high baseflow rates. The deep aquifer spring Q1, which produces runoff throughout the year from a fractured siltstone layer is not exposed to the local effect of riparian vegetation, hence diel variations of the outflow were never observed (Figure 5a and Figure 6).

Based on the magnitudes of the observed relative amplitudes and the process reasoning above, the tributaries in the HOAL were separated into two distinct groups, which were lumped into two virtual gauges by taking the sum of the measured streamflow rates. Virtual gauge LF consists of three gauges (A1 and A2 wetlands, Sys3 tile drain/wetland) with large amplitudes relative to the streamflow rates. Virtual gauge SF consists of four gauges (Sys4 inlet pipe, Frau2, Sys1 and Sys2 tile drainage systems) with small or no diel fluctuations in streamflow (Figure 5a).

MW catchment outlet integrates the characteristics of the different runoff mechanisms and therefore it can be regarded as a mixture of the different systems. However, Figure 6 suggests that the relative fluctuations at MW (about one fifth of the average low flow, solid line) were much larger than the average of all the tributaries (dashed line). This means there must be significant additional mechanisms that give rise to the observed fluctuations at MW. About 40% of the streamflow observed at MW enters the stream laterally in a diffuse way (Table 1), mostly through the subsurface, and is not captured by the tributary gauges, which suggests that the remaining part of the fluctuations is related to these diffusive inflows. These are the areas of significant riparian vegetation.

For comparison, Figure 5 (b) shows the groundwater levels normalized to the minimum and maximum levels during the episode for five piezometers (Appendix A7), one of them located on the left bank, 1.6 m from the stream (BP07) and four of them on the right bank (BP02, H04, H02 and H01, which are installed 0.3 m, 1.4 m, 7.4 m and 14.7 m from the stream respectively). The mean amplitude of the measured groundwater level fluctuations during the five-day recession period in August 2013 (Figure 5) was 16.9 cm at BP07 piezometer, while no diel fluctuations were observed further away from the stream at H01 piezometer. The amplitude of the groundwater table fluctuations was 13 times larger at BP02 piezometer (0.3 m from the stream), than the water level fluctuations at MW Outlet which corresponds well to Gribovszki et al. (2013) and Szilagy et al. (2011).

The impact of distance from the stream on the magnitude of the diel fluctuations was further investigated (Figure 7). Similarly to Figure 5b, the mean amplitude of the measured groundwater level fluctuations through the episodes decreased with larger distance from the stream. While the mean amplitude was 6.0 cm at BP02 piezometer 0.3 m from the stream, at H01 piezometer (14.7 m from the stream) it was 0.1 cm (Figure 7a). These results indicate that the signal which is propagated into the streamflow decreased with larger distances from the stream, where the depth to the groundwater level was larger (Figure 7b).

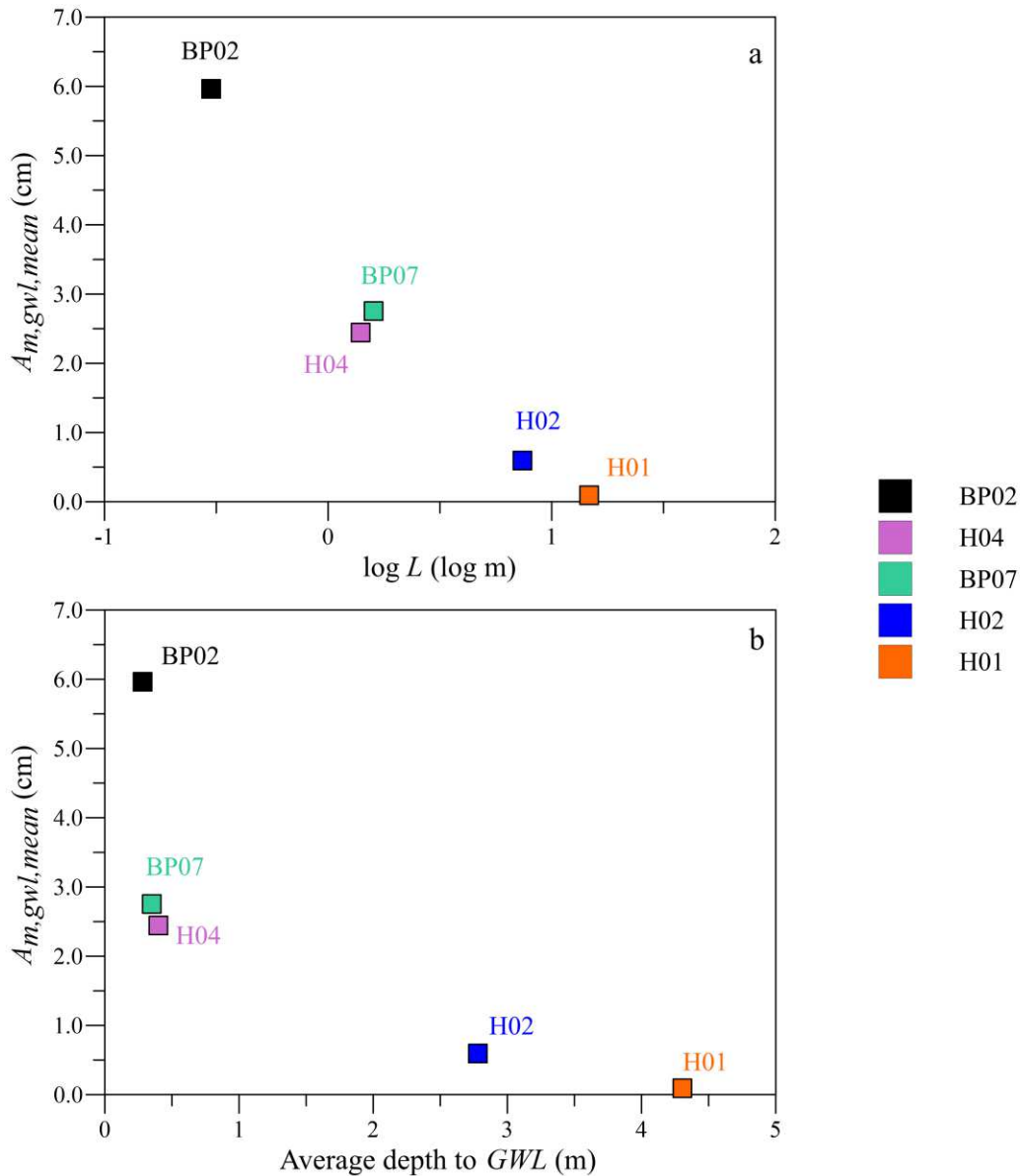


Figure 7. Mean amplitude of the measured groundwater level fluctuations ($A_{m,gwl,mean}$) as a function of (a) logarithmic distance ($\log L$) from the stream and (b) average depth to the groundwater level.

2.7.2. Timing of the observed diel signals

The timing of the diel fluctuations, i.e. the time of the day when the minimum streamflow or groundwater level occurs, showed a clear seasonal pattern (Figure 4b and d). In February and March the daily minimum discharge and groundwater level occurred in the early afternoon. As the season progressed, the daily minimum discharge and groundwater level occurred later in the afternoon, at around 16:00 at MW outlet and around 18:00 at the piezometers. These changes may be related to the response times of the system to the energy input. This phenomena was analyzed in detail later by the model (section 4.5).

The timing of streamflow fluctuations was compared with the timing of latent heat of vaporization measured by the Eddy covariance technique at the weather station (Figure 5a at the top) and groundwater level fluctuations (Figure 5b). The latent heat of vaporization representative of the crop fields and grassland surrounding the weather station lagged behind the incoming solar radiation by 0.5-2 hours (Figure 5a at the top). The phase shift between streamflow and groundwater table fluctuations varied between locations. Piezometers were either in phase with MW, such as closest to the stream at BP02 piezometer (0.3 m away), while further away they changed earlier (BP07) or later (H02 and H04) than MW. According to other studies (e.g. Nachabe et al., 2005; Szilagyi et al., 2008), the signal is observed first in streamflow, followed by groundwater level, and finally by soil moisture.

2.7.3. Model performance

For the 185 recession periods between 2002 and 2015 (Table 2), the median of the model efficiency, i.e. Nash-Sutcliffe coefficient, for MW was 0.89, the 25th and 75th percentiles were 0.68 and 0.95 respectively (see Appendix A8). For autumn, the model efficiency was lower due to the smaller amplitudes and signal-to-noise ratios. Figure 8 and Figure 9 show examples of the model fit. A 17 day recession period in June 2009 had several overcast days (Figure 8). The significant differences in the amplitudes between the days were captured very well by the model. For example, the 11th of June was an overcast day with a diel amplitude of streamflow at MW of only 0.3 l/s compared to the other days with amplitudes of about 1.1 l/s. During a five-day period in August 2013 (Figure 9), diel signals were observed at most of the tributaries. Virtual gauge LF (A1 and A2 wetlands and Sys3 tile drain/wetland) was characterized by large amplitudes, while amplitudes at virtual gauge SF (Sys4 inlet, Frau2, Sys1 and Sys2 tile drainage systems) were small.

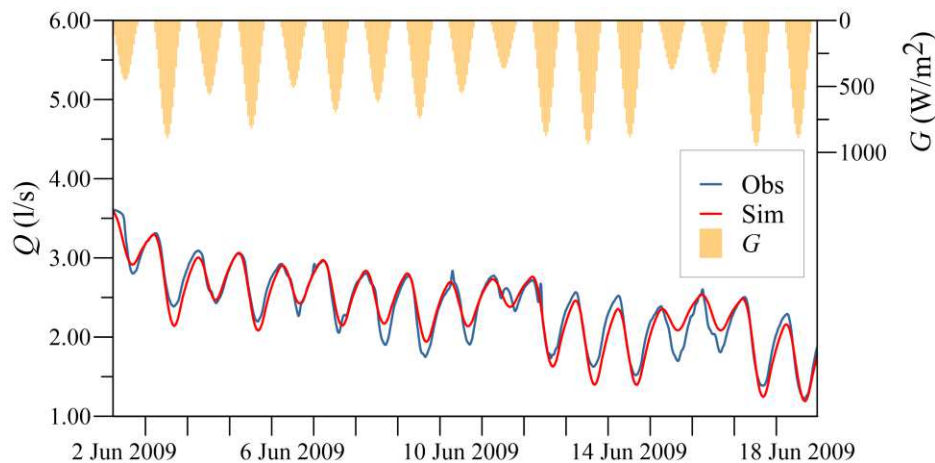


Figure 8. Observed (blue line) and simulated (red line) streamflow fluctuations Q at the MW catchment outlet in the period 2-18 June 2009. Incoming shortwave radiation G is shown at the top.

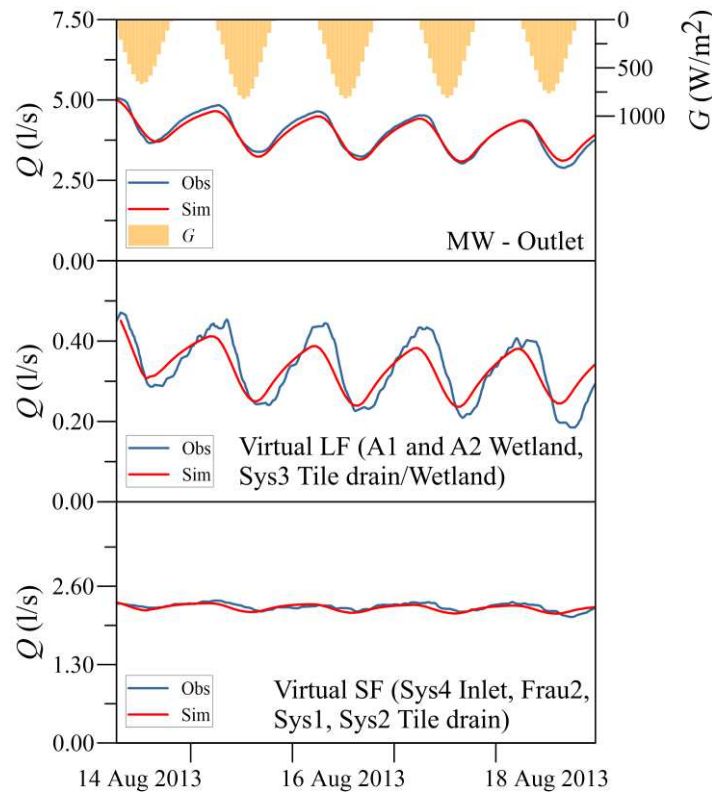


Figure 9. Observed (blue line) and simulated (red line) streamflow Q for 14-18 August 2013. The tributaries are lumped into virtual gauge LF (A1 and A2 wetlands, Sys3 tile drain/wetland) and virtual gauge SF (Sys4 inlet, Frau2, Sys1 and Sys2 tile drainage systems). Incoming shortwave radiation G is shown at the top.

2.7.4. Process controls on amplitude of diel streamflow variation

In order to describe the dynamics of the diel low flow fluctuations, we analyzed the seasonal evolution of the calibrated model parameters. Figure 10 (a, c) shows the seasonal variability of the amplitude factor f at MW catchment outlet (a) and at Virtual gauge LF (c). Amplitude factor f expresses the proportion of maximum available energy in the catchment which affects the diel streamflow variations, if the energy consumed by evapotranspiration is equal to incoming shortwave solar radiation. The amplitude factor increased from early spring until the beginning of summer. Assuming that all the entire incoming shortwave radiation is allocated to transpiration, $f=0.004$ in April and 0.008 in the summer would imply that only a small portion of the available energy in the catchment (0.004 and 0.008 respectively) contributes to the component of transpiration that causes the diel streamflow fluctuations.

The seasonal variability of the amplitude factor f at Virtual gauge LF (A1, A2 wetlands and Sys3 tile drain/wetland) was similar to the value for the MW catchment outlet (Figure 10c). It increased from early spring (median of 0.003) until the summer (median of 0.009) and decreased in the autumn (median of 0.002). At Virtual gauge LF the mean of the amplitude factor f was 0.0081 , five times larger than at Virtual gauge SF (0.0018), where the relative amplitude of the fluctuations was one magnitude smaller.

Figure 10 (b, d) shows the calibrated amplitude factor f as a function of the maximum measured discharge Q_0 on the first day of each recession period. Amplitude factor f increased with discharge both for MW catchment outlet (Figure 10b) and even more clearly for Virtual gauge LF (Figure 10d).

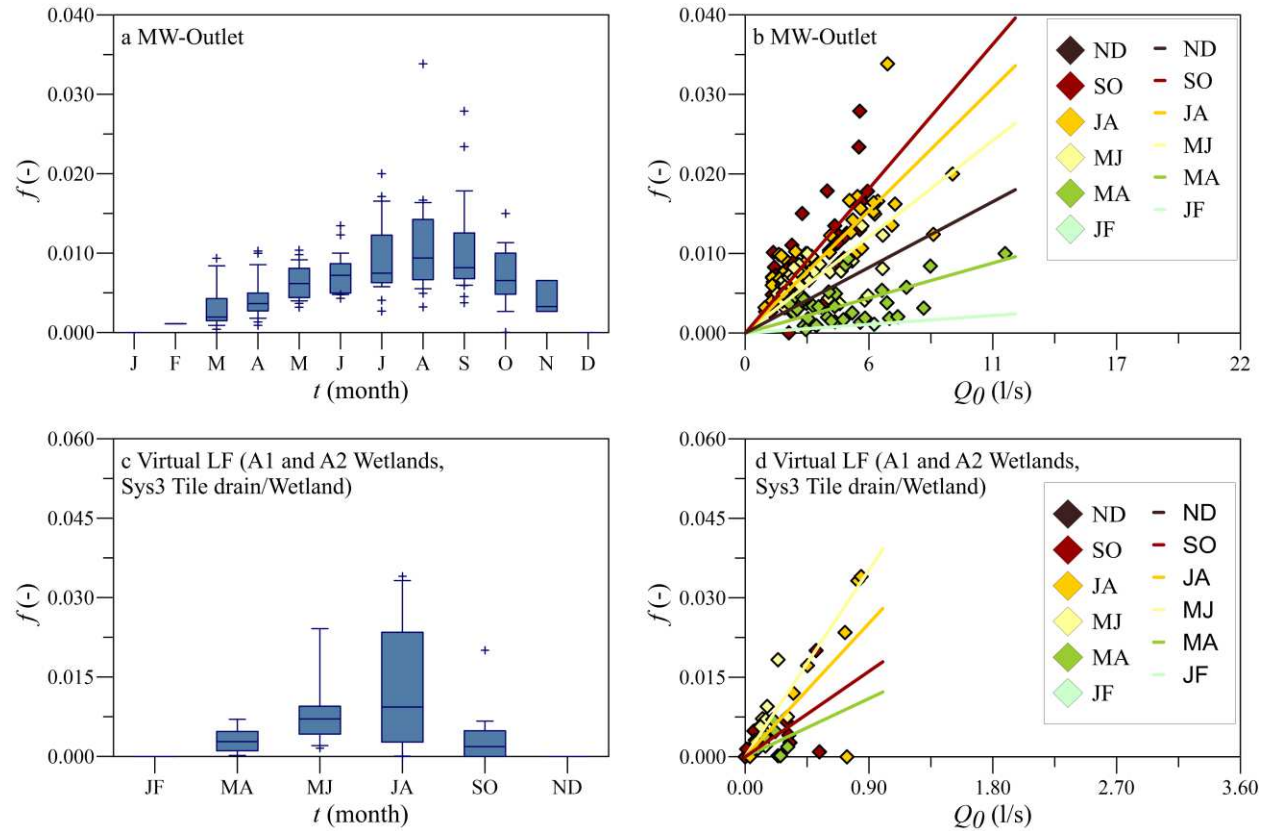


Figure 10. Seasonal evolution of the calibrated amplitude factor f (panels a, c), and f as a function of the maximum measured discharge Q_0 on the first day of each recession period with fitted linear regression lines (panels b, d) for MW catchment outlet (panels a, b) and for Virtual gauge LF comprising the three gauges (A1 and A2 wetlands, Sys3 tile drain/wetland) with large amplitudes (panels c, d). For LF the results have been lumped into bimonthly bins because of the small sample size.

The amplitude factor of the model comprises several factors influencing the diel streamflow fluctuations. One of these factors is the efficiency of the vegetation in root water uptake, which we estimated using Eddy covariance measurements. Although the flux footprints of the Eddy covariance stations located in the crop fields do not cover the riparian zone close to the stream, we estimated the proportion of the incoming solar radiation which is allocated for transpiration in the crop fields under the climatic conditions in the HOAL. The seasonal variability of the ratio of daily evapotranspiration (expressed as latent heat) and incoming shortwave radiation is shown in Figure 11 (a). Analyzing only the clear sky days between June and August (2012-2014), the average of the incoming shortwave radiation was approximately $25.3 \text{ MJ/m}^2/\text{d}$ (or 293 W/m^2 , 10 mm/d), and the latent heat measured in the crop fields was around $7.4 \text{ MJ/m}^2/\text{d}$ (or 86 W/m^2 , 3 mm/d). The LE/G ratio, i.e. the plants' efficiency of allocating energy to transpiration, increased from 0.1 in January to about 0.30 in summer (Figure 11a). This means that in summer only about 30% of the

incoming shortwave radiation was balanced by latent heat in the crop fields. Assuming similarity between tree and crop transpiration, the ratio of the calibrated model parameter f and the median of the LE/G ratios represents a measure of how well the trees are connected to the stream (Figure 11b). Amplitude factor f is essentially the ratio of evapotranspiration seen by the stream and G , so $f/(LE/G)$ is the ratio of evapotranspiration seen by the stream and the evapotranspiration measured by the eddy covariance method, and hence a measure of the proportion of the available energy in the entire catchment influencing the streamflow fluctuations at the diel time scale. Figure 11 (b) suggests that this proportion slightly increased from spring to summer, but the change was much smaller than that of the transpiration efficiency.

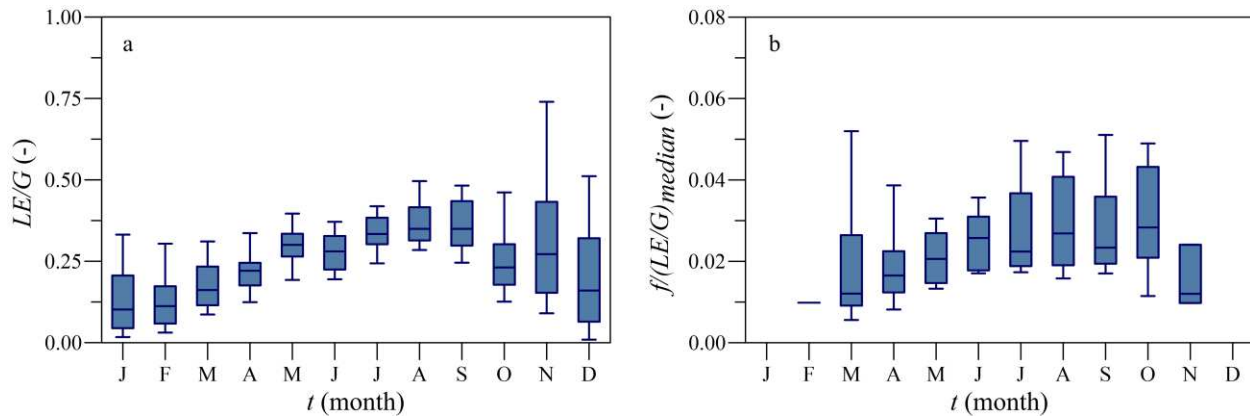


Figure 11. Seasonal variability of the ratio of daily latent heat LE measured by eddy covariance and daily incoming shortwave radiation G (panel a) and the ratio of the calibrated model parameter f and the median of the LE/G ratios (panel b).

2.7.5. Process controls on lag times of diel streamflow variation

The calibrated values of time lag λ at MW catchment outlet and Virtual gauge LF including only those periods when the model accurately reproduced the observed time series (Nash Sutcliffe coefficient >0.2 ; 90% and 69% of the modelled episodes were considered for MW outlet and Virtual gauge LF respectively) gradually increased through the year (Figure 12). In March the median of the time lag λ was around 3 hours, it increased to 8 hours in May and 11 hours in October. For Virtual gauge LF (A1, A2 wetlands and Sys3 tile drain/wetland) there was a similar increase (Figure 12c). The mean value of the time lag λ was higher at Virtual gauge SF (11.1 hours) than at Virtual gauge LF (10.3 hours) (Table 3).

The second lag time of the model is the recession time scale α . It operates at longer time scales than λ and is related to the catchment drainage and evapotranspiration as the catchment dries out during the recession periods. Figure 12 (b, d) shows the seasonal evolution of the calibrated values of recession time scale α at MW catchment outlet (b) and Virtual gauge LF (d) for only those periods when the Nash Sutcliffe Coefficient was larger than 0.2 and there was a clear recession in the hydrographs at MW catchment outlet (66% and 38% of the modelled episodes were considered for MW outlet and Virtual gauge LF respectively). A few episodes had increasing baseflow, which would imply a rainfall input not considered. Therefore these episodes were not considered in estimating α . There was a tendency for time scale α to gradually increase from spring to autumn, from a median value of 21 days in March to 54 days in October. According to Table 3 the mean

Separation of scales in transpiration effects on low flows – A spatial analysis in the Hydrological Open Air Laboratory

recession time scale α was four times larger at Virtual gauge SF than at Virtual gauge LF, where the systems are fed by deeper subsurface flow.

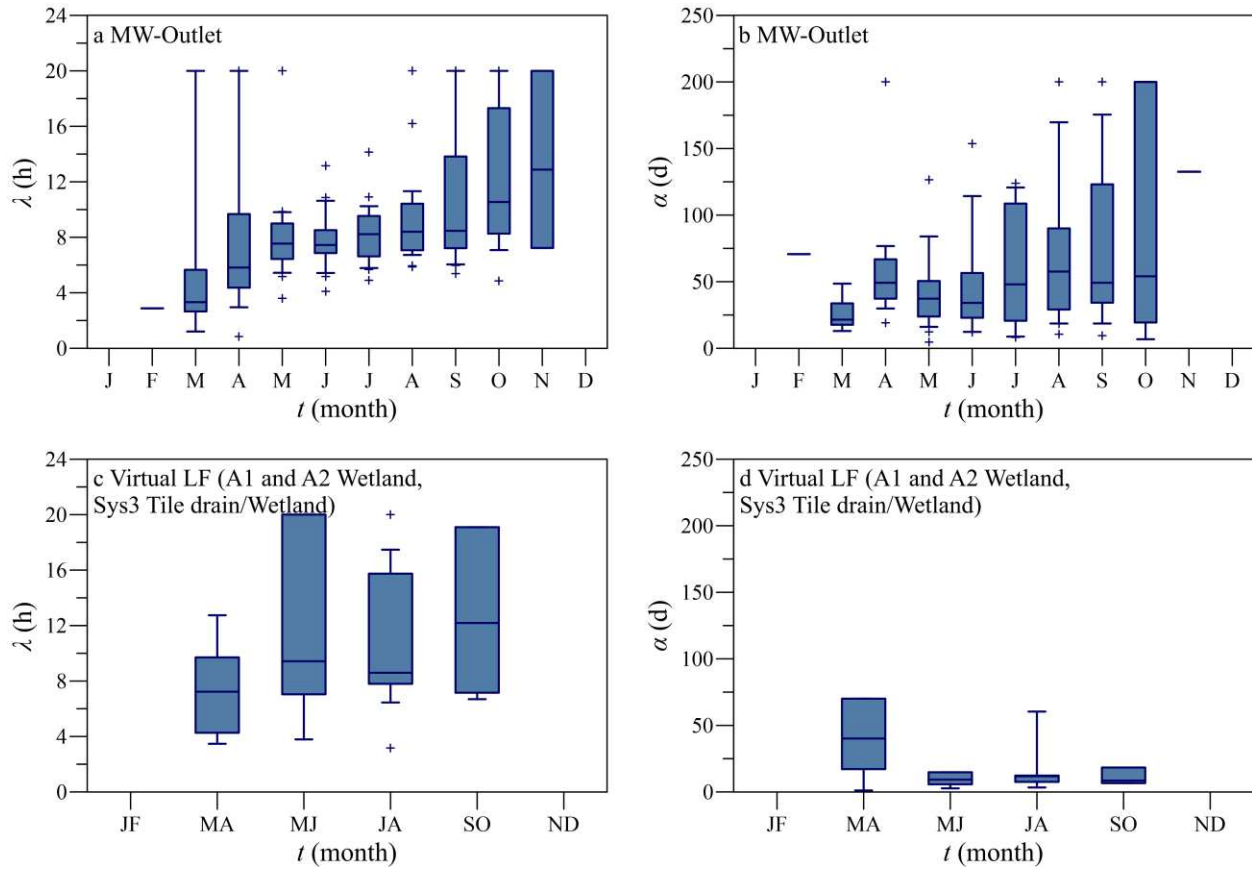


Figure 12. Seasonal evolution of the calibrated model parameter time lag λ (panels a, c) and α (panels b, d) for MW catchment outlet (panels a, b) and for Virtual gauge LF comprising the three gauges (A1 and A2 wetlands, Sys3 tile drain/wetland) with large amplitudes (panels c, d). For LF the results have been lumped into bimonthly bins because of the small sample size.

Table 3. Mean of the three calibrated model parameters for MW catchment outlet and the virtual gauges in the HOAL (Location of the gauges is shown in Figure 1).

Gauge	Time period	f (-)	λ (hrs)	α (days)
MW Catchment Outlet	2013-2015	0.0096	11.3	69.3
MW Catchment Outlet	2002-2015	0.0076	8.9	58.4
Virtual LF (A1 and A2 Wetland, Sys3 Tile drain/Wetland)	2013-2015	0.0081	10.3	19.8
Virtual SF (Sys4 Inlet, Frau2, Sys1, Sys2 Tile drain)	2013-2015	0.0018	11.1	77.6

2.8. Discussion

2.8.1. Spatio-temporal patterns of streamflow fluctuations

Several studies analyzed the spatial and temporal differences of diel streamflow fluctuations between catchments and in single or nested catchments, but focused only on spatially uniform runoff generation mechanisms. For instance, Lundquist and Cayan (2002) analyzed diel variations in 100 rivers in the western United States and found a weak correlation between the amplitude of the diel cycle and mean monthly temperature, discharge, basin area and mean basin elevation. They concluded that each catchment was unique and that the spatio-temporal patterns of diel streamflow signals were determined by local physiographic and hydrological characteristics.

The observation and modelling results presented in this study show that the spatial differences in runoff generation greatly influence the magnitude of the diel streamflow fluctuations. We found that the amplitude of the diel streamflow signal relative to the streamflow at the wetlands was one magnitude larger than those of the tile drains and the deep aquifers (Figure 5a and Figure 6). The wetlands are the areas in the catchment with high wetness conditions, shallow groundwater table and large riparian forest cover (Table 1). In contrast, the inlet pipe, the tile drainage systems and deep aquifers are fed by deeper water sources, covered mainly by crop fields and the depth to the groundwater level is larger than in the wetlands.

The modelling results also reflected the differences in the relative fluctuations between the tributaries. The amplitude factor f of the solar radiation driven model was larger in the wetlands, i.e. a larger proportion of the maximum catchment energy was allocated for transpiration which influenced the diel streamflow fluctuations. The recession time scale α of the model was longer for systems fed by deeper subsurface flow (tile drainage systems and deep aquifers) than for the shallower systems as would be expected.

2.8.2. Simplified process representation

In this study incoming shortwave radiation proved to be a useful proxy for representing streamflow fluctuations due to riparian transpiration (see section 4.2). Dominant controls on transpiration, i.e. stomatal (vapor pressure deficit and conductances) versus boundary layer (radiation) control, depend on the vegetation type, scale and meteorological conditions (Jarvis and McNaughton, 1986; Martin et al., 2001). Due to the energy which is advected in the form of vapor pressure deficit, daily latent heat flux can exceed the daily sum of net radiation (e.g. Hall et al. [1998] observed such a phenomena on a few, dry summer days). Therefore using net radiation as a driver of the diel streamflow fluctuations might introduce a bias on longer time scales. Some studies suggested the atmospheric moisture deficit as the main driver of transpiration (e.g. Granier et al., 2000; Szeftel, 2010) or a combination of both energy and vapor pressure deficit (e.g. different versions of the Penman-Monteith equation). However, other studies showed that transpiration of various tree species was closely related to solar radiation (e.g. Dragoni et al., 2005; Granier et al., 2000; Kume et al., 2008; Oguntunde and Oguntuase, 2007). For example, Pieruschka et al. (2010) suggested that the absorbed solar energy by the leaves influenced the stomatal control of transpiration, therefore solar radiation was an important control on transpiration. Using 110 FLUXNET eddy covariance sites, Boese et al. (2017) observed a substantial transpiration component, also termed as equilibrium transpiration (Jarvis and McNaughton, 1986), which was independent of stomatal conductance and driven by incoming solar radiation. Phillips et al. (1999) showed that diel sap flow in a Panamanian humid forest was more correlated with radiation than with atmospheric moisture deficit and Williams et al. (2004) found that transpiration was not correlated with atmospheric vapor pressure deficit in an olive orchard in Morocco. Renner et al. (2016) argued that vapor pressure deficit and wind speed, two variables widely used in evapotranspiration estimations, only slightly increased the predictability of atmospheric demand in a beech forest in Luxembourg.

In this study the evapotranspiration pattern was convoluted with an exponential response function which resulted in a hydrograph shape with convex rising limbs and concave recessions. Kovar and Bacinova (2015) used similar methods when they simulated the diel streamflow fluctuations with the Fourier Series Model. They applied both a linear and an exponential regression to simulate the depletion process and found a very small difference between the two approaches. Dvorakova et al. (2012, 2014) developed and calibrated a Linear Storage Model to describe the recession process, where the actual evapotranspiration was reproduced by simplified Fourier series or sine curves (Dvorakova et al., 2012, 2014). Similar to our results, these studies also showed that the depletion of the catchment storage during low flow conditions could be captured by an exponential function.

The diel fluctuations of the groundwater levels and streamflow can be caused by melting and freezing-thawing processes at the end of winter and early spring in this climatic region (Gribovszki et al., 2006; Gribovszki et al., 2010b; Lundquist & Cayan, 2002). On these days the soil temperature fluctuates around 0°C and the diel signal in this case is asymmetric to the transpiration induced diel signal (Lundquist & Cayan, 2002). Gribovszki et al. (2010b) pointed out that the diel fluctuations caused by melting was more pronounced in the streamflow compared to the groundwater levels. However, this was not the case in the HOAL (Figure 4). In the timing of the daily minimum streamflow there is no shift, i.e. in late winter and early spring the minimum streamflow rates did not have their minimum in the very early morning shown by Gribovszki et al. (2006) (Figure 4 b). Although at the piezometers this shift could be observed (e.g. at H04 in

March and April, and at both H04 and BP07 in November and December on Figure 4 d), this was most probably not caused by freezing-thawing processes considering that the water temperature at the piezometers was between 8-13°C in these months according to the drivers' water temperature measurements. This means that the late winter and early spring events could be analyzed together, and with the same methods as the other, transpiration induced events during the year.

