Ingenieurfakultät Bau Geo Umwelt

Lehrstuhl für Hydrologie und Flussgebiets-management

Application of the Soil Water Assessment Tool (SWAT) in alpine catchments: pitfalls and solutions

Ye Tuo

Vollständiger Abdruck der von der Ingenieurfakultät für Bau Geo Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines **Doktor-Ingenieurs (Dr.-Ing.)** genehmigten Dissertation.

Vorsitzender: Prof. Dr.-Ing. Michael Manhart

Prüfer der Dissertation:

1. Prof. Dr.-Ing. Markus Disse
2. Prof. Dr. rer. nat. Gabriele Chiogna (University of Innsbruck)
3. Prof. Dr. rer. nat. Niels Schütze (Technical University of Dresden)

 Die Dissertation wurde am 29.03.2018 bei der Technischen Universität München eingereicht und durch die Ingenieurfakultät Bau Geo Umwelt am 07.06.2018 angenommen.

ABSTRACT

Hydrological behavior of large scale mountainous catchment is difficult to model properly due to complex topographic and climatic conditions and the sparse available data especially in high elevation areas. Improving the reliability of modeling results has important implications for water management and hydrological hazard prevention in mountainous region. Therefore, this work aims to improve the application of Soil and Water Assessment Tool (SWAT) in alpine catchments. It focuses on investigating the major influencing factors, including precipitation, snow processes and hydropower plant operation, which complicate the modeling applications in many alpine catchments such as the Adige River Basin, Italy. The leading question is how to effectively utilize the data related to precipitation, snow, and hydropower plant operation during the modeling processes in order to improve the reliability of the SWAT model results.

Evaluations have been performed to study the selection of the optimal precipitation input for SWAT, not only by comparing the performance of discharge estimations, but also by considering the parameter and output uncertainty propagated from different precipitation datasets. Indicated by the results, selecting an optimal precipitation input is of fundamental importance for SWAT applications in these Alpine catchments, since different precipitation inputs result in different model performances, parameter uncertainties and output uncertainties. These results can lead to distinct judgments about the most relevant hydrological processes and hence result in different water management strategies or policies.

To cope with SWAT model structure, snow water equivalent (SWE) time series of both subbasin and point scales have been converted based on the snow depth measurements. Furthermore, the effects of involving SWE data into SWAT model calibration are investigated, by assessing different single objective calibration approaches and a multi-objective calibration approach. Revealed by the outcomes, involving snow information is important for constraining the snow parameters to respect the real snow behaviors of the Alpine catchments. Moreover, using the multi-objective calibration against both SWE and river discharge, SWAT model could obtain river discharge simulation that is composed of more reliable contributions of different hydrological components.

For simulating the effect of hydropower plant operation, Support Vector Machine (SVM) has been jointly used with SWAT model to take use of energy price information that is relevant to hydropower plant operation but cannot be utilized by SWAT directly. As consequences, employing SVM improves the simulation of the affected river discharge by learning from the day of the week and the energy price

data which drive the operation of hydropower plant. This helps to fill the gap of SWAT model structure to model the hydropeaking effect by utilizing alternative data.

To conclude, SWAT applications in Alpine catchments have been improved by rational utilization of the available data related to precipitation, snow and hydropower plant operation, considering the corresponding model structure. The selection of optimal precipitation input, utilization of snow data during model calibration, and consideration of energy price data and the day of the week to reproduce hydropeaking, have the common effect of reducing the uncertainty of model results by considering more observed information. This effect leads to improvement on the reliability of model results and could benefit the water management practices and even hydrological hazard forecast in Alpine catchments.

ZUSAMMENFASSUNG

Das hydrologische Verhalten von mesoskaligen alpinen Einzugsgebieten ist aufgrund ihrer Komplexität, sowohl hinsichtlich der gegebenen Topographie als auch der klimatischen Bedingungen sowie der größtenteils spärlichen Datenlage, schwierig zu modellieren. Speziell die Heterogenität von Niederschlägen, Schneeprozessen und der Betrieb von Wasserkraftanlagen erschwert die Anwendung hydrologischer Modelle in solchen Gebieten. Um jedoch wichtige Erkenntnisse, sowohl für die Wasserwirtschaft, als auch zur Prävention von hydrologisch-bedingten Naturgefahren in alpinen Gebieten zu gewinnen, müssen die Modelle hydrologische Prozesse zuverlässig wiedergeben. Das Ziel dieser Arbeit ist es daher, die Anwendbarkeit des Soil and Water Assessment Tool (SWAT) in alpinen Flussgebieten, speziell im Einzugsgebiet der Etsch in Italien, zu verbessern. Dabei soll untersucht werden, wie Niederschlags-, Schnee- und der Betriebsdaten von Wasserkraftanlagen im Modell effektiver genutzt werden können, um die damit verbundenen natürlichen Prozesse besser abzubilden und somit die Zuverlässigkeit der SWAT-Modellergebnisse zu verbessern.

Zunächst wurden verschiedene Niederschlagsdatensätze in das SWAT-Modell integriert und die simulierten Abflüsse bezüglich ihrer Parameter- und Ergebnisunsicherheit verglichen und ausgewertet. Die Auswahl eines optimalen Niederschlagsdatensatzes für SWAT-Anwendungen in alpinen Einzugsgebieten ist von grundlegender Bedeutung, da verschiedene Niederschlagsdatensätze stark unterschiedliche Modellgüten, Parameterwerte und Ergebnisunsicherheiten verursachen. So können die Modellergebnisse zu verschiedenen Simulationen der hydrologischen Prozesse und anschließend zu unterschiedlichen Wassermanagementstrategien oder –richtlinien führen.

Anschließend wurden die Schneetiefenmessungen der Schneestationen sowohl für die jeweiligen Messpunkte, als auch für Teileinzugsgebiete in Schneewasseräquivalente (SWÄ) umgerechnet und in die Modellstruktur von SWAT integriert. Darüber hinaus wurde durch die Verwendung verschiedener ‚single objective'-Kalibrierungsansätze und eines ‚multi-objective'- Kalibrierungsansatzes untersucht, wie sich die Einbeziehung von SWÄ-Daten im Kalibrierungsprozess auf die Modellergebnisse auswirkt. Die Ergebnisse zeigen, dass die Integration von Schneedaten im Modellevaluierungsprozess unabdingbar ist, um die tatsächlich stattfindenden Schneeprozesse in alpinen Einzugsgebieten adäquat zu berücksichtigen. Durch eine ‚multi-objective'-Kalibrierung für SWÄ **und** Abfluss konnten schließlich zuverlässigere Abflusszeitreihen im SWAT-Modell generiert werden.

Energiepreisinformationen gelten als guter ‚predictor' für den Betrieb von Wasserkraftanlagen. Da diese jedoch nicht direkt von SWAT verwendet werden können, wurde im letzten Schritt der Dissertation eine Support Vector Machine (SVM) mit dem SWAT-Modell gekoppelt, um die Auswirkungen des Betriebs

von Wasserkraftanlagen auf das Abflussverhalten zu simulieren. Es konnte gezeigt werden, dass die Simulationsergebnisse des von der Wasserkraft beeinflussten Abflusses durch die Berücksichtigung des Energiepreises verbessert wurden. Diese gekoppelte Verwendung der SVM hilft, die Lücke der SWAT-Modellstruktur zu füllen und alternative Daten zur Modellierung des Hydropeakingeffekts zu verwenden.

Zusammenfassend kann gesagt werden, dass die Anwendung des SWAT-Modells in alpinen Einzugsgebieten durch die intelligente Nutzung von verfügbaren Niederschlags-, Schnee- und Wasserkraftbetriebsdaten und unter Berücksichtigung der entsprechenden Modellstruktur verbessert wurde. Die Auswahl eines optimalen Niederschlagsdatensatzes, die Nutzung von Schneedaten während der Modellkalibrierung und die Berücksichtigung von Energiepreisinformationen zur Simulation des Hydropeakingeffekts tragen dazu bei, die Unsicherheit der Modellresultate zu reduzieren. Dies wiederum führt zu einer allgemeinen Verbesserung der Zuverlässigkeit von Modellergebnissen und kann die Planung wasserwirtschaftlicher Maßnahmen, sowie die Vorhersage hydrologisch-bedingter Naturgefahren in alpinen Einzugsgebieten verbessern.

AFFIDAVIT

I hereby affirm that I wrote this PhD thesis independently and on my own without illegal assistance of third parties. To the best of my knowledge, all sources that I used to prepare that thesis are labeled as such. This thesis has not been received by any examination board, neither in this nor in a similar form.

_____

Ye Tuo

March, 20, 2018

ACKNOWLEDGEMENTS

Last, but not least I would like to thank my family for their constant support and encouragement over these years. Especially, I would like to thank Ying Zhu, the person who is always there to support me and share my feelings and experiences with endless patience.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

%MAE	Relative Mean Absolute Error
ANN	Artificial Neural Network
BIAS	Relative Bias
CCD	Cold Cloud Duration
CDF	Cumulative Density Function
CFS	Climate Forecast System
CGIAR-CSI	Consortium for Spatial Information
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station data [mm]
CHPclim	CHIRPS product include the monthly precipitation climatology
CLC2006	Corine Land Cover 2006
CMORPH	Climate Prediction Center MORPHing technique
CMORPH_BLD	CMORPH Blended product [mm]
CMORPH_CRT	CMORPH bias corrected product [mm]
CMORPH_RAW	CMORPH Raw product [mm]
CPC	Climate Prediction Center
CRU	Climatic Research Unit
CSAVs	Cloud System Advection Vectors
CSI	Critical Success Index
ECMWF	European Centre for Medium-Range Weather Forecasts
FAO	Food and Agriculture Organization
FAR	False Alarm Ratio
GHCN	Global Historical Climatology Network
GLUE	Generalized Likelihood Uncertainty Estimation
GPCC	Global Precipitation Climatology Centre
GPM	Global Precipitation Measurement
GridSat-B1	Gridded Satellite Infrared data
GSMaP_MVK	Global Satellite Mapping of Precipitation project Moving Vector with Kalman-filter product [mm]
HRU	Hydrologic Response Unit
HS	Snow depth data [m]

IDW	Inverse Distance Weighting interpolated product [mm]
IR	Infrared
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ME	Mean Error
MODFLOW	Modular finite-difference flow model
ModSpa	Modèle Spatialisé
MV	Microwave
NCDC	National Climate Data Center
NCEP–NCAR	National Centers for Environmental Prediction–National Center for Atmospheric Research
NOAA	National Oceanic and Atmospheric Administration
NS/NSE	Nash–Sutcliffe coefficient/Nash-Sutcliffe Efficiency
OI	Optimal interpolation
OP	Observed precipitation [mm]
PBIAS	Percent Bias
PCDR	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record [mm]
PDF	Probability Density Function
PERSIANN	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks
PERSIANN-CDR	PERSIANN Climate Data Record
PGF	Princeton University global meteorological dataset [mm]
POD	probability of detection
PREVAH	Precipitation-Runoff-Evapotranspiration-Hydrotope model
R^2	Coefficient of determination
RHESSys	Regional Hydro-Ecologic Simulation System
RMSE	Root Mean Square Error
RSR	Ratio of RMSE to the standard deviation of observations
RT3D	Reactive Transport in 3 Dimensions
SCD	Snow Cover Duration
SMAP	Soil Moisture Active Passive

SPD	Snow Peak Day
SPV	Snow Peak Value
SRTM	Shuttle Radar Topography Mission
SUFI-2	Sequential Uncertainty Fitting algorithm version 2
SVM	Support Vector Machine
SWAT	Soil and Water Assessment Tool
SWE	Snow Water Equivalent [mm]
SWOT	Surface Water and Ocean Topography Mission
TMPA	TRMM multi-satellite precipitation analysis
Topkapi	Topographic Kinematic Approximation and Integration
TRMM	Tropical Rainfall Measuring Mission product [mm]
TUM	Technical University of Munich
WaSiM	Wasserhaushalts-Simulations-Modell
WMO	World Meteorological Organization
WPS	Wavelet Power Spectrum
WTC	Wavelet Coherence

AUTHOR CONTRIBUTIONS

Chapters 2 to 6 in this work have been published in a similar form in the following peer-reviewed journals. Two additional publications are not relevant to the topic of the dissertation, but they are displayed in the list of the peer-reviewed publications below.

Chapter 2

Duan Z., Liu J., Tuo Y., Chiogna G., Disse M., 2016. **Evaluation of eight spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales**. Science of The Total Environment, 573: 1536-1553. DOI: 10.1016/j.scitotenv.2016.08.213

Ye Tuo prepared observed precipitation data and DEM map, took part in conceiving the study, and reviewed the paper. Zheng Duan wrote the manuscript, designed the study, and carried out the data processing of remote sensed products. Junzhi Liu took part in the data processing of remote sensed products. Gabriele Chiogna took part in conceiving the study, collected the observed precipitation data and reviewed the manuscript. Markus Disse took part in conceiving the study and reviewed the manuscript. All authors contributed to the modification of the manuscript.

Chapter 3

Tuo Y., Duan Z., Disse M., Chiogna G., 2016. **Evaluation of precipitation input for SWAT modeling in Alpine catchment: A case study in the Adige river basin (Italy)**. Science of The Total Environment, 573: 66-82. DOI: 10.1016/j.scitotenv.2016.08.034

Ye Tuo wrote the manuscript, prepared the geographic and climate inputs and the observed and IDW precipitation datasets, designed the study, and carried out the modeling works. Zheng Duan provided the remote sensed datasets and the corresponding information, took part in in conceiving the study, and reviewed the manuscript. Markus Disse took part in in conceiving the study and reviewed the manuscript. Gabriele Chiogna collected the hydrological data, conceived the study, and reviewed the manuscript. All authors contributed to the modification of the manuscript.

Chapter 4

Tuo, Y., Marcolini, G., Disse, M., Chiogna, G., 2018. **Calibration of Snow Parameters in SWAT: Comparison of Three Approaches in the Upper Adige River Basin**. Hydrological Sciences Journal. DOI: 10.1080/02626667.2018.1439172

Ye Tuo wrote the manuscript, prepared the geographic and climate inputs, designed the study, and carried out the modeling works. Giorgia Marcolini provided the homogenized snow depth dataset and

reviewed the manuscript. Markus Disse took part in conceiving the study and reviewed the manuscript. Gabriele Chiogna collected the hydrological data, prepared the snow water equivalent data, conceived the study, and reviewed the manuscript. All authors contributed to the modification of the manuscript.

Chapter 5

Tuo, Y., Marcolini, G., Disse, M., Chiogna, G., 2018. **A Multi-Objective Approach to Improve SWAT Model Calibration in Alpine Catchments**. Journal of Hydrology, 559: 347-360. DOI: 10.1016/j.jhydrol.2018.02.055

Ye Tuo wrote the manuscript, prepared the geographic and climate inputs, designed the study, and carried out the modeling works. Giorgia Marcolini provided the homogenized snow depth dataset and reviewed the manuscript. Markus Disse took part in conceiving the study and reviewed the manuscript. Gabriele Chiogna collected the hydrological data, prepared the snow water equivalent data, conceived the study, and reviewed the manuscript. All authors contributed to the modification of the manuscript.

Chapter 6

Chiogna, G., Marcolini, G., Liu W.Y., Ciria T.P., Tuo, Y., 2018. **Coupling Hydrological Modeling and Support Vector Regression to Model Hydropeaking in Alpine Catchments**. Science of the Total Environmental, 633: 220-229. DOI: 10.1016/j.scitotenv.2018.03.162

Ye Tuo designed and carried out the SWAT modeling work, took part in conceiving the study, and reviewed the manuscript. Gabriele Chiogna wrote the manuscript, collected the energy price data, and designed the study. Giorgia Marcolini took part in conceiving the study and guided the use and analysis of the support vector machines (SVM) work. Wanying Liu carried out the SVM work. Teresa Perez Ciria carried out the wavelet analysis. All authors contributed to the modification of the manuscript.

List of Peer-Reviewed Publications

Chiogna, G., Marcolini, G., Liu W.Y., Ciria T.P., <u>Tuo, Y.</u>, 2018. **Coupling Hydrological Modeling and Support Vector Regression to Model Hydropeaking in Alpine Catchments.** Science of the Total Environmental, 633: 220-229. DOI: 10.1016/j.scitotenv.2018.03.162

<u>Tuo, Y.</u>, Marcolini, G., Disse, M., Chiogna, G., 2018. **A Multi-Objective Approach to Improve SWAT Model Calibration in Alpine Catchments.** Journal of Hydrology, 559: 347-360. DOI: 10.1016/j.jhydrol.2018.02.055

<u>Tuo, Y.</u>, Marcolini, G., Disse, M., Chiogna, G., 2018. **Calibration of Snow Parameters in SWAT: Comparison of Three Approaches in the Upper Adige River Basin.** Hydrological Science Journal, DOI: 10.1080/02626667.2018.1439172

Vigiak, O., Lutz, S., Mentzafou, A., Chiogna, G., <u>Tuo, Y.</u>, Majone, B., Beck, H., de Roo, A., Malagó, A., Bouraoui, F., Kumar, R., Samaniego, L., Merz, R., Gamvroudis, C., Skoulikidis, N., Nikolaidis, N.P., Bellin, A., Acuňa, V., Mori, N., Ludwig, R., Pistocchi, A., 2018. **Uncertainty of modelled flow regime for flow-ecological assessment in Southern Europe.** Science of the Total Environment, 615: 1028-1047. DOI: 10.1016/j.scitotenv.2017.09.295

<u>Tuo, Y.</u>, Duan, Z., Disse, M., Chiogna, G., 2016. **Evaluation of precipitation input for SWAT modeling in Alpine catchment: A case study in the Adige river basin (Italy).** Science of the Total Environment, 573: 66-82. DOI: 10.1016/j.scitotenv.2016.08.034

Duan, Z., Liu, J., <u>Tuo, Y.</u>, Chiogna, G., Disse, M., 2016. **Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales.** Science of the Total Environment, 573: 1536-1553. DOI: 10.1016/j.scitotenv.2016.08.213

<u>Tuo, Y.</u>, Chiogna, G., Disse, M., 2015. **A multi-criteria model selection protocol for practical applications to nutrient transport at the catchment scale.** Water, 7: 2851-2880. DOI: 10.3390/w7062851