2.8.3. Estimated evapotranspiration volumes and rates

The daily transpiration rates depend on the type of vegetation, structure, age and leaf area index (Farid et al., 2008; Schaeffer et al., 2000). Numerous studies compared evapotranspiration or transpiration rates of different vegetation cover using eddy covariance or sap flow measurements. For example, Granier et al. (2000) found summer transpiration rates of about 4.5 mm/day in a beech forest in France. Water abstraction of willows next to the stream can reach 5.6 mm/day (Marttila et al., 2017). Crop evapotranspiration, typically, is also on the order of 4 mm/day, depending on crop type (e.g. Delzon and Loustau, 2005).

The volumes of the diel streamflow fluctuations may be interpreted in two ways. The first and most common interpretation, as presented e.g. in Gribovszki et al. (2010b), is to attribute the entire missing volume in the hydrograph to evapotranspiration. This is the assumption underlying our estimations. In terms of volumes, independently from the LE/G ratio, by integrating equation (1) over the day and taking the average over the episode, the daily average summer (between June and August) evapotranspiration at MW, LF and LS was 69 m³/d, 8 m³/d and 8 m³/d respectively (Table 4). The tributary influences are expected to propagate synchronously along the main stream due to the short, approx. 50 and 80 min long lag times of celerity and velocity respectively between the upstream tributaries and MW outlet (Eder et al., 2014). Therefore this result means that 53 m³/d are not accounted for by the tributaries (LF+LS), so needs to be due to the diffusive subsurface inflow to the stream. In other words about 77% (53/69) of the volumes associated with the streamflow fluctuations are related to (hyporheic) exchange along the riparian zone of the main stream.

Table 4. Estimated average summer (June, July, August) evapotranspiration volumes.

Method	Reference gauge/piezometer	Time period	ET (m ³ /d)
Incoming shortwave radiation driven model	MW Outlet	2002-2015	55
Incoming shortwave radiation driven model	MW Outlet	2013-2015	69
Incoming shortwave radiation driven model	Virtual gauge LF (A1 and A2 Wetland, Sys3 Tile drain/Wetland)	2013-2015	8
Incoming shortwave radiation driven model	Virtual gauge SF (Sys4 Inlet, Frau2, Sys1, Sys2 Tile drain)	2013-2015	8
Upscaling literature based evapotranspiration values for the entire riparian zone	-	-	81

The second interpretation of the diel streamflow fluctuations is that subsurface flow is controlled by slight changes in the potential gradients that, in turn, are controlled by the diel cycle of evapotranspiration as the suction of the roots changes during the day. If this is the case, the summer evapotranspiration rates causing the diel streamflow fluctuations are smaller than the estimated ones. Szilágyi et al. (2008) distinguished between a local and an overall hydraulic gradient driving the water transport in the vadose and saturated zones during recession flow periods using a 2D finite element numerical model. Voltz et al. (2013) observed an overall relatively small response of the hydraulic gradients during a summer recession period in a steep headwater catchment in Oregon, where the ratio of the cross-to down-valley hydraulic gradient showed the largest diel fluctuations in wells closest to the stream. Given the topography, where the roots of riparian trees can easily reach into the groundwater, and groundwater levels in the riparian zones themselves fluctuate, this mechanisms is not likely important in the HOAL.

In order to compare the mean daily evapotranspiration rates from the main outlet with evapotranspiration rates estimated from the shallow groundwater level fluctuations (ET_G), the estimated evapotranspiration volumes (equation 1) were divided by the product of the catchment area and the calibrated amplitude factor f . The comparison indicates that the daily rates from diel streamflow fluctuations from the groundwater levels were 7 and 5 mm/day based on the empirical method of Gribovszki et al. (2008) and the White (1932) method respectively, which was slightly lower than the simulated rates for MW catchment outlet (9.5 mm/day). If the estimated evapotranspiration volumes (equation 1) were divided by the effective size of the riparian zone, which could be around 20000 m² (based on the finding that diel fluctuations were observed in the piezometers 7 meters away from the stream but not at 14 meters, which meant that the effective width of the riparian zone was around 10 meters in summer), the simulated rates for MW would be around 3.5 mm/d. These results are consistent with the range (8-11 mm/day) found by

Gribovszki et al. (2008) in an alder forest in Hungary. As the depth to the groundwater did not exceed 1 m (2013-2015) in the riparian zone within 0.5 m distance from the stream in the driest summer months, the root system of the trees (especially the 40-70 year old, >10 m species) are likely in direct contact with the saturated zone and evapotranspiration estimated from the diel groundwater level fluctuations could be close to groundwater evapotranspiration (Dawson, 1996; Shah et al., 2007; Williams et al., 2006).

Because the estimated evapotranspiration volumes could not be validated against measurements of sapflow and stomatal conductances, we estimated the transpired volumes for the entire riparian forest using aerial photographs, a tree survey and literature based transpiration values of different tree species (Table 4). The estimated summer transpiration rates ($81 \text{ m}^3/\text{d}$) show a good agreement with the modelling results (e.g. incoming shortwave radiation driven model set up for MW Outlet, for the time period 2013-15: $69 \text{ m}^3/\text{d}$). The slightly larger values are expected as the literature based estimate was calculated for the entire riparian zone (Table 4) which may not fully contribute to the streamflow fluctuations.

2.8.4. Separation of scales in time implies a separation of scales in space

The analysis of the streamflow fluctuations during low flow conditions at MW catchment outlet shows that the time lag λ between radiation and the diel low flow fluctuations gradually increased from 3 to 11 hours as the season progressed. The time lag represents the total response consisting of a cascade of responses, which includes the time lags between radiation and evaporation from the stomata, sap flow in the branches and the stem, root water movement, groundwater movement and groundwater – stream interactions (Figure 13). Each component has its own time lag. A number of studies found that the time lag between the diel fluctuations of radiation and sap flow in the tree was approximately 30 min for species such as apple trees (Dragoni et al., 2005), beech (Granier et al., 2000), and Japanese cedar (Kumagai et al., 2009). Gartner et al. (2009) found that the sap flow of birch and spruce in the south-eastern part of Austria lagged solar radiation by one hour during early August. When the soils dried out during a significant drought, the time lags increased to approximately 2.5 hours. In an alder forest in Western Hungary Gribovszki et al. (2010a) reported an increase in time lags from 90 to 120 minutes between Penman-Monteith evapotranspiration and evapotranspiration calculated from streamflow fluctuations. Hence, it is likely that the time lag of the mixed vegetation in the riparian zone of the HOAL catchment is also on the order of 0.5 to 2.5 h. In some of the studies above, the time lag of sap flow from stem to branch was included in the estimates, therefore this value is also small. Similarly, it is likely that the time lags for root water uptake are small in the riparian zone, where the groundwater table is high, which is typical in the HOAL, especially on the left side of the stream. The remaining time lag components are associated with subsurface processes (groundwater movement and groundwater-stream interactions). Assuming that the lag components are additive one would estimate lags of the subsurface processes of about 1.5 hrs in early spring to about 9 hrs in autumn for MW catchment outlet. The time lags are shorter on those tributaries that are located on the left side of the stream with a western aspect where the groundwater levels are shallow and the riparian forest cover is more dense (Virtual gauge LF).

Separation of scales in transpiration effects on low flows – A spatial analysis in the Hydrological Open Air Laboratory

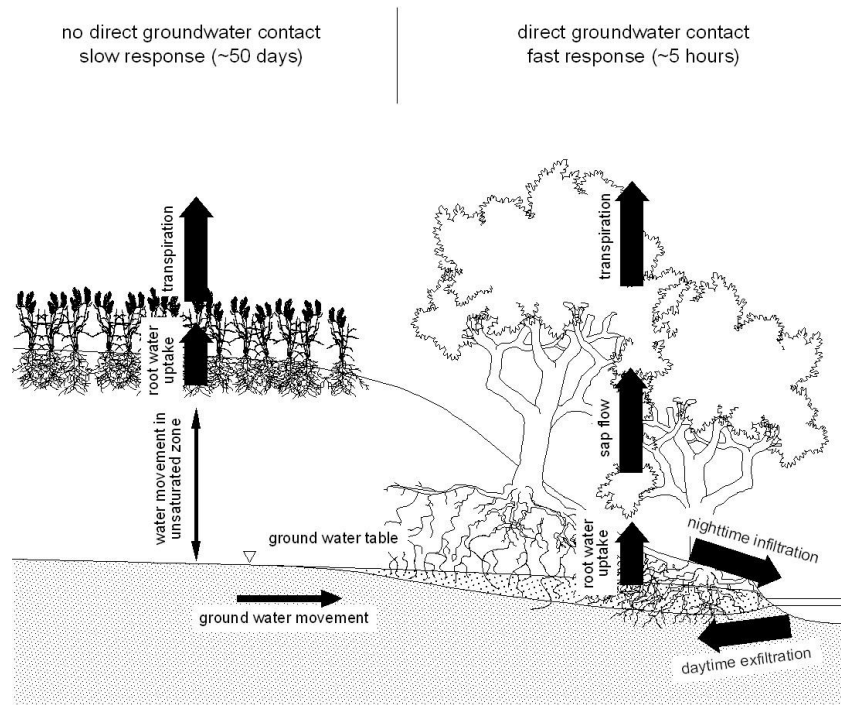


Figure 13. A possible conceptualization of the catchment during low flow periods: schematic of the decoupling of riparian zone fluxes and catchment zone fluxes for the catchment outlet. Width of the arrows indicates the time scale. The thick arrows represent time scales of hours (i.e. well-connected at the diel scale), the thin arrows represent time scales of weeks or months (i.e. decoupled at the diel scale).

The reason for the longer subsurface lag times in the summer months are presumably a consequence of the soil moisture status of the riparian zone. Along the stream in the HOAL there is probably a mixture of infiltration and exfiltration due to the heterogeneity of the topography in the near-stream zone. Previous studies (e.g. Caldwell et al., 1998; Voltz et al., 2013) found a reversal of groundwater gradients in the riparian zone due to night time infiltration of groundwater into the stream and reversed flow during the day as the trees take up groundwater and part of the stream water to satisfy their transpiration needs (Figure 13). Although we did not observe the complete reversal of the groundwater gradient in the riparian zone with the current measurement setup, a diel change in the magnitude of the gradient was observed. The fact that the main part of the time lag is related to the movement of water in the subsurface suggests that the increase in lag times during the summer months is a consequence of the amount of water stored in the near-stream zone. The subsurface water storage, both soil moisture and groundwater, is largest in spring, which is also indicated by the seasonal maximum in the discharge. Therefore, it is likely that the short lag times in spring are a consequence of the fact that the roots are well-connected to the water sources and the celerities are higher. As the season progresses, the catchment gradually dries out and the roots become less connected to the subsurface water storage. When the soils get drier, the unsaturated zone becomes thicker and flow paths get longer and the overall hydraulic conductivity gets smaller. These mechanisms cause an increase in the lag times from 3 to 11 hours for MW catchment outlet. This interpretation is consistent with the findings of Bond et al. (2002), Fonley et al. (2016), Moore et al. (2011) and Wondzell et al. (2007, 2010).

Similarly to time lag λ , the recession time scale α could be also described by a cascade of responses. This cascade includes the time lags between transpiration of the crop fields in the catchment and root water uptake, water movement in the unsaturated zone, groundwater movement and groundwater – stream interactions. While it is difficult to separate the individual components, it is clear that the dynamics of root water uptake and transpiration of crops are faster than the water movement in the unsaturated zone and the groundwater movement. The former operates at a daily scale with significant diel variations, while the latter operates at time scales of weeks and months. The differences in the magnitudes of the time scales are indicated in Figure 13 by the thickness of the arrows. Thick arrows represent time scales of hours, thin arrows represent time scales of weeks or months. The recession time scale shows a gradual seasonal increase from 21 to 54 days for MW catchment outlet. This increase is presumably related to the amount of water stored in the subsurface at the catchment scale. As the catchment dries out, groundwater flow seems to follow deeper flow paths, which are associated with longer response times consistent with the findings of Bond et al. (2002). A similar, increasing trend of the recession time scale α was found for the subcatchments with shallow groundwater levels (Virtual gauge LF) although, overall, the time scales are smaller, as would be expected because of the shorter distances. At subcatchments fed by deeper subsurface flow (Virtual gauge SF) the recession time scale α was longer indicating a reduced dependence on riparian processes. However, some studies found a decrease in the recession time scale between spring and summer (e.g. Federer, 1973; Kalicz et al., 2011). They explained this phenomenon by the larger root water uptake of the vegetation in summer which caused faster recession.

Observations and model simulations showed that the diel signal observed in streamflow at different outlet points of the HOAL mainly originates from diel fluctuations in the riparian evapotranspiration. There is a clear spring onset, a late autumn offset (Figure 3b) of the diel diel signal in streamflow and a clear seasonal pattern in the amplitudes and timings (Figure 4). While the first harvest of the crop fields in July does not influence the amplitude of the diel signal in the stream (Figure 4), the amplitudes start to decrease only in the autumn months and the first frost when the trees in the riparian zone next to the stream drop their leaves terminates the diel streamflow fluctuations. Streamflow from tributaries with large riparian forest cover fluctuated more than streamflow from tributaries which were covered mainly by crop fields and fed by deep aquifers. These observations imply that the riparian zone is the main driver of the diel streamflow fluctuations. The results of a solar radiation driven model showed that only a small proportion of the maximum available catchment energy induced the diel streamflow fluctuations (Figure 10). This proportion was larger during higher baseflow conditions and it increased from spring to summer and decreased in autumn. The daily minimum streamflow occurred later as the season progressed, which was also reflected in the time lags of the solar radiation driven model. The change in the timing of the daily minimum streamflow and the increase in lag time are likely related to the drying of the catchment during the summer and autumn which may lead to a partial disconnect of the riparian zone and the stream. This is consistent with the findings of previous studies, such as Bond et al. (2002) and Wondzell et al. (2007, 2010) who showed that the area contributing to streamflow fluctuations decreased as the catchment gradually dried out.

A clear separation of scales exists in the time domain, which is apparent in the streamflow signal. Conceptually, time scales and space scales of variability are linked through their characteristic velocities, therefore a separation of scales in the space domain would be expected, if a separation of scales in the time domain exists (Skøien et al., 2003). There is also interaction across the diel

and seasonal time scales. The diel low flow fluctuations due to riparian and near-riparian transpiration are modulated by the soil moisture state at the seasonal time scales. Conversely, daily transpiration contributes to the seasonal totals. This interaction across time scales is reminiscent of the effects of climate variability on floods - in a direct way through the seasonal variability of storm characteristics and indirectly through the seasonality of rainfall and evapotranspiration that affect the antecedent catchment conditions for individual storm events (Sivapalan et al., 2005).

The streamflow and piezometer observations imply that at the daily time scale, riparian evapotranspiration induces the diel streamflow fluctuations, while most of the catchment evapotranspiration, such as evapotranspiration from the crop fields further away from the stream does not contribute to these fluctuations. This implication was confirmed by piezometer data. We found that the amplitudes of the fluctuations in the riparian zone within two meter distance from the stream were significant, and did not find significant fluctuations 15 m from the stream. This is consistent with the findings of Reigner (1966), who also found that the amplitudes were significant within 2 m distance from the stream. This behavior is also apparent in the time domain. As the first frost occurs, the low flow fluctuations at MW stop within days (Figure 3b). A similar behavior has been observed by Goodrich et al. (2000) and for groundwater fluctuations by Lautz (2008). Deutscher et al. (2016) distinguished between two distinct parts of the catchment with different connectivity, and connectivity may also exist in terms of soil moisture patterns (Western et al., 1998). Conversely, there is a quick onset of the fluctuations in spring at the beginning of the growing season, when the vegetation is growing. Preferred states and switching behavior associated with thresholds seem to be more common characteristics than what is usually assumed (Blöschl and Zehe, 2005; Zehe and Sivapalan, 2009). Spatial patterns of hydrological dynamics may help identify preferred states (Blöschl, 2006; Grayson et al., 1997). Analyzing spatial patterns may also help to address the difficulties in predicting the whole-catchment water balance from observations at the local scale (Thompson et al., 2011).

2.9. Conclusions

This study investigated the spatial and temporal patterns of diel low flow fluctuations for different runoff generation mechanisms in a 66 ha Austrian experimental catchment, the Hydrological Open Air Laboratory. Our results showed that:

- i. The ratio of streamflow fluctuations and mean streamflow was around 0.3 for wetlands, where the riparian forest cover is the largest and the depth to the groundwater table does not exceed one meter. The amplitudes were much smaller for tile drainage systems and springs that are fed by deeper subsurface flow, where the dominant land cover is crop and the ratio is around 0.04.
- ii. The separation of scales in the time domain could be reproduced by a solar radiation driven model. Lag times between radiative forcing and diel streamflow fluctuations increased from 3 to 11 hours from spring to autumn as the catchment became more disconnected from the stream. The recession time scales increased from 25 days in spring to 60 days in autumn, which was likely a consequence of the decreasing storage of subsurface water at the catchment scale.
- iii. A separation of scales in the time domain is apparent in the streamflow signal, i.e. diel and seasonal fluctuations induced by transpiration, imply a separation of scales in the space domain: the diel streamflow fluctuation are driven by the riparian zone along the main stream, while most of the catchment (the crop fields located further away from the stream)

did not affect the diel signals. This interpretation is supported by the groundwater level data.

2.10. Acknowledgements

The authors would like to acknowledge financial support provided by the Austrian Science Funds (FWF) as part of the Vienna Doctoral Program on Water Resource Systems (DK W1219-N28). The data used are listed in the tables and the data necessary to reproduce the reported findings is available on the following ftp site: <ftp://ftp.hydro.tuwien.ac.at/> (Username: Szeles_etal_2017_Data; Password: 20cWRs17).

The added value of different data types for calibrating and testing a hydrological model in a small catchment

3. The added value of different data types for calibrating and testing a hydrological model in a small catchment

3.1. General

The goal of this chapter was to better understand the catchment behavior on longer time scales with both dry and wet periods.

The aim of this chapter was to investigate how the variety of field observations in the catchment can be linked to hydrological model simulations. We aimed to understand which type of observations and on which degree improve runoff simulations and the internal consistency of the model.

The present chapter corresponds to the following scientific publication in its original form:

Széles, B., Parajka, J., Hogan, P., Silasari, R., Pavlin, L., Strauss, P., & Blöschl, G. (2020). The added value of different data types for calibrating and testing a hydrologic model in a small catchment. *Water Resources Research*. Under review.

3.2. Key points

1. A new framework is presented for stepwise runoff model parameter estimation from observed runoff and model state variables and fluxes.
2. For the study catchment, correlation coefficient of monthly runoff in the validation period is 0.82 and the relative volume error is -1%.
3. Combination of soil moisture and evapotranspiration observations had the largest influence on parameter estimation.

3.3. Abstract

This study investigated the added value of different data for calibrating a runoff model for small basins. The analysis was performed in the 66 ha Hydrological Open Air Laboratory, in Austria. An HBV type, spatially lumped hydrological model was parameterized following two approaches. First, the model was calibrated using only runoff data. Second, a step-by-step approach was followed, where the modules of the model (snow, soil moisture and runoff generation) were calibrated using measurements of runoff and model state variables and output fluxes. These measurements comprised laser-based measurements of precipitation, satellite and camera observations of snow, ultrasonic measurements of snow depth, eddy covariance measurements of evapotranspiration, time domain transmissometry based soil moisture measurements, time lapse photography of overland flow and groundwater level measurements by piezometers. The two model parameterizations were evaluated on annual, seasonal and daily time scales, in terms of how well they simulated snow, soil moisture, evapotranspiration, overland flow, storage change in the saturated zone and runoff. Using the proposed step-by-step approach, the relative runoff volume errors in the calibration and validation periods were 0.00 and -0.01, the monthly Pearson correlation coefficients were 0.92 and 0.82, and the daily logarithmic Nash Sutcliffe efficiencies were 0.59 and 0.18, respectively. By using different sources of data besides runoff, the overall process consistency improved, compared to the case when only runoff was used for calibration. Soil moisture and evapotranspiration observations had the largest influence on simulated runoff, while the parameterization of the snow and runoff generation modules had a smaller influence.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

3.4. Introduction

Observed runoff represents the overall, aggregated catchment behavior. Therefore, runoff observations are the most common information used for identifying the parameters of hydrological models. However, parameters of conceptual and physically based hydrological models usually cannot be identified accurately using observed runoff alone as it is difficult to decide, whether the model performs well for the right reasons (Beven & Freer, 2001; Savenije, 2001; Viglione et al., 2018). One way of dealing with this issue is to use additional measurements of input and output fluxes and model states in hydrological modeling. Additional information on hydrological processes helps to constrain and validate hydrological models and testing whether they get the right answers for the right reasons (Grayson et al., 1992).

Most of the studies use other measurements of fluxes and states in multiple objective calibration as a part of the objective function. Previous studies used ground-based, or alternatively remote sensing products or their combination as such additional information on hydrological processes. Soil moisture (Parajka et al., 2006; Kundu et al., 2017; Kunnath-Poovakka et al., 2016; Rajib et al., 2016; Shahrban et al., 2018), evapotranspiration (Gui et al., 2019; Herman et al., 2018; Immerzeel & Droogers, 2008; Kunnath-Poovakka et al., 2016) and groundwater level data (Demirel et al., 2019; Seibert, 2000) were often used for model calibration to improve the model's internal consistency. These studies showed the added value of different observations besides runoff, e.g. for soil moisture, evapotranspiration, groundwater levels. But only a few studies combined the different type of observations (e.g. Avanzi et al., 2020; Kuras et al., 2011). In the past, these data were used mainly in connection with calibration of physically based hydrological models, where measurements could be more explicitly linked to the simulations than for conceptual hydrological models. The performance of the distributed hydrology soil vegetation model (DHSVM) was evaluated by Thyer et al. (2004) and Kuras et al. (2011) using field data. Their study area was located in British Columbia, Canada. While Thyer et al. (2004) focused mainly only on the micro-meteorological part of the process based model (such as snowmelt and energy balance) and also used observed hydrograph from another, nearby catchment, Kuras et al. (2011) also evaluated the subsurface and surface runoff dynamics with a spatially extensive database. Thyer et al. (2004) and Kuras et al. (2011) achieved a daily Nash Sutcliffe efficiency for runoff of approximately 0.90 and Thyer et al. (2004) stated that runoff simulations were most sensitive to snowmelt characteristics as runoff was driven by spring snowmelt in their high elevated, forested study region. In another study, Wei et al. (2016) used measurements of snow water equivalent, snow depth, transpiration, stomatal feedback to vapor pressure, soil and forest properties, soil moisture to parameterize a physically based model without a flow routing module to simulate the water balance in a 4 km² catchment in the US. Without using runoff data for model calibration, they could reproduce the annual and monthly variability of potential runoff (combined outputs of surface runoff and deep drainage) with a Nash Sutcliffe efficiency of 0.62 and 0.56, respectively. A similar study was performed by Kuppel et al. (2018) in the Scottish Highlands with a fully distributed eco-hydrological model. They could simulate daily runoff reasonably well without using runoff observations, but the model performance was substantially better, when runoff was also included in the calibration. Kuppel et al. (2018) argued that the spatio-temporal footprint of the observations involved in model calibration had to be carefully considered. They found that certain variables could only be well simulated, when the model was calibrated to measurements of that variable, for instance soil moisture in gley soils and transpiration in Scots pine stands.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

When ground-based monitoring data are not available, remote sensing products may be a useful alternative for parameter estimation (López et al., 2017; Nijzink et al., 2018; Silvestro et al., 2015). López et al. (2017) found on a Moroccan catchment with Mediterranean and semi-arid climate, that runoff could be better estimated when both remotely-sensed evapotranspiration and remotely-sensed soil moisture products were involved in calibrating a large-scale hydrological model compared to a scenario, when these products were used independently. Silvestro et al. (2015) also found that using data from both ground stations and remotely sensed products, that is land surface temperature and surface soil moisture estimates, improved the model's internal consistency in two Italian catchments with temperate climates. Nijzink et al. (2018) comprehensively tested 9 remotely sensed products on 27 European catchments with diverse landscapes and climates. The products included remotely-sensed soil moisture, evaporation, total water storage and snow accumulation. They found that two soil moisture products and GRACE total water storage anomalies, and in snow dominated catchments, the MODIS snow cover products helped the most in constraining model parameters, when runoff data were not available. Remotely sensed surface water extent and water levels have also been found to be useful proxies on large river basins (Liu et al., 2015; Revilla-Romero et al., 2015; Sun et al., 2012; 2015), while Corbari et al. (2015) used satellite land surface temperature data for distributed hydrological model calibration. Ruiz-Perez et al. (2017) used only MODIS NDVI data to calibrate an eco-hydrological model and obtained good runoff estimates at the catchment outlet. The resolution of these remote sensing products, both in time and space, is usually too coarse for small catchments. Therefore, small catchment scale processes require field observations (e.g. Avanzi et al., 2020).

An alternative is to use stepwise calibration. Most of the studies used runoff signatures, e.g. low flows, high flows, annual runoff, etc. to calibrate their models step-by-step (e.g. Fenicia et al., 2007; Gelleszun et al., 2017; Hogue et al., 2000; Lu & Li, 2015). For instance, Hogue et al. (2000) separated the parameterization of the low and high flow simulations. Gelleszun et al. (2017) separately calibrated the parameters influencing runoff volume and peaks, seasonality and low flows, and the shape of the hydrograph. Fenicia et al. (2007) compared two multi-objective model calibration approaches. One of these approaches was a stepped calibration approach, where they separately calibrated certain parameter sets associated with different processes. These processes influenced distinct aspects of the system response, low flows, high flows and lag time. Lu & Li (2015) proposed a different calibration strategy, grouping the model parameters according to time scales (annual, seasonal and daily) where they are the most sensitive. Only a few studies used measurements on different fluxes and states in a stepwise fashion. These studies performed stepwise model calibration by looking at the internal state variables and processes of the model (e.g. Arheimer et al., 2020; Avanzi et al., 2020; Hay et al., 2006; Kuras et al., 2011; Ning et al., 2015). Hay et al. (2006) calibrated solar radiation, potential evapotranspiration, water balance and daily runoff using measurements of fluxes and states and runoff, while Ning et al. (2015) calibrated the water storage and runoff generation. Inspired by these studies, which used stepwise model calibration approaches, in this study we aimed to link model simulations with field observations focusing on the different hydrological processes. We used a large set of field observations of input and output fluxes and states besides runoff to calibrate our model in a step-by-step way.

The objective of this study was to investigate the added value of different data types of hydrological processes for calibrating and testing a hydrological model in a small catchment. Our goal was to propose a stepwise approach for constraining hydrological model by using runoff data and observations of snow, soil moisture, evapotranspiration, overland flow and groundwater

The added value of different data types for calibrating and testing a hydrological model in a small catchment

levels. Our goal was to link field observations with lumped, conceptual hydrological model simulations by using all the available data in a stepwise mode. The model performance was evaluated at the annual, seasonal and daily time scales. The analysis was performed in the Hydrological Open Air Laboratory (HOAL) in Austria, a 66 ha experimental catchment, where a large variety of long-term field observations are available (Blöschl et al., 2016).

3.5. Study area and data

3.5.1. Study area

The study was conducted in a small experimental catchment, the Hydrological Open Air Laboratory (HOAL) in Petzenkirchen, located in the western part of Lower Austria (Figure 14, chapter 2.5.1.) (Blöschl et al., 2016; Eder et al., 2010, 2014; Exner-Kittridge et al., 2016; Széles et al., 2018).

Mean annual air temperature, precipitation, runoff and evapotranspiration in the study period (1991-2017) were 9.6 °C, 782 mm/yr, 184 mm/yr and about 598 mm/yr, respectively. The amount of snow falling in winter is small, and quickly melts in the catchment. On average (2013-2017), snow is observed on less than 10% of the days in a year. A significant amount of snow was observed only in 2016 and 2017 in the catchment, while the winters of 2014 and 2015 were almost snow-free.

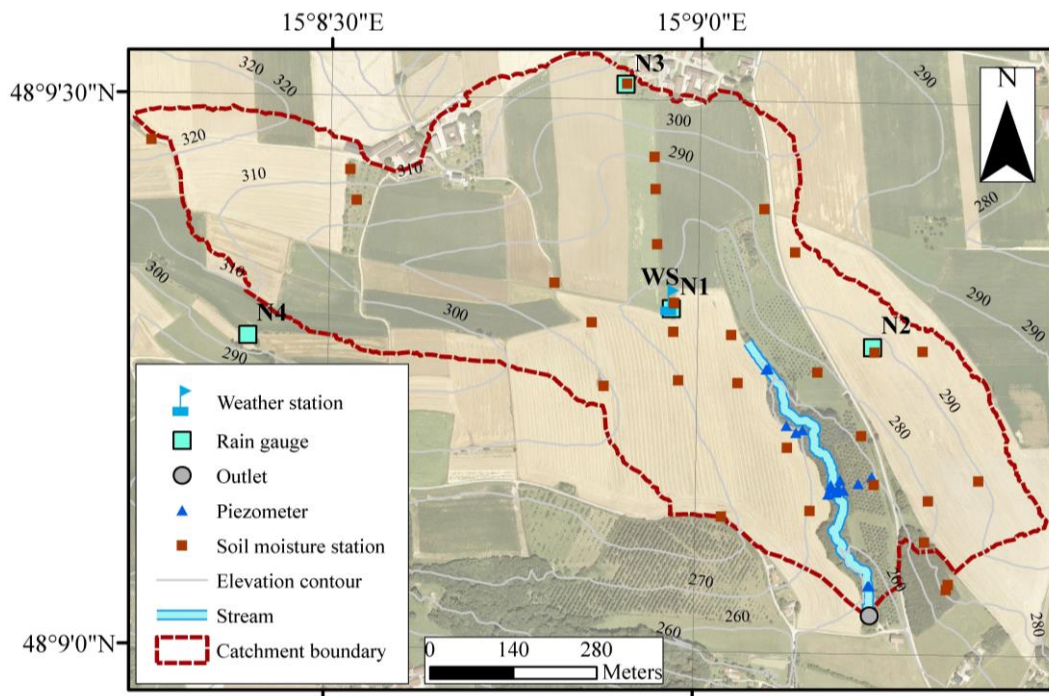


Figure 14. Study area: Hydrological Open Air Laboratory (HOAL) in Petzenkirchen, Lower Austria and location of devices for precipitation and evapotranspiration (weather station), soil moisture, groundwater level (piezometer) measurements.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

3.5.2. Data

Rainfall has been measured with high temporal resolution (one minute) by a weighing rain gauge (PESA) situated 200 m from the catchment outlet since 2002. In 2012, four additional weighing rain gauges (OTT Pluvio) were installed in the HOAL. Since 2013, a laser-based present weather sensor located at the weather station (Campbell PWS100) has measured the size and velocity of water droplets in the air with one minute temporal resolution and time lapse photographs have been taken by the weather station camera (Sanyo VCC-MCH5600P) every minute during daylight. Since 1986, air temperature (Pt 1000) has been recorded at 7, 14 and 19h by a thermometer. Since October 2012 air temperature at 2 m height (HMP 155) and snow depth (SR50AT) have been measured at the HOAL weather station with half hourly temporal resolution (Figure 14).

Twenty-nine time domain transmission soil moisture stations, 19 permanent and 10 temporary stations, with sensors at 0.05, 0.10, 0.20 and 0.50 m depths have measured the water content in the unsaturated zone since 2013 with hourly temporal resolution. Since August 2012 grass evapotranspiration has been measured by a closed-path eddy covariance station (Campbell EC155) at the weather station and crop evapotranspiration has been measured by two open-path eddy covariance stations (Campbell IRGASON) at various locations according to the agricultural crop rotation.

Nineteen piezometers (SWS Mini Driver) located in the riparian forest close to the stream have monitored the groundwater level at a resolution of five minutes since 2013.

Runoff has been monitored at the outlet of the catchment by a calibrated H-flume with a pressure transducer since 2001, and additionally with an ultrasonic probe since mid 2010 with one-minute temporal resolution (Figure 14). Details on the sensors are given in Blöschl et al. (2016).

Three time periods were selected for the analysis, a 22-year-long period when only runoff measurements (1991-2012), and a 3-year-long (2013-15) and a 2-year-long (2016-17) period when runoff measurements and additional sources of data were available (Figure 15). The 3-year-long period was used for model calibration (Calib), the 22-year-long (Val1) and 2-year-long (Val2) periods for model validation. One year preceding each period was used as warm-up period.

A significant amount of snow was observed only in 2016 and 2017 in the catchment, while the winters of 2014 and 2015 were almost snow-free. 2014 and 2015 were exceptionally dry years, while, 2013, 2016 and 2017 were more wet (Figure 15).

The added value of different data types for calibrating and testing a hydrological model in a small catchment

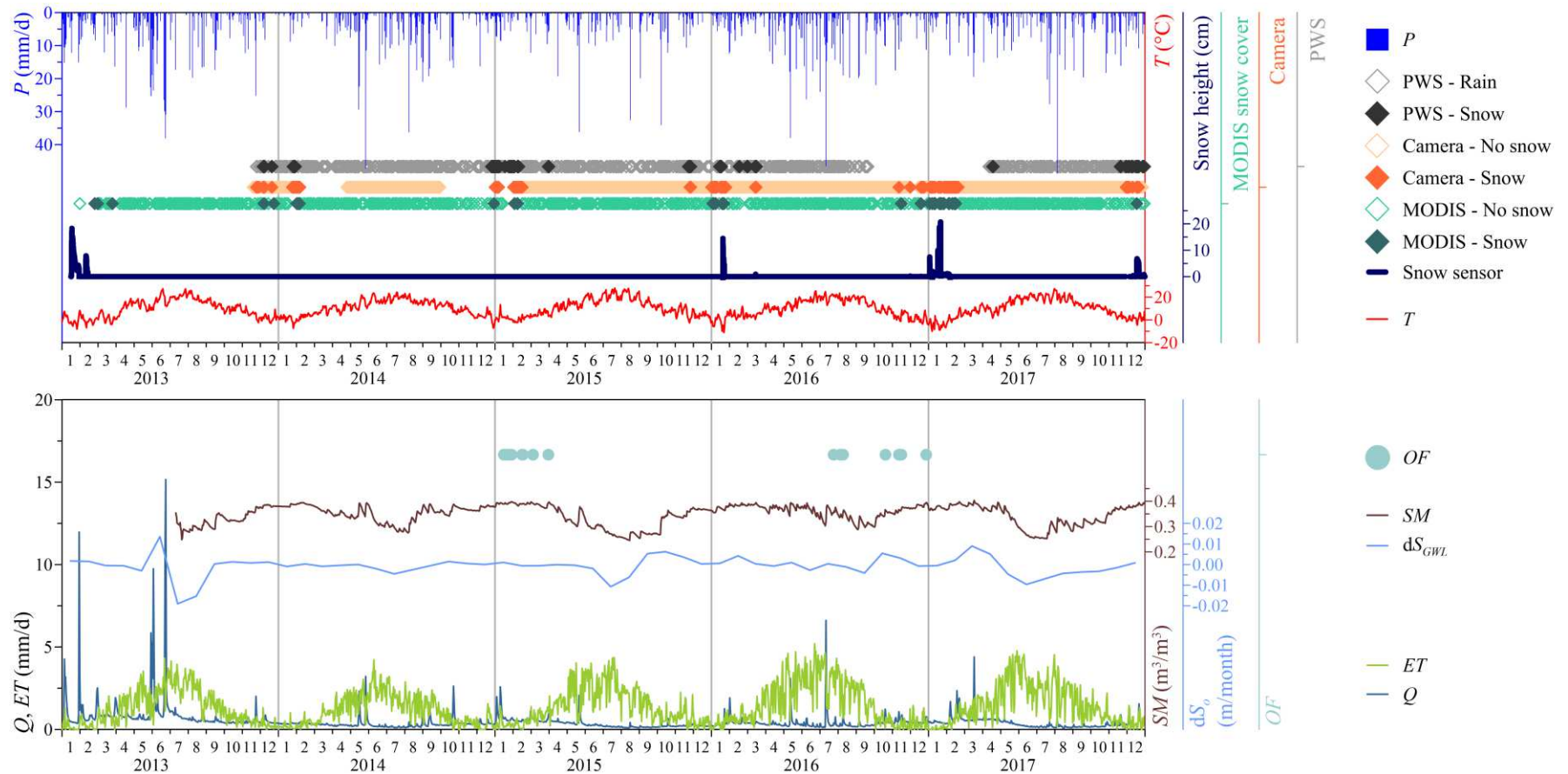


Figure 15. Field observations since 2013, used for model calibration and validation. Top: air temperature T , precipitation P (average of four rain gauges), snow depth measured by snow sensor, reclassified categories from MODIS snow cover images, snow cover based on time lapse photos from the digital camera located at the weather station, reclassified categories from present weather sensor PWS. Bottom: runoff Q , evapotranspiration ET , soil moisture SM (average of all stations over all depths), monthly average storage change in the saturated zone dS_o (catchment average of spatially interpolated raster map based on piezometers in the riparian zone), occurrence of overland flow OF .

The added value of different data types for calibrating and testing a hydrological model in a small catchment

3.6. Methodology

3.6.1. Hydrological model

A lumped, conceptual hydrological model, the TUWmodel was used in this study (Parajka et al., 2007). The model follows the structure of the HBV model (Bergström, 1976; Bergström & Lindström, 2015; Lindström et al., 1997). Numerous studies have shown that this type of model structure works well in Austrian catchments (e.g. Parajka et al., 2007; Slezziak et al., 2018) and worldwide (Bergström & Lindström, 2015). The model consists of three main modules, a snow module, a soil moisture accounting module and a runoff generation module (Merz & Blöschl, 2004; Parajka et al., 2007) (Figure 16 and Table 5). It has 14 free model parameters, which need to be calibrated.

While we could have modified the model structure to tailor it to the runoff processes in the HOAL, we chose not to do this. Instead, we used a more general model structure that could be used in a wider range of catchments.

Within the snow module according to (6)-(10) in Table 5, precipitation P (mm/d) is separated into P_S solid and P_R liquid precipitation depending on the wet bulb temperature parameter T_{wb} ($^{\circ}\text{C}$) (Bergström, 1976; Blöschl et al., 1991; Jennings et al., 2018; Rohrer, 1992; Steinacker, 1983). The catch deficit of precipitation gauges during snowfall is corrected by a snow correction factor SCF (-). Snowmelt M (mm/d) is simulated based on the degree-day concept, using a degree-day factor DDF (mm/ $^{\circ}\text{C}/\text{d}$) and a melt temperature parameter T_m ($^{\circ}\text{C}$).

Within the soil moisture module according to (11)-(13) in Table 5, the fraction of precipitation producing runoff and evapotranspiration are simulated as a function of the soil moisture state SM (mm) of the catchment. If the soil moisture storage exceeds a threshold parameter FC (mm), all rainfall and melt contribute to runoff. The characteristics of runoff production are controlled by the non-linearity parameter β (-). If the soil moisture state exceeds the limit for potential evapotranspiration LP (mm), which is the product of FC and parameter $LPrat$ (-), actual evapotranspiration ET (mm/d) reaches its potential rate ET_P (mm/d). The potential evapotranspiration ET_P was calculated with the modified Blaney-Criddle method (Parajka et al., 2003; Schrödter, 1985).

Within the runoff generation module according to (14)-(20) in Table 5, an upper and a lower reservoir represent hillslope routing. The proportion of rainfall and melt contributing to runoff enters the upper reservoir and leaves it through three paths. The first path is very fast runoff q_0 (mm/d) with very fast storage time k_0 (d), if a threshold of the storage state LS_{UZ} (mm) is exceeded in the upper reservoir. The other two paths are an outflow from the upper reservoir q_1 (mm/d) with a fast storage time k_1 (d) and percolation to the lower reservoir with a constant percolation rate c_P (mm/d). Water leaves the lower zone as baseflow q_2 (mm/d) with a slow storage time k_2 (d). The outflow from the two reservoirs is routed by a triangular transfer function representing runoff routing in the stream, where B_{MAX} (d) is the maximum base at low flows and c_R (d²/mm) is a free scaling parameter.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

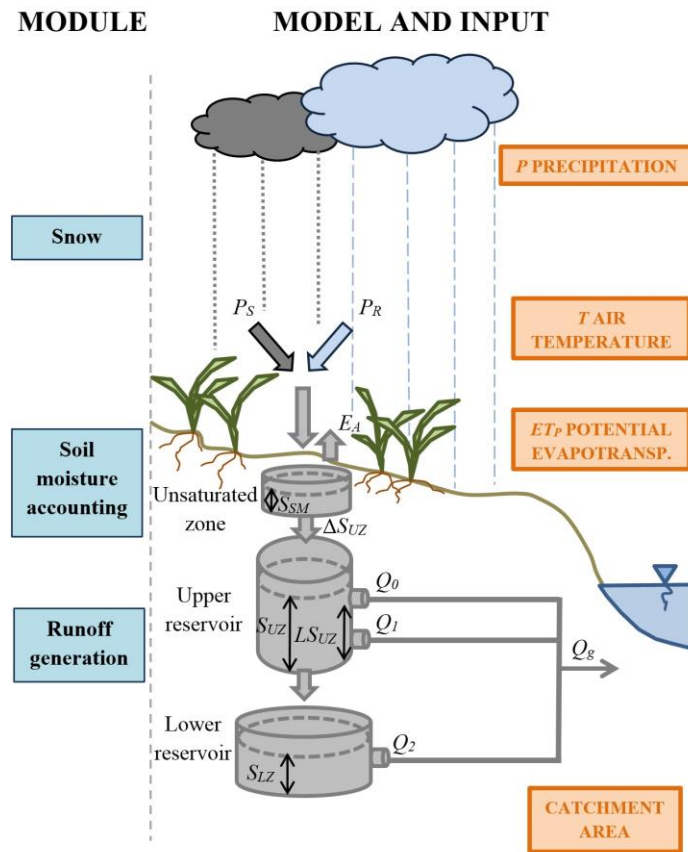


Figure 16. Model structure with the three modules (snow, soil moisture and runoff generation) and input shown in orange rectangles on the right. Table 5 contains further information on the equations and notations.

Table 5. Model equations and notations.

Module	Equation	Notation
Snow module	Separation of solid and liquid precipitation:	Model input:
	$P_S = x \cdot P$	(6) $-P$ (mm/d) Precipitation
	$P_R = (1 - x) \cdot P$	(7) $-T$ (°C) Air temperature
	where:	Model parameter:
	$\begin{cases} x = 0 & \text{if } T_{wb} \leq T \\ x = 1 & \text{if } T < T_{wb} \end{cases}$	(8) $-T_{wb}$ (°C) Wet bulb temperature
		$-T_m$ (°C) Threshold temperature, above which melt starts
		$-DDF$ (mm/°C/d) Degree day factor
		$-SCF$ (-) Snow correction factor

The added value of different data types for calibrating and testing a hydrological model in a small catchment

	<p>Snow melt:</p> $P_m = (T - T_m) \cdot DDF \text{ if } T_m < T \text{ and } 0 < SWE$ <p>Otherwise $P_m = 0$</p> <p>Snow storage:</p> $SWE_i = SWE_{i-1} + (SCF \cdot P_S - P_m) \cdot \Delta t$	<p>(9) State variable:</p> <p>-SWE_i (mm) Snow water equivalent at time step i</p> <p>Model output:</p> <p>(10) -P_S (mm/d) Solid precipitation -P_R (mm/d) Liquid precipitation -P_m (mm/d) Snow melt</p> <p>Other:</p> <p>-Δt (d) Time step</p>
<p>Soil moisture accounting module</p>	<p>Change in soil moisture:</p> $S_{SM,i} = S_{SM,i-1} + P_R + P_m - ET - \Delta S_{UZ}$ <p>Fraction of precipitation generating runoff:</p> $\Delta S_{UZ} = \left(\frac{S_{SM}}{FC}\right)^\beta \cdot (P_R + P_m)$ <p>Actual evapotranspiration:</p> $ET = ET_P \frac{S_{SM}}{LPrat \cdot FC} \text{ if } S_{SM} < LPrat \cdot FC$ $ET = ET_P \text{ if } S_{SM} \geq LPrat \cdot FC$	<p>(11) Model input:</p> <p>-ET_P (mm/d) Potential evapotranspiration</p> <p>Model parameter:</p> <p>-FC (mm) Field capacity, maximum soil moisture storage -β (-) Nonlinear parameter for runoff production -$LPrat$ (-) Parameter related to the limit for potential evapotranspiration</p> <p>(12) State variable:</p> <p>-$S_{SM,i}$ (mm) Soil moisture at time step i -ΔS_{UZ} (mm) Fraction of precipitation generating runoff</p> <p>Model output:</p> <p>(13) -ET (mm/d) Actual evapotranspiration</p>

The added value of different data types for calibrating and testing a hydrological model in a small catchment

Runoff generation module

Hillslope routing:

$$q_0 = (S_{UZ} - LS_{UZ}) \frac{\exp\left(-\frac{1}{k_0}\right)}{k_0} \text{ if } LS_{UZ} \leq S_{UZ}$$

$$q_0 = 0 \text{ otherwise} \quad (14)$$

$$q_1 = -c_P + \left(c_P + \frac{S_{UZ}}{k_1}\right) \cdot \exp\left(-\frac{1}{k_1}\right) \quad (15)$$

$$q_2 = c_P - \left(c_P - \frac{S_{LZ}}{k_2}\right) \cdot \exp\left(-\frac{1}{k_2}\right) \quad (16)$$

$$Q_g = q_0 + q_1 + q_2 \quad (17)$$

$$S_{UZ,i} = S_{UZ,i-1} + \Delta S_{UZ,i} - q_1 - c_P \quad (18)$$

$$S_{LZ,i} = S_{LZ,i-1} - q_2 + c_P \quad (19)$$

Routing in the river – transfer function:

$$B_Q = B_{MAX} - c_R \cdot Q_g \text{ if } (B_{MAX} - c_R \cdot Q_g) \geq 1 \quad (20)$$

$$B_Q = 1 \text{ otherwise}$$

Model parameter:

- k_0 (d) Storage time for very fast response

- k_1 (d) Storage time for fast response

- k_2 (d) Storage time for slow response

- LS_{UZ} (mm) Threshold storage state, very fast runoff q_0 starts, if it is exceeded

- c_P (mm/d) percolation rate

- B_{MAX} (d) Maximum base at low flows

- c_R (d²/mm) Free scaling parameter

State variable:

- S_{UZ} (mm) Storage state in upper reservoir

- S_{LZ} (mm) Storage state in lower reservoir

- B_q (d) Duration of convolution

Model output:

- q_0 (mm/d) Surface runoff

- q_1 (mm/d) Subsurface runoff

- q_2 (mm/d) Baseflow

Following a sensitivity analysis (Appendix A9 and A10), the study proceeded along two modelling approaches. In the first approach, the model was calibrated in one step on a daily temporal resolution, using only runoff data in the objective function. In the second approach, the model was parametrized step-by-step using additional data besides runoff, starting with (1) snow accumulation and melt processes; (2) evapotranspiration and soil moisture changes; and (3) runoff generation and storage changes in the saturated zone. The abbreviations of the scenarios are listed in Table 6. The calibrated model parameters are found in Table 7 for each scenario. For simulating snow accumulation, a half hourly temporal resolution was used, while a daily time step was used to simulate snowmelt. For the other two modules, a daily time step was used for model calibration. The model performance was evaluated at the annual, seasonal and daily time scales.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

Table 6. Summary of the scenarios.

Scenario	Details
Obs	Observations
R	Estimation of all model parameters using only observed runoff data
R+Snowacc	Estimation of all model parameters using runoff and precipitation phase data
R+Snowmelt	Estimation of all model parameters except T_{wb} wet bulb temperature using runoff and snow cover data
R+ET	Estimation of soil moisture accounting and runoff generation parameters using runoff and actual evapotranspiration data
R+SM	Estimation of soil moisture accounting and runoff generation parameters using runoff and soil moisture data
R+50ET+50SM	Estimation of soil moisture accounting and runoff generation parameters using runoff, soil moisture ($w_{ET}=50\%$) and actual evapotranspiration ($w_{SM}=50\%$) data
R+SM+G	Estimation of runoff generation parameters using runoff and runoff generation data, the soil moisture accounting parameters were fixed in scenario R+SM
R+ET+G	Estimation of runoff generation parameters using runoff and runoff generation data, the soil moisture accounting parameters were fixed in scenario R+ET
R+50ET+50SM+G	Estimation of runoff generation parameters using runoff and runoff generation data, the soil moisture accounting parameters were fixed in scenario R+50ET+50SM

The added value of different data types for calibrating and testing a hydrological model in a small catchment

Table 7. Scenarios (according to Table 6) and corresponding, calibrated model parameter sets. Parameters which were fixed at a certain scenario are shown in bold.

Scenario name	Parameters													
	SCF	DDF	T_{wb}	T_m	LP _{rat}	FC	β	k_0	k_1	k_2	LS _{UZ}	c_P	B_{MAX}	c_R
R	1.3	4.3	-0.1	0.0	0.0	319.4	0.6	1.1	4.7	84.7	9.2	1.5	5.3	20.0
R+Snowacc	1.0	2.5	0.6	-0.9	0.1	109.4	1.3	1.0	5.2	97.5	9.0	1.6	7.5	40.1
R+Snowmelt	0.9	3.1	0.6	-0.3	0.0	118.5	0.9	1.0	4.2	100.1	8.6	1.5	2.3	32.3
R+ET	0.9	3.1	0.6	-0.3	1.0	476.8	6.3	1.0	6.8	123.2	10.4	0.9	4.8	21.3
R+SM	0.9	3.1	0.6	-0.3	0.9	149.5	8.9	0.3	3.9	142.2	1.2	3.6	12.3	40.2
R+50ET+50SM	0.9	3.1	0.6	-0.3	1.0	177.8	19.6	1.1	4.8	168.1	14.1	3.5	12.6	33.9
R+ET+G	0.9	3.1	0.6	-0.3	1.0	476.8	6.3	0.1	26.6	156.4	2.0	0.4	12.7	16.7
R+SM+G	0.9	3.1	0.6	-0.3	0.9	149.5	8.9	0.2	17.1	169.4	1.0	1.3	13.2	36.2
R+50ET+50SM+G	0.9	3.1	0.6	-0.3	1.0	177.8	19.6	0.2	15.1	215.7	1.1	1.4	12.2	37.5

3.6.2. Model calibration with only runoff data

According to the first calibration approach (scenario R), the model was calibrated to observed runoff by maximizing Z_1 (-) multi-objective function (21). The main idea was to capture the water balance by minimizing the relative volume error and to mimic the recession parts of the hydrographs. Z_1 was maximized for calibrating all 14 model parameters, using the DEoptim R package for parameter optimization (Ardia et al., 2010a, 2010b, 2016; Mullen et al., 2011).

$$Z_1 = 0.5 \cdot \ln Nash_Q + 0.5 \cdot VE_Q \quad (21)$$

where $VE_Q = -VE_Q$ if $0 < VE_Q$

where $\ln Nash_Q$ (-) is the logarithmic Nash Sutcliffe efficiency for daily runoff according to (22)

$$\ln Nash_Q = 1 - \frac{\sum (\log Q - \log Q_o)^2}{\sum (\log Q_o - \overline{\log Q_o})^2} \quad (22)$$

where Q (mm/d) is daily simulated runoff, Q_o (mm/d) is daily observed runoff, $\overline{Q_o}$ (mm/d) is the mean of the observed runoff.

VE_Q (-) (Criss & Winston, 2008) is the relative volume error for runoff according to (23)

$$VE_Q = \frac{\sum Q - \sum Q_o}{\sum Q_o} \quad (23)$$

The modelling results were evaluated on three time scales, annual, seasonal and daily. On the annual time scale, the volumes of observed and simulated runoff were compared and the relative volume error was calculated according to (23). On the seasonal time scale, monthly average observed and simulated runoff was compared. The monthly Pearson correlation coefficient for runoff was calculated according to (24)

The added value of different data types for calibrating and testing a hydrological model in a small catchment

$$r_{Q,m} = \frac{cov(Q_m, Q_{m,o})}{\sigma_{Q_m} \sigma_{Q_{m,o}}} \quad (24)$$

where σ_{Q_m} (mm/d) and $\sigma_{Q_{m,o}}$ (mm/d) are the standard deviation of simulated and observed monthly average runoff, Q_m and $Q_{m,o}$, respectively. On the daily time scale, daily logarithmic Nash Sutcliffe efficiency for runoff $lNash_Q$ was calculated according to (22).

3.6.3. Model calibration with runoff and additional data

We divided the model parameter estimation procedure into separate steps, looking at the processes associated with the three modules of the model (snow, soil moisture and runoff generation), which were linked to field observations. Using runoff and additional data, the 14 free parameters were gradually fixed, step-by-step, proceeding along the three modules of the model. First, all 14 parameters were calibrated using runoff and snow data. In the next step, the snow parameters were fixed, and only the soil moisture and runoff generation parameters were calibrated using runoff and actual evapotranspiration and/or soil moisture data. In the last step, the snow and soil moisture parameters were fixed and only the runoff generation parameters were calibrated using runoff, overland flow and storage change data.

3.6.3.1. Snow parameters

Simulation of the snow accumulation, i.e. the phase of precipitation (scenario R+Snowacc), and snowmelt (scenario R+Snowmelt) was optimized using observations of runoff, four precipitation gauges, the present weather sensor (PWS), time lapse photos of snow cover in the catchment, MODIS snow cover images and data of the snow depth sensor (Figure 17).

Depending on the size and velocity of the precipitation measured by the present weather sensor, the observed precipitation P , which was calculated as the average of the four gauges, was assigned an output code describing its type (Figure 17). The output codes with one minute temporal resolution were resampled to half hourly resolution, using the median, to match the half hourly temperature measurements. The resampled output codes were assigned to one of the four categories: no data (0), no precipitation (1), rain (2) and snow (3).

All 14 free parameters were optimized to fit the modelled runoff and phase of the precipitation to the observed one. Z_2 (-) multi-objective function was maximized using the DEoptim R package for parameter optimization, according to (25)

$$Z_2 = 0.1 \cdot lNash_Q + 0.1 \cdot VE_Q - 10 \cdot Z_{Snowacc} \quad (25)$$

where $VE_Q = -VE_Q$ if $0 < VE_Q$

where $Z_{Snowacc}$ (-) is the number of half hours with false phase simulations according to (26)

$$Z_{Snowacc} = n_{false} \quad (26)$$

where n_{false} (-) is the number of those half hours, when the model simulated precipitation phase (rain, snow, or no precipitation) did not agree with the observed precipitation phase. The weights on the single objectives were found with sensitivity analysis.

Out of the optimum parameter set, the wet bulb temperature parameter T_{wb} was kept constant in the following optimization steps.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

In order to optimize the snowmelt simulations, an observed snow cover index SCI_o was set up, showing if snow was observed (1) or not (0) in the catchment on a daily basis, based on 3 types of observations (Figure 17):

- 1 Time lapse photos were manually checked: if snow cover was visible in the catchment, SCI_o was assigned 1 (Snow), otherwise 0 (No Snow),
- 2 During periods, when time lapse photos were not available due to camera malfunction or power outage, daily MODIS Normalized Difference Snow Index (NDSI) images were checked (Hall & Riggs, 2016a, 2016b). The territory of the HOAL was extracted and reprojected (MRT, 2004) from the h19v06 MODIS tile with 500 m spatial resolution. Out of the eight pixels, five cover more than 50% of the catchment area, therefore the average of these five pixels was calculated. The catchment average MODIS NDSI values were reclassified to three categories: snow ($40 \leq \text{NDSI} \leq 100$), no snow ($0 \leq \text{NDSI} < 40$) and no data ($100 < \text{NDSI}$), choosing 40 as a threshold based on Dozier (1989). If the NDSI values were classified to category snow or no snow, the composite snow cover index was assigned 1 or 0, respectively,
- 3 If the NDSI values were classified as no data, the snow sensor data were checked. If the recorded snow depth was above 0 cm, the composite snow cover index SCI_o was assigned 1, otherwise 0.

The simulated snow cover index SCI was assigned 1 (snow observed in the catchment), if the modelled snow water equivalent SWE (mm) was larger than 2 mm. Otherwise it was assigned 0. The 2mm threshold was chosen based on sensitivity analyses.

The 13 free parameters (all parameters except T_{wb} wet bulb temperature, which was fixed in the previous step) were optimized to fit the modelled runoff and snow cover index SCI to the observed one. Z_3 (-) multi-objective function was maximized using the DEoptim R package for parameter optimization, according to (27)

$$Z_3 = 0.1 \cdot \ln \text{Nash}_Q + 0.1 \cdot VE_Q - 10 \cdot Z_{\text{Snowmelt}} \quad (27)$$

where $VE_Q = -VE_Q$ if $0 < VE_Q$

where Z_{Snowmelt} (-) is the number of days with false snow cover index simulations similarly to (26).

Out of the optimum parameter set, the remaining three snow parameters (SCF , DDF , T_m) were kept constant in the following optimization steps.

The modelling efficiency in terms of simulating snow accumulation and snowmelt was evaluated by analyzing Z_{Snowacc} and Z_{Snowmelt} for the scenario, when only runoff was used for model calibration (scenario R) and for the scenarios when additional information on snow was used besides runoff (scenarios R+Snowacc and R+Snowmelt, respectively).

The added value of different data types for calibrating and testing a hydrological model in a small catchment

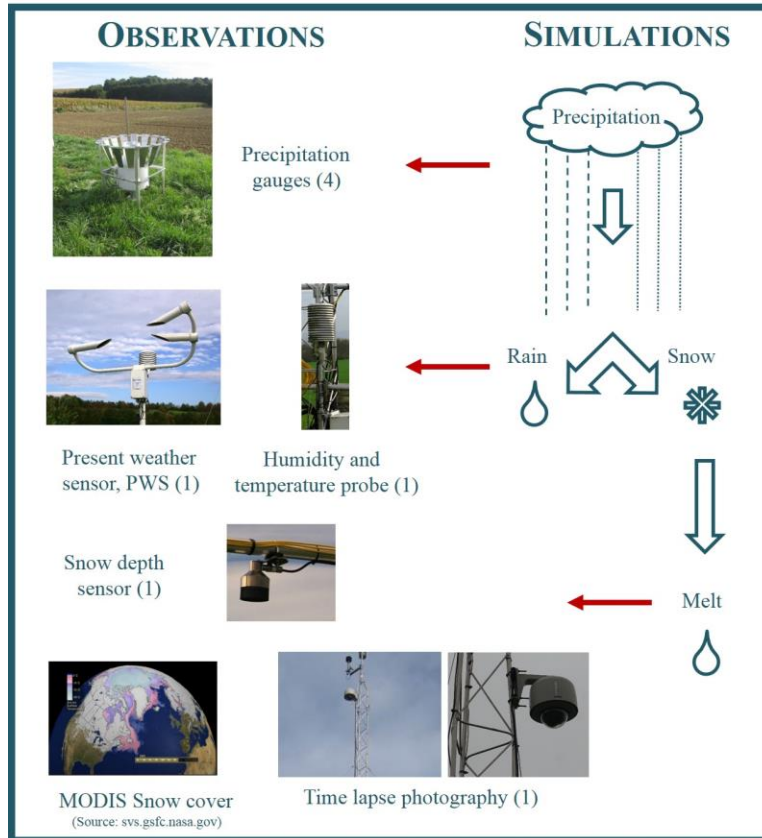


Figure 17. Linking observations with hydrological model simulations I: Snow module (number of sensors is indicated in brackets).

3.6.3.2. Soil moisture and evapotranspiration parameters

Simulation of catchment average soil moisture and evapotranspiration was optimized using runoff data and observations of the soil moisture sensors, the eddy covariance systems and results from a soil survey (Figure 18).

For actual evapotranspiration, a catchment average evapotranspiration was calculated for each day. Measurements of the closed path eddy at the weather station (measuring grass evapotranspiration) and the two mobile eddies (measuring evapotranspiration from different crop types, such as maize, winter wheat, winter barley or bare soil) were weighted with the area of the different land use types. For the riparian forest close to the stream, due to the lack of evapotranspiration measurements, a crop coefficient was used. The grass evapotranspiration from the weather station eddy was then multiplied with this crop coefficient (Allen et al., 1998). A similar method was used in case of data gaps, or in case the evapotranspiration rates from a certain main crop type was not measured in a certain year.

For soil moisture, the average of all stations over all depths was considered. To make the simulated soil moisture comparable with the measurements, the soil moisture time series were standardized according to (28)

The added value of different data types for calibrating and testing a hydrological model in a small catchment

$$SM_s = \frac{SM - \overline{SM}}{\sigma_{SM}} \quad (28)$$

where SM_s (-) is the simulated standardized soil moisture, SM (mm) is the simulated soil moisture, \overline{SM} (mm) and σ_{SM} (mm) are the average and the standard deviation of the simulated soil moisture over each period (calibration and validation periods separately). The observed standardized soil moisture SM_{s_o} (-) was calculated similarly to (28).

A multi-objective function Z_4 (-) (29) was maximized for calibrating 10 model parameters, using the DEoptim R package. The snow module parameters (optimized according to chapter 3.6.3.1) were not changed in this step. Z_4 combined runoff efficiency, the daily Nash Sutcliffe efficiency for evapotranspiration Z_{ET} and standardized soil moisture Z_{SM} with different weights according to (29)

$$Z_4 = 0.25 \cdot lNash_Q + 0.25 \cdot VE_Q + 0.5 \cdot (w_{ET}Nash_{ET} + w_{SM}Nash_{SM_s}) \quad (29)$$

where $VE_Q = -VE_Q$ if $0 < VE_Q$

where w_{ET} (-) is the weight on the evapotranspiration objective, ranging between 0 and 1, and w_{SM} (-) is the weight on the soil moisture objective according to (30)

$$w_{SM} = 1 - w_{ET} \quad (30)$$

$Nash_{ET}$ is the daily Nash Sutcliffe efficiency for actual evapotranspiration (31)

$$Nash_{ET} = 1 - \frac{\sum(ET - ET_o)^2}{\sum(ET_o - \overline{ET_o})^2} \quad (31)$$

where ET (mm/d) is the simulated actual evapotranspiration, ET_o (mm/d) is observed actual evapotranspiration, and $\overline{ET_o}$ (mm/d) is the mean of the observed actual evapotranspiration. The Nash Sutcliffe efficiency for the daily standardized soil moisture $Nash_{SM_s}$ (-) was calculated similarly to (31) with $\overline{SM_{s_o}} = 0$.

Results of a soil survey performed on a 50x50m raster were used to constrain the field capacity FC . The upper boundary of FC was set to 450 mm considering that the maximum of the observed, depth averaged field capacity was 430 mm in the catchment (Murer et al., 2004).

This optimization step was repeated ten times, to check the stability of the optimized model parameters. Out of the optimum parameter set, the three soil moisture parameters (LP_{rat} , FC and β) were kept constant in the following optimization steps. Three main scenarios were chosen for further analysis. In scenario R+ET w_{ET} was chosen as 1, therefore only evapotranspiration information was used besides runoff in the multi-objective function Z_4 . In scenario R+SM w_{ET} was chosen as 0, therefore only soil moisture information was used besides runoff in the multi-objective function Z_4 . In scenario R+50ET+50SM w_{ET} was chosen as 0.5, therefore both evapotranspiration and soil moisture information was used besides runoff in the multi-objective function Z_4 . These parameters were not changed in the following optimization step.

The modelling efficiency in terms of simulating soil moisture and actual evapotranspiration was evaluated by analyzing the relative volume error for actual evapotranspiration VE_{ET} similarly to (23), and the monthly Pearson correlation coefficient according to (24) for actual evapotranspiration and standardized soil moisture, $r_{m,ET}$ and r_{m,SM_s} , respectively, as a function of

The added value of different data types for calibrating and testing a hydrological model in a small catchment

weight on the evapotranspiration objective w_{ET} . Furthermore, the annual (relative volume error for actual evapotranspiration) and monthly performances (monthly average actual evapotranspiration rates and standardized soil moisture) were also compared for selected scenarios (scenarios R, R+ET, R+SM, and R+50ET+50SM).