Chapter 1

Introduction and research questions

Mountainous catchments (e.g. Alpine catchments) are vital sources of fresh water for agricultural production and drinking water of all continents (Farinotti et al., 2012; Fatichi et al., 2015; Nepal et al., 2014; Shrestha et al., 2004; Viviroli et al., 2007; Zierl and Bugmann, 2005). They play an important role in regulating regional water cycle and seasonal water resource distributions and maintaining sustainable conservation of the regional ecosystem (Shrestha et al., 2004; Verbunt et al., 2003; Viviroli et al., 2007). The topography and climate conditions of mountainous regions are complex and highly variable in both time and space (Bertoldi et al., 2010; Gurtz et al., 2003; Laghari et al., 2012; Moussa et al., 2007; Nepal et al., 2014). These complicated topography and climate characteristics lead to complex behaviors of different hydrological processes and make it difficult to quantify the interactions in-betweens (Chauvin et al., 2011; Tobin et al., 2011; Verbunt et al., 2003). What is more, erosion, flood and drought are not uncommon hydrological phenomena in mountainous areas. In particular, ascribed to the impact of climate change and global warming, they have become prevalent in recent decades and brought noticeable damages and threatens to water security, economy, and the regional ecosystem (Blöschl et al., 2017; Bohm et al., 2015; Fatichi et al., 2015; Iida et al., 2012; Meißl et al., 2017; Moussa et al., 2007; Poulenard et al., 2012; Schaefli et al., 2007; Van Loon and Laaha, 2015; Zappa and Kan, 2007). Therefore, it is necessary to model hydrological processes in mountainous areas, which will facilitate the formulation of reasonable water management strategies and policies for hydropower operation, drinking water supply and agriculture activity of the region (Farinotti et al., 2012; Tobin et al., 2011). In addition, the model results could also offer reliable and timely references to forecast hydrological hazards including flood, drought, and erosion (Blöschl et al., 2017; Poulenard et al., 2012; Van Loon and Laaha, 2015).

A successful hydrological modeling work starts with appropriate understandings of the key influence factors. In mountainous catchments such as the Alpine regions, precipitation, snow processes (snow accumulation and snowmelt) and hydropower plant operation play crucial roles in affecting the regional water resources. Precipitation is a major component and a key driving force for hydrology cycle (Masih et al., 2011; Price et al., 2014), having decisive impacts on the total water quantity. Therefore, reliable precipitation data are important bases to investigate hydrological processes. Impacted by orographic effect and specific local climate condition such as wind effect (Tobin et al., 2011), precipitation is complex and highly variable in both temporal and spatial scales with large uncertainty in mountainous region (Moussa et al., 2007; Tobin et al., 2011; Tuo et al., 2016). Due to the orographic effect, high elevation areas generally receive much higher precipitation than the lower areas (Horton et al., 2006; Laghari et al., 2012; Shinohara et al., 2009; Tuo et al., 2016). However, in mountainous catchments, in particular, in the Alpine catchments, climate stations are generally sparse compared to the scale of the area and they are normally located at low elevation areas and rarely available at high elevation parts. As

a consequence, precipitation in mountainous region is normally underestimated by the observed data. In other words, the precipitation data of mountainous catchments cannot represent the real regional total precipitation. Large uncertainty is expected in the hydrological information derived from these data. In addition to precipitation, snow is another important hydrological components of the water cycle in high mountains, mainly affecting the seasonal distribution of water quantity in space and time (Farinotti et al., 2012; Horton et al., 2006; Shinohara et al., 2009; Shrestha et al., 2004; Warscher et al., 2013). It brings delays in runoff by storing water through snow accumulation and later releases water through snowmelt (Zierl and Bugmann, 2005). Similar to precipitation, snow processes of the mountainous region are also highly variable in both time and space (Chauvin et al., 2011; Warscher et al., 2013), due to the impact of the complex topography such as shadowing effect (Tobin et al., 2011) and the influence of local climate like wind effect (Hiemstra et al., 2002). Snow information is normally available in terms of snow depth data with sparse spatial distribution. Moreover, snowmelt data are rare in large scale catchments. These snow depth data are valuable and should be handled properly for deriving useful information of the snow processes.

Apart from the impacts of these natural driving forces, the water cycle of mountainous regions also gets affected by hydropower plant operation. Hydropower is one major water use in the mountainous regions, benefiting from large elevation differences of the topography. The operation of hydropower plant interrupts the natural behaviors of streamflow, which impacts the available water quantity downstream and threatens the ecosystem (Person, 2013). The operation information is therefore very important for a reliable description of water resources in the mountainous regions (Majone et al. 2016). Furthermore, a reasonable understanding about the regional water cycle in return helps to formulate a reasonable hydropower plant operation strategy to cope with different hydrological situations properly. However, operation data are normally not publically available and kept as commercial secret by hydropower company. Fortunately, since the operation of major hydropower plants now generally depends on the fluctuations of energy price, the energy price data could be an alternative reference to imply the operation information and they are generally open access.

Based on their important implications for the water resources of mountainous regions, precipitation, snow processes and the effect of hydropower plant operation require primary attentions during the hydrological modeling applications. Well understanding and reliable representation of these principal factors indicate a successful grasp of general hydrological behaviors.

Hydrological models have been developed and applied to simulate and estimate the hydrological processes such as precipitation and snow dynamics and the effect of hydropower plant operation in mountainous catchments. These models, e.g. J2000 (Nepal et al., 2014), ModSpa (Modèle Spatialisé)

(Moussa et al., 2007), PREVAH (Precipitation-Runoff-Evapotranspiration-Hydrotope model) (Laghari et al., 2012), RHESSys (Regional Hydro-Ecologic Simulation System) (Zierl and Bugmann, 2005), SWAT (Soil and Water Assessment Tool) (Arnold et al., 1998), Topkapi (Topographic Kinematic Approximation and Integration) (Fatichi et al., 2015), and WASIM (Wasserhaushalts-Simulations-Modell) (Warscher et al., 2013), differ from each other in model structure, e.g. distinct representations of processes and different temporal resolutions and spatial scales, depending on the application purposes. Among them, SWAT is one of the most frequently used around the world (Abbaspour et al., 2015; Dile et al., 2016; Francesconi et al., 2016; Golmohammadi et al., 2017; Monteiro et al., 2016; Shen et al., 2012), because of its catchment scale based structure, diverse modules, relative low data requirement, user friendly interface, and the consideration of physical processes as well. The SWAT model is a comprehensive, time-continuous, semi-distributed, process-based model developed by the Agricultural Research Service of the United States Department of Agriculture (Arnold and Fohrer, 2005; Arnold et al., 1998). It divides river basins into subbasins and subsequently into Hydrologic Response Units (HRUs), characterized by different combinations of land use, soil characteristics, and topography. The hydrological cycle is calculated based on water balance, which is controlled by climate inputs such as daily precipitation and maximum/minimum air temperature. Using daily input time series, SWAT simulates the daily, monthly and yearly fluxes of water and solutes in river basins.

Although SWAT is originally developed to quantify the effect of land management practices in agricultural catchments, it has also been frequently applied in hydrological modeling researches in the mountainous regions (Ahl et al., 2008; Grusson et al. 2015; Rahman et al., 2014; Troin and Caya, 2014; Wagner et al., 2012). Because SWAT's catchment scale based structure matches the current data availability of mountainous catchments, and its diverse modules such as nutrient transport and water management help to predict the local water quality dynamics and to explore the effects of different driving forces such as water use on the regional hydrological system. However, the aforementioned key factors (precipitation, snow processes and hydropower plant operation) play decisive roles in affecting the performance and reliability of the SWAT modeling applications (Debele et al. 2010; Rahman et al., 2013; Shen et al., 2012; Shope and Maharjan, 2015; Troin and Caya, 2014; Zhang et al., 2011). Handling them properly during the modeling applications has fundamental implications for improving the SWAT model performance in mountainous catchments, with the necessity and feasibility discussed below:

1) Reliable precipitation data are important inputs for SWAT (Galván et al., 2014; Monteiro et al., 2016; Strauch et al., 2012). Since SWAT only uses data from one precipitation gauging station that is nearest to the centroid of each subbasin (Galván et al., 2014; Masih et al., 2011), the sparse and heterogeneous spatial distribution of rain gauges at mountainous regions often results in

inaccurate/unrepresentative precipitation inputs for SWAT. Without representative precipitation for the subbasins, the hydrological simulations of SWAT cannot be accurate and are highly uncertain. Therefore, evaluating and taking use of different available precipitation inputs which consider regional spatial variations could be a useful way to increase the reliability of SWAT modeling results (Tobin and Bennett, 2009). Evaluation of precipitations inputs for SWAT has been attempted (Shen et al., 2012; Shope and Maharjan, 2015; Vu et al., 2012; Wagner et al., 2012; Zhang and Srinivasan, 2009), but only a few case studies have investigated mountainous catchments with high elevation gradients (Shope and Maharjan, 2015; Xu et al., 2010). Besides, the choice of the best precipitation input data for SWAT models is basin-specific and unusual in Alpine region. More importantly, the uncertainties propagated from precipitation inputs to model parameters and outputs are rarely evaluated. A systematic investigation considering different available precipitation datasets for SWAT model in Alpine catchments is useful.