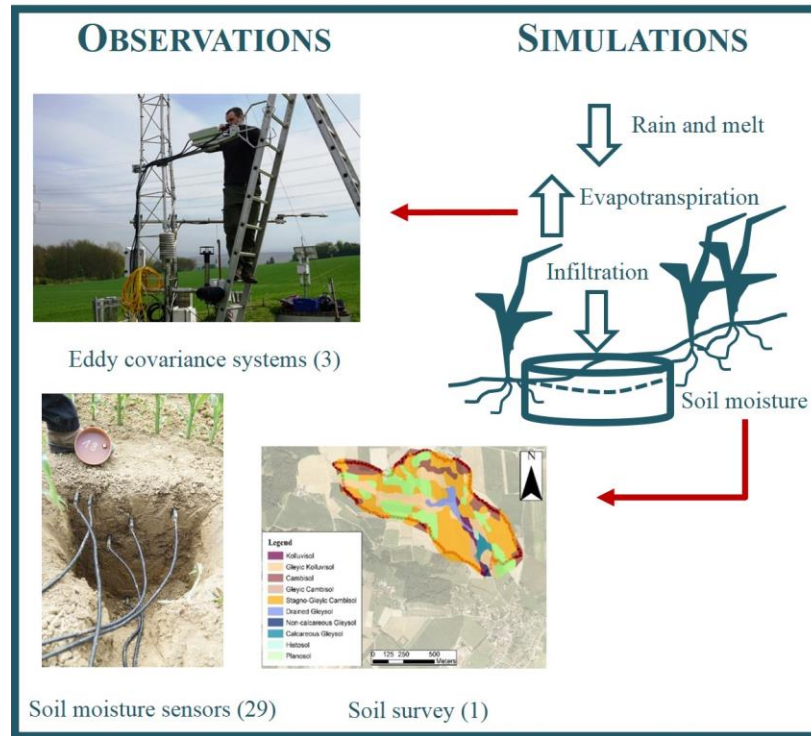


Figure 18. Linking observations with hydrological model simulations II: Soil moisture accounting module (number of sensors is indicated in brackets).

3.6.3.3. Runoff generation parameters

The model parameters related to the subsurface dynamics were optimized using runoff data, time lapse photos of saturation patterns and piezometer measurements of groundwater levels (Figure 19).

In order to optimize the very fast runoff q_0 simulations, saturation excess runoff events were identified in the valley bottom located in the centroid of the catchment by a digital camera at the weather station according to Silasari et al. (2017). The days when the model simulated very fast runoff (q_0) were calibrated to the dates when overland flow events were observed by the camera.

Monthly average storage change in the lower zone reservoir dS_{LZ} (mm/month) was optimized using piezometer observations of groundwater levels GWL , located in the riparian zone. Riparian groundwater level data with 5 minute temporal resolution were averaged to daily values, and missing data were spline interpolated for each riparian piezometer. The daily, gap-filled time series were aggregated to monthly values. The monthly average groundwater level data were differentiated in time and the difference was multiplied by drainable porosity 0.036 to get the monthly storage change values dS_0 (mm/month). The storage change calculated for each piezometer was then spatially interpolated using Thiessen polygons, and a catchment average storage change was calculated from the interpolated raster map. The simulated monthly average

The added value of different data types for calibrating and testing a hydrological model in a small catchment

standardized lower zone storage change dS_s (-) was calibrated to the observed catchment average standardized storage change $dS_{s,o}$ (-). The observed and simulated standardized storage change time series were calculated analogously to (28).

A multi-objective function Z_5 (-) (32) was maximized for calibrating 7 model parameters, using the DEoptim R package. The snow module and soil moisture accounting module parameters (optimized according to chapters 3.6.3.1 and 3.6.3.2) were not changed in this step. Z_5 combined runoff efficiency, the relative number of days with correctly simulated very fast runoff Z_{OF} (-) and the relative number of months with correctly simulated sign of the standardized storage change Z_{dS_s} (-) with different weights

$$Z_5 = 0.25 \cdot lNash_Q + 0.25 \cdot VE_Q + 0.5 \cdot (w_{OF}Z_{OF} + w_{dS_s}Z_{dS_s}) \quad (32)$$

where $VE_Q = -VE_Q$ if $0 < VE_Q$

where w_{OF} (-) is the weight on the overland flow OF objective, ranging between 0 and 1, and w_{dS_s} (-) is the weight on the storage change objective according to (33)

$$w_{dS_s} = 1 - w_{OF} \quad (33)$$

Z_{OF} (-) is the relative number of days, when very fast runoff q_0 was correctly simulated (34)

$$Z_{OF} = \frac{n_{q_0}}{n_{OF}} \quad (34)$$

where n_{q_0} (-) is the number of those days when the model simulated very fast runoff q_0 and overland flow was simultaneously observed by time lapse photos taken by the camera located at the weather station, and n_{OF} (-) is the total number of days, when overland flow was observed in the catchment.

Z_{dS_s} (-) is the relative number of months, when the model correctly simulated the sign of the standardized storage change (35)

$$Z_{dS_s} = \frac{n_{sgn(dS_s)}}{n_{dS_s,o}} \quad (35)$$

where $n_{sgn(dS_s)}$ (-) is the number of those months when the sign of the model simulated standardized lower zone storage change was the same as the sign of the observed standardized storage change, and $n_{dS_s,o}$ (-) is the total number of months when storage change was observed.

This optimization step was repeated 10 times to check the stability of the optimized model parameters. Three main scenarios were chosen for detailed analysis, R+ET+G, R+SM+G, and R+50ET+50SM+G (G refers to runoff generation). In each scenario, equal weights were put on the overland flow ($w_{OF}=0.5$) and storage change ($w_{dS_s}=0.5$) objectives.

The modelling efficiency in terms of simulating overland flow was evaluated by analyzing Z_{OF} as a function of w_{OF} . The modelling efficiency in terms of simulating storage change in the saturated zone was assessed by Z_{dS_s} as a function of w_{OF} .

The added value of different data types for calibrating and testing a hydrological model in a small catchment

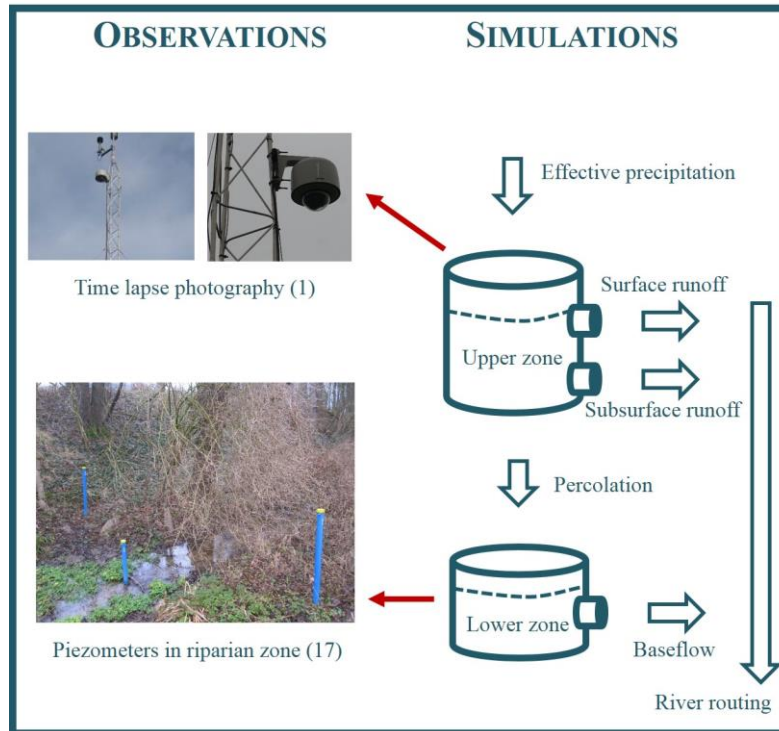


Figure 19. Linking observations with hydrological model simulations III: Runoff generation module (number of sensors is indicated in brackets).

3.6.3.4. Runoff simulation

The modelling efficiency in terms of simulating runoff was evaluated on annual, monthly and daily time scales. On the annual time scale, the volumes of observed and simulated runoff were compared. On the seasonal time scale, monthly average observed and simulated runoff time series were compared. On the daily time scale, daily runoff time series were compared.

3.7. Results

3.7.1. Model calibration with only runoff data

When the model was calibrated using only runoff data, the performance of runoff simulations was very good during the calibration period on each time scale (Table 8). During model validation, the annual and seasonal performance of runoff simulations were still good. The relative volume error was below 20% and the monthly Pearson correlation coefficient was 0.75. But the daily performance was very low, the daily logarithmic Nash Sutcliffe efficiency for runoff was around 0.2 during both periods.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

Table 8. Model calibration using only runoff data: performance of runoff Q (relative volume error VE_Q , monthly Pearson correlation coefficient $r_{m,Q}$, daily logarithmic Nash Sutcliffe efficiency $lNash_Q$) during model calibration (2013-15), first and second validation (1991-2012 and 2016-17, respectively) periods. Scenario according to Table 6.

Scenario	Calibration period 2013-15			Validation period 1 1991-2012			Validation period 2 2016-17		
	VE_Q (-)	$r_{m,Q}$ ()	$lNash_Q$ (-)	VE_Q (-)	$r_{m,Q}$ ()	$lNash_Q$ (-)	VE_Q (-)	$r_{m,Q}$ ()	$lNash_Q$ (-)
R	0.00	0.98	0.81	0.18	0.75	0.22	0.15	0.75	0.17

3.7.2. Model calibration with runoff and additional data

3.7.2.1. Snow simulation

The first step of the step-by-step model parameter estimation was to find the optimal snow parameters.

Calibrating the model parameters to runoff and the observations from the Present Weather Sensor (scenario R+Snowacc) gave 0.31 and 0.40% of poor simulation times steps, i.e. the phase of the model simulated precipitation differed from the observed one, in the calibration and validation periods, respectively. This was slightly better than the simulations that used only runoff data for calibration instead (scenario R) (Table 9). Compared to scenario R, in scenario R+Snowacc the calibrated wet bulb temperature T_{wb} increased from -0.1°C to 0.6°C (Table 7). This value is more realistic considering that it is closer to 1.0°C , which is the median of the wet bulb temperature observed in the catchment during precipitation events with a shift in precipitation phase (calculated according to Appendix A11).

Calibrating the snowmelt parameters to the observed snow cover index SCI_o (scenario R+Snowmelt) gave 4.4 and 6.3% of poor simulation time steps which, again, was better than the R scenario (Table 9). In scenario R, the calibrated snow correction factor SCF indicated 30% increase in snowfall precipitation, which was higher than expected for lowland catchments (Table 7). The lower SCF values in scenarios R+Snowacc and R+Snowmelt were much more realistic (Table 7). The higher snow correction factor SCF in scenario R was then compensated by more intense snowmelt, i.e. higher degree day factor DDF model parameter, which was much higher than expected and found in flatland catchments in Austria (Merz et al., 2011; Slezniak et al., 2020).

The added value of different data types for calibrating and testing a hydrological model in a small catchment

Table 9. Performance of snow accumulation and snowmelt simulations: number of time steps with poor snow accumulation simulations and snow cover simulations relative to the number of time steps with observations. Scenarios according to Table 6: Scenario R where only runoff was used for model calibration, Scenarios R+Snowacc and R+Snowmelt where runoff and snow data were used for model calibration.

Scenario	Relative number of time steps with poor snow accumulation simulations (%)		Scenario	Relative number of time steps with poor snowmelt simulations (%)	
	Calibration period 2013-15	Validation period 2 2016-17		Calibration period 2013-15	Validation period 2 2016-17
R	0.45	0.52	R	4.66	7.25
R+Snowacc	0.31	0.40	R+Snowmelt	4.38	6.29
Number of half hourly time steps	35626	23972	Number of daily time steps	1095	731

The tradeoff between simulating runoff and snowmelt accurately was not large, i.e. using additional information on snow did not have a large influence on runoff simulations during the calibration period (Table 10). But it changed the simulation results during the 22-year-long validation period by causing larger relative volume error and lower logarithmic Nash Sutcliffe efficiency for runoff. During the second, 2-year-long validation period the runoff simulation performance on the seasonal and daily time scales improved (Table 10).

The added value of different data types for calibrating and testing a hydrological model in a small catchment

Table 10. Model calibration using runoff and snow data: performance of runoff Q (Volume error VE_Q , monthly Pearson correlation coefficient $r_{m,Q}$, daily logarithmic Nash Sutcliffe efficiency $lNash_Q$) during model calibration (2013-15), first and second validation (1991-2012 and 2016-17, respectively) periods. Scenarios according to Table 6: R where only runoff was used for model calibration, R+Snowacc and R+Snowmelt where runoff and snow data were used for model calibration.

Scenario	Calibration period 2013-15			Validation period 1 1991-2012			Validation period 2 2016-17		
	VE_Q (-)	$r_{m,Q}$ ()	$lNash_Q$ (-)	VE_Q (-)	$r_{m,Q}$ ()	$lNash_Q$ (-)	VE_Q (-)	$r_{m,Q}$ ()	$lNash_Q$ (-)
R+Snowacc	0.00	0.94	0.81	0.28	0.74	0.18	0.15	0.85	0.27
R+Snowmelt	0.00	0.96	0.82	0.28	0.74	0.19	0.17	0.87	0.27

3.7.2.2. Soil moisture and evapotranspiration simulation

The second step of the step-by-step parameter estimation was to fix the parameters of the soil moisture accounting module.

The smallest relative volume errors of evapotranspiration were achieved for soil moisture weights of $w_{SM}=0.3$ and $w_{SM}=0.0$ in the calibration and validation periods based on the median of 10 model runs (Figure 20 a and b). Generally, the model tended to overestimate evapotranspiration (Figure 20 c and d). This could be a consequence of using the Nash Sutcliffe efficiency for evapotranspiration in the objective function, where the model was fitted to the peaks and not lower values of evapotranspiration rates. Furthermore, there is a mismatch between field observations and the HBV type, soil moisture dependent evapotranspiration calculations. For example, during precipitation events measured evapotranspiration drops to zero, while model simulations increase due to the higher soil moisture content. During model validation, the relative volume error for evapotranspiration was closer to zero, especially when using only evapotranspiration information in the objective function (Figure 20 b and d).

The added value of different data types for calibrating and testing a hydrological model in a small catchment

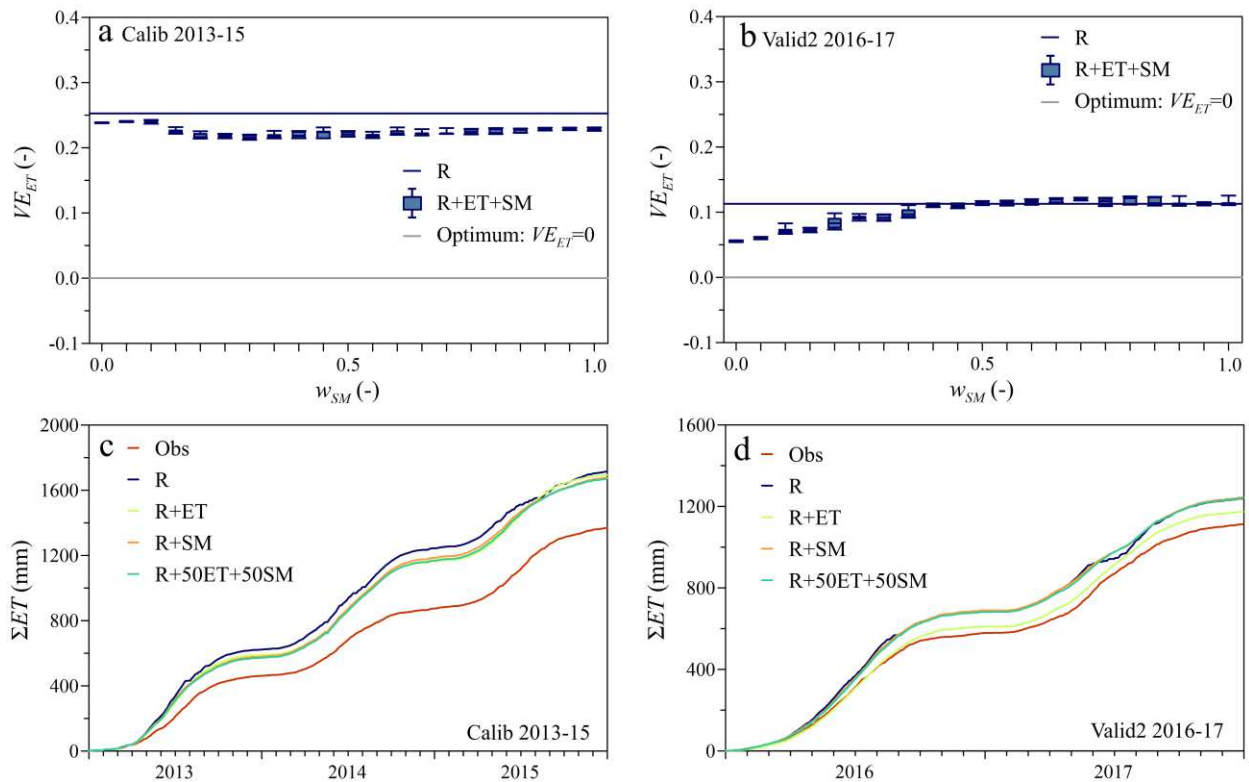


Figure 20. Soil moisture and evapotranspiration simulations: (a and b) Volume error for evapotranspiration VE_{ET} as a function of weight on soil moisture w_{SM} in the compound objective function shown as boxplots (R+ET+SM) and R scenario, when only runoff was used for calibration. (c and d) Simulated cumulative actual evapotranspiration ET when the model was calibrated only with runoff and evapotranspiration (R+ET), only with runoff and standardized soil moisture (R+SM), and a combination of $w_{ET}=0.5$ evapotranspiration and $w_{SM}=0.5$ standardized soil moisture (R+ET+SM). For comparison, observations (Obs) and R scenario are shown. Scenarios according to Table 6.

The monthly correlation coefficients for evapotranspiration and standardized soil moisture were above 0.8 both during model calibration and validation. The monthly correlation coefficient for standardized soil moisture was slightly lower, if only evapotranspiration was used in the objective function (Figure 21 b and d). When both evapotranspiration and soil moisture objectives were involved in the objective function, the proposed approach outperformed scenario R (Figure 21 a-d). The root mean squared error between simulated and observed monthly average evapotranspiration and standardized soil moisture was generally lower, when besides runoff extra information on evapotranspiration and/or soil moisture were involved in the model calibration (Figure 21 e-h).

Compared to scenario R, the simulated soil moisture and actual evapotranspiration. In scenario R, the parameter related to the limit for potential evapotranspiration $LPrat$ was very close to zero, meaning that actual evapotranspiration was almost always reaching its potential rate, even when the soil was very dry (Table 7). If evapotranspiration and soil moisture were used for model calibration, $LPrat$ was closer to 1.0. This means, that a certain wetness in the soils, i.e. more water was needed to reach potential evapotranspiration (Table 7). The nonlinear parameter for runoff

The added value of different data types for calibrating and testing a hydrological model in a small catchment

production β was below 1.0, if only runoff was used for model calibration (scenario R, Table 7), meaning that more water was allocated for runoff and less for soil moisture storage. Considering the clayish soil types in the HOAL, this is highly unrealistic. β was well above 1.0, if actual evapotranspiration and/or soil moisture were also used for model calibration (scenarios R+ET, R+50ET+50SM, R+SM). This refers to a more nonlinear runoff generation, which would be expected in the catchment considering the soil types. According to the soil survey, the catchment average field capacity FC was around 400 mm. This means, that the calibrated field capacities for scenarios R+ET and R seemed more realistic. However, when checking the simulated storage time series, the scenarios R+50ET+50SM and R+SM stood much closer to reality. In scenario R, when only runoff was used for model calibration, the root zone soil storage remained well below saturation and it even reached 0 mm in the summer months. This was not realistic and it was not supported by the soil moisture observations in the catchment. When soil moisture and actual evapotranspiration data were used together for model calibration, the root zone soil storage never dried out. In scenario R+ET, most of the water was stored in the soil, the upper zone storage was only a few mm during events and the lower zone storage was around 4 times smaller. When also soil moisture was used for model calibration, the water was more evenly distributed between the storage elements (scenarios R+SM, R+50ET+50SM).

For further analysis, calibrated model parameters from three main scenarios were chosen, when a weight of $w_{SM}=0.0$ (R+ET), $w_{SM}=1.0$ (R+SM), and $w_{SM}=0.5$ (R+50ET+50SM) was used on the soil moisture objective in the objective function when calibrating the soil moisture accounting routine.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

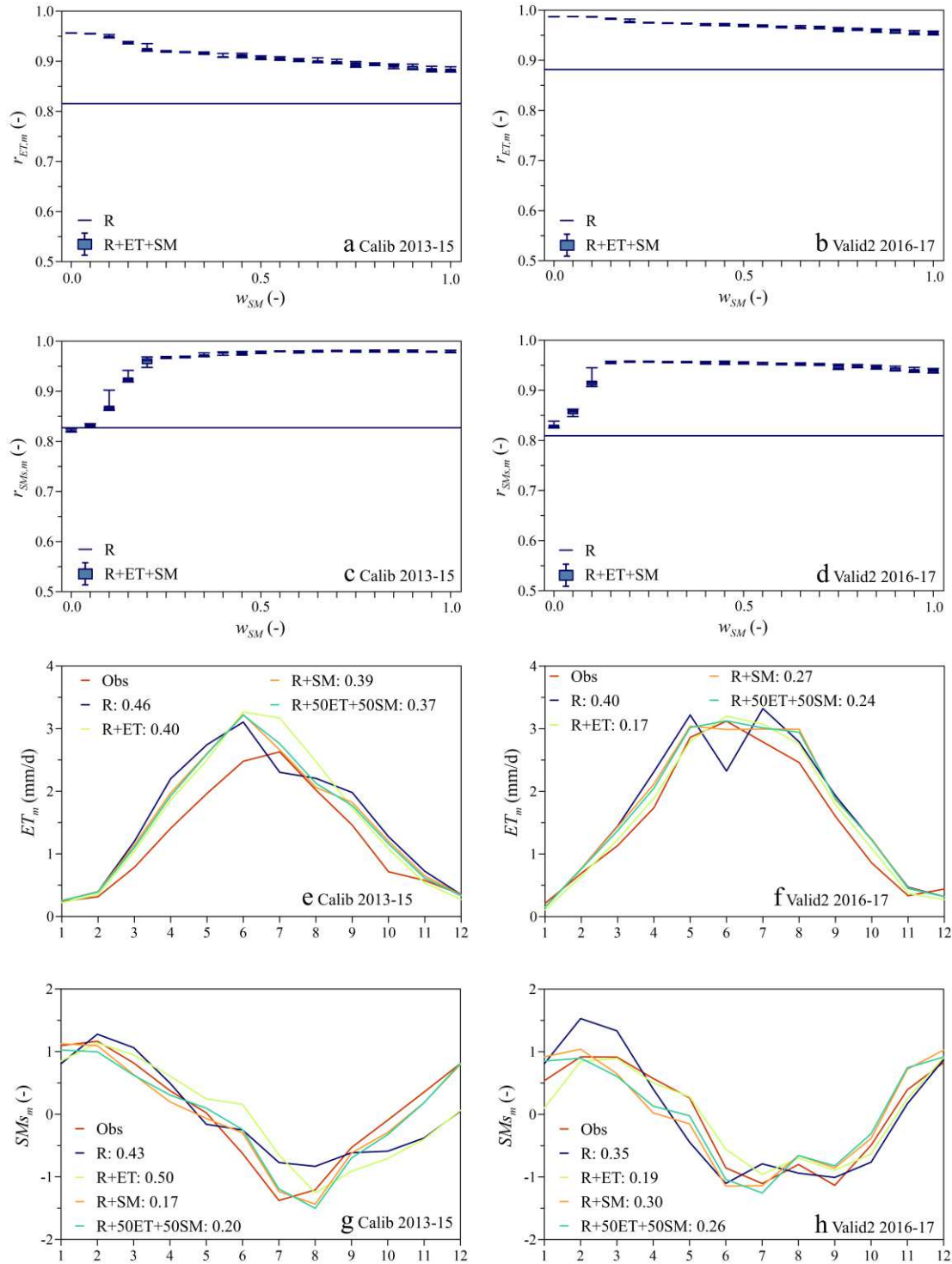


Figure 21. Soil moisture and evapotranspiration simulations: (a-d) Monthly Pearson correlation coefficients for evapotranspiration (ET) and standardized soil moisture (SMs) as a function of weight on soil moisture w_{SM} in the compound objective function shown as boxplots (R+ET+SM) and R scenario. (e-h) Monthly averages for ET and SMs for different scenarios as in Figure 20, according to Table 6. Numbers in legend indicate root mean squared error.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

The results indicate a clear tradeoff between accurately simulating actual evapotranspiration, soil moisture and runoff. Using additional information unequivocally improved evapotranspiration and soil moisture simulations, especially on the daily time scale (Table 11). Using additional information on evapotranspiration and soil moisture besides runoff generally improved runoff simulations during the validation periods, while it only slightly deteriorated the results during model calibration (Table 12). The relative volume error for runoff became one magnitude smaller during the 2-year-long validation period. The daily performance of runoff simulations became slightly worse during model calibration, but it improved during the 2-year-long validation if soil moisture was also used in the objective function (Table 12). The difference between the scenarios in terms of runoff efficiency was small during model calibration and the 22-year-long validation. While soil moisture information improved runoff efficiency and evapotranspiration information deteriorated it during the 2-year-long validation period (Table 12).

Table 11. Model calibration using runoff, soil moisture and evapotranspiration data: performance of evapotranspiration ET and standardized soil moisture SMs simulations (Volume error VE , monthly Pearson correlation coefficient r_m , daily Nash Sutcliffe efficiency $Nash$) during model calibration (2013-15) and second validation (2016-17) periods. Scenarios according to Table 6: R where only runoff, R+ET where runoff and evapotranspiration, R+SM where runoff and soil moisture, R+50ET+50SM where runoff, evapotranspiration and soil moisture were used for model calibration.

Variable	Scenario	Calibration period 2013-15			Validation period 2 2016-17		
		VE (-)	r_m ()	$Nash$ (-)	VE (-)	r_m ()	$Nash$ (-)
ET	R	0.25	0.82	-0.54	0.11	0.88	0.15
	R+ET	0.24	0.96	0.52	0.05	0.99	0.78
	R+SM	0.23	0.88	0.39	0.12	0.95	0.68
	R+50ET+50SM	0.22	0.91	0.45	0.11	0.97	0.72
SMs	R		0.83	0.55		0.81	0.60
	R+ET		0.82	0.66		0.83	0.64
	R+SM		0.98	0.93		0.94	0.86
	R+50ET+50SM		0.98	0.93		0.95	0.88

The added value of different data types for calibrating and testing a hydrological model in a small catchment

Table 12. Model calibration using runoff, soil moisture and evapotranspiration data: performance of runoff Q simulations (Volume error VE_Q , monthly Pearson correlation coefficient $r_{m,Q}$, daily logarithmic Nash Sutcliffe efficiency $lNash_Q$) during model calibration (2013-15), first and second validation (1991-2012 and 2016-17, respectively) periods. Scenarios according to Table 6: R+ET where runoff and evapotranspiration, R+SM where runoff and soil moisture, R+50ET+50SM where runoff, evapotranspiration and soil moisture were used for model calibration.

Scenario	Calibration period 2013-15			Validation period 1 1991-2012			Validation period 2 2016-17		
	VE_Q (-)	$r_{m,Q}$ ()	$lNash_Q$ (-)	VE_Q (-)	$r_{m,Q}$ ()	$lNash_Q$ (-)	VE_Q (-)	$r_{m,Q}$ ()	$lNash_Q$ (-)
R+ET	0.00	0.96	0.67	0.18	0.76	0.16	0.03	0.62	-0.52
R+SM	0.00	0.95	0.72	0.21	0.74	0.15	0.05	0.83	0.25
R+50ET+50SM	0.01	0.95	0.70	0.21	0.74	0.17	0.01	0.85	0.20

3.7.2.3. Runoff generation and routing simulation

The third step of the step-by-step parameter estimation was to fix the parameters of the runoff generation and routing module.

Using the parameters from the R+50ET+50SM scenario from the previous step, the model was run 10 times (Figure 22). The median of the relative number of days with good overland flow simulations, i.e. days when the model simulated very fast runoff (q_0) and overland flow was simultaneously observed, immediately exceeded 0.5 as soon as the weight on the overland flow part in the compound objective function was larger than zero (Figure 22 a). In terms of overland flow simulations, the results outperformed scenario R for each w_{OF} weight, as the model did not simulate overland flow at all when only runoff was used for model calibration. The median of the relative number of months with good storage change simulations, i.e. months when the sign of the model simulated storage change agreed with the sign of the observed storage change, became gradually 10-20% lower as the weight on the overland flow objective increased and during calibration the results also underperformed scenario R (Figure 22 c and d).

Compared to scenario R, when runoff generation data were also used for model calibration, the calibrated very fast storage time k_0 became smaller (1.1 day for scenario R, and below 0.5 day if runoff generation data were also used for model calibration, Table 7), while the fast and slow storage times, k_1 and k_2 , respectively, increased (Table 7). Faster overland flow and slower subsurface runoff correspond well with field observations. Overland flow events usually last a few hours according to camera observations, while the outflow from the subsurface reservoirs take several months due to the heterogeneous subsurface properties of the catchment.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

For further analysis, the runoff generation parameters of the three main scenarios (R+ET+G, R+SM+G, R+50ET+50SM+G) were calibrated by choosing the weight on the overland flow objective as $w_{OF}=0.5$.

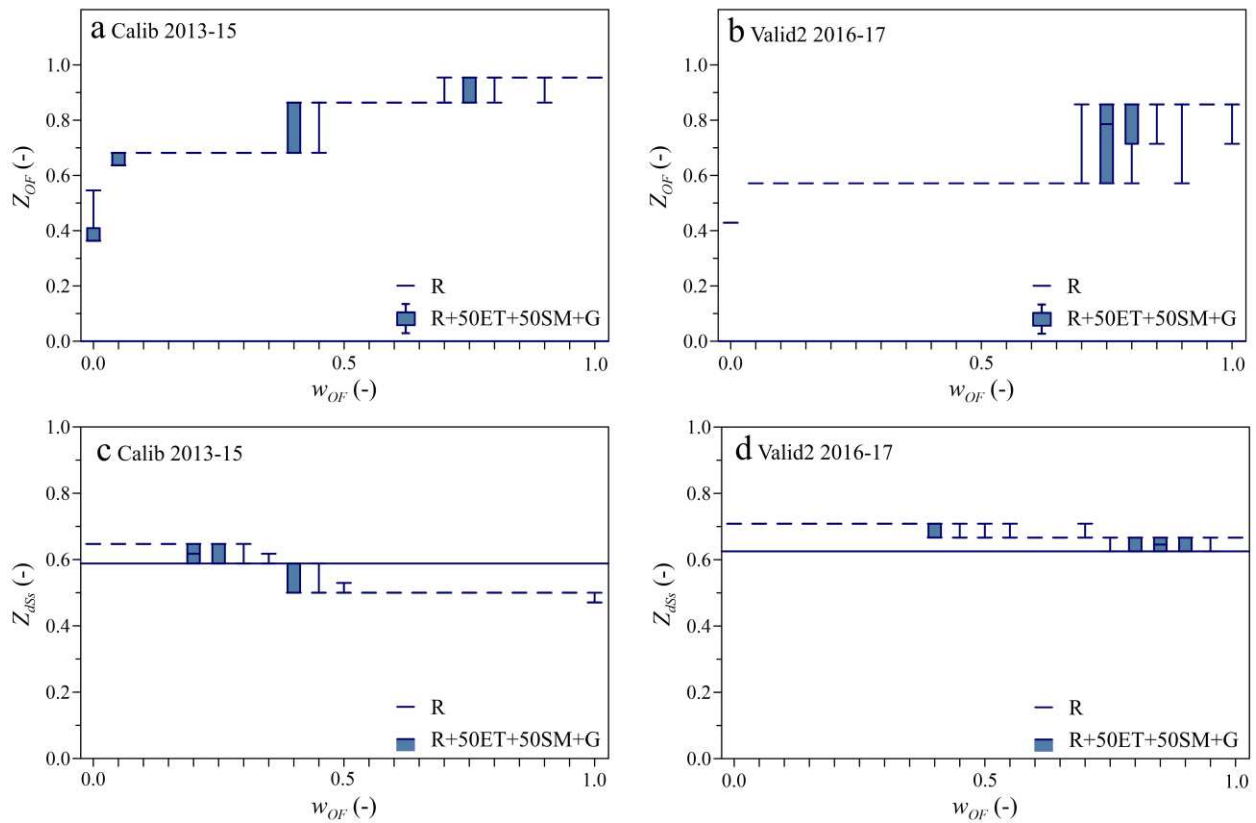


Figure 22. Overland flow and storage change simulations: (a and b) Relative number of days with good overland flow simulations Z_{OF} as a function of weight on overland flow part w_{OF} in the compound objective function (containing overland flow and storage change objectives) shown as boxplots. R scenario shown as line. (b and c) Relative number of months when the model correctly simulated the sign of the standardized storage change Z_{dSS} as a function of weight on overland flow part w_{OF} in the compound objective function shown as boxplots. R scenario shown as line.

Regarding the tradeoff between simulating runoff and runoff generation processes accurately, by using additional information on overland flow and storage change besides runoff further improved runoff simulations during the 2-year-long validation period, while it slightly deteriorated the daily performance during model calibration and validation (Table 13). Generally, the scenarios when either only soil moisture or both soil moisture and evapotranspiration were used in the model calibration performed the best (Table 13).

The added value of different data types for calibrating and testing a hydrological model in a small catchment

Table 13. Performance of runoff Q (Volume error VE_Q , monthly Pearson correlation coefficient $r_{m,Q}$, daily logarithmic Nash Sutcliffe efficiency $lNash_Q$) during model calibration (2013-15), first and second validation (1991-2012 and 2016-17, respectively) periods. Scenarios according to Table 6. Bold indicates the better performing scenario.

Scenario	Calibration period 2013-15			Validation period 1 1991-2012			Validation period 2 2016-17		
	VE (-)	r_m ()	$lNash$ (-)	VE (-)	r_m ()	$lNash$ (-)	VE (-)	r_m ()	$lNash$ (-)
R+ET+G	-0.01	0.85	0.52	0.18	0.70	0.05	0.02	0.57	-0.62
R+SM+G	0.00	0.91	0.63	0.21	0.75	0.10	0.03	0.81	0.26
R+50ET+50SM+G	0.00	0.92	0.59	0.20	0.76	0.09	-0.01	0.82	0.18

3.7.2.4. Runoff simulation

The final step was to evaluate the different scenarios in terms of simulating runoff on annual, seasonal and daily time scales (Figure 23-Figure 25).

When comparing the evolution of the runoff simulation efficiencies through the different scenarios during the 2-year-long validation period (2016-17), it is clear that using additional data besides runoff for model calibration improved the runoff simulation efficiencies (Figure 23). The relative volume error slightly increased when snow data were also used for model calibration. But it definitely decreased, when soil moisture, evapotranspiration, overland flow and storage change information was also used for model calibration. The monthly and daily runoff performance improved when additional data were also used for model calibration. Evapotranspiration and soil moisture data had the largest influence on runoff simulations, which also agreed with the results of the sensitivity analysis (Appendix A10). The sensitivity analysis showed that the most sensitive parameter of the model belonged to the soil moisture accounting module. During the two-year-long validation period, the results improved the best if either both soil moisture and evapotranspiration or only soil moisture information was used during model calibration besides runoff.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

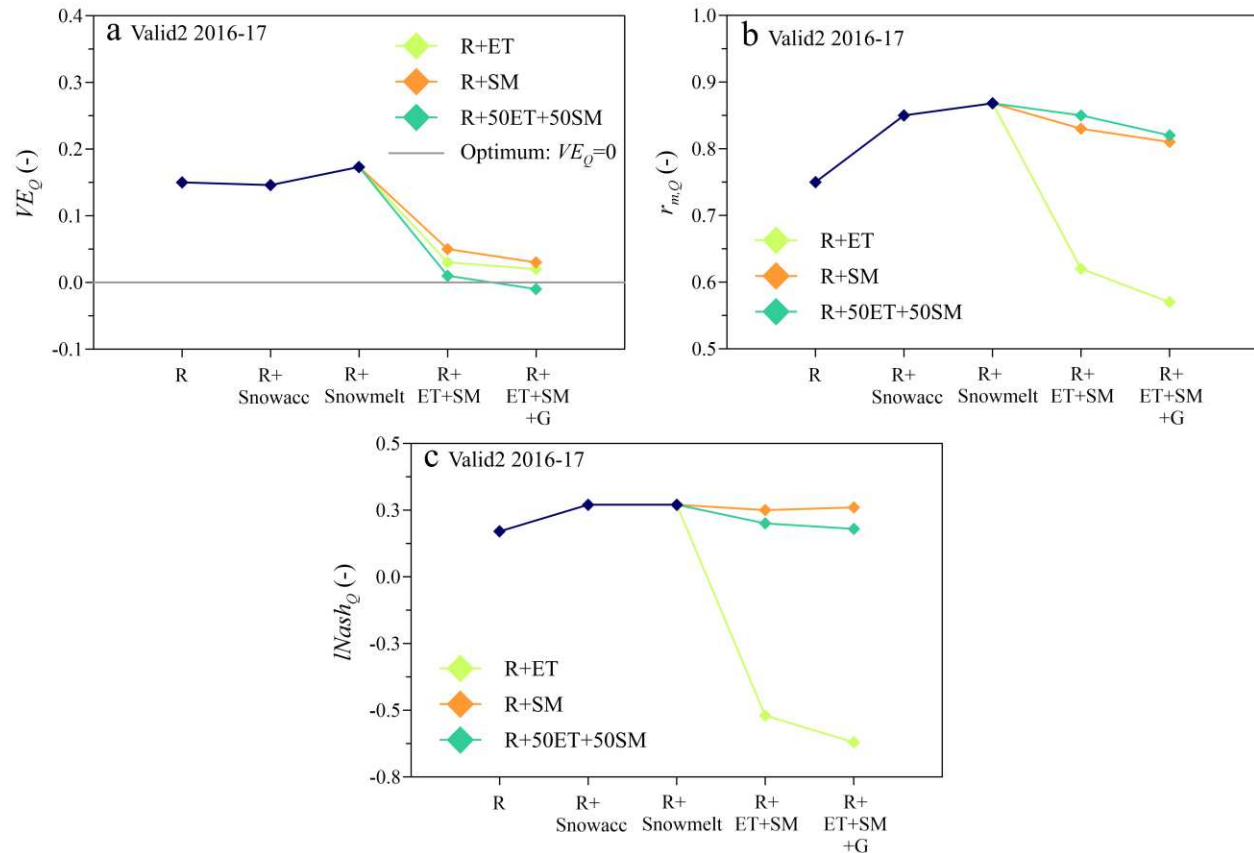


Figure 23. Runoff simulation efficiencies for each scenario during model validation (2016-17): (a) Volume error VE_Q . (b) Monthly Pearson correlation coefficient $r_{m,Q}$. (c) Daily logarithmic Nash Sutcliffe efficiency $lNash_Q$. Scenarios according to Table 6.

On the annual time scale, the performance of the proposed step-by-step approach was similar to the performance of the traditional scenario R, when only runoff was used for model calibration. The proposed method even outperformed scenario R during the shorter, two-year long validation period (Val2, 2016-17). Runoff was overestimated during the longer, 22-year-long validation period both by scenario R and the proposed calibration approach (Figure 24, a-c).

Similarly, on the seasonal time scale the proposed method could efficiently model runoff. The root mean squared error between simulated and observed monthly average runoff was lower during the validation periods (Figure 24, d-f).

The added value of different data types for calibrating and testing a hydrological model in a small catchment

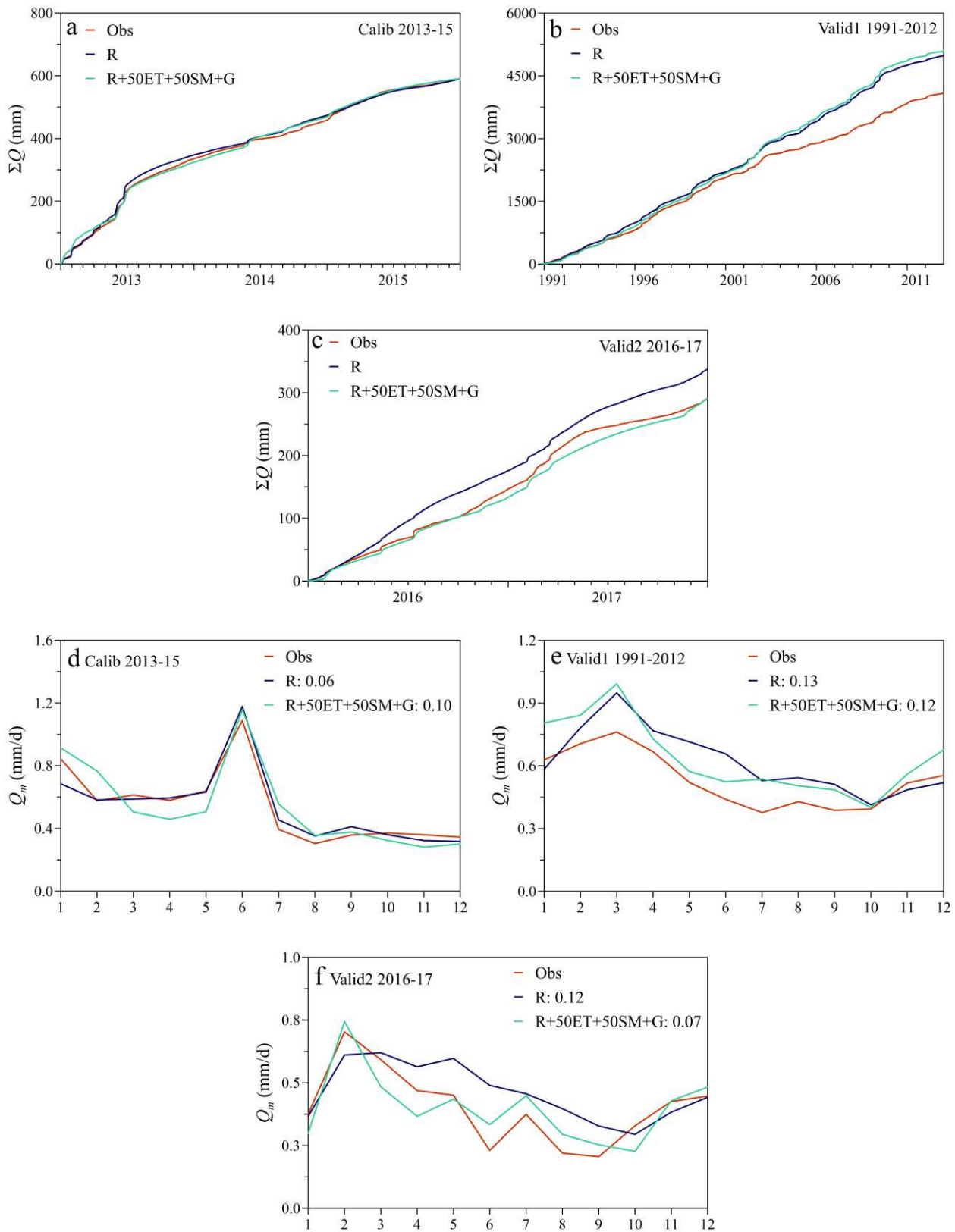


Figure 24. Runoff simulation: (a-c) Cumulative runoff. (d-f) Monthly average runoff Q_m . Scenarios according to Table 6. Numbers in legend indicate root mean squared error.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

On the daily time scale, the model performed worse during model calibration and the 22-year-long validation, if additional data besides runoff were used for model calibration. The peaks were underestimated and the recession was more delayed, when additional data were involved in the model calibration. But the daily performance slightly improved during the 2-year-long validation period by using additional data, and not only runoff (Figure 25).

The added value of different data types for calibrating and testing a hydrological model in a small catchment

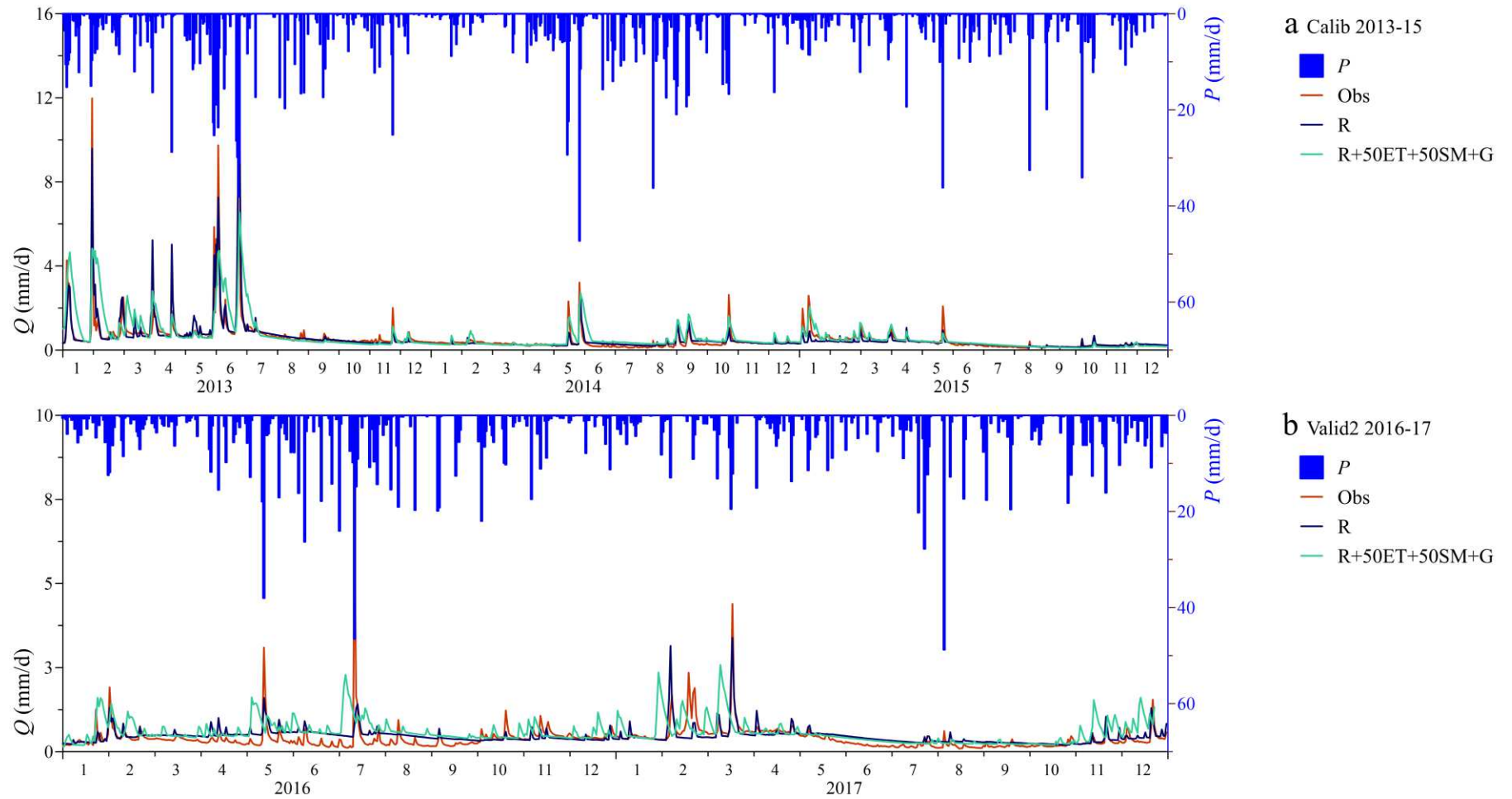


Figure 25. Runoff simulation: (a-b) Daily runoff. Scenarios according to Table 6.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

3.8. Discussion

Actual field measurements are of utmost importance to understand and model catchment processes. In this study, we showed that by using a variety of different field observations, we were able to simulate not only runoff, but also other hydrological processes generally efficiently using a lumped conceptual hydrological model and a stepwise model calibration approach. This means that the model simulated runoff well for the right reasons (Grayson et al., 1992) on the annual and seasonal time scales. Avanzi et al. (2020) drew similar conclusions, when they found that ground-based measurements allowed identifying more hydrological model parameters than runoff alone. Rakovec et al. (2016) also noted that although it was necessary to constrain the model against runoff, this was not sufficient to simulate other variables, for instance soil moisture, accurately.

In previous studies, which used additional data, not only runoff, to calibrate a hydrological model, the simulation of certain state variables usually improved. But generally, these were only the targeted state variables, such as snow water equivalent only, or soil moisture only, or evapotranspiration only. For instance, Kundu et al. (2017) reported that although using remotely sensed soil moisture improved the rainfall-runoff response, but not the routing simulations. Rajib et al. (2016) also pointed out that using remotely sensed soil moisture improved the simulation of surface soil moisture, but runoff, evapotranspiration and root zone soil moisture were less affected. In this study, we managed to improve snow accumulation, snowmelt, soil moisture, evapotranspiration, overland flow and storage change simulations at the same time.

While simulation of different state variables can improve, a common drawback of calibrating a rainfall-runoff model not only to runoff but also to other data (e.g. snow, soil moisture, evapotranspiration, groundwater levels) is the slight deterioration of the efficiency in simulating runoff (e.g. Gui et al., 2019; Parajka et al., 2007; Seibert, 2000). In this study, when additional data besides runoff were used for model calibration, runoff simulation efficiencies also slightly deteriorated during the calibration period, mainly on the daily time scale. The change was smaller during the validation periods. Moreover, we could improve runoff simulations during the second, 2-year-long validation period, on all time scales, i.e. annual, monthly and daily.

The sensitivity of runoff simulations to changing model parameters depended on the module of the model, as shown by the sensitivity analysis (Appendix A10). Although parameterizing the snow module in this catchment is not as relevant as in an Alpine catchment due to the small amount of snow, the proposed method might be useful for other studies, where the precipitation partitioning is a crucial modeling step but usually lacks validation (e.g. Jennings et al., 2018). Parameters describing the soil moisture accounting routine of the model were the most sensitive according to the sensitivity analysis. Therefore, in this study observations of soil moisture and evapotranspiration played the most important role in parameter estimation. Using either both evapotranspiration and soil moisture or only soil moisture observations, the model's most sensitive routine could be parameterized well. Previous studies are mixed in terms of whether soil moisture and/or evapotranspiration observations improve the efficiency of runoff simulations. Nijzink et al. (2018) found that soil moisture satellite products were more effective than evaporation products for deriving more constrained parameter distributions. López et al. (2017) showed that estimating runoff was more efficient, if both soil moisture and evapotranspiration satellite products were involved in the model calibration. Bergström & Lindström (2015) argued that the relative importance of these observations of course depends on the time step of the model considering that evapotranspiration volumes are less than storage in the unsaturated zone on a daily basis. Similarly,

The added value of different data types for calibrating and testing a hydrological model in a small catchment

Baroni et al. (2019) pointed out that different data collection strategies should be considered for different variables, depending on their degree of coupling to the atmosphere. Considering that changes in the atmosphere are faster, evapotranspiration observations are necessary on a daily basis, while changes in soil moisture and groundwater levels are more consistent over time, as these are more decoupled from the atmosphere.

Linking observations with model simulations is not a straightforward task and local experience with the catchment processes may be a substantial advantage (Avanzi et al., 2020; Holländer et al., 2009). For physically-based and spatially distributed models, modelers can often directly and explicitly compare measured and simulated volumes of water and energy fluxes (e.g. Kuras et al., 2011; Thyer et al., 2004). For conceptual hydrological models, especially if the model is spatially lumped, a balance between the lumped conceptual model concept and spatial and temporal variability of processes has to be found. The spatial organization of soil moisture and other variables within the catchment may matter for runoff generation which cannot be captured by lumped models as they aggregate spatial processes variability (Blöschl et al., 1995; Viglione et al., 2018; Western et al., 1998;). In this study, our aim was to use such objective functions and compare such quantities that helped bridging the gap between the model and observations. Therefore, instead of comparing volumes, we compared the standardized values of observed and simulated soil moisture and storage change. For the snow module, we used a binomial snow cover index based on different types of measurements to decide if there was snow in the catchment and a threshold for simulated snow water equivalent, only above which the catchment was considered snow covered. For the subsurface module, we compared the sign of the standardized storage change. When different objectives were tested instead of using these, the model often gave unrealistic results.

It is important to note that the aim of this study was not to modify or optimize the structure of the model, but to use a conceptual model structure that has been used in various climatic regions in previous studies both in gauged and ungauged basins (e.g. Blöschl et al., 2013; Parajka et al., 2007; Slezziak et al., 2018), and that is general enough so that it could be used to simulate runoff in arbitrarily chosen basins. Previous studies proved that this type of model structure works well in Austrian catchments, both in Alpine and lowland environments (e.g. Parajka et al., 2007; Slezziak et al., 2018). Based on our current understanding on the HOAL, i.e. how this catchment works conceptually, the HBV model structure is a very good representation of the catchment. Regarding the runoff generation in the catchment, the saturation excess overland flow events in the valley bottom identified by the weather station camera according to Silasari et al. (2017) could be well matched with the very fast runoff response of the HBV-based TUWmodel. If a certain storage state threshold LS_{UZ} is reached in the upper reservoir of the model, i.e. filling up from the bottom, which corresponds to saturation excess runoff mechanisms, very fast runoff starts. Regarding the subsurface mechanisms, Exner-Kittridge et al. (2016) distinguished between two main aquifers in the HOAL, a shallow and a deep one. Their contribution to runoff depends on the hydrological conditions, e.g. low or high flow conditions, seasonality, etc. The runoff generation module of the HBV-based TUWmodel consists of two subsurface reservoirs, the upper reservoir contributes to rainfall events, while the outflow from the lower reservoir takes place on much longer, monthly time scales.

Although we had a certain understanding how the catchment works conceptually, on small catchment scale the role of inhomogeneous surface and subsurface properties (e.g. complex geology, cracks and earthworm paths resulting in preferential flow paths) which result in nonlinear

The added value of different data types for calibrating and testing a hydrological model in a small catchment

responses and processes with different thresholds (e.g. flow paths which are only activated above a certain groundwater level) can be more dominant than on large catchment scale where the response is averaged out. On small catchments, there are several exceptional cases, where the model does not work. The shorter the time scale, the more pronounced and visible these exceptions. The model in this study also underperformed on the daily time scale for simulating runoff during the validation periods. This underperformance slightly improved during the second, 2-year-long validation period where additional data were also used for model calibration. A possible reason for this daily underperformance for runoff could be that the model was trained on three years (2013-15), which were extreme years including a very wet year (2013) and two very dry years (2014-15). During extreme hydrological conditions, i.e. wet or dry years, the area contributing to runoff, the active flow paths and the catchment rainfall response can be very different. A lumped model often cannot handle these differences. A possible solution might be to use a distributed model, where a difference in model structure could better take into account the inhomogeneities of a small catchment, such as different runoff generation mechanisms, groundwater level dependent switches in flow paths. In this study we used only the lumped version of the model, which performed well on the annual and seasonal time scales. In order to improve the daily performance of the model, a spatial distribution and modified model structure might be necessary.

This study showed that by using field measurements of input and output fluxes and states besides runoff, we were able to simulate these fluxes and states (snow, soil moisture, evapotranspiration, storage), moreover also the annual and seasonal runoff more efficiently. This finding suggests that hydrological models which are constrained only by runoff might be simulating runoff well for the wrong reasons, that is, with wrong parameters. But this can be only revealed, if other state variables of the model are tested against observations. For small basins, where satellite information may be too coarse, field observations play a crucial role. For instance, if data on evapotranspiration, soil moisture or saturation areas are available, these may be complementary to existing runoff data, or surrogates of runoff data if no runoff measurements exist. They can all help in constraining a hydrological model.

3.9. Conclusions

This study presented a new framework for estimating the parameters of runoff model components in a stepwise fashion from field observations of input and output fluxes and states besides runoff and investigated the value of different these data for simulating runoff well for the right reasons in a small agricultural catchment. Our results showed that:

- By using the proposed step-by-step model calibration approach with different field observations of input and output fluxes and states besides runoff for parameter estimation, we were able to efficiently simulate these fluxes and states correctly. This means that we simulated runoff well for the right reasons on the annual and seasonal time scales.
- For this catchment, field observations of soil moisture and evapotranspiration had the largest influence on runoff simulations.
- Future research and possibly a spatially distributed and modified model structure might be necessary on a small catchment to take into account the role of small scale inhomogeneities and to improve the daily performance of the model.

The added value of different data types for calibrating and testing a hydrological model in a small catchment

3.10. Acknowledgements

The authors would like to acknowledge financial support provided by the Austrian Science Funds (FWF) as part of the Vienna Doctoral Program on Water Resource Systems (DK W1219-N28). The data used are listed in the tables and the data necessary to reproduce the reported findings is available on the following site: <https://owncloud.tuwien.ac.at/index.php/s/uQfz97O14srsW50> (Password: 20cWRs19).

4. Stepwise prediction of runoff using proxy data in a small agricultural catchment

4.1. General

The aim of this chapter was to test the method presented in the previous chapter on an ungauged catchment. We aimed to investigate how efficiently the variety of field observations can be used alone, without using runoff data, to predict runoff in a small agricultural catchment.

The present chapter corresponds to the following scientific publication in its original form:

Széles, B., Parajka, J., Hogan, P., Silasari, R., Pavlin, L., Strauss, P., & Blöschl, G. (2020). Stepwise prediction of runoff using proxy data in a small agricultural catchment. *Journal of Hydrology and Hydromechanics*. Under review.

4.2. Key points

1. Using only snow and soil moisture information for calibration, the runoff model performance was comparable to the scenario when the model was calibrated in one step, using only runoff measurements.
2. By using proxy data for model calibration, the simulation of state variables and therefore the process consistency improved, implying that the model represents reality better than the scenario when only runoff was used for model calibration.

4.3. Abstract

In this study, the value of proxy data was explored for calibrating a conceptual hydrological model for small ungauged basins, i.e. ungauged in terms of runoff. The study site was a 66 ha Austrian experimental catchment dominated by agricultural land use, the Hydrological Open Air Laboratory (HOAL). The three modules of a conceptual, lumped hydrological model (snow, soil moisture accounting and runoff generation) were calibrated step-by-step using only proxy data, and no runoff observations. Using this stepwise approach, the relative runoff volume errors in the calibration and first and second validation periods were -0.04, 0.19 and 0.17, and the monthly Pearson correlation coefficients were 0.88, 0.71 and 0.64, respectively. By using proxy data, the simulation of state variables improved compared to model calibration in one step using only runoff data. Using snow and soil moisture information for model calibration, the runoff model performance was comparable to the scenario when the model was calibrated using only runoff data. While the runoff simulation performance using only proxy data did not considerably improve compared to a scenario when the model was calibrated on runoff data, the more accurately simulated state variables imply that the process consistency improved.

4.4. Introduction

Runoff is a reflection of the aggregated hydrological catchment behavior. Therefore, in most cases, runoff observations are used for calibrating hydrological models. However, in many catchments runoff observations are not available (Blöschl et al., 2013) and therefore, other measurements on the hydrological processes, i.e. proxy data, are used to calibrate the model.

There are only a few studies that used only proxy data for parameter estimation in ungauged basins (Parajka et al., 2013). These studies were mainly focusing on physically based hydrological

models, where model simulations can be explicitly linked to field measurements. For instance, Thyer et al. (2004) evaluated the performance of the distributed hydrology soil vegetation model (DHSVM) using field data. Their study site was a high elevated and forested catchment, where they found that the simulated runoff was most influenced by snowmelt characteristics (Thyer et al., 2004). While Thyer et al. (2004) focused on the micro-meteorological part of the model, in a follow-up study, Kuras et al. (2011) completed this evaluation by testing the subsurface and surface runoff dynamics. Both studies achieved a daily Nash Sutcliffe efficiency for runoff of above 0.75. Kuppel et al. (2018) performed a similar study in the Scottish Highlands with a distributed process-based eco-hydrological model. They found that the model performance was better, when runoff was also used for calibration and they also noted that certain state variables can be only well simulated when the model was calibrated for them.

If in-situ measurements are unavailable, an alternative could be to use remote sensing products for model calibration (López et al., 2017; Nijzink et al., 2018; Silvestro et al., 2015). Nijzink et al. (2018) tested nine remotely sensed products and found that without using runoff data, remotely sensed soil moisture products and the GRACE total water storage anomalies constrained the most the model parameters. López et al. (2017) found that remotely sensed evapotranspiration and remotely sensed soil moisture should be used together, and not independently, to predict runoff. Generally, due to the coarse spatio-temporal resolution of these products, they cannot be used for small catchments. For small catchments, in-situ observations are necessary.

With or without runoff data an efficient way of model calibration is stepwise parameter estimation which reduces the dimensionality of the problem. With runoff data, model parameters can be grouped according to which runoff signatures they influence (e.g. Fenicia et al., 2007; Gelleszun et al., 2017; Hogue et al., 2000), or on which time scales the model parameters are sensitive (e.g. Lu & Li, 2015). Some of the studies also used proxy data to perform a step-by-step model calibration (e.g. Avanzi et al, 2020; Hay et al, 2006; Kuras et al., 2011; Ning et al., 2015), which is very useful in order to understand possible mismatches between model simulations and measurements (Rogger et al., 2012). In a recent study, Széles et al. (2020) proposed a stepwise model calibration approach, where they aimed to calibrate a conceptual hydrological model according to the simulated processes, such as snow accumulation and snowmelt, soil moisture and evapotranspiration, and runoff generation. They linked the simulated processes with a variety of in-situ field observations. These proxy data and runoff data were together used in their study to estimate the parameters of their conceptual hydrological model. However, it was not yet clear whether this method could be potentially used in ungauged catchments, to predict runoff without using runoff observations.

The objective of this study was to test whether the stepwise model calibration approach proposed by Széles et al. (2020) could be used for predicting runoff without using runoff observations. Without incorporating runoff in the objective functions, we aimed to test how well we can predict runoff and various state variables of the model on the annual and seasonal time scales. The analysis was performed in the 66 ha Austrian Hydrological Open Air Laboratory (HOAL), where long-term field observations are available (Blöschl et al., 2016).

4.5. Study area and data

4.5.1. Study area

The study site was a 66 ha experimental catchment, the Hydrological Open Air Laboratory (HOAL) in Petzenkirchen, Lower Austria (chapter 3.5.1.).

4.5.2. Data

In this study, we used the same data presented by Széles et al. (2020). The measurements included precipitation amount, precipitation type, runoff observations and time lapse photographs with one minute temporal resolution. Air temperature has been measured at 7, 14, 19h until October 2012, since then it has been measured with half hourly time step. Snow depth, soil moisture and actual evapotranspiration have been monitored with half hourly temporal resolution. Groundwater levels have been measured every five minutes. Details on the instruments, their location and spatio-temporal resolution are given in Blöschl et al. (2016) and Széles et al. (2020).

Three time periods were selected for the analysis, a 22-year-long period when only runoff measurements (1991-2012), and a 3-year-long (2013-15) and a 2-year-long (2016-17) period when runoff measurements and additional sources of data were available. The 3-year-long period was used for model calibration (Calib), the 22-year-long (Val1) and 2-year-long (Val2) periods for model validation. One year preceding each period was used as warm-up period. Snow accumulation was simulated with half hourly temporal resolution, while other processes were simulated with daily time step.

4.6. Methodology

4.6.1. Hydrological model

In this study we used a conceptual hydrological model, the TUWmodel (Parajka et al., 2007), which follows the structure of the HBV model (Bergström, 1976; Bergström & Linström, 2015; Lindström et al., 1997). The model has three modules (snow, soil moisture accounting and runoff generation) and 14 free parameters (Merz & Blöschl, 2004; Parajka et al., 2007; Széles et al., 2020). The free parameters according to the three modules and their calibration ranges are shown in Table 14. The ranges were specified based on literature values (Merz et al., 2011; Viglione et al., 2013), except for field capacity FC , which was constrained according to a soil survey (Murer et al., 2004).

In this study, we followed the stepwise calibration approach from Széles et al. (2020) but without using runoff in the optimization steps. The results were compared to a scenario, when the model was calibrated in one step, using only runoff data.

Table 14. 14 free parameters of the hydrological model according to the three modules and their calibration range.

Module	Parameter (unit)	Parameter name	Calibration range: min÷max
Snow	SCF (-)	Snow correction factor	0.9÷1.5
	DDF (mm/°C/d)	Degree day factor	0.0÷5.0
	T_{wb} (°C)	Wet bulb temperature, i.e. threshold temperature below which precipitation is snow	-3.0÷1.0
	T_m (°C)	Threshold temperature above which melt starts	-2.0÷2.0
Soil moisture accounting	LP_{rat} (-)	Limit for potential evapotranspiration	0.0÷1.0
	FC (mm)	Field capacity	0.0÷450.0
	β (-)	Non-linear parameter for runoff production	0.0÷20.0
Runoff generation	k_0 (d)	Storage time for very fast response	0.0÷2.0
	k_1 (d)	Storage time for fast response	2.0÷30.0
	k_2 (d)	Storage time for slow response	30.0÷250.0
	LS_{UZ} (mm)	Threshold storage state for very fast runoff	1.0÷100.0
	c_P (mm/d)	Constant percolation rate	0.0÷8.0
	B_{MAX} (d)	Maximum base at low flows	0.0÷30.0
	c_R (d ² /mm)	Free scaling parameter	0.0÷50.0

4.6.2. Model calibration without runoff data

The three modules of the model were calibrated step-by-step. The separate steps focused on the rainfall-runoff processes, which were calibrated using field measurements. In this way, the free parameters were step-by-step fixed, according to the modules of the model. The scenarios are listed in Table 15.