2) The influence of snow processes on the hydrological cycle of mountainous catchments has been frequently evaluated by SWAT models (e.g. Ahl et al. 2008, Debele et al. 2010, Grusson et al. 2015, Rostamian et al. 2008, Troin and Caya 2014). Most of the previous studies estimated snow dynamics with large model uncertainty, mainly owing to snow data scarcity and the lack of utilization of snow data during model calibration. In fact, SWAT adopts a specific approach to describe snow processes: the same Snow Water Equivalent (SWE) value is assigned to all HRUs in a subbasin, and this value is derived from the weighted sum of the SWE computed for each elevation band in a subbasin. Therefore, output time series of SWE are available at two spatial scales: for the entire subbasin (the average subbasin SWE) and for each elevation band. Only a few studies have taken these two snow outputs into account to evaluate the performance of the SWAT snow module (Fu et al., 2014; Fu et al., 2015; Grusson et al., 2015; Pradhanang et al., 2011) and no study has included the snow information of subbasin and elevation band scales into model calibration processes. Most of the previous studies assigned a unique set of snow parameters to an entire basin (e.g., Fu et al. 2014, Fu et al. 2015, Rahman et al. 2013), although snow distribution can be highly heterogeneous (Egli and Jonas 2009). More importantly, in snow-dominated catchments, SWAT model parameters are generally estimated using discharge records alone (e.g., Levesque et al., 2008; Rahman et al., 2013; Wang and Melesse, 2005), but snow data, which are very important information in snow-dominated catchments to represent snow dynamics, have rarely been used for calibration, ascribed to the typical lack of SWE time series. Because river discharge is the combination of the contributions of different hydrological components, calibration based on discharge alone can easily lead to the problem of model equifinality (Beven, 2006; Hanzer et al., 2016). Therefore, snow based calibration would be

helpful to get representative snow parameters according to the real snow information. Moreover, a multi-objective model calibration, involving river discharge observation and available snow information of different spatial scales that fit the specific snow outputs of SWAT, could be a promising approach to increase the reliability of SWAT modeling results in mountainous region.

3) SWAT model has been successfully applied in several case studies to reproduce the hydrograph impacted by hydropower plant operation (e.g., Zhang et al., 2011; Rahman et al., 2013; Rahman et al., 2014), However, data availability is one of the main challenges in reproducing streamflow when considering reservoir operation for hydropower production (Gaudard et al., 2014), because SWAT needs detailed operation data, such as inflow/outflow from the dams, to reproduce the impact such as hydropeaking. Unfortunately, as discussed above, reliable operation data are normally not accessible. Energy price is an alternative available dataset closely related to hydropower plant operation, but it cannot be used by SWAT model. Without considering the operation information, the simulations of SWAT are not reliable. Therefore, involving the relevant data (i.e. energy price and the day of the week) could be helpful to reduce the uncertainty of the simulations and improve the SWAT model results. Instead of complicate source code modification, a simpler and feasible way is to couple SWAT with machine learning methods which have been highlighted as a valuable and accurate tool for modeling complex river basin systems in support of water management information demands (Karamouz et al., 2009; Khalil et al., 2005; Ticlavilca and McKee, 2011).

In this work, we aimed to improve the SWAT modeling applications in large scale Alpine catchments, based on the consideration of both model structure and available data. The Adige River Basin located in North-East Italy was selected as a study area. As a feasible solution, the concept is to utilize the important data available at hand in a suitable way that matches the model structure so as to reduce the model uncertainty and hence increase the reliability of the model results. The major uncertainty sources affecting SWAT model results are 1) input data, 2) model parameters, 3) model structure (Abbaspour et al., 2007; Nepal et al., 2014). To clarify, standard methods are available to perform uncertainty analysis for modeling approaches (e.g. GLUE (Beven and Binley, 1992)), but we are not aiming at implementing statistical uncertainty analysis. What we concern is about how to improve the practical SWAT model application under the limitation of model structure and available data. Based on their important impacts on alpine hydrology aforementioned, we consider precipitation, snow processes and hydropower plant operation as the major factors which complicate hydrological modeling in alpine catchment and increase uncertainty of model results. The specific research questions are:

1) How can we select the suitable precipitation input that leads to less uncertainty for SWAT modeling in an alpine catchment?

2) How can we take use of the snow data in a reasonable way that matches the model structure of SWAT to improve the reliability of the results?

3) How can we take use of relevant data to improve the streamflow simulations affected by hydropower plant operations?

To address these questions, this work explored the possibility of improving reliability of SWAT simulations in alpine applications by: selecting the suitable precipitation dataset considering the uncertainty propagated from precipitation inputs to model parameters and outputs; involving snow data for calibration to constrain the parameters; and introducing an additional tool to take into account hydropower relevant data so as to fill the model structure gap.

In chapters 2 and 3, we answered the research question 1). First, due to their advantage of availability, eight remoted sensed precipitation data were collected and compared with ground observations in daily, monthly and yearly scale. As results, the remote sensed precipitation products that are closest to the ground observations are selected as inputs for SWAT models. Furthermore, SWAT models were set up with different precipitation inputs (observation of rain gauges, IDW (inverse distance weighting) data and the remote sensed products). The model performances were evaluated and compared in terms of output uncertainty, parameter uncertainty and performances of the best simulations. In this way, we aimed at improving the SWAT applications at the studied Alpine catchments by selecting the best precipitation input that results in both good reproduction of river discharge and less uncertainty of parameter and output.

In chapter 4 and 5, snow processes were investigated to address the research question 2). Considering the SWAT model structure that produces SWE outputs of subbasin and elevation band scales, we computed two types of SWE data based on snow depth observations: the mean subbasin scale SWE and the SWE of each station. To calculate the subbasin scale SWE values, daily snow depth measurements were first interpolated using a simple power law model to establish a correlation between snow depth and elevation. The interpolated values were then used to compute a mean subbasin snow depth, and finally these values were converted into SWE data using the equation suggested by Pistocchi (2016). For calculating the SWE for each station, snow depth observations were multiplied by temporal evolution of snow density (Pistocchi, 2016). These SWE data were applied in snow model calibration of SWAT first to evaluate the feasibility of such snow data based procedure. Furthermore, the SWE data were utilized together with discharge data as a multi-objective calibration approach to constrain the snow

parameters, which was compared with the common used discharge-based calibration approach. The performances of discharge, subbasin averaged SWE, and elevation band scale SWE simulations were evaluated and compared, as well as the simulated major water components including precipitation, snowmelt, and streamflow. We aimed to investigate the possibility of increasing model result reliability by constraining parameters according to both snow data and discharge information.

In chapter 6, we focused on improving the simulation of streamflow under the impact of hydropower plant operation, which addressed the research question 3). The streamflow affected by hydropower plant operation was firstly simulated by SWAT model. However, the simulated streamflow time series was highly uncertain, because the operation input required by SWAT is not available and other relevant information cannot be considered by SWAT due to the lack of certain components in the model structure. To fill this gap of SWAT model, Support Vector Machine (SVM) tool was introduced to reproduce hydropeaking based on the SWAT modeling results and the available information (energy price and the day of the week) closely related to operation of hydropower plant. By coupling SVM and SWAT, we aimed to improve simulation of the streamflow affected by hydropower plant operation, when specific data required by SWAT are missing but additional relevant data that cannot be handled by SWAT are available.

Chapter 2

Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales[1]

This study provides a comprehensive evaluation of eight high spatial resolution gridded precipitation products in Adige Basin located in Italy within 45-47.1°N. The Adige Basin is characterized by a complex topography, and independent ground data are available from a network of 101 rain gauges during 2000-2010. The eight products include the Version 7 TRMM (Tropical Rainfall Measuring Mission) Multi-satellite Precipitation Analysis 3B42 product, three products from CMORPH (the Climate Prediction Center MORPHing technique), i.e., CMORPH_RAW, CMORPH_CRT and CMORPH_BLD, PCDR (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record), PGF (Global Meteorological Forcing Dataset for land surface modeling developed by Princeton University), CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) and GSMaP_MVK (Global Satellite Mapping of Precipitation project Moving Vector with Kalman-filter product). All eight products are evaluated against interpolated rain gauge data at the common 0.25° spatial resolution, and additional evaluations at native finer spatial resolution are conducted for CHIRPS (0.05°) and GSMaP_MVK (0.10°). Evaluation is performed at multiple temporal (daily, monthly and annual) and spatial scales (grid and watershed). Evaluation results show that in terms of overall statistical metrics the CHIRPS, TRMM and CMORPH_BLD comparably ranks as the top three best performing products, while the PGF performs worst. All eight products underestimate and overestimate the occurrence frequency of daily precipitation for some intensity ranges. All products tend to show higher error in the winter months (December-February) when precipitation is low. Very slight difference can be observed in the evaluation metrics and aspects between at the aggregated 0.25° spatial resolution and at the native finer resolutions (0.05°) for CHIRPS and (0.10°) for GSMaP_MVK products. This study has implications for precipitation product development and the global view of the performance of various precipitation products, and provides valuable guidance when choosing alternative precipitation data for local community.