4.6.2.1. Calibration of snow module

First, all the model parameters were calibrated and the temperature threshold parameter (wet bulb temperature T_{wb}) was fixed (Scenario Sim-Snowacc, Table 15) by fitting the modelled phase of the precipitation to the observed one. The precipitation phase was measured by a present weather sensor at the weather station. The number of half hours with false precipitation phase simulations was minimized using the DEoptim R package for parameter optimization (Ardia et al., 2010a, 2010b, 2016; Mullen et al., 2011). The wet bulb temperature parameter was fixed. Based on the results from Széles et al. (2020), we used wet bulb temperature instead of air temperature for defining threshold of the precipitation phase.

In the next step, the remaining three snowmelt parameters (scenario Sim-Snowmelt, Table 15) were fixed. A daily snow cover index was created (showing 1 if there was snow in the catchment, otherwise 0) using 3 types of measurements. First, time lapse photos were checked to decide if there was snow in the catchment. If these were unavailable, daily MODIS Normalized Difference Snow Index images were analyzed (Hall & Riggs, 2016a, 2016b). Finally, if the MODIS images were also unavailable, the snow sensor measurements were examined. The modelled snow cover index was chosen to be 1, if the snow water equivalent exceeded 2 mm. The modelled snow cover index was fitted to the observed one by minimizing the number of days with false snow cover index simulations and using the DEoptim R package for parameter optimization. Out of the 13 calibrated parameters, the three snowmelt parameters were fixed.

4.6.2.2. Calibration of the soil moisture accounting module

In scenarios Sim-ET+SM (Table 15), the soil moisture accounting module parameters were fixed. For soil moisture, in-situ soil moisture measurements were used. In order to describe the temporal dynamics of soil moisture in the catchment, measurements of all stations were averaged. To compare measured and modelled soil moisture, we compared standardized soil moisture values according to equation (36)

$$SM_s = \frac{SM - \overline{SM}}{\sigma_{SM}} \quad (36)$$

where SM_s (-) is the simulated standardized soil moisture, SM (mm) is the simulated soil moisture, \overline{SM} (mm) and σ_{SM} (mm) are the average and the standard deviation of the simulated soil moisture. Observed soil moisture was standardized in a similar way to equation (36). For actual evapotranspiration, average evapotranspiration was calculated over the catchment based on measurements of three eddy covariance stations. According to the land use types, an area weighted evapotranspiration was calculated using the measurements of an eddy covariance system at the weather station (representing grass evapotranspiration) and two mobile systems (representing different crop evapotranspiration). To estimate the evapotranspiration from the riparian forest next to the stream, crop coefficients were introduced. A multi-objective function ZI (-) according to equation (37) was maximized by optimizing 10 model parameters with the help of the DEoptim R package. ZI consisted of the daily Nash Sutcliffe efficiency for standardized soil moisture Z_{SM} and evapotranspiration Z_{ET} with different weights

$$Z1 = w_{SM}Z_{SM} + w_{ET}Z_{ET} \quad (37)$$

where w_{SM} (-) is the weight on the soil moisture objective, between 0 and 1. The weight on the evapotranspiration objective w_{ET} (-) is the difference between 1 and w_{SM} . This optimization step was repeated ten times to check the stability of the optimized model parameters. The results were examined on two time scales, annual and seasonal. On the annual time scale, the volumes of observed and simulated actual evapotranspiration were compared and the relative volume error VE_{ET} (-) was calculated (Criss & Winston, 2008). On the seasonal time scale, monthly average observed and simulated daily actual evapotranspiration and standardized soil moisture were compared and the monthly Pearson correlation coefficient for evapotranspiration $r_{ET,m}$ (-) and standardized soil moisture $r_{SMs,m}$ (-) were calculated. Three main scenarios were chosen and the soil moisture accounting module parameters were fixed according to these. In these scenarios w_{SM} was chosen to be 0, 0.8 (where the relative volume error for ET was the smallest during the second

validation period), and 1.0 (scenarios Sim-ET, Sim-ET20-SM80, and Sim-SM, respectively, Table 15).

4.6.2.3. Calibration of the runoff generation module

In order to optimize the very fast runoff q_0 simulations, saturation excess runoff events were identified according to Silasari et al. (2017). The days with very fast runoff simulations were calibrated to the days when saturation excess runoff events were observed. Storage change in the lower zone dS_{LZ} (mm/month) was calibrated using piezometer measurements. Based on the observed groundwater levels monthly storage change values dS_o (mm/month) were calculated for each piezometer. A catchment average storage change was calculated from spatially interpolated storage change values. The storage change values were standardized, and the simulated monthly average standardized storage change dS_s (-) was fitted to the observed one dS_{s_o} (-). A multi-objective function $Z2$ (-) according to equation (38) was maximized for calibrating the remaining 7 model parameters with the help of the DEoptim R package. $Z2$ consisted of the relative number of days with correctly modelled very fast runoff Z_{OF} (-) and the relative number of months with correctly modelled sign of the standardized storage change Z_{dS} (-) with different weights

$$Z2 = w_{OF}Z_{OF} + w_{dS}Z_{dS} \quad (38)$$

where w_{OF} (-) is the weight on the overland flow OF objective, ranging between 0 and 1. The weight on the storage change objective w_{dS} (-) is the difference between 1 and w_{OF} . This optimization step was repeated 10 times to check the stability of the optimized model parameters. The modelling results were evaluated on a daily time scale for overland flow, by analyzing Z_{OF} as a function of w_{OF} . The modelling results for storage change simulations were assessed on the monthly time scale, by analyzing Z_{dS} as a function of w_{OF} . For the selected scenarios (scenarios Sim-ET+G, Sim-SM+G, Sim-ET20+SM80+G, Table 15), w_{OF} was chosen to be 0.5.

The modelling efficiency in terms of simulating runoff was evaluated on annual and monthly time scales. On the annual time scale, the volumes of observed and simulated runoff were compared and the relative volume error for runoff VE_Q (-) was calculated. On the seasonal time scale, monthly average observed and simulated runoff time series were compared and the monthly Pearson correlation coefficient for runoff $r_{Q,m}$ (-) was calculated.

4.6.3. Model calibration with runoff data

Simulation results were compared with a scenario, when only runoff was used for model calibration and the model parameters were estimated in one step (Scenario Sim-R, Table 15). The model was calibrated to observed runoff by minimizing the daily root mean square error between observed and simulated runoff using the DEoptim R package.

Table 15. Scenarios presented in the study.

Scenario name	Details
Sim-Snowacc	Calibration of all model parameters and fixing the temperature threshold parameter (wet bulb temperature T_{wb}) using snow accumulation data
Sim-Snowmelt	Calibration of all model parameters except T_{wb} and fixing the snowmelt parameters (snow correction factor SCF , degree day factor DDF , snowmelt temperature T_m) using snow cover data
Sim-ET+SM	Calibration of soil moisture accounting and runoff generation parameters and fixing the soil moisture accounting module parameters (field capacity FC , nonlinear parameter for runoff production β , limit for potential evapotranspiration $LPrat$) using evapotranspiration and soil moisture objectives
Sim-ET	Calibration of soil moisture accounting and runoff generation parameters and fixing the soil moisture accounting module parameters (see scenario Sim-ET+SM) using only evapotranspiration objective
Sim-SM	Calibration of soil moisture accounting and runoff generation parameters and fixing the soil moisture accounting module parameters (see scenario Sim-ET+SM) using only soil moisture objective
Sim-ET20+SM80	Calibration of soil moisture accounting and runoff generation parameters and fixing the soil moisture accounting module parameters (see scenario Sim-ET+SM) using a combination of $w_{ET}=20\%$ evapotranspiration and $w_{SM}=80\%$ soil moisture objectives
Sim-ET+G	Calibration of runoff generation parameters, the soil moisture accounting module parameters were fixed in scenario Sim-ET
Sim-SM+G	Calibration of runoff generation parameters, the soil moisture accounting module parameters were fixed in scenario Sim-SM
Sim-ET20+SM80+G	Calibration of runoff generation parameters, the soil moisture accounting module parameters were fixed in scenario Sim-ET80+SM20
Sim-R	Calibration of all model parameters in one step using only runoff observations

4.7. Results

4.7.1. Snow module simulations

Calibrating the wet bulb temperature T_{wb} , to the observations from the present weather sensor (Scenario Sim-Snowacc) gave 0.31 and 0.40% of poor simulation times steps in the calibration and validation periods, respectively. This was slightly better than the simulations that used only runoff for calibration (Scenario Sim-R) (Table 16).

Fixing the snowmelt parameters, such as snowmelt temperature T_m , degree day factor DDF and snow correction factor SCF , to the observed snow cover index (Scenario Sim-Snowmelt) gave 4.38 and 6.29% of poor simulation time steps which, again, was better than the Sim-R scenario (Table 16).

Table 16. Performance of snow accumulation and snowmelt simulations for three scenarios (Sim-R, Sim-Snowacc, Sim-Snowmelt) in the calibration and validation periods. Snow simulation efficiency is described by the number of time steps with poor (i.e. when the simulated phase of the precipitation and simulated snow cover index, respectively, mismatched the observed one) snow accumulation and snowmelt simulations relative to the number of time steps with observations. Scenarios are described in Table 15.

Scenario	Relative number of time steps with poor snow accumulation simulations (%)		Scenario	Relative number of time steps with poor snowmelt simulations (%)	
	Calibration period 2013-15	Validation period 2 2016-17		Calibration period 2013-15	Validation period 2 2016-17
Sim-R	0.45	0.52	Sim-R	4.66	7.25
Sim-Snowacc	0.31	0.40	Sim-Snowmelt	4.38	6.29
Number of half hourly time steps	35626	23972	Number of daily time steps	1095	731

4.7.2. Soil moisture accounting module simulations

The relative volume error and the cumulative values of actual evapotranspiration calculated for 4 scenarios (Table 15) are compared in Figure 26. The smallest relative volume errors of actual evapotranspiration were achieved for soil moisture weights of $w_{SM}=0.3$ and $w_{SM}=0.8$ in the calibration and validation periods based on the average of 10 model runs (Figure 26 a and b). The model tended to overestimate evapotranspiration significantly (Sim-ET, Figure 26 c and d), except when evapotranspiration was used in the objective function with a higher weight. Compared to the

Sim-R scenario, the annual performance of actual evapotranspiration simulations improved for all w_{SM} weights, except for $w_{SM}=1$ during validation.

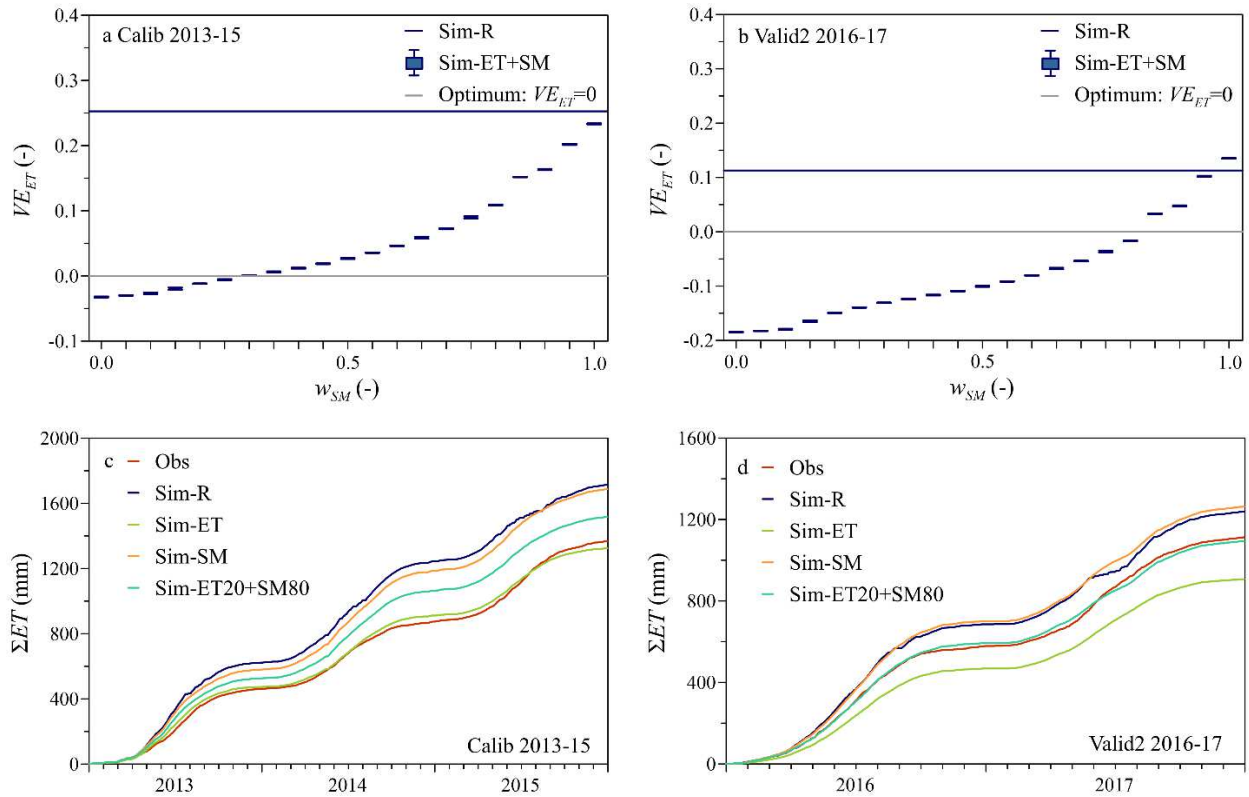


Figure 26. Performance of model simulations in terms of relative volume error for actual evapotranspiration VE_{ET} as a function weight on soil moisture objective w_{SM} (panels a and b) and cumulative actual evapotranspiration ΣET for 4 scenarios (panels c and d) in the calibration (panels a and c) and validation periods (panels b and d). Scenarios are described in Table 15.

The monthly correlation coefficient for evapotranspiration was above 0.75 both during model calibration and validation. While the monthly correlation coefficient for standardized soil moisture was low if only evapotranspiration was used in the objective function, the monthly correlation for evapotranspiration was almost constant independently from the weight on the evapotranspiration objective (Figure 27 a-d). When both evapotranspiration and soil moisture objectives were involved in the objective function, the proposed approach generally outperformed the Sim-R scenario (Figure 27 a-d). Among the analyzed scenarios for the proposed step-by-step model calibration, the monthly average evapotranspiration rates and the standardized soil moisture could be best simulated, when both evapotranspiration and soil moisture were involved in the objective function (Figure 27 e-h).

For further analysis, three main scenarios were chosen, when a weight of $w_{SM}=0.0$ (Sim-ET), $w_{SM}=1.0$ (Sim-SM), and $w_{SM}=0.8$ (Sim-ET20-SM80) was used on the soil moisture objective in the objective function when calibrating the soil moisture accounting routine.

Stepwise prediction of runoff using proxy data in a small agricultural catchment

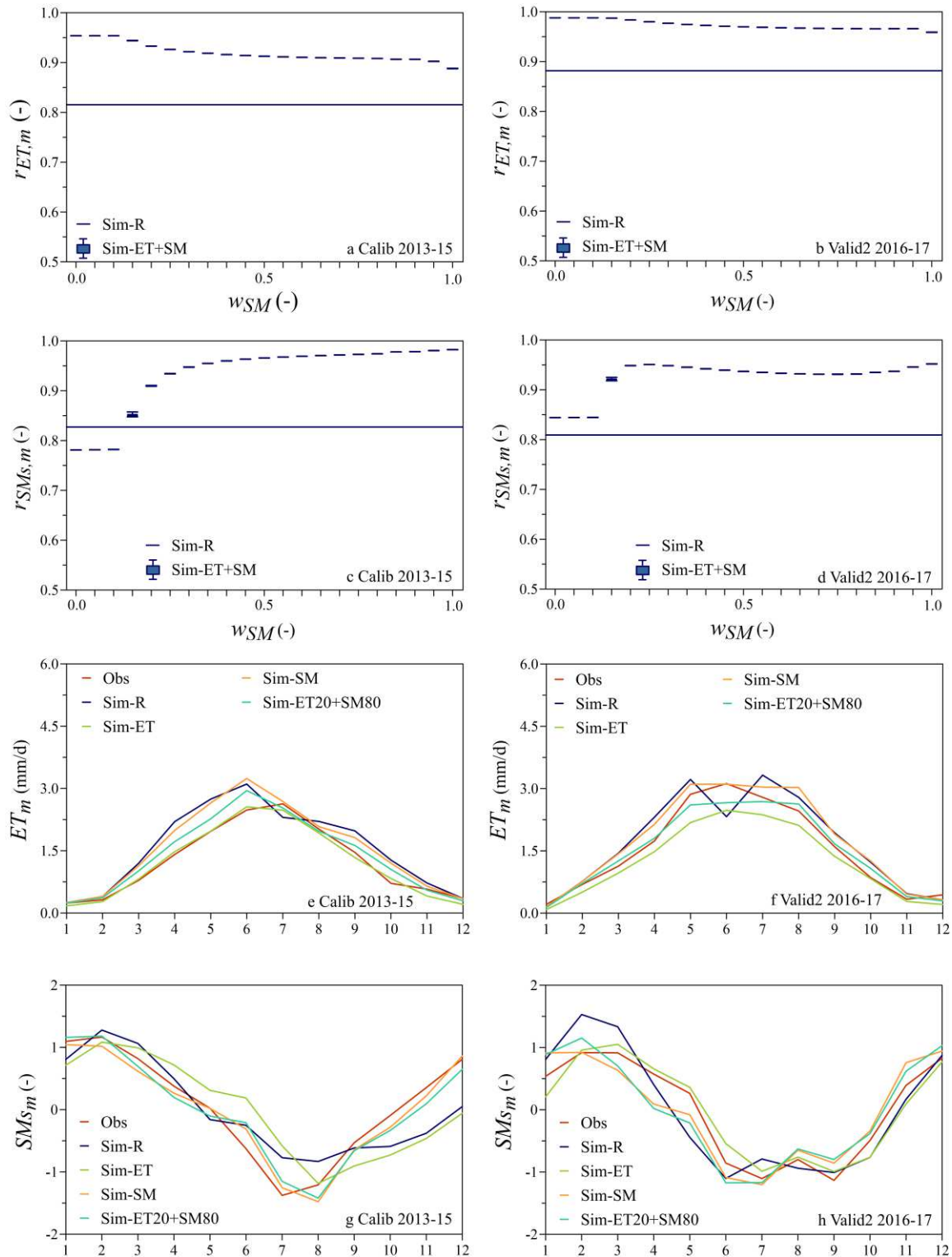


Figure 27. Performance of model simulations in terms of monthly Pearson correlation coefficients for evapotranspiration (ET) and standardized soil moisture (SMs) as a function of weight on soil moisture objective w_{SM} (panels a-d) and monthly averages for ET and SMs for different scenarios according to Table 15.

4.7.3. Runoff generation module simulations

The performance of the runoff generation module simulations in terms of the overland flow Z_{OF} and storage change Z_{dss} objectives as a function of weight on the overland flow objective are shown in Figure 28. The median of the relative number of days with good overland flow simulations immediately exceeded 0.6 as soon as the weight on the overland flow part in the compound objective function was larger than zero (Figure 28 a). The standardized groundwater storage change objective gradually deteriorated as the weight on the overland flow objective increased. Generally, the results outperformed the Sim-R scenario, except for standardized monthly average storage change during validation (Figure 28 d).

For further analysis, the subsurface module parameters of the three main scenarios (Sim-ET+G, Sim-SM+G and Sim-ET20+SM80+G) were calibrated by choosing the weight on the overland flow objective as $w_{OF}=0.5$.

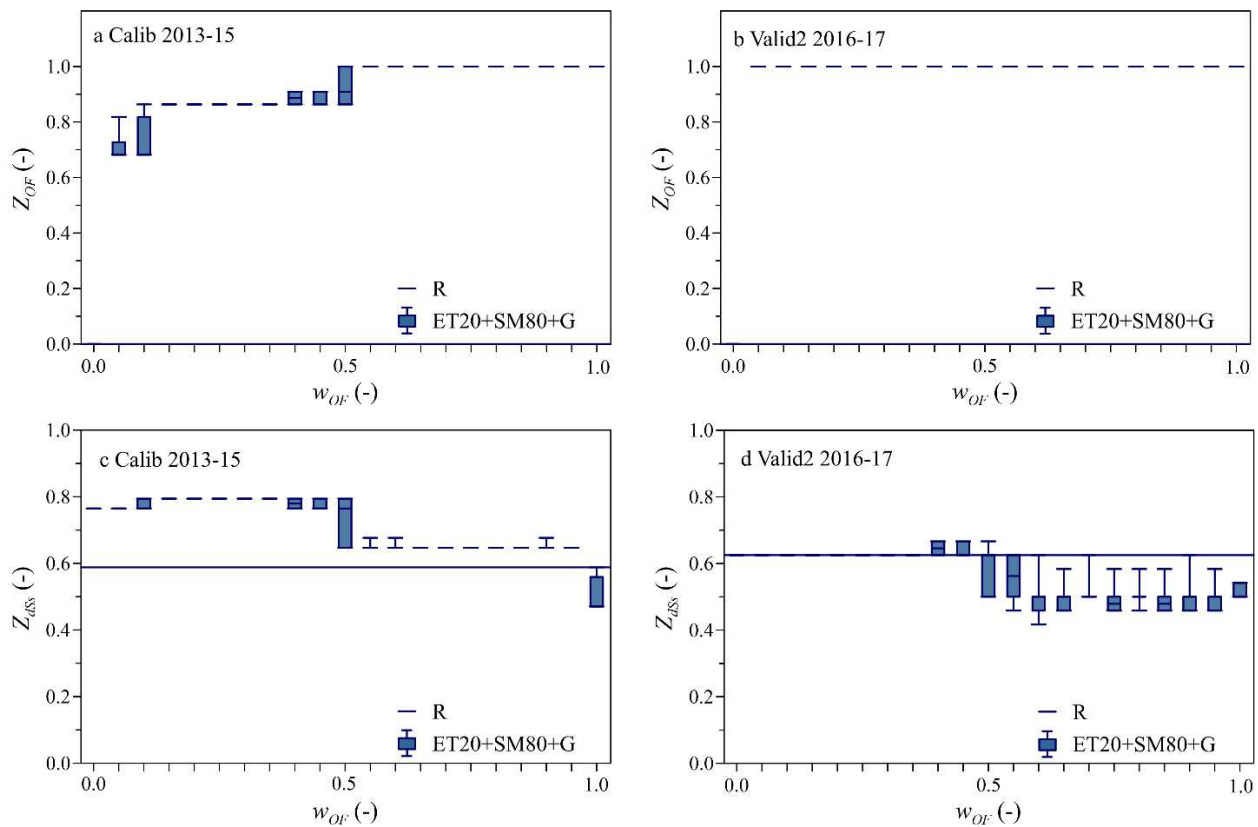


Figure 28. Performance of model simulations in terms of the relative number of days with good overland flow simulations Z_{OF} (when the model simulated very fast runoff simultaneously with the observed overland flow events) as a function of weight on overland flow part w_{OF} (panels a and b) and the relative number of months with correctly simulated sign of the standardized groundwater storage change as a function of weight on overland flow part w_{OF} (panels c and d) during model calibration (panels a and c) and validation (panels b and d). Scenarios are described in Table 15.

4.7.4. Runoff simulations

The final step was to evaluate the different scenarios in terms of simulating runoff on annual and seasonal time scales (Figure 29 and Table 17). On the annual time scale, the performance of the proposed step-by-step approach was similar to the performance of the traditional Sim-R scenario. The proposed method (scenario Sim-SM) even outperformed the Sim-R scenario during the shorter validation period (Val2, 2016-17). For all periods, the model performance was the best, when only soil moisture was used in the optimization of the soil moisture accounting module and the runoff generation module was not optimized (Sim-SM).

Similarly, on the seasonal time scale, the proposed method could efficiently model runoff. In terms of the monthly correlation coefficients (Table 17), the Sim-SM scenario performed the best during the validation periods by slightly outperforming the Sim-R scenario.

Stepwise prediction of runoff using proxy data in a small agricultural catchment

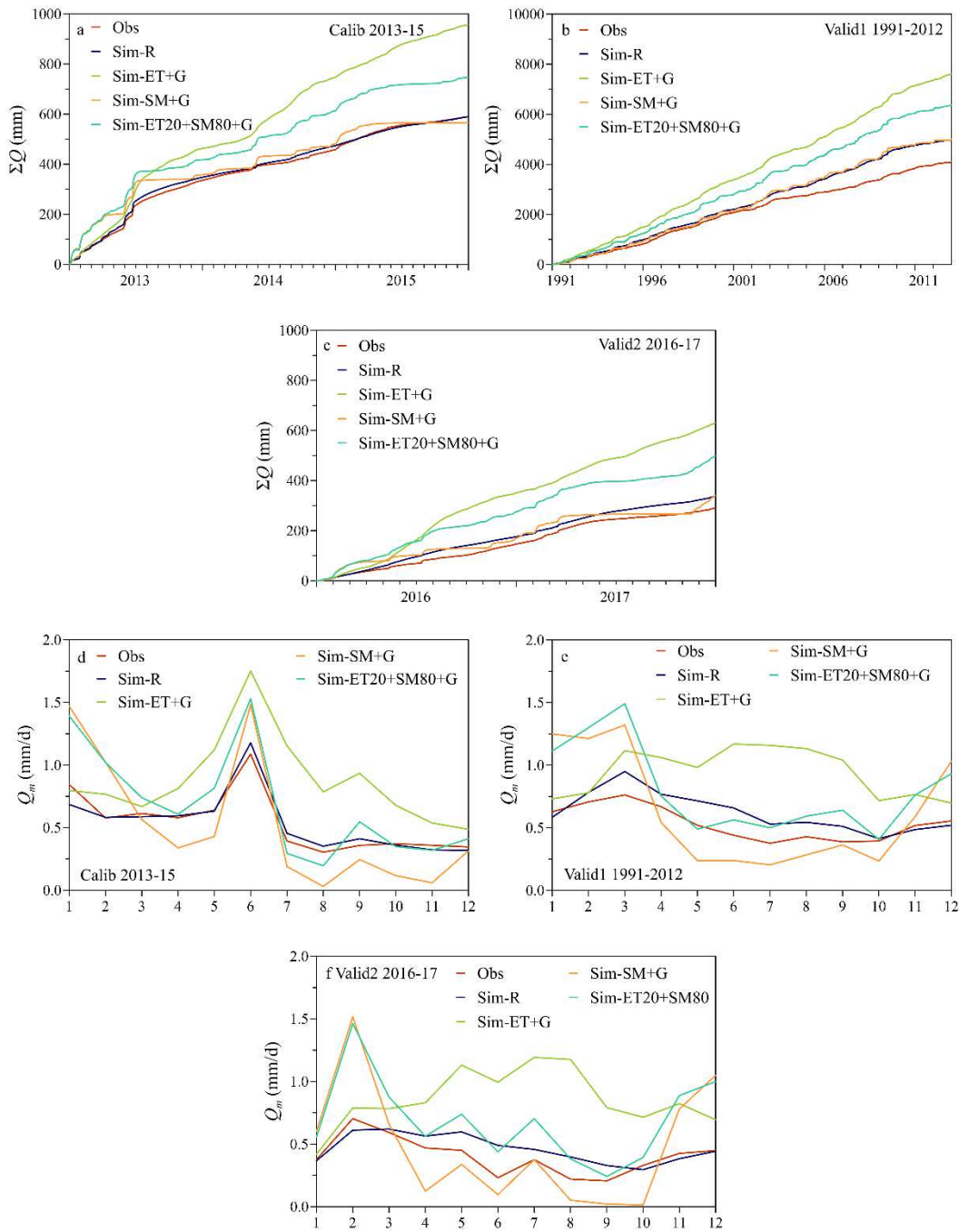


Figure 29. Performance of model simulations in terms of cumulative runoff (panels a-c) and monthly average runoff Q_m (panels d-f). Scenarios are shown in Table 15.

Table 17. Performance of runoff simulations in terms of relative volume error VE and monthly Pearson correlation coefficient r_m for runoff (as in Figure 29) for model calibration and validation periods. Best performing scenario for the proposed approach is shown in bold.

Scenario	Calibration period 2013-15		Validation period 1 1991-2012		Validation period 2 2016-17	
	VE (-)	r_m (-)	VE (-)	r_m (-)	VE (-)	r_m (-)
Sim-R	0.00	0.98	0.18	0.75	0.15	0.75
Sim-Snowacc	-0.97	0.00	-0.71	0.28	-0.99	-0.03
Sim-Snowmelt	-0.23	0.78	-0.03	0.62	-0.31	0.43
Sim-ET	0.65	0.81	0.79	0.53	1.09	-0.02
Sim-SM	0.00	0.93	0.19	0.76	0.00	0.79
Sim-ET20+SM80	0.31	0.96	0.50	0.73	0.50	0.68
Sim-ET+G	0.62	0.82	0.80	0.58	1.15	0.16
Sim-SM+G	-0.04	0.88	0.19	0.71	0.17	0.64
Sim-ET20+SM80+G	0.27	0.91	0.52	0.74	0.70	0.69

4.8. Discussion

In this study, we followed the stepwise model calibration approach proposed by Széles et al. (2020) but without using runoff observations. The aim was to test if their method could be efficiently used in a quasi ungauged catchment case, i.e. ungauged in terms of runoff. We aimed to investigate how accurately we could simulate runoff and other state variables, whether there might be a tradeoff between these. Our results showed that additional measurements can help to efficiently predict runoff on the annual and seasonal time scales. This finding suggests that there is room for expanding the usual focus on runoff predictions (Blöschl et al., 2013; Hrachowitz et al., 2013) to other components of the hydrological cycle. For instance, if snow, soil moisture and actual evapotranspiration observations are available, these can significantly help to constrain hydrological models and to improve the process consistency, if no runoff measurements exist. This

is useful, because this implies that the model simulates runoff for the right reasons and the model indeed represents reality (Beven & Freer, 2001; Rogger et al., 2012; Savenije, 2001; Viglione et al., 2018). Although a limitation of this method might be that such field observations and well-equipped experimental catchments are rare (Blöschl et al., 2016).

In our study we found that in iterative model calibration, soil moisture measurements were the most important to obtain runoff efficiency comparable to runoff efficiency achieved with a model calibrated to runoff only. This result might be expected considering that the most sensitive parameter of the model was the field capacity in this catchment (Széles et al., 2010), which was influenced by the soil moisture and actual evapotranspiration measurements during model calibration. The soil moisture dynamics in this catchment followed better the runoff dynamics compared to actual evapotranspiration, therefore soil moisture proved to be a better proxy to predict runoff. At the annual and seasonal time scales runoff model performance was similar in the 22-year-long validation period and even slightly larger in the 2-year-long validation period. These results are consistent with the results of Thyer et al. (2004) and Kuras et al. (2011) who reported similarly good results on the annual (relative volume error below 13% and 23%, respectively) and on the daily time scales for a physically-based eco-hydrological model.

Results from other studies in terms of whether soil moisture and/or evapotranspiration measurements improve runoff simulations are less conclusive. Similarly to our results, Nijzink et al. (2018) also found that soil moisture satellite products were more effective than evaporation products for deriving more constrained parameter distributions. On the other hand, López et al. (2017) showed that estimating runoff was more efficient, if both soil moisture and evapotranspiration satellite products were involved in the model calibration. Bergström & Lindström (2015) and Baroni et al. (2019) pointed out that the relative importance of the measurements is influenced by the time step of the model considering that more water is stored in the unsaturated zone compared to the evapotranspired volumes. Furthermore, the changes in the different processes depend on how much they are decoupled from the atmosphere. For example, storage in the unsaturated and saturated zones changes more slowly than evapotranspiration which is coupled to the atmosphere.

In our study actual evapotranspiration, overland flow and groundwater level measurements did not help much to constrain the conceptual hydrological model. Possible explanation why this was the case is an apparent mismatch between field observations and the HBV type, soil moisture dependent evapotranspiration calculations. For example, during precipitation events measured evapotranspiration drops to zero, while model simulations increase due to the higher moisture content in the soil. Another issue is the possible overestimation of actual evapotranspiration considering that the Nash Sutcliffe efficiency was used during calibration. This means that actual evapotranspiration was fitted to the higher values and not the lower ones. The third possibility is the difficulty in upscaling evapotranspiration to the catchment scale using point measurements. The aim of the three different eddy covariance stations was to capture the difference between crop types, which could be used for an area-based upscaling. Still, what the model sees as a “catchment average” evapotranspiration rate might be different from the upscaled values. Finally, a fourth explanation might be that water for evapotranspiration especially in the summer months might be extracted from deeper soil layers, and not the layers which are monitored by the soil moisture sensors. Using overland flow and groundwater observations generally did not improve runoff simulations. Although, simulation of runoff generation processes improved, the runoff simulation efficiencies deteriorated. Several studies also found that runoff can no longer be simulated that

efficiently, if additional data (e.g. snow, evapotranspiration, groundwater levels) were also used for model calibration besides runoff (e.g. Gui et al., 2019; Parajka et al., 2007; Seibert, 2000). This is the cost of improving model consistency.

4.9. Conclusions

In this study, we calibrated the parameters of a conceptual hydrological model step-by-step using all the available field observations except runoff. We investigated the value of proxy data for predicting runoff in a small agricultural catchment. Our results showed that:

- Using only snow and soil moisture information for calibration, the runoff model performance was comparable to the scenario when the model was calibrated in one step, using only runoff measurements.
- By using proxy data for model calibration, the simulation of state variables and therefore the process consistency improved, implying that the model represents reality better than the scenario when only runoff was used for model calibration.

4.10. Acknowledgements

The authors would like to acknowledge financial support provided by the Austrian Science Funds (FWF) as part of the Vienna Doctoral Program on Water Resource Systems (DK W1219-N28).

5. Summary and conclusions

Exploring and modelling hydrological processes in small, experimental catchments provide a way to better understand the overall catchment behavior, which is useful not only for the wider scientific community, but also for practitioners working in water resources management, risk and operational forecasting.

The aim of this thesis was to explore the links between the observed hydrological process patterns in a small experimental catchment, the Hydrological Open Air Laboratory (HOAL) with the help of comparative data analyses and hydrological modelling. The main aim was to conceptually model the catchment behavior during different hydrological conditions, such as short dry periods and longer periods with a mixture of dry and wet periods. To achieve this goal, a variety of hydrological and meteorological field observations were linked with model simulations.

The aim of Chapter 2 was to understand the interaction between streamflow, the riparian zone, and the crop fields during rainless periods. Due to catchment heterogeneities (e.g. land use, soil types, aspect, etc.), capturing the spatial differences of evapotranspiration within a catchment and giving an estimate of the catchment average evapotranspiration via upscaling the point measurements is challenging. Previous studies showed that there are no robust methods for measuring the evapotranspiration rates in riparian forests with mixed vegetation types. This is mainly due to technical challenges of using eddy covariance systems in narrow riparian corridors, upscaling uncertainties of point measurements, and heterogeneity of species. Therefore, the idea of this study was to link the diel streamflow fluctuations, i.e. the daily variation of streamflow rates during low flow periods induced by evapotranspiration, with radiation and eddy covariance measurements of evapotranspiration. The spatio-temporal variability of the streamflow fluctuations observed at twelve locations in the Hydrological Open Air Laboratory was explained by differences in the vegetation cover and runoff generation mechanisms. Wetlands, which are covered mainly by trees and are highly saturated due to the shallow groundwater table, featured the largest diel streamflow fluctuations relative to the streamflow rates, while these rates were smaller for tile drains and springs. Most of the volume associated with diel streamflow fluctuations at the catchment outlet was explained by transpiration by the riparian forest along the main stream. A solar radiation driven model was proposed to estimate the evapotranspiration rates and the time lags between the radiative forcing and diel streamflow fluctuations. The time lags showed a strong seasonality, increasing from spring to summer and decreasing from summer to autumn. These findings suggest that a separation of scales in transpiration effects on low flows exists both in time and space. This means that the diel streamflow fluctuations are induced by transpiration from the riparian vegetation, while most of the catchment evapotranspiration, such as evapotranspiration from the crop fields further away from the stream, does not influence the diel signal in streamflow.

Chapter 3 investigated the added value of different field observations besides runoff for calibrating a conceptual rainfall runoff model for a small catchment. The field observations were linked with model simulations to reproduce the runoff rates at the main outlet of the catchment. An HBV type, spatially lumped hydrological model was calibrated using a new step-by-step parametrization approach, where the three modules of the model (snow, soil moisture accounting and runoff generation modules) were calibrated using different data types besides runoff. By using the new approach, the overall process consistency improved, compared to model simulations when only runoff was used for model calibration. Soil moisture and evapotranspiration observations had the

Summary and conclusions

largest influence on simulated runoff, while the parameterization of the snow and runoff generation modules had a smaller influence.

Chapter 4 explored the value of proxy data for calibrating a conceptual hydrological model for small ungauged basins, i.e. ungauged in terms of runoff. In this study, it was explored how accurately the hydrological variables can be modelled by using all of the available field observations in the catchment, except runoff for calibrating the same runoff model. Using snow and soil moisture information for model calibration, the runoff model performance was found to be comparable with the scenario when the model was calibrated using only runoff data. Similarly to Chapter 3, by using this stepwise approach and proxy data, the simulation of state variables improved compared to model calibration in one step using only runoff data.

This thesis has advanced the understanding of the dynamic catchment behavior and the interconnection between different storages of water within the catchment. This work has developed a new technique for estimating the evapotranspiration rates of the riparian zone during rainless periods and a better understanding on which part of the catchment contributes to evapotranspiration during dry periods and which does not. In this thesis a new step-by-step model parametrization framework has been developed, which helped to link field measurements with model simulations. With the new technique, a lumped, conceptual hydrological model can be parameterized using field observations including or excluding runoff. In both cases, the process consistency has improved. This is useful, because process consistency implies that the model performs well for the right reasons, and represents reality accurately – which is the general goal of each modelling study.

For each research chapter, several open questions remain and can be explored in future work. For Chapter 2, based on the available measurements, the distinction between flow path connectedness (propagation of particles, i.e. velocity) and signal contribution (propagation of the signal, i.e. celerity) is not yet well understood. Further measurements, for instance installation of sap flow sensors, will help to improve the understanding of the role of flow path connectivity on streamflow response. In the future, discharge gauges with narrower cross sections could be installed which are more sensitive in low flow periods. Regarding Chapter 3 and 4, little is known on whether the presented spatial and temporal aggregation of the field observations is adequate, and whether these measurements represent the average catchment behavior well. Furthermore, it will be interesting to explore whether the current HBV-type model structure does represent the catchment well or whether the structure of the model should be changed to better adapt to small scale variability in an agricultural landscape. Another question is whether and to what extent can a spatially distributed model structure improve the runoff model performance.

Hydrological models are a useful tool among others for operational flood forecasting, design flood estimation, agricultural and urban water management, climate change prediction, and runoff prediction in ungauged basins. The solar radiation driven model presented in this thesis in Chapter 2 can be used for evapotranspiration estimation in headwater catchments, which can be useful to improve hydrological model simulations considering that a shortcoming of several hydrological modelling studies is the inappropriate representation of evapotranspiration. A better and more accurate modelling of evapotranspiration can be useful for agricultural purposes, e.g. for irrigation planning. Chapter 3 introduced an approach to parameterizing a hydrological model using additional measurements besides runoff. The stepwise approach can be used in other modelling studies to simulate not only runoff but other state variables more accurately. This is an important

Summary and conclusions

step considering that hydrological models often give good runoff estimations with wrong parameter combinations that is inaccurately simulated state variables. Finally, Chapter 4 presented a way to predict runoff for ungauged catchments using only proxy data, out of which soil moisture proved to be the most useful. This finding can be used in other ungauged small agricultural catchments for runoff prediction.

References

References

Allen, R.G., Pereira, L.S., Raes, D., & Smith, M. (1998). Crop evapotranspiration – Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56. FAO - Food and Agriculture Organization of the United Nations, Rome. ISBN: 92-5-104219-5

Ardia, D., Ospina Arango, J.D., & Giraldo Gomez, N.D. (2010a). Jump-diffusion calibration using differential evolution. *Wilmott Magazine*, 55, 76-79.

Ardia, D., Boudt, K., Carl, P., Mullen, K.M., & Peterson, B.G. (2010b). Differential evolution with 'DEoptim': An application to non-convex portfolio optimization. *The R Journal*, 3(1), 27-34.

Ardia, D., Mullen, K.M., Peterson, B.G., & Ulrich, J. (2016). 'DEoptim': Differential evolution in 'R'. version 2.2-4.

Arheimer, B., Pimentel, R., Isberg, K., Crochemore, L., Andersson, J.C.M., Hasan, A., et al. (2020). Global catchment modelling using World-Wide HYPE (WWH), open data, and stepwise parameter estimation. *Hydrology and Earth System Sciences*, 24, 535-559. doi: 10.5194/hess-24-535-2020

Avanzi, F., Maurer, T., Glaser, S.D., Bales, R.C., & Conklin, M.H. (2020). Information content of spatially distributed ground-based measurements for hydrologic-parameter calibration in mixed rain-snow mountain headwaters. *Journal of Hydrology*, 582, 124478. doi: 10.1016/j.jhydrol.2019.124478

Barbeta, A., & Peñuelas, J. (2017). Relative contribution of groundwater to plant transpiration estimated with stable isotopes. *Scientific Reports*, 7(1), 10580. doi: 10.1038/s41598-017-09643-x

Barnard, H.R., Graham, C.B., Van Verseveld, W.J., Brooks, J.R., Bond, B.J., & McDonnell, J.J. (2010). Mechanistic assessment of hillslope transpiration controls of diel subsurface flow: A steady-state irrigation approach. *Ecology*, 91, 114-124. doi: 10.1002/eco.114

Baroni, G., Schalge, B., Rakovec, O., Kumar, R., Schüler, L., Samaniego, L., et al. (2019). A comprehensive distributed hydrological modeling intercomparison to support process representation and data collection strategies. *Water Resources Research*, 55, 990-1010. doi: 10.1029/2018WR023941

Beeson, R. (2011). Evapotranspiration of woody landscape plants. In G. Gerosa (Ed.): *Evapotranspiration - From Measurements to Agricultural and Environmental Applications* (pp.347-370, p 410). London: Intech Open. doi: 10.5772/18762

Bergström, S. (1976). Development and application of a conceptual runoff model for Scandinavian catchments. Department of Water Resources Engineering, Lund Institute of Technology/University of Lund, Bulletin Series A, no. 52.

Bergström, S. (1995). The HBV model. In V. P. Singh (Ed.): *Computer Models of Watershed Hydrology* (pp.443-476). Highlands Ranch, Colorado: Water Resour. Publ.

Bergström, S., & Lindström, G. (2015). Interpretation of runoff processes in hydrological modelling – experience from the HBV approach. *Hydrological Processes*, 29, 3535-3545. doi: 10.1002/hyp.10510

Beven, K.J. (2001a). How far can we go in distributed hydrological modelling? *Hydrological Earth System Sciences*, 5(1), 1-12. doi: 10.5194/hess-5-1-2001

References

- Beven, K.J. (2001b). *Rainfall-Runoff Modelling. The Primer*. Chichester: Wiley. ISBN: 978-0-470-71459-1
- Beven, K.J. (2007). Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process. *Hydrological Earth System Sciences*, 11(1), 460-467. doi:10.5194/hess-11-460-2007
- Beven, K.J., & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249(1-4), 11-29. doi: 10.1016/S0022-1694(01)00421-8
- Beven, K.J., & Kirkby, M.J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences*, 24, 43-69. doi: 10.1080/02626667909491834
- Beven, K.J., Kirkby, M.J., Schofield, N., & Tagg, A.F. (1984). Testing a physically-based flood forecasting model (TOPMODEL) for three U.K. catchments. *Journal of Hydrology*, 69(1-4), 119-143. doi: 10.1016/0022-1694(84)90159-8
- Blöschl, G. (2006). Hydrologic synthesis - across processes, places and scales, Special section on the vision of the CUAHSI National Center for Hydrologic Synthesis (NCHS). *Water Resources Research*, 42, W03S02. doi: 10.1029/2005WR004319
- Blöschl, G., Blaschke, A.P., Broer, M., Bucher, C., Carr, G., Chen, X., et al. (2016). The Hydrological Open Air Laboratory (HOAL) in Petzenkirchen: a hypothesis-driven observatory. *Hydrology and Earth System Sciences*, 20(1), 227-255. doi: 10.5194/hess-20-227-2016
- Blöschl, G., Kirnbauer, R., & Gutknecht, D. (1991). Distributed snowmelt simulations in an Alpine catchment: 1. Model evaluation on the basis of snow cover patterns. *Water Resources Research*, 27(12), 3171-3179. doi: 10.1029/91WR02250
- Blöschl, G., Reszler, C., & Komma, J. (2008). A spatially distributed flash flood forecasting model. *Environmental Modelling & Software*, 23, 464-478. doi:10.1016/j.envsoft.2007.06.010.
- Blöschl, G., & Sivapalan, M. (1995). Scale issues in hydrological modelling: A review. *Hydrological Processes*, 9(3-4), 251-290. doi: 10.1002/hyp.3360090305
- Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., & Savenije, H. (2013). *Runoff prediction in ungauged basins: Synthesis across processes, places and scales*. Cambridge University Press, United Kingdom, 465 p. ISBN: 9781107028180
- Blöschl, G., & Zehe, E. (2005). On hydrological predictability. *Hydrological Processes*, 19(19), 3923-3929. doi: 10.1002/hyp.6075
- Boese, S., Jung, M., Carvalhais, N., & Reichstein, M. (2017). The importance of radiation for semiempirical water-use efficiency models. *Biogeosciences*, 14(12), 3015-3026. doi: 10.5194/bg-14-3015-2017
- Bond, B.J., Jones, J.A., Moore, G., Phillips, N., Post, D., & McDonnell, J.J. (2002). The zone of vegetation influence on baseflow revealed by diel patterns of streamflow and vegetation water use in a headwater basin. *Hydrological Processes*, 16(8), 1671-1677. doi: 10.1002/hyp.5022
- Bren, L. J. (1997). Effects of slope vegetation removal on the diurnal variations of a small mountain stream. *Water Resources Research*, 33(2), 321-331. doi: 10.1029/96WR02648

References

- Butler, J.J.Jr., Kluitenberg, G.J., Whittemore, D.O., Loheide, S.P.II., Jin, W., Billinger, M.A., & Zhan, X. (2007). Effects of slope vegetation removal on the diurnal variations of a small mountain stream. *Water Resources Research*, 43, W02404. doi: 10.1029/2005WR004627
- Cadol, D., Kampf, S., & Wohl, E. (2012). Effects of evapotranspiration on baseflow in a tropical headwater catchment. *Journal of Hydrology*, 462-463, 4-14. doi: 10.1016/j.jhydrol.2012.04.060
- Caldwell, M.M., Dawson, T.E., & Richards, J.H. (1998). Hydraulic lift: consequences of water efflux from the roots of plants. *Oecologia*, 113(2), 151-161. doi: 10.1007/s004420050363
- Cermák, J.N., Kučera, J., & Nadezhkina, N. (2004). Sap flow measurements with some thermodynamic methods, flow integration within trees and scaling up from sample trees to entire forest stands. *Trees*, 18(5), 529-546. doi: 10.1007/s00468-004-0339-6
- Cernohous, V., & Šach, F. (2008). Daily baseflow variations and forest evapotranspiration. *Ekológia (Bratislava)*, 27(2), 189-195.
- Chu, W., Gao, X., & Sorooshian, S. (2011). A new evolutionary search strategy for global optimization of high-dimensional problems. *Information Sciences*, 181(22), 4909-4927. doi:10.1016/j.ins.2011.06.024.
- Corbari, C., Mancini, M., Li, J., & Su, Z. (2015). Can satellite land surface temperature data be used similarly to river discharge measurements for distributed hydrological model calibration? *Hydrological Sciences Journal*, 60(2), 202-217. doi: 10.1080/02626667.2013.866709
- Criss, R.E., & Winston, W.E. (2008). Do Nash values have value? Discussion and alternate proposals. *Hydrological Processes*, 22, 2723-2725. doi:10.1002/hyp.7072
- Crow, P. (2005). The influence of soils and species on tree root depth. Forestry Commission, Information Note, Edinburgh, United Kingdom.
- Dawson, T.E. (1996). Determining water use by trees and forests from isotopic, energy balance and transpiration analyses: The role of tree size and hydraulic lift. *Tree Physiology*, 16(1-2), 263-272. doi: 10.1093/treephys/16.1-2.263
- Delzon, S., & Loustau, D. (2005). Age-related decline in stand water use: Sap flow and transpiration in a pine forest chronosequence. *Agricultural and Forest Meteorology*, 129(3-4), 105-119. doi: 10.1016/j.agrformet.2005.01.002
- Demirel, M.C., Özen, A., Orta, S., Toker, E., Demir, H.K., Ekmekcioglu, Ö., et al. (2019). Additional value of using satellite-based soil moisture and two sources of groundwater data for hydrological model calibration. *Water*, 11(10), 2083. doi: 10.3390/w11102083
- Deutscher, J., Kupec, P., Dundek, P., Holik, L., Machala, M., & Urban, J. (2016). Diurnal dynamics of streamflow in an upland forested micro-watershed during short precipitation-free periods is altered by tree sap flow. *Hydrological Processes*, 30(13), 2042-2049. doi: 10.1002/hyp.10771
- DHI (Danish Hydraulic Institute) (1998). MIKE-SHE Water Movement User's Guide and Technical Reference Manual. Denmark: Danish Hydraulic Institute.
- Dozier, J. (1989). Spectral signature of alpine snow cover from the landsat thematic mapper. *Remote Sensing of Environment*, 28, 9-22. doi: 10.1016/0034-4257(89)90101-6

References

- Dragoni, D., Lakso, A.N., & Piccioni, R.M. (2005). Transpiration of apple trees in a humid climate using heat pulse sap flow gauges calibrated with whole-canopy gas exchange chambers. *Agricultural and Forest Meteorology*, 130(1-2), 85-94. doi: 10.1016/j.agrformet.2005.02.003
- Drexler, J.Z., Snyder, R.L., Spano, D., & Paw, U.K.T. (2004). A review of models and micrometeorological methods used to estimate wetland evapotranspiration. *Hydrological Processes*, 18(11), 2071-2101. doi: 10.1002/hyp.1462
- Duffie, J.A., & Beckman, W.A. (2013). *Solar Engineering of Thermal Processes*. Hoboken, NJ: John Wiley. doi: 10.1002/9781118671603
- Dunford, E.G., & Fletcher, P.W. (1947). Effect of removal of stream-bank vegetation upon water yield. *Transactions of the American Geophysical Union*, 28(1), 105. doi: 10.1029/TR028i001p00105
- Dvorakova, S., Kovar, P., & Zeman, J. (2012). Implementation of conceptual linear storage model of runoff with diurnal fluctuations in rainless periods. *Journal of Hydrology and Hydromechanics*, 60(4), 217-226. doi: 10.2478/v10098-012-0019-y
- Dvorakova, S., Kovar, P., & Zeman, J. (2014). Impact of evapotranspiration on discharge in small catchments. *Journal of Hydrology and Hydromechanics*, 62(4), 285-292. doi: 10.2478/johh-2014-0039
- Eder, A., Exner-Kittridge, M., Strauss, P., & Blöschl, G. (2014). Re-suspension of bed sediment in a small stream – Results from two flushing experiments. *Hydrology and Earth System Sciences*, 18(3), 1043-1052. doi: 10.5194/hess-18-1043-2014
- Eder, A., Strauss, P., Kreuger, T., & Quinton, J. (2010). Comparative calculation of suspended sediment loads with respect to hysteresis effects. *Journal of Hydrology*, 389(1-2), 168-176. doi: 10.1016/j.jhydrol.2010.05.043
- European Space Agency – ESA (2020). Sentinel 2 Level-2A products. Accessed: 16 May 2020.
- Exner-Kittridge, M., Strauss, P., Blöschl, G., Eder, A., Saracevic, E., & Zessner, M. (2016). The seasonal dynamics of the stream sources and input flow paths of water and nitrogen of an Austrian headwater agricultural catchment. *Science of the Total Environment*, 542(Pt A), 935-945. doi: 10.1016/j.scitotenv.2015.10.151
- FAO, ISRIC, & ISSS (1998). *World Reference Base for Soil Resources*. 84 World Soil Resources Reports, Rome.
- Farid, A., Goodrich, D.C., Bryant, R., & Sorooshian, S. (2008). Using airborne lidar to predict Leaf Area Index in cottonwood trees and refine riparian water-use estimates. *Journal of Arid Environments*, 72(1), 1-15. doi: 10.1016/j.jaridenv.2007.04.010
- Federer, C.A. (1973). Forest transpiration greatly speeds streamflow recession. *Water Resources Research*, 9(6), 1599–1604. doi: 10.1029/WR009i006p01599
- Fenicia, F., Savenije, H.H.G., Matgen, P., & Pfister, L. (2007). A comparison of alternative multiobjective calibration strategies for hydrological modeling. *Water Resources Research*, 43, W03434, doi:10.1029/2006WR005098

References

- Fonley, M., Mantilla, R., Small, S.J., & Curtu, R. (2016). On the propagation of diel signals in river networks using analytic solutions of flow equations. *Hydrology and Earth System Sciences*, 20(7), 2899-2912. doi:10.5194/hess-20-2899-2016
- Fukushima, Y. (1988). A model of river flow forecasting for a small forested mountain catchment. *Hydrological Processes*, 2, 167-185. doi: 10.1002/hyp.3360020207
- Gartner, K., Nadezhdina, N., Englisch, M., Cermak, J., & Leitgeb, E. (2009). Sap flow of birch and Norway spruce during the European heat and drought in summer 2003. *Forest Ecology and Management*, 258(5), 590-599. doi: 10.1016/j.foreco.2009.04.028
- Gelleszun, M., Kreye, P., & Meon, G. (2017). Representative parameter estimation for hydrological models using a lexicographic calibration strategy. *Journal of Hydrology*, 553, 722-734. doi: 10.1016/j.jhydrol.2017.08.015
- Glenn, E.P., Neale, C.M.U., Hunsaker, D.J., & Nagler, P.L. (2011). Vegetation index-based crop coefficients to estimate evapotranspiration by remote sensing in agricultural and natural ecosystems. *Hydrological Processes*, 25(26), 4050-4062. doi: 10.1002/hyp.8392
- Goodrich, D.C., Scott, R., Qi, J., Goff, B., Unkrich, C.L., Moran, M.S., et al. (2000). Seasonal estimates of riparian evapotranspiration using remote and in situ measurements. *Agricultural and Forest Meteorology*, 105(1-3), 281-309. doi: 10.1016/S0168-1923(00)00197-0
- Götzinger, J., & Bardossy, A. (2007). Comparison of four regionalisation methods for a distributed hydrological model. *Journal of Hydrology*, 333(2-4), 374-384. doi: 10.1016/j.jhydrol.2006.09.008.
- Graham, C.B., Barnard, H.R., Kavanagh, K.L., & McNamara, J.P. (2013). Catchment scale controls the temporal connection of transpiration and diel fluctuations in streamflow. *Hydrological Processes*, 27(18), 2541-2556. doi: 10.1002/hyp.9334
- Granier, A., Biron, P., & Lemoine, D. (2000). Water balance, transpiration and canopy conductance in two beech stands. *Agricultural and Forest Meteorology*, 100(4), 291-308. doi: 10.1016/S0168-1923(99)00151-3
- Grayson, R.B., Moore, I.D., & McMahon, T.A. (1992). Physically based hydrologic modeling: 2. Is the concept realistic? *Water Resources Research*, 28(10), 2659-2666. doi: 10.1029/92WR01259
- Grayson, R.B., Western, A.W., Chiew, F.H.S., & Blöschl, G. (1997). Preferred states in spatial soil moisture patterns: Local and non-local controls. *Water Resources Research*, 33(12), 2897-2908. doi: 10.1029/97WR02174
- Gribovszki, Z. (2014). Diurnal method for evapotranspiration estimation from soil moisture profile. *Acta Silvatica et Lignaria Hungarica*, 10(1), 67-75. doi: 10.2478/aslh-2014-0005
- Gribovszki, Z. (2018). Validation of diurnal soil moisture dynamic-based evapotranspiration estimation methods. *Időjárás*, 122(1), 15-30. doi: 10.28974/idojaras.2018.1.2
- Gribovszki, Z., Kalicz, P., Kucsara, M. (2006). Streamflow characteristics of two forested catchments in the Sopron Hills. *Acta Silvatica et Lignaria Hungarica*, 2, 81-92.
- Gribovszki, Z., Kalicz, P., & Szilágyi, J. (2013). Does the accuracy of fine-scale water level measurements by vented pressure transducers permit for diurnal evapotranspiration estimation? *Journal of Hydrology*, 488, 166-169. doi:10.1016/j.jhydrol.2013.03.001

References

- Gribovszki, Z., Kalicz, P., Szilágyi, J., & Kucsara, M. (2008). Riparian zone evapotranspiration estimation from diurnal groundwater level fluctuations. *Journal of Hydrology*, 349(1-2), 6-17. doi: 10.1016/j.jhydrol.2007.10.049
- Gribovszki, Z., Kalicz, P., Szilágyi, J. (2010a). Talajvíz evapotranszspiráció számítása a vízhozamok napi periódusú ingadozása alapján. *Hidrológiai Közlöny*, 90(5), 19-28.
- Gribovszki, Z., Kalicz, P., Szilágyi, J. (2011). Numerical validation of a diurnal streamflow-pattern-based evapotranspiration estimation method. *Acta Silvatica et Lignaria Hungarica*, 7, 63-74.
- Gribovszki, Z., Szilágyi, J., & Kalicz, P. (2010b). Diurnal fluctuations in shallow groundwater levels and streamflow rates and their interpretation - A review. *Journal of Hydrology*, 385(1-4), 371-383. doi:10.1016/j.jhydrol.2010.02.001
- Gui, Z., Liu, P., Cheng, L., Guo, S., Wang, H., & Zhang, L. (2019). Improving runoff prediction using remotely sensed actual evapotranspiration during rainless periods. *Journal of Hydrologic Engineering*, 24(12), 04019050. doi: 10.1061/(ASCE)HE.1943-5584.0001856
- Hall, R.L., Allen, S.J., Rosier, P.T.W., & Hopkins, R. (1998). Transpiration from coppiced poplar and willow measured using sap-flow methods. *Agricultural and Forest Meteorology*, 90(4), 275-290. doi: 10.1016/S0168-1923(98)00059-8
- Hall, D.K., & Riggs, G.A. (2016a). MODIS/Terra Snow Cover Daily L3 Global 500m Grid, Version 6. [January 2013 – December 2017]. Boulder, Colorado USA, NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: 10.5067/MODIS/MOD10A1.006. [24 January 2018].
- Hall, D.K., & Riggs, G.A. (2016b). MODIS/Aqua Snow Cover Daily L3 Global 500m Grid, Version 6. [January 2013 – December 2017]. Boulder, Colorado USA, NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: 10.5067/MODIS/MYD10A1.006. [24 January 2018].
- Hay, L.E., Leavesley, G.H., Clark, M.P., Markstrom, S.L., Viger, R.J., & Umemoto, M. (2006). Step wise, multiple objective calibration of a hydrologic model for a snowmelt dominated basin. *Journal of the American Water Resources Association*, 42(4), 877-890. doi: 10.1111/j.1752-1688.2006.tb04501.x
- Herman, M.R., Nejadhashemi, A.P., Abouali, M., Hernandez-Suarez, J.S., Daneshvar, F., Zhang, Z., et al. (2018). Evaluating the role of evapotranspiration remote sensing data in improving hydrological modeling predictability. *Journal of Hydrology*, 556, 39-49. doi: 10.1016/j.jhydrol.2017.11.009
- Hinckley, T.M., Brooks, J.R., Cermák, J., Ceulmans, R., Kučera, J., Meinzer, F.C., & Roberts, D.A. (1994). Water flux in a hybrid poplar. *Tree Physiology*, 14, 1005-1018.
- Hogue, T.S., Sorooshian, S., Gupta, H., Holz, A., & Braatz, D. (2000). A multistep automatic calibration scheme for river forecasting models. *Journal of Hydrometeorology*, 1(6), 524-542.
- Holländer, H.M., Blume, T., Bormann, H., Buytaert, W., Chirico, G.B., Exbrayat, J.-F., et al. (2009). 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

References

- Hrachowitz, M., Savenije, H.H.G., Blöschl, G., McDonnell, J.J., Sivapalan, M., Pomeroy, J.W., et al. (2013). A decade of Predictions in Ungauged Basins (PUB) - a review. *Hydrological Sciences Journal*, 58(6), 1198-1255. doi: 10.1080/02626667.2013.803183
- Immerzeel, W.W., & Droogers, P. (2008). Calibration of a distributed hydrological model based on satellite evapotranspiration. *Journal of Hydrology*, 349(3-4), 411-424. doi: 10.1016/j.jhydrol.2007.11.017
- Jarvis, P.G., & McNaughton, K.G. (1986). Stomatal control of transpiration: Scaling up from leaf to region. *Advances in Ecological Research*, 15, 1-49. doi: 10.1016/S0065-2504(08)60119-1
- Jennings, K.S., Winchell, T.S., Livneh, B., & Molotch, N.P. (2018). Spatial variation of the rain-snow temperature threshold across the Northern Hemisphere. *Nature Communications*, 9, 1148. doi: 10.1038/s41467-018-03629-7
- Kalicz, P., Gribovszki, Z., & Király, G. (2011). Galériaerdők hatása a vízfolyások recessziós görbéire és ennek információ tartalma. *Erdészettudományi közlemények*, 1(1), 45-57.
- Kirchner, J.W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research*, 45, W02429. doi: 10.1029/2008WR006912
- Knauer, J., El-Madany, T.S., Zaehle, S., & Migliavacca, M. (2018). Bigleaf - An R package for the calculation of physical and physiological ecosystem properties from eddy covariance data. *PLOS ONE*, 13(8), e0201114. doi: 10.1371/journal.pone.0201114
- Köcher, P., Gebauer, T., Horna, V., & Leuschner, C. (2008). Leaf water status and stem xylem flux in relation to soil drought in five temperate broad-leaved tree species with contrasting water use strategies. *Annals of Forest Science*, 66, 101. doi: 10.1051/forest/2008076
- Kovar, P., & Bacinova, H. (2015). Impact of evapotranspiration on diurnal discharge fluctuations determined by the Fourier Series Model in dry periods. *Soil and Water Research*, 10(4), 210-217. doi: 10.17221/122/2015-SWR
- Kumagai, T., Aoki, S., Otsuki, K., & Utsumi, Y. (2009). Impact of stem water storage on diurnal estimates of whole-tree transpiration and canopy conductance from sap flow measurements in Japanese cedar and Japanese cypress trees. *Hydrological Processes*, 23(16), 2335-2344. doi: 10.1002/hyp.7338
- Kume, T., Komatsu, H., Kuraji, K., & Suzuki, M. (2008). Less than 20-min time lags between transpiration and stem sap flow in emergent trees in a Bornean tropical rainforest. *Agricultural and Forest Meteorology*, 148(6-7), 1181-1189. doi: 10.1016/j.agrformet.2008.02.010
- Kundu, D., Vervoort, R.W., & van Ogtrop, F.F. (2017). The value of remotely sensed surface soil moisture for model calibration using SWAT. *Hydrological Processes*, 31(15), 2764-2780. doi: 10.1002/hyp.11219
- Kunnath-Poovakka, A., Ryu, D., Renzullo, L.J. & George, B. (2016). 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
- Kuppel, S., Tetzlaff, D., Maneta, M.P., & Soulsby, C. (2018). What can we learn from multi-data calibration of a process-based ecohydrological model? *Environmental Modelling & Software*, 101, 301-316. doi: 10.1016/j.envsoft.2018.01.001

References

Kuras, P.K., Alila, Y., Weiler, M., Spittlehouse, D., & Winkler, R. (2011). Internal catchment process simulation in a snow-dominated basin: performance evaluation with spatiotemporally variable runoff generation and groundwater dynamics. *Hydrological Processes*, 25, 3187-3203. doi: 10.1002/hyp.8037

Landon, M.K., Rus, D.L., Dietsch, B.J., Johnson, M.R., & Eggemeyer, K.D. (2009). Evapotranspiration rates of riparian forests, Platte River, Nebraska, 2002-06. U.S. Geological Survey Scientific Investigations Report 2008-5228, 65 pp.