[1] Duan Z., Liu J., Tuo Y., Chiogna G., Disse M., 2016. Evaluation of eight spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. Science of The Total Environment, 573: 1536-1553. DOI: 10.1016/j.scitotenv.2016.08.213

2.1 Introduction

Precipitation is a major component of the hydrological cycle and is a crucial forcing data for many applications such as hydrology, weather forecasts, meteorology and agriculture (Ebert et al., 2007). Precipitation is characterized by a significant variability both in time and space, and accurate precipitation data at high spatial and temporal resolution is highly desirable (Duan and Bastiaanssen, 2013). Conventional observations from rain gauge stations can generally provide the most direct and most accurate measurements at the gauge locations. As a drawback, the relatively sparse distribution of rain gauges often leads to poor spatial representation of precipitation patterns (Javanmard et al., 2010). Besides, due to the frequent occurrence of gaps and inhomogeneities in precipitation time series collected at rain gauges (Brunetti et al., 2006), it is difficult to force hydrological models that require a continuous time-series of precipitation data.

In contrast, satellite remote sensing data offers a new way in detecting spatial and temporal variability of precipitation at high spatial and temporal resolutions (Xie and Xiong, 2011). Regardless of the processing methods used, satellite-based precipitation estimates use either mainly the infrared (IR) information frequently measured from geostationary satellites, the microwave (MV) information less frequently measured from low earth orbiting satellites, or a combination of IR and MV information. Overviews of principles and various techniques in use for satellite-based precipitation estimation can be found in the review by Kidd and Huffman (2011). A brief introduction is given here as follows: In general, the IR information can be used to derive the cloud-top information which is used to estimate precipitation using the established relationship between cloud-top temperature and rainfall. One typical algorithm using IR information for estimation of precipitation is detailed in Arkin and Meisner (1987). The MV information is able to present more information about cloud through different layers and is more physically related to precipitation, and different retrieval algorithms (e.g. emission-based, scattering-based) can be applied to convert the MV information to precipitation estimates (Ferraro, 1997; Wilheit et al., 1994). Precipitation estimates from IR (frequent sampling and at high resolution, but less accurate) and MV (more accurate, but at coarse resolution and infrequent sampling) are often combined to complement each other to achieve a better estimate at high temporal and spatial resolution and coverage. For this purpose, various combination techniques have been developed, but they can be essentially classified into two categories: (1) using the MV-based precipitation estimates to calibrate the IR-based precipitation estimates; (2) using the derived cloud motion from IR data to propagate MV-based precipitation estimates (Kidd and Huffman, 2011). Evaluation studies showed the satellite precipitation estimates can still contain substantial biases and errors, and a further merging or blending satellite precipitation estimates with rain gauge data can result in improved precipitation products (Ebert et al., 2007; Xie and Xiong, 2011).

10

Over the past decades, great efforts have been made to generate gridded precipitation products, thereby leading to the increasing availability of precipitation datasets at different spatial and temporal resolutions over the global or quasi-global scale (Tapiador et al., 2012). They can be broadly classified into four categories: (1) gauge-only products that build only on observations from rain gauge stations using different interpolation methods, these widely used products for example include the Global Precipitation Climatology Centre (GPCC) monthly precipitation product (Schneider et al., 2014; 2015), the Climatic Research Unit (CRU) monthly precipitation (Harris et al., 2014) and the Climate Prediction Center (CPC) unified gauge-based analysis of global daily precipitation (Chen et al., 2008). These products are often available at a coarser spatial resolution than 0.5°; (2) precipitation products from numerical weather predictions or atmospheric models that uses a combination of satellite and in-situ observations of various atmospheric properties as inputs (Ebert et al., 2007), these products include for example reanalysis products from National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) (Kalnay et al., 1996) and European Centre for Medium-Range Weather Forecasts (ECMWF) (Balsamo et al., 2015); (3) satellite-only products that builds on using IR, MV or IR-MV combined information; (4) satellite-gauge products that combine two individual (gauge-only and satellite-only) products together through different bias correction or blending procedures. Products belonging to the latter two categories include for example the TRMM (Tropical Rainfall Measuring Mission) multi-satellite precipitation analysis (TMPA) (Huffman et al., 2007), the CMORPH (CPC MORPhing technique) (Joyce et al., 2004), PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks) (Hsu et al., 1997), and such products are often available at the spatial resolution of 0.25° or finer. It is particularly worth noting that the recently released "satellite-gauge" type CHIRPS product (Climate Hazards Group InfraRed Precipitation with Station data) provides precipitation at the finest spatial resolution of 0.05° (Funk et al., 2015). Since merged precipitation products generally have a more improved quality than individual dataset (Xie and Xiong, 2011), different merged precipitation products (reanalysis-gauge and satellite-gauge) are increasingly being developed.

The available gridded precipitation products have global (or quasi-global) orientation, and their performance could vary from region to region. Evaluation of these precipitation products with reliable and independent measurements is important for both product developers and users. For developers, evaluation results provide beneficial feedback which enables them to identify the problems and direction for development of improved products; for local users, evaluation results enable users to take into account the possible uncertainty associated with a particular product and make the appropriate choice of a suitable product for specific applications in a given study area. Recognizing the importance of product evaluation, many studies have been carried out to evaluate a single or multiple precipitation products in

scales from the global or quasi-global (Yong et al., 2015), continental (Awange et al., 2016; Negrón Juárez et al., 2009), country-wide (Prakash et al., 2015; Shen et al., 2010; Tan et al., 2015; Tian et al., 2010), regional (Duan et al., 2012, Khandu et al. 2015), and basin scales (Liu et al., 2015; Yong et al., 2010). A few evaluation studies have been conducted in Italy. For example, Lo Conti et al. (2014) evaluated six satellite precipitation products (CMORPH, two products of TMPA and three products of PERSIANN) in the island of Sicily. Nikolopoulos et al. (2013) evaluated three products (TRMM 3B42, CMORPH and PERSIANN) for a single major flood event in the Friuli-Venezia Giulia region. Mei et al. (2014) evaluated four products (two products of TMPA, CMORPH and PERSIANN) in the upper Adige Basin with focus on only the heavy precipitation events, and evaluation was made only at the basin-average scale. Compared to Mei et al. (2014), our current study covers almost the whole Adige Basin and evaluates eight precipitation products at a range of spatial and temporal scales, thereby providing new insights into the overall performance of various precipitation products in the Adige Basin.

Most available satellite-only and satellite-gauge gridded precipitation products have the quasi-global coverage (latitude bands of 50° N-S or 60° N-S). More evaluation of precipitation products in the high latitude and elevation areas are particularly needed to gain more insights into the global view of the performance of various precipitation products (Yong et al., 2010). Therefore, the objective of this study is to evaluate eight high resolution gridded precipitation products in the Adige Basin, Italy with dense rain gauge data. The Adige Basin is located with 45-47.1° N latitude and characterized by complex topography, which serves as a representative and a good testing site. The evaluated eight precipitation products are Version 7 TRMM 3B42, three different products of CMORPH, PERSIANN-CDR (Climate Data Record) (Ashouri et al., 2015), GsMaP-MVK (Global Satellite Mapping of Precipitation Moving Vector with Kalman-filter product) (Ushio et al., 2009), PGF (the Princeton Global Forcings) (Sheffield et al., 2006) and CHIRPS. This study conducted the comprehensive evaluation of these precipitation products at various spatial resolutions (0.25° all eight products, and additional original finest resolution of 0.10° for GsMaP_MVK and 0.05° for CHIRPS) and temporal resolutions (daily, monthly and yearly) for the period 2000-2010. The eight products represent a wide range of advanced retrieval algorithms and data availability at high spatial resolution. Each of them has been found with good performance in some other study areas (e.g. Maggioni et al., 2016; Tan et al., 2015). It is therefore relevant and interesting to evaluate these products in Adige Basin. We aim at identifying the products with best performance that can be merged with available rain gauge data to achieve a better quality than individual dataset. Such improved precipitation dataset is essential for many ongoing research and operational applications in Adige Basin such as hydrological modeling, sediment and pollutant transport modeling, and water resources assessment and management (Chiogna et al., 2016). It is particularly worth noting that because of recent (2015) availability of the CHIRPS product, very few studies have been done to

evaluate and compare this product against other products. Katsanos et al. (2016) compared only the correction between the CHIRPS product and rain gauge data over Cyprus at monthly and annual temporal scales and only one spatial scale, and they found good correction between the two datasets. In this study, we extend the evaluation of the CHIRPS product besides other seven products to the Adige Basin for the first time in a more comprehensive way in terms of evaluated aspects and temporal and spatial scales.

The remainder of this paper is organized as follows: Section 2.2 introduces the study area and provides a brief description of the eight evaluated gridded precipitation products and used rain gauge station data. The evaluation methods and statistical metrics are also presented in Section 2.2. Section 2.3 presents the evaluation results and discussion. Finally, the conclusions drawn are summarized in Section 2.4.

2.2 Materials and methods

2.2.1   Study area

The Adige Basin is located in the northern part of Italy within the latitude of 45°-47.1° N and longitude of 10.2°-12.5° E (Figure 2.1). The Adige is the second longest river in Italy, with a length of 410 km and a drainage area of 12000 km^2 and is one of the six study sites investigated in the FP7 GLOBAQUA Project (Navarro-Ortega et al., 2015). The Adige Basin has a complex topography with elevation ranging from -7 to 3865 m above the mean sea level. The north part of Adige Basin belongs to the central and southern Alps. Climate in the Adige Basin is characterized by dry winter, snowmelt in the spring and humid summer and fall. Because of its morphology and humid climate, the river basin is well suited for hydroelectric production, and to date 30 major reservoirs exist in the catchment, with a total storage capacity of 571 10^6 m^3 (8.5% of the long-term mean annual runoff). Streamflow is severely impacted by hydropower production, particularly at intermediate and low flow regimes (Zolezzi et al., 2009). Earlier snow melting is already affecting the Adige Basin reducing water resources availability during the irrigation period (roughly June-August), while the higher temperature recorded in the summer months in the last decades is expected to cause an increase of water demand in this period. This is expected to increase the deficit of water resources in summer, when agricultural and recreational uses reach the highest demand, thereby exacerbating the conflict between different uses of water resources (Majone et al., 2016). More details about Adige Basin and its hydrological, chemical and ecological status can be found in (Chiogna et al., 2016).
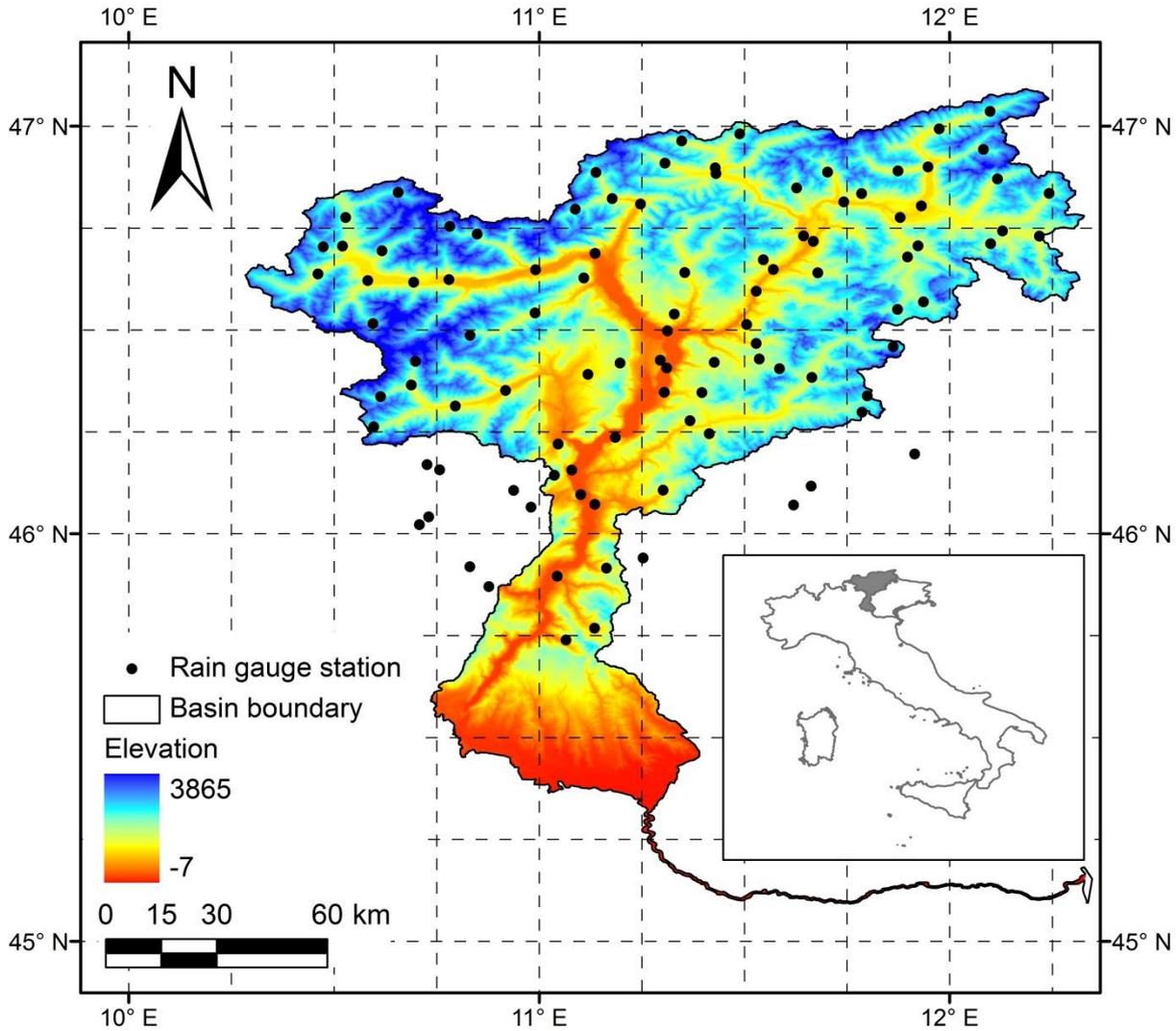
Figure 2.1 Locations of the Adige Basin and rain gauge stations. The dashed rectangle represents the 0.25° x 0.25° grid.

2.2.2   Datasets and preprocessing

This section briefly describes the eight high resolution gridded precipitation products evaluated in this study and the reference rain gauge data available. Table 2.1 summarizes the spatial and temporal resolutions, available periods and coverage of the eight products.

Table 2.1 Summary of eight gridded precipitation products to be evaluated in this study.

Dataset	Period	Spatial resolution	Finest temporal resolution	Coverage	Category
TRMM 3B42	1998-present	0.25°	3-hourly	50°N-50°S	Satellite-gauge
CMORPH_RAW	1998-present	0.25°	3-hourly	50°N-50°S	Satellite-only

CMORPH_CRT	1998-2015	0.25°	3-hourly	50°N-50°S	Satellite-gauge
CMORPH_BLD	1998-2015	0.25°	Daily	50°N-50°S	Satellite-gauge
PCDR	1983-present	0.25°	Daily	60°N-60°S	Satellite-gauge
PGF	1948-2010	0.25°	3-hourly	90°N-90°S	Reanalysis-gauge
CHIRPS	1981-present	0.05°	Daily	50°N-50°S	Satellite-gauge
GSMaP_MVK	2000-present	0.10°	1-hourly	60°N-60°S	Satellite-only

2.2.2.1 TRMM 3B42

The TRMM 3B42 product is one type of the TMPA (TRMM Multi-satellite Precipitation Analysis) products (Huffman et al., 2007). TRMM 3B42 product provides 3-hourly precipitation at a spatial resolution of 0.25° for the quasi-global coverage of 50° N-50° S from 1998 to present. The latest product version is Version 7 and the applied algorithm is the TMPA algorithm that combines precipitation estimates from microwave and infrared satellites, as well as the GPCC monthly gauge analysis. Passive microwave data from a number of satellites are collected and converted to precipitation estimates at 3-hourly time scale using the Goddard Profiling Algorithm (Kummerow et al., 2001). The microwave precipitation estimates are combined and calibrated to the TRMM Combined Instrument (TCI) estimate (TRMM product 2B31). The infrared-based precipitation estimates from multiple geostationary satellites are computed through histogram matching of monthly microwave precipitation estimates. Then the microwave and the infrared-based precipitation estimates are merged with infrared-based estimates being used only to fill in the missing data where microwave estimates are not available. The merged 3-hourly precipitation estimates are then added up to monthly totals and further combined with GPCC monthly rain gauge analysis products using the inverse error variance weighting to generate the monthly best precipitation product (TRMM 3B43). Finally, all the 3-hourly precipitation estimates are adjusted for each month to make their sums equal to the TRMM 3B43 monthly value. The adjusted precipitation time series are the final TRMM 3B42 product. More details about TMPA algorithms can be found in Huffman et al. (2007) and Huffman and Bolvin (2015). All TRMM products including 3B42 are available from Goddard Earth Sciences Data and Information Services Center at http://mirador.gsfc.nasa.gov. Two types TRMM 3B42 product are available: 3-hourly and daily accumulated products. The daily accumulated TRMM 3B42 product for the period 2000-2010 were used in this study. The TRMM 3B42 product belongs to the "satellite-gauge" category.

2.2.2.2 Three different products from CMORPH

The CMORPH product stands for the precipitation product estimated by the NOAA Climate Prediction Center MORPHing technique. This technique is a flexible method to combine existing passive microwave-based precipitation estimates from multiple low orbit satellites and the infrared data from multiple geostationary satellites (Joyce et al., 2004). This technique uses the infrared data to firstly derive the cloud system advection vectors (CSAVs). The CSAVs are then used to propagate the passive microwave-based precipitation estimates in two directions (forward and backward) in time for periods when passive microwave data are not available at a location. The two propagated precipitation estimates are finally computed using a time-weighting. More details about CMORPH technique can be found in Joyce et al. (2004). Initially, no bias correction and no rain gauge data were used in the CMORPH technique for previous Version 0.x product. The latest (Version 1.0) CMORPH products include three different products: the raw satellite-only precipitation product (CMORPH_RAW), bias corrected product (CMORPH_CRT) and satellite-gauge blended product (CMORPH_BLD). The CRT product is generated through adjusting the RAW product against the CPC unified daily gauge analysis over land and the pentad GPCP over ocean using the probability density function (PDF) matching bias correction method (Xie et al., 2011). The CRT product is further combined with the gauge analysis through an optimal interpolation (OI) technique to generate the BLD product (Xie et al., 2013). The RAW product belongs to the "satellite-only" category while the CRT and BLD products belong to the "satellite-gauge" category. All the three products provide precipitation for the quasi-global coverage of 60° N-60° S from 1998 to a slightly different ending period (2016 for RAW, 2015 for both CRT and BLD). For RAW and CRT products, three combinations of spatial-temporal resolutions are available, that is, 8 km-30 min, 0.25°-3 hourly, and 0.25°-daily. The BLD product is available only at the 0.25°-daily resolution, and the definition of a day (daily ending time) is different from country to country due to the limitation in the input gauge data. All three products are available at ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/. By design, the BLD is considered as the product with the best quality followed by CRT and RAW. In this study, all the three products at the 0.25°-daily spatial-temporal resolution for the period 2000-2010 were evaluated and compared.