Lautz, L.K. (2008). Estimating groundwater evapotranspiration rates using diurnal water-table fluctuations in semi-arid riparian zone. *Hydrogeology Journal*, 16(3), 483-497. doi: 10.1007/s10040-007-0239-0

Lawrence, R.E. (1990). The interaction between the environment, land use, and hydrology of the Bogong High Plains area from 1850 to 1985, (PhD. thesis), 798 pp. AB: University of Melbourne, Parkville, Australia.

Leenhouts, J.M., Stromberg, J.C., & Scott, R.L. (2006). Hydrologic requirements of and consumptive groundwater use by riparian vegetation along the San Pedro River, Arizona. U.S. Geological Survey Scientific Investigations Report 2005-5163, 154 pp.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., & Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, 201(1-4), 272-288. doi: 10.1016/S0022-1694(97)00041-3

Liu, G., Schwartz, F.W., Tseng, K.-H., & Shum, C.K. (2015). Discharge and water-depth estimates for ungauged rivers: Combining hydrologic, hydraulic, and inverse modeling with stage and water-area measurements from satellites. *Water Resources Research*, 51, 6017-6035. doi: 10.1002/2015WR016971

Loheide, S.P.II., Butler, J.J.Jr., & Gorelick, S.M. (2005). Estimation of groundwater consumption by phreatophytes using diurnal water table fluctuations: A saturated-unsaturated flow assessment. *Water Resources Research*, 41, W07030. doi: 10.1029/2005WR003942

López, L.P., Sutanudjaja, E.H., Schellekens, J., Sterk, G., & Bierkens, F.P. (2017). 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

Lu, M., & Li, X. (2015). Strategy to automatically calibrate parameters of a hydrological model: a multi-step optimization scheme and its application to the Xinanjiang model. *Hydrological Research Letters*, 9(4), 69-74. doi: 10.3178/hrl.9.69

Lundquist, J.D., & Cayan, D.R. (2002). Seasonal and spatial patterns in diurnal cycles in streamflow in the western United States. *Journal of Hydrometeorology*, 3(5), 591-603.

Martin, T.A., Brown, K.J., Kucera, J., Meinzer, F.C., Sprugel, D.G., & Hinckley, T.M. (2001). Control of transpiration in a 220-year-old *Abies amabilis* forest. *Forest Ecology and Management*, 152(1-3), 211-224. doi: 10.1016/S0378-1127(00)00604-6

Marttila, H., Dudley, B.D., Graham, S., & Srinivasan, M.S. (2017). Does transpiration from invasive stream side willows dominate low-flow conditions? An investigation using hydrometric and isotopic methods in a headwater catchment. *Ecohydrology*, 11(2), e1930. doi: 10.1002/eco.1930

References

- McKay, M.D., Beckman, R.J., & Conover, W.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2), 239–245. doi: 10.2307/1268522
- Merz, R., & Blöschl, G. (2004). Regionalisation of catchment model parameters. *Journal of Hydrology*, 287, 95–123. doi: 10.1016/j.jhydrol.2003.09.028
- Merz, R., Parajka, J., & Blöschl, G. (2011). Time stability of catchment model parameters: Implications for climate impact analyses. *Water Resources Research*, 47, W02531. doi:10.1029/2010WR009505
- Meyboom, P. (1965). Three observations on streamflow depletion by phreatophytes. *Journal of Hydrology*, 2(3), 248–261. doi: 10.1016/0022-1694(65)90040-5
- Montanari, A., Young, G., Savenije, H.H.G., Hughes, D., Wagener, T., Ren, L.L., et al. (2013). “Panta Rhei - Everything Flows”: Change in hydrology and society - The IAHS Scientific Decade 2013–2022. *Hydrological Sciences Journal*, 58(6), 1256–1275. doi: 10.1080/02626667.2013.809088
- Moore, G.W., Jones, J.A., & Bond, B.J. (2011). How soil moisture mediates the influence of transpiration on streamflow at hourly to interannual scales in a forested catchment. *Hydrological Processes*, 25(24). doi: 10.1002/hyp.8095
- Moradkhani, H., & Sorooshian, S. (2009). General Review of Rainfall-Runoff Modeling: Model Calibration, Data Assimilation, and Uncertainty Analysis. In: Sorooshian, S., Hsu, K.L., Coppola, E., Tomassetti, B., Verdecchia, M., Visconti, G. (Eds.): *Hydrological Modelling and the Water Cycle*. Water Science and Technology Library, 63, Springer, Berlin, Heidelberg. doi: 10.1007/978-3-540-77843-1_1
- Morris, M.D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2), 161–174. doi: 10.2307/1269043
- MRT (MODIS Reprojection Tool) (2004). Users guide. Release 3.2a. Department of Mathematics and Computer Science, South Dakota School of Mines and Technology, USGS EROS Data Center, 58 pp.
- Mullen, K.M., Ardia, D., Gil, D.L., Windover, D., & Cline, J. (2011). 'DEoptim': An R Package for global optimization by differential evolution. *Journal of Statistical Software*, 40(6), 1–26. doi: 10.18637/jss.v040.i06
- Murer, E., Wagenhofer, J., Aigner, F., & Cline, J. (2004). Die nutzbare Feldkapazität der mineralischen Böden der landwirtschaftlichen Nutzfläche Österreichs. Schriftenreihe BAW, Band 20, 72–78.
- Nachabe, M., Shah, N., Ross, M., & Vomacka, J. (2005). Evapotranspiration of two vegetation covers in a shallow water table environment. *Soil Science Society of America Journal*, 69, 492–499. doi: 10.2136/sssaj2005.0492
- Nagler, P.L., Scott, R.L., Westenburg, C., Cleverly, J.R., Glenn, E.P., & Huete, A.R. (2005). Evapotranspiration on western U.S. rivers estimated using the Enhanced Vegetation Index from MODIS and data from eddy covariance and Bowen ratio flux towers. *Remote Sensing of Environment*, 97(3), 337–351. doi: 10.1016/j.rse.2005.05.011

References

- Nester, T, Komma, J., & Blöschl, G. (2016). Real time flood forecasting in the Upper Danube basin. *Journal of Hydrology and Hydromechanics*, 64(4), 404 - 414. doi: 10.1515/johh-2016-0033
- Nielsen, S.A., & Hansen, E. (1973). Numerical simulation of the rainfall runoff process on a daily basis. *Nordic Hydrology*, 4, 171 -190.
- Nijzink, R.C., Almeida, S., Pechlivanidis, I.G., Capell, R., Gustafssons, D., Arheimer, B., et al. (2018). Constraining conceptual hydrological models with multiple information sources. *Water Resources Research*, 54, 8332-8362. doi: 10.1029/2017WR021895
- Ning, S., Ishidaira, H., & Wang, J. (2015). Calibrating a hydrologic model by step-wise method using GRACE TWS and discharge data. *Journal of Japan Society of Civil Engineers, Ser. B1 (Hydraulic Engineering)*, 71(4), I_85-I_90. doi: 10.2208/jscejhe.71.I_85
- Oguntunde, P.G., & Oguntuase, A.M. (2007). Influence of environmental factors on the sap flux density of mango trees under rain-fed cropping systems in West Africa. *International Journal of Plant Production*, 1(2), ISSN 1735-6814.
- Oishi, A.C., Oren, R., & Stoy, P.C. (2008). Estimating components of forest evapotranspiration: A footprint approach for scaling sap flux measurements. *Agricultural and Forest Meteorology*, 148(11), 1719-1732. doi: 10.1016/j.agrformet.2008.06.013
- O'Loughlin, E.M., Cheney, N.P., & Burns, J. (1982). The Bushrangers Experiment: hydrological responses of a eucalypt catchment to fire, 'First national symposium on forest hydrology', National Conference Publication (Vol 82, pp. 132-138). Australia: Institution of Engineers.
- Parajka, J., Merz, R., & Blöschl, G. (2003). Estimation of daily potential evapotranspiration for regional water balance modeling in Austria. In: 11th International Poster Day and Institute of Hydrology Open Day "Transport of Water, Chemicals and Energy in the Soil - Crop Canopy - Atmosphere System", Slovak Academy of Sciences, Bratislava, 299-306.
- Parajka, J., Naeimi, V., Blöschl, G., Wagner, W., Merz, R., & Scipal, K. (2006). Assimilating scatterometer soil moisture data into conceptual hydrologic models at the regional scale. *Hydrology and Earth System Sciences*, 10, 353-368. doi: 10.5194/hess-10-353-2006
- Parajka, J., Merz, R., & Blöschl, G. (2007). Uncertainty and multiple objective calibration in regional water balance modelling: case study in 320 Austrian catchments. *Hydrological Processes*, 21, 435-446. doi: 10.1002/hyp.6253
- Parajka, J., Viglione, A., Rogger, M., Salinas, J.L., Sivapalan, M., & Blöschl, G. (2013). Comparative assessment of predictions in ungauged basins - Part 1: Runoff hydrograph studies. *Hydrology and Earth System Sciences*, 17, 1783-1795. doi: 10.5194/hess-17-1783-2013
- Perrin, C., Michel, C., & Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 279(1-4), 275-289. doi: 10.1016/S0022-1694(03)00225-7
- Phillips, N., Oren, R., Zimmermann, R., & Wright, S.J. (1999). Temporal patterns of water flux in trees and lianas in a Panamanian moist forest. *Trees*, 14(3), 0116-0123. doi: 10.1007/s004680050216
- Pieruschka, R., Huber, G., & Berry, J.A. (2010). Control of transpiration by radiation. *PNAS*, 107(30), 13372-13377. doi: 10.1073/pnas.0913177107

References

- Rajib, M.A., Merwade, V., & Yu, Z. (2016). 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
- Rakovec, O., Kumar, R., Mai, J., Cuntz, M., Thober, S., Zink, M., et al. (2016). Multiscale and Multivariate Evaluation of Water Fluxes and States over European River Basins. *Journal of Hydrometeorology*, 17(1), 287–307. doi: 10.1175/JHM-D-15-0054.1
- Reigner, I.C. (1966). A method for estimating streamflow loss by evapotranspiration from the riparian zone. *Forest Science*, 12(2), 130–139.
- Renner, M., Hassler, S.K., Blume, T., Weiler, M., Hildebrandt, A., Guderle, M., et al. (2016). Dominant controls of transpiration along a hillslope transect inferred from ecohydrological measurements and thermodynamic limits. *Hydrological Earth System Sciences*, 20(5), 2063-2083. doi: 10.5194/hess-20-2063-2016
- Revilla-Romero, B., Beck, H.E., Burek, P., Salamon, P., de Roo, A., & Thielen, J. (2015). Filling the gaps: Calibrating a rainfall-runoff model using satellite-derived surface water extent. *Remote Sensing of Environment*, 171, 118-131. doi: 10.1016/j.rse.2015.10.022
- Rogger, M., Kohl, B., Pirkl, H., Viglione, A., Komma, J., Kirnbauer, R., Merz, R., & Blöschl, G., (2012). Runoff models and flood frequency statistics for design flood estimation in Austria - Do they tell a consistent story? *Journal of Hydrology*, 456-457, 30-43. doi: 10.1016/j.jhydrol.2012.05.068
- Rohrer, M.B. (1992). Die Schneedecke im Schweizer Alpenraum und ihre Modellierung. Zürcher Geographische Schriften, Heft 49, Geographisches Institut, Eidgenössische Technische Hochschule Zürich, Zürich.
- Ruiz-Perez, G., Koch, J., Manfreda, S., Caylor, K., & Frances, F. (2017). Calibration of a parsimonious distributed ecohydrological daily model in a data-scarce basin by exclusively using the spatio-temporal variation of NDVI. *Hydrology and Earth System Sciences*, 21, 6235-6251. doi: 10.5194/hess-21-6235-2017
- Savenije, H.H.G. (2001). Equifinality, a blessing in disguise? *Hydrological Processes*, 15, 2835-2838. doi: 10.1002/hyp.494
- Schaeffer, S.M., Williams, D.G., & Goodrich, D.C. (2000). Transpiration of cottonwood/willow forest estimated from sap flux. *Agricultural and Forest Meteorology*, 105(1-3), 257-270. doi: 10.1016/S0168-1923(00)00186-6
- Scharffenberg, W.A. (2001). Hydrologic Modeling System User's Manual. US Army Corps of Engineers Hydrologic Engineering Center, Davis, California, United States.
- Schrödter, H. (1985). Verdunstung - Anwendungsorientierte Messverfahren und Bestimmungsmethoden, Springer, 186 pages, ISBN: 978-3-642-70434-5.
- Seibert, J. (2000). Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. *Hydrology and Earth System Sciences*, 4, 215-224. doi: 10.5194/hess-4-215-2000
- Shah, N., Nachabe, M., & Ross, M. (2007). Extinction depth and evapotranspiration from ground water under selected land covers. *Ground Water*, 45(3), 329-338. doi: 10.1111/j.1745-6584.2007.00302.x

References

- Shahrban, M., Walker, J.P., Wang, Q.J., & Robertson, D.E. (2018). On the importance of soil moisture in calibration of rainfall-runoff models: two case studies. *Hydrological Sciences Journal*, 63(9), 1292-1312. doi: 10.1080/02626667.2018.1487560
- Shaw, S.B., & Riha, S.J. (2012). Examining individual recession events instead of a data cloud: Using a modified interpretation of $dQ/dt-Q$ streamflow recession in glaciated watersheds to better inform models of low flow. *Journal of Hydrology*, 434-435, 46-54. doi: 10.1016/j.jhydrol.2012.02.034
- Silasari, R., Parajka, J., Ressler, C., Strauss, P., & Blöschl, G. (2017). Potential of time - lapse photography for identifying saturation area dynamics on agricultural hillslopes. *Hydrological Processes*, 31, 3610-3627. doi: 10.1002/hyp.11272
- Silvestro, F., Gabellani, S., Rudari, R., Delogu, F., Laiolo, P., & Boni, G. (2015). 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
- Singh, V.P. (1988). Hydrologic Systems. Volume 1: Rainfall-runoff modelling. Prentice Hall, New Jersey, USA.
- Sivapalan, M., Blöschl, G., Merz, R., & Gutknecht, D. (2005). Linking flood frequency to long-term water balance: incorporating effects of seasonality. *Water Resources Research*, 41, W06012. doi: 10.1029/2004WR003439
- Skøien, J.O., Blöschl, G., & Western, A.W. (2003). Characteristic space scales and timescales in hydrology. *Water Resources Research*, 39(10), 1304. doi: 10.1029/2002WR001736
- Sleziak, P., Szolgay, J., Hlavcova, K., Danko, M., & Parajka, J. (2020). The effect of the snow weighting on the temporal stability of hydrologic model efficiency and parameters. *Journal of Hydrology*, 583, 124639. doi: 10.1016/j.jhydrol.2020.124639
- Sleziak, P., Szolgay, J., Hlavcova, K., Duethmann, D., Parajka, J., & Danko, M. (2018). Factors controlling alterations in the performance of a runoff model in changing climate conditions. *Journal of Hydrology and Hydromechanics*, 66(4), 381-392. doi: 10.2478/johh-2018-0031
- Snyder, K.A., & Williams, D.G. (2000). Water sources used by riparian trees varies among stream types on the San Pedro River, Arizona. *Agricultural and Forest Meteorology*, 105(1-3), 227-240. doi: 10.1016/S0168-1923(00)00193-3
- Steinacker, R. (1983). Diagnose und Prognose der Schneefallgrenze. *Wetter Leben*, 35, 81-90.
- Stensrud, D.J. (2007). Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models. Cambridge: Cambridge University Press. ISBN: 9780511812590
- Sun, W., Ishidaira, H., & Bastola, S. (2012). Calibration of hydrological models in ungauged basins based on satellite radar altimetry observations of river water level. *Hydrological Processes*, 26(23), 3524-3537. doi: 10.1002/hyp.8429
- Sun, W., Ishidaira, H., Bastola, S., & Yu, J. (2015). Estimating daily time series of streamflow using hydrological model calibrated based on satellite observations of river water surface width: Toward real world applications. *Environmental Research*, 139, 36-45. doi: 10.1016/j.envres.2015.01.002

References

- Szeftel, P. (2010). Stream-catchment connectivity and streamflow dynamics in a Montane landscape (PhD thesis). AB: The University of British Columbia, Vancouver.
- Széles, B., Broer, M., Parajka, J., Hogan, P., Eder, A., Strauss, P., & Blöschl, G. (2018). Separation of scales in transpiration effects on low flows – A spatial analysis in the Hydrological Open Air Laboratory. *Water Resources Research*, *54*, 6168-6188. doi: 10.1029/2017WR022037
- Széles, B., Parajka, J., Hogan, P., Silasari, R., Pavlin, L., Strauss, P., & Blöschl, G. (2020). The added value of different data types for calibrating and testing a hydrologic model in a small catchment. *Water Resources Research*. Under review.
- Szilágyi, J., Gribovszki, Z., & Kalicz, P. (2007). Estimation of catchment-scale evapotranspiration from baseflow recession data: Numerical model and practical application results. *Journal of Hydrology*, *336*(1-2), 206-217. doi: 10.1016/j.jhydrol.2007.01.004
- Szilágyi, J., Gribovszki, Z., & Kalicz, P. (2011). Comment on „Interference of river level changes on riparian zone evapotranspiration estimates from diurnal groundwater level fluctuations” by J. Zhu, J., Young, M., Healy, J., Jasoni, R., Osterberg, J. [*Journal of Hydrology*, *403*(3-4), 381-389]. *Journal of Hydrology*, *409*, 578-579. doi:10.1016/j.jhydrol.2011.08.006
- Szilágyi, J., Gribovszki, Z., Kalicz, P., & Kucsara, M. (2008). On diurnal riparian zone groundwater-level and streamflow fluctuations. *Journal of Hydrology*, *349*(1-2), 1-5. doi: 10.1016/j.jhydrol.2007.09.014
- Thompson, S.E., Harman, C.J., Troch, P.A., Brooks, P.D., & Sivapalan, M. (2011). Spatial scale dependence of ecohydrologically mediated water balance partitioning: A synthesis framework for catchment ecohydrology. *Water Resources Research*, *47*, W00J03. doi: 10.1029/2010WR009998
- Thyer, M., Beckers, J., Spittlehouse, D., Alila, Y., & Winkler, R. (2004). Diagnosing a distributed hydrologic model for two high-elevation forested catchments based on detailed stand- and basin-scale data. *Water Resources Research*, *40*, W01103. doi: 10.1029/2003WR002414
- van Griensven, A., & Meixner, T. (2003). LH-AT sensitivity analysis tool. Department of Environmental Sciences, University of California, Riverside, USA.
- van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., & Srinivasan, R. (2006). A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology*, *324*(1-4), 10-23. doi: 10.1016/j.jhydrol.2005.09.008
- Viglione, A., Parajka, J., Rogger, M., Salinas, J.L., Laaha, G., Sivapalan, M., & Blöschl, G. (2013). Comparative assessment of predictions in ungauged basins – Part 3: Runoff signatures in Austria. *Hydrology and Earth System Sciences*, *17*, 2263–2279. doi: 10.5194/hess-17-2263-2013
- Viglione, A., Rogger, M., Pirkl, H., Parajka, J., & Blöschl, G. (2018). Conceptual model building inspired by field-mapped runoff generation mechanisms. *Journal of Hydrology and Hydromechanics*, *66*(3), 303-315. doi: 10.2478/johh-2018-0010
- Voltz, T., Gooseff, M., Ward, A.S., Singha, K., Fitzgerald, M., & Wagener, T. (2013). Riparian hydraulic gradient and stream-groundwater exchange dynamics in steep headwater valleys. *Journal of Geophysical Research: Earth Surface*, *118*, 953-969. doi: 10.1002/jgrf.20074
- Wei, L., Link, T.E., Hudak, A.T., Marshall, J.D., Kavanagh, K.L., Abatzoglou, J.T., et al. (2016). Simulated water budget of a small forested watershed in the continental/maritime hydroclimatic region of the United States. *Hydrological Processes*, *30*, 2000-2013. doi: 10.1002/hyp.10769

References

- Western, A.W., Blöschl, G., & Grayson, R.B. (1998). How well do indicator variograms capture the spatial connectivity of soil moisture? *Hydrological Processes*, *12*(12), 1851-1868.
- White, W.N. (1932). A method of estimating ground-water supplies based on discharge by plants and evaporation from soil: Results of investigations in Escalante Valley, Utah. *U.S. Geol. Surv. Water Supply Pap.*, 659-A.
- Williams, D.G., Cable, W., Hultine, K., Hoedjes, J.C.B., Yezpe, E.A., Simonneaux, V., et al. (2004). Evapotranspiration components determined by stable isotope, sap flow and eddy covariance techniques. *Agricultural and Forest Meteorology*, *125*(3-4), 241-258. doi: 10.1016/j.agrformet.2004.04.008
- Williams, D.G., Scott, R.L., Huxman, T.E., Goodrich, D.C., & Lin, G. (2006). Sensitivity of riparian ecosystems in arid and semiarid environments to moisture pulses. *Hydrological Processes*, *20*(15), 3191-3205. doi: 10.1002/hyp.6327
- Wondzell, S.M., Gooseff, M.N., & McGlynn, B.L. (2007). Flow velocity and the hydrologic behaviour of streams during baseflow. *Geophysical Research Letters*, *34*, L24404. doi: 10.1029/2007GL031256
- Wondzell, S.M., Gooseff, M.N., & McGlynn, B.L. (2010). An analysis of alternative conceptual models relating hyporheic exchange flow to diel fluctuations in discharge during baseflow recession. *Hydrological Processes*, *24*(6), 686-694. doi: 10.1002/hyp.7507
- Yue, W., Wang, T., Franz, T.E., & Chen, X. (2016). Spatiotemporal patterns of water table fluctuations and evapotranspiration induced by riparian vegetation in a semiarid area. *Water Resources Research*, *52*, 1948-1960. doi: 10.1002/2015WR017546
- Zehe, E., & Sivapalan, M. (2009). Threshold behaviour in hydrological systems as (human) geoecosystems: Manifestations, controls, implications. *Hydrological Earth System Sciences*, *13*(7), 1273-1297. doi: 10.5194/hess-13-1273-2009

Appendix A1

Appendix A1

Appendix A1 shows the thresholds, i.e. minimum difference between the daily minimum and maximum streamflow rate, for each gauge used for the automatic episode identification (Table A1.1). These thresholds were chosen for each gauge after manually checking the streamflow and groundwater level time series.

Table A1.1. Minimum difference between the daily minimum and maximum streamflow and water level for each gauge, automatic episode identification.

Gauge	Threshold
MW - Outlet	0.1300 l/s
Sys4 - Inlet pipe	0.0700 l/s
Frau2 - Tile drain	0.0006 l/s
Sys1 - Tile drain (deep aquifer)	0.0130 l/s
Sys2 - Tile drain	0.0045 l/s
Sys3 - Tile drain/Wetland	0.0018 l/s
A1 - Wetland	0.0120 l/s
A2 - Wetland	0.0040 l/s
Piezometers	0.5 cm

Appendix A2

Appendix A2 contains details on the calculations applied to quantify the amplitudes of the diel fluctuations.

A difference was made between the falling and the rising limbs of the diel signals when the amplitudes of the diel fluctuations were calculated. For each day in each episode the daily minimum $Q_{m,min}$ (L^3T^{-1}) and maximum $Q_{m,max}$ (L^3T^{-1}) measured discharge values were selected. Those measured amplitudes a_m (L^3T^{-1}) which describe the falling limbs of the diel fluctuations were calculated according to (A2.1) starting from the first day

$$a_m(2j - 1) = |Q_{m,max}(j) - Q_{m,min}(j)| \quad (A2.1)$$

where j denotes the day within an episode.

Measured amplitudes a_m (L^3T^{-1}) describing the rising limbs of the diel fluctuations were calculated according to (A2.2) starting from the second day

$$a_m(2j - 2) = |Q_{m,max}(j) - Q_{m,min}(j - 1)| \quad (A2.2)$$

Measured amplitude A_m describing an episode was calculated as the mean of a_m . Simulated amplitudes a and the mean value A were obtained in a similar way.

The average measured amplitude of the groundwater level fluctuations $A_{m,gwl}$ was evaluated similarly based on the daily minimum and maximum measured groundwater levels.

Appendix A3

Appendix A3

Appendix A3 contains information on the optimization procedure.

Before the optimization procedure, we applied a simple smoothing algorithm on the Q_m observed discharge data according to (A3.1)

$$Q_m'(t_i) = \frac{1}{2}(1-w)Q_m(t_{i-1}) + wQ_m(t_i) + \frac{1}{2}(1-w)Q_m(t_{i+1}) \quad (\text{A3.1})$$

where w is 0.5, i is the hourly time step.

The multiple objective calibration approach is based on the shuffled complex evolution method with principal components analysis (Chu *et al.*, 2011). The compound objective function incorporates additional information on the timing and amplitudes of the fluctuations, thereby providing a better fit between the measured and simulated streamflow. The model was fitted to each of the recession periods independently by minimizing the objective function Z according to (A3.2)

$$Z = w_1Z_E + w_2Z_A + w_3Z_T \quad (\text{A3.2})$$

where Z_E (L^3T^{-1}) is the root mean square error (A3.3), Z_A (L^3T^{-1}) is the amplitude error (A3.4), Z_T (T) is the error of timing (A3.5). w_1, w_2, w_3 weights were assigned in test simulations and sensitivity analyses as $w_1=10, w_2=1, w_3=0.1$ for MW catchment outlet and $w_1=100, w_2=10, w_3=0.1$ for the tributaries. The results were only moderately sensitive to the selection of weights. Z_E was calculated as the root of the average of the squared difference between the simulated and observed discharge according to (A3.3)

$$Z_E = \sqrt{\frac{\sum_{i=1}^N (Q(t_i) - Q_m'(t_i))^2}{N}} \quad (\text{A3.3})$$

where Q is the simulated streamflow, Q_m' is the observed and smoothed streamflow, N is the number of time steps in one episode. Z_A amplitude error expresses the difference between the a (L^3T^{-1}) simulated amplitudes and a_m' (L^3T^{-1}) amplitudes of the measured and smoothed discharge time series according to (A3.4)

$$Z_A = \sum_{k=1}^{2M-1} |a(k) - a_m'(k)| \quad (\text{A3.4})$$

where M is the number of days within one episode.

Z_T is the error of timing which is defined as the difference between the time of the simulated and measured minimum and maximum discharge within one day according to (A3.5)

Appendix A4

$$Z_T = \sum_{j=1}^M |t_{max}(j) - t_{m,max}'(j)| + \sum_{j=1}^M |t_{min}(j) - t_{m,min}'(j)| \quad (A3.5)$$

where t_{max} is the time of the maximum simulated discharge within one day, $t_{m,max}'$ is the time of the maximum observed and smoothed discharge within one day, t_{min} is the time of the minimum simulated discharge within one day, $t_{m,min}'$ is the time of the minimum observed and smoothed discharge within one day.

In certain cases (e.g. when the tributaries did not show the diel signals and the procedure weakly converged) the compound objective function (A3.2) was simplified and only the root mean square error according to (A3.3) was applied.

Appendix A4

Appendix A4 contains information on a literature based evapotranspiration estimation method.

Based on the heterogeneity of the vegetation, i.e. the dominant tree types, the riparian zone was divided into three parts (upstream area is dominated by field maple and black alder, middle section by poplar, downstream by ash). For each part, a 10x10 m representative area was selected, the dominant tree types were counted and based on tree height and the diameter at breast height the transpiration rates for the different tree species for the growing season were estimated based on literature values (see Appendix A5). According to previous studies the dominant trees (one third of the total tree number) account for about two-thirds of the stand level water loss (Čermák et al., 2004), therefore the water loss of the dominant tree types was multiplied by a factor of 1.5 in order to take into account the transpiration of the understory vegetation. The water use of the representative 100 m² patch was upscaled to the entire tree stand. The spatial extent of the tree stand was estimated based on aerial photographs.

Appendix A5

Appendix A5 contains information on literature based transpiration values for different tree species.

Table A5.1. Literature based transpiration values

Tree type	Reference	Diameter at breast height (cm)	Water use (l/d)
Ash	Köcher et al. (2008)	39.5	8.7
Poplar	Hinckley et al. (1994)	16.0	50.0
Field maple	Beeson (2011)	11.0	55.0

Appendix A6

Appendix A6

Appendix A6 contains information on a literature based evapotranspiration estimation method.

Evapotranspiration was calculated based on the White method (White, 1932) and the empirical method of Gribovszki et al. (2008) using piezometer measurements from the left side of the stream (piezometer BP07). Specific yield was estimated using Table 1 from Loheide et al. (2005). According to a detailed soil survey in the HOAL, the surrounding area of BP07 is dominated by silt (silt: 76.70%, clay: 14.45%, sand: 8.85%), therefore the readily available specific yield was chosen to be $S_y=0.037$.

Appendix A7

Appendix A7 shows the minimum and maximum groundwater level and the corresponding depth to a certain level at five piezometers during a five-day period, 14-18 August 2013.

Table A7.1. Minimum and maximum groundwater level and corresponding depth to groundwater level during a five-day episode, 14-18 August 2013.

Piezometer	Minimum		Maximum	
	<i>GWL</i> (m asl)	Depth to <i>GWL</i> (m)	<i>GWL</i> (m asl)	Depth to <i>GWL</i> (m)
BP02	258.23	0.76	258.50	0.49
BP07	270.87	0.41	270.97	0.31
H01	261.70	4.80	261.79	4.72
H02	258.04	3.57	258.61	3.00
H04	265.07	0.97	265.25	0.79

Appendix A8

Appendix A8 contains information on the model performance statistics.

Table A8.1. Model performance statistics.

Nash Sutcliffe Coefficient	MW Outlet 2002-2015	MW Outlet 2013-2015	Virtual gauge LF (A1 and A2 Wetland, Sys3 Tile drain/Wetland)	Virtual gauge SF (Sys4 Inlet, Frau2, Sys1, Sys2 Tile drain)
Median	0.89	0.72	0.45	0.48
25 th percentile	0.68	0.31	0.04	0.01
75 th percentile	0.95	0.89	0.73	0.73

Appendix A9

Appendix A9 contains details on the sensitivity analysis.

The influence of changes in the model parameters on the daily runoff hydrograph was assessed by the LH-OAT method, a global sensitivity analysis method (van Griensven & Meixner, 2003; van

Appendix A10

Griensven et al., 2006), which combines the Latin-Hypercube (McKay et al., 1979) with the One-Factor-At-a-Time sampling (Morris, 1991). The number of intervals in the Latin Hypercube was chosen as 500 and the parameter change for OAT fraction as 0.05. The changes in the runoff hydrograph relative to the base case (LH point) was measured by the Nash-Sutcliffe coefficient. The mean sensitivities of the 14 model parameters were compared.

Comparing the mean sensitivities of the model parameters assessed by the daily Nash Sutcliffe coefficient for runoff, the two most sensitive parameters were the field capacity FC and the storage time for very fast response k_0 (Appendix A10). The same parameters were found to be the most sensitive for all time periods (2013-15 calibration period in Appendix A10, 1991-2012 and 2016-17 validation periods not shown here).

Appendix A10

Table A10.1 contains the results of the sensitivity analysis.

Table A10.1. Results of the LH-OAT sensitivity analysis for the calibration period 2013-15: mean sensitivities (%) of 14 free parameters, assessed by the influence of model parameters on the change in daily Nash-Sutcliffe coefficient for runoff. The most sensitive model parameter is FC (field capacity) of the soil moisture accounting module as would be expected in a humid climate.

Module	Parameter (unit)	Mean sensitivity (%)
Snow	SCF (-)	0.7
	DDF (mm/°C/d)	5.1
	T_{wb} (°C)	0.7
	T_m (°C)	1.8
Soil moisture accounting	LP_{rat} (-)	2.4
	FC (mm)	37.7
	β (-)	6.2
Runoff generation	k_0 (d)	18.6
	k_1 (d)	8.2
	k_2 (d)	0.6
	LS_{UZ} (mm)	0.6
	c_P (mm/d)	11.2
	B_{MAX} (d)	1.7
	c_R (d ² /mm)	4.4

Appendix A11

Appendix A11

Appendix A11 contains details on the calculation of the wet bulb temperature during precepitation events with a shift in precipitation phase.

Those precipitation events were selected during the model calibration period (2013-15), when a shift was observed in the phase of the precipitation measured by the Present Weather Sensor (from snow to rain, i.e. from category 3 to 2, or from rain to snow, i.e. from category 2 to 3). The wet bulb temperature was calculated with the R package bigleaf (Knauer et al., 2018) for each two half hourly time steps during the precipitation phase shifts (one half hour with category 3, and the following half hour with category 2, or vice versa). The average of the wet bulb temperature for each two half hourly time steps was calculated. And finally, the median of the average wet bulb temperature was extracted.