2.2.2.3 PERSIANN-CDR (PCDR)

The PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks) Precipitation Climate Data Record (PCDR) provides daily precipitation estimates at the spatial resolution of 0.25° for the quasi-global coverage of 60° N-60° S from 1983 to present. The product is available at ftp://eclipse.ncdc.noaa.gov/pub/cdr/persiann/files/. The construction of PCDR uses mainly the stage IV radar data, gridded satellite infrared data (GridSat-B1), the GPCP monthly

precipitation data at 2.5° degree. The main algorithm is the PERSIANN algorithm that uses an artificial neural network (ANN) model to convert the infrared information into rain rates in which the ANN model is trained using concurrent infrared-based and passive microwave-based precipitation estimates from multiple satellites (Hsu et al., 1997). The trained model parameters using the stage IV radar data are kept fixed and then used to estimate precipitation (called as PERSIANN-B1) using the GridSat-B1 data as inputs. The estimated PERSIANN-B1 data are further adjusted through a bias removal procedure using the GPCP 2.5° monthly precipitation data to generate the PCDR product. Therefore, the PCDR maintains the consistency with the GPCP precipitation at the monthly and 2.5° scales. More details about this product can be found in (Ashouri et al., 2015). Because the precipitation gauge data product from the Global Precipitation Climatology Centre (GPCC) are used as inputs to the development of GPCP monthly precipitation (Adler et al., 2003), the PCDR product belongs to the "satellite-gauge" precipitation estimate category. In this study, the PCDR daily precipitation product at the 0.25°spatial resolution for the period 2000-2010 were used.

2.2.2.4 Princeton Global Forcings product (PGF)

The Princeton Global Forcings (PGF) is the short name of the Global Meteorological Forcing Dataset for land surface modeling developed by Princeton University. Details about the development of this product are presented in Sheffield et al. (2006). This product provides precipitation and other meteorological variables including surface air temperature, radiation, surface air pressure, specific humidity, wind speed for the global coverage at the spatial resolutions of 0.25°, 0.5° and 1.0° and temporal resolutions of 3-hourly, daily and monthly for the period 1948-2010. An updated version (Version 2) of PGF product is currently available for an extended time period 1901-2012, but the finest spatial resolution is at 0.5°. The PGF product is available at http://hydrology.princeton.edu/data.pgf.php. The PGF product is constructed by merging the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (referred to as the NCEP reanalysis hereafter) and a suite of global observation-based datasets. For the construction of the final PGF precipitation product, four datasets NCEP reanalysis, the Climatic Research Unit (CRU) monthly precipitation data, GPCP daily precipitation data, and TRMM 3-hourly real-time data are used in a number of processing procedures including correction of rain day anomaly, spatial and temporal downscaling and monthly bias correction to make the monthly totals of PGF precipitation match those of the CRU data (Sheffield et al., 2006). Since CRU monthly precipitation data are based on analysis of rain gauge data (Harris et al., 2014), thus PGF precipitation product belongs to "reanalysis-gauge" category. It should be noted that the use of satellite product (TRMM) in the generation of the PGF product make it inevitably vulnerable to all uncertainties that could affect satellite products. The nature of longest time period, the high resolution and global coverage clearly makes PGF product have many

potential applications, thus it is interesting to evaluate the performance of PGF and compare with other products. In this study, the PGF daily precipitation data at the spatial resolution of 0.25° for the period 2000-2010 were used.

2.2.2.5 CHIRPS

The CHIRPS product stands for the Climate Hazards Group InfraRed Precipitation with Station data. The CHIRPS product provides daily precipitation data at spatial resolution of 0.05° for the quasi-global coverage of 50°N-50°S from 1981 to present. The latest product is the Version 2.0 product that was released in February 2015. The CHIRPS product and its supporting data are available at: http://chg.geog.ucsb.edu/data/chirps/. The main used datasets for the construction of CHIRPS product include the monthly precipitation climatology (CHPclim) that is created using rain gauge stations collected from FAO and GHCN, the Cold Cloud Duration (CCD) information based on thermal infrared data archived from CPC and NOAA National Climate Data Center (NCDC), the Version 7 TRMM 3B42 data, the Version 2 atmospheric model rainfall field from the NOAA Climate Forecast System (CFS), and the rain gauge stations data from multiple sources. First, the CCD data are calibrated with TRMM 3B42 to generate the 5-daily CCD-based precipitation estimates which are further converted to the fractions of the long-term mean precipitation estimates. The fractions are then multiplied with CHPclim data to remove the systematic bias and the derived product is called CHIRP product. Finally, the CHIRP product is blended with rain gauge stations data using a modified inverse distance weighting algorithm to produce the CHIRPS. All the processing mentioned above are performed at the 5-daily timescales. The daily CCD data and daily CFS data are finally used to disaggregate the 5-daily products to daily precipitation estimates using a simple redistribution method. More detailed information on CHIRPS can be found in Funk et al. (2015). CHIRPS product belongs to the "satellite-gauge" category. Daily CHIRPS products at the spatial resolution of 0.05° and 0.25° are available from the downloading website mentioned above. CHIRPS products at both 0.05° and 0.25° for the period 2000-2010 were used and evaluated in this study, which enables us to investigate the effect of spatial resolutions on the performance of the products. The evaluate results were referred to as "CHIRPS005" and "CHIRPS025", respectively.

2.2.2.6 GSMap_MVK product

The GSMaP refers to the Global Satellite Mapping of Precipitation project that aims to develop precise high resolution global precipitation product from passive microwave and infrared satellites. The GSMaP_MVK product stands for the GSMaP Moving Vector with Kalman-filter product. The GSMaP_MVK product provides hourly precipitation at the spatial resolution of 0.10° for the quasi-global coverage of 60° N-60° S. At the time of writing, the GSMaP_MVK product is at the transition of

two versions, the Version 5 product covers the period from March 2000 to November 2010 and the latest Version 6 product starts from March 2014 to present. The GSMaP_MVK product at hourly and daily timescales and others are available at http://sharaku.eorc.jaxa.jp/GSMaP/. The creation of GSMaP_MVK product uses two main data sources: the infrared data archived by CPC from multiple geostationary satellites, and the passive microwave-based precipitation estimates from multiple satellite sensors. The GSMaP_MVK uses the hourly infrared data to compute the moving vectors using the same morphing technique as in CMORPH. The derived vectors are then used to propagate forward and backward the passive microwave-based precipitation estimates, and the propagate estimates are further refined using the Kalman filter based on the relationship between infrared brightness temperature and surface precipitation rate. The two propagated precipitation estimates are then weight averaged to generate the final precipitation estimate for the area and time where passive microwave data are not available. More details on GSMaP_MVK can be found in Ushio et al. (2009). The GSMaP_MVK belongs to the "satellite-only" category. The daily accumulated (00Z-23Z) GSMaP-MVK product at the spatial resolution of 0.10° for the period 2000-2010 were used and evaluated in this study, the evaluated result at this 0.10° resolution was referred to as "GSMaP_MKV010". The 0.10° daily GSMaP_MVK product was aggregated by pixel averaging (simply the area-weighted integral of all 0.10° grids that intersect a targeted 0.25° grid) to the same spatial resolution of 0.25° for inter-comparison with other products in this study, and the evaluated result was referred to as "GSMaP_MKV025".

2.2.2.7 Rain gauge data

A relatively dense network of rain gauge stations are maintained by the meteorological surveys of the Autonomous Province of Trento (data available at www.meteotrentino.it) and of the Autonomous Province of Bolzano (data available at http://www.provincia.bz.it/meteo/home.as) in the Adige Basin. The daily measured rainfall data from a total number of 101 stations with complete temporal coverage for the period 2000-2010 were finally used in this study. The rain gauge data have gone through the quality control checks. Distributions of these 101 stations are shown in Figure 2.1. Rain gauges are particularly sparse over the southern part near the outlet of Adige Basin. This highlights the need for reliable alternative precipitation datasets to enhance understanding of water-related aspects of the whole river basin, which is one motivation of this current study. As described above, most gridded precipitation products to be evaluated incorporate the rain gauge data from a variety of sources and use different algorithms, namely the GPCC monthly rain gauge analysis (used for TRMM 3B42, PCDR), CPC unified daily gauge analysis (used for CMORPH_CRT and BLD), and CUR data (used for PGF) and rain gauge data from multiple sources used for CHIRPS. In order to achieve an independent evaluation, it is important to ensure that rain gauge stations used as ground truth should not be used in the creation of the products to be evaluated. The rain gauge-based data mentioned above provide the number of used

stations for each grid, although the locations of used rain gauges are usually not available. We have checked the number of rain gauge stations within each grid in the respective rain gauge-based data (at their finest spatial resolutions) used in the seven precipitation products. We found that for grids covering the Adige Basin for the period 2000-2010, on average the number of used rain gauges is around 10, 7, 1 and 3 for GPCC, CPC, CRU and CHIRPS, respectively. Therefore, more than 90% (at least) of our 101 stations are not used in the creation of each of the eight precipitation products to be evaluated, justifying the independent evaluation in this study.

2.2.3   Evaluation methods

There is a typical scale mismatch issue between point-based rain gauge data and the gridded precipitation products. The common practice in the evaluation studies is to upscale the point-based rain gauge data to the same grid scale with the precipitation products. Many interpolation techniques have been used to achieve this upscaling, such as simple algorithmic averaging (Xie and Xiong, 2011), the Thiessen polygon (Liu et al., 2015), the inverse distance weighting (IDW) (Hu et al., 2014; Yong et al., 2010), and the Kriging method (Khandu et al., 2015). Each interpolation method has its advantages and disadvantages, and its performance depends on various factors (Hofstra et al., 2008) and also varies from region to region. It is therefore practically impossible to identify an optimal method applicable in all study areas. A detailed comparison of different interpolation methods for interpolation of precipitation can be found for example in (Hofstra et al., 2008). It is interesting to investigate the effect of the different interpolation methods used in upscaling rain gauge data on the evaluation of gridded precipitation products; this is a topic that remains not well understood and is one of our ongoing studies. In this study, the IDW method was used because it is relatively easy to implement and its popular application in evaluation studies (Hu et al., 2014; Yong et al., 2010). The daily rain gauge data was first interpolated by IDW to the finer 1 km grid scale which is then aggregated to the larger grid scales (0.25°, 0.10° 0.05°) through spatial pixel averaging following (Hu et al., 2014). The interpolated 0.25° gridded precipitation data from rain gauge data were used as ground truth to evaluate all gridded precipitation products at the common 0.25° grid, while the interpolated gridded precipitation data from rain gauge data at 0.10° and 0.05° were used to evaluate the GSMaP_MVK and CHIRPS at their original finer grid scale, respectively. For evaluation at all three grid scales (0.25°, 0.10° and 0.05°), only grids containing at least one rain gauge were considered. We evaluated the performance of selected precipitation products at three temporal scales (daily, monthly, and annual). For each temporal scale, evaluation was conducted at two spatial scales: the grid scale and the watershed scale. It should be noted that the averaged gridded precipitation (from either interpolated from rain gauge data or gridded precipitation products) from all grids containing at least one rain gauge was considered to represent the watershed precipitation during the evaluation at watershed scale.

For evaluation purpose, eight commonly used statistical metrics were computed following (Liu et al., 2015; Tan et al., 2015; Yong et al., 2010). The equations and optimal values of the eight metrics are summarized in Table 2.2. The coefficient of determination (R^2) describes the degree of collinearity between rain gauge data and gridded precipitation product and also describes the proportion of the variance in rain gauge data explained by the gridded precipitation product. The four metrics describing the error and bias are the relative bias (BIAS), mean error (ME), mean absolute error (MAE) and root mean square error (RMSE). The last three are contingency table metrics: the probability of detection (POD), false alarm ratio (FAR) and critical success index (CSI). The POD describes the fraction of the observed precipitation events detected correctly by the evaluated product. The FAR describes the fraction of events detected by the product but not observed. The CSI describes the overall skill of the product relative to observed precipitation. Calculation of these three metrics is well explained in AghaKouchak and Mehran (2013). The occurrence frequency of precipitation with different intensities is an important feature and has significant effects on surface runoff and flooding modeling (Tian et al., 2007). Therefore, in this study the occurrence frequency for the daily precipitation was also computed to evaluate the probability of different rainfall intensities for rain gauge data and each gridded precipitation product. Daily precipitation were classified into seven categories with different rain intensity following the World Meteorological Organization (WMO) standard, namely 0-1; 1-2, 2-5, 5-10; 10-20, 20-50, and $\geq$ 50 mm/d (Tan et al., 2015).

Table 2.2 List of the statistical metrics used in the evaluation of precipitation products.

Statistical metric	Equation	Optimal value
Coefficient of determination (R^2)	$R^2 = \left(\dfrac{\sum_{i=1}^{n}(G_i - \overline{G})(P_i - \overline{P})}{\sqrt{\sum_{i=1}^{n}(G_i - \overline{G})^2}\sqrt{\sum_{i=1}^{n}(P_i - P)^2}} \right)^2$	1
Relative Bias (BIAS)	$BIAS = \dfrac{\sum_{i=1}^{n} P_i}{\sum_{i=1}^{n} G_i} - 1$	0
Mean Error (ME)	$ME = \dfrac{\sum_{i=1}^{n}(P_i - G_i)}{n}$	0
Mean Absolute Error (MAE)	$MAE = \dfrac{\sum_{i=1}^{n}\left|P_i - G_i\right|}{n}$	0
Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\dfrac{\sum_{i=1}^{n}(P_i - G_i)^2}{n}}$	0
Probability of Detection (POD)	$POD = \dfrac{H}{H + M}$	1

False Alarm Ratio (FAR)	$FAR = \dfrac{F}{H + F}$	0
Critical Success Index (CSI)	$CSI = \dfrac{H}{H + M + F}$	1

Notation: n refers to the number of samples; G_i means observed precipitation from rain gauge stations, P_i means the precipitation estimates from the evaluated products; H refers to the number of observed precipitation correctly detected by respective products; M means the number of the precipitation observed by rain gauge stations but not detected by products; F means the number of the precipitation detected by the products but not observed by rain gauge.

2.3 Results and discussion

2.3.1 Evaluation at daily scale

Scatterplots of daily precipitation from all selected gridded precipitation products against rain gauge data at the grid scale and watershed scale are shown in Figure 2.2 and Figure 2.3, respectively. Large scatters in Figure 2.2 and Figure 2.3 clearly show the poor agreements between all gridded precipitation products and rain gauge data for the daily precipitation. As described in Section 2.2, all eight gridded products are generated using multiple datasets and a number of different procedures for combination, blending and/or bias correction. Therefore, poor agreements between evaluated products and rain gauge data are attributed to several factors including satellite sampling error, errors in algorithms for estimation of precipitation from individual platforms (satellite, weather prediction model, rain gauge analysis), errors in algorithms for combing or blending individual estimates (Shen et al., 2010), errors in bias correction algorithms and the unrepresentative of gauge data used in the bias correction (the very limited number (less than 10) of rain gauges used in the relevant gauge analysis datasets for Adige Basin, see Section 2.2.7). Improvements in the error sources mentioned above are needed to adapt remote sensing precipitation data to the Adige Basin and increase the accuracy of all eight gridded precipitation products at the daily scale.

Table 2.3 summarizes the statistical metrics for daily precipitation. At the grid scale, the CMORPH_BLD product had the highest R^2 value (0.48), followed by TRMM and CHIRPS025 with the same value of 0.15. The lowest R^2 value of 0.01 was observed for PGF product. Overall, four products (CHIRPS, TRMM, PCDR, PGF) overestimated precipitation with a BIAS value ranging from 0.04 to 0.63, while all three different products from CMORPH and the GSMAP_MVK025 underestimated gauge precipitation with a BIAS ranging from -0.11 to -0.55. The results are consistent with other studies that have found similar overall overestimation by TRMM and underestimation by CMORPH across Europe (Maggioni et al. 2016). The CHIRPS025 showed the lowest ME of 0.09 mm/d. and lowest BIAS. The lowest MAE (1.86 mm/d) and RMSE (4.43 mm/d) values were achieved by the CMORPH_BLD

product. It is worth noting that among three different products from CMORPH, the RAW was consistently the worst performing product and improvements were achieved by CRT and further remarkably by BLD. This confirms the effectiveness of used bias correction and blending algorithms using daily rain gauge analysis by CPC. The PGF product had the worst performance in terms of all seven statistical metrics except the POD and CSI for which it can be ranked the second worst performing product.

All eight products had quite small FAR value ranging from 0.02 to 0.08, indicating that all products detected a very limited number of unrealistic precipitation events. The GSMaP_MVK025 showed the highest POD (0.67) and CSI (0.64), and moderate FAR (0.06), suggesting the better performance in capturing daily precipitation events than other products. Tian et al. (2007) also found that GSMaP_MVK025 had slightly higher POD than TRMM 3B42, PERSIANN and other evaluated products over the contiguous United States. They attributed this better performance to two aspects in the GSMaP_MVK algorithm: effective application of infrared interpolation technique and the rain/no-rain classification scheme in the microwave algorithm. However, it should be noted that GSMaP_MVK product had missing values for several days and grids, this resulted in 4% less data points than other products during evaluation which will add slight uncertainty for computation of statistical metrics. In addition, as shown in Figure 2.2, the GSMaP_MVK product had extremely high daily precipitation (> 200 mm/d) for several days (less than 10 days) during the whole period, while in reality the maximum daily precipitation from interpolated gauge data were not larger than 150 mm/d. The CHIRPS025 had the lowest POD (0.34) and CSI (0.33). Since none of the products is consistently the closest to perfect value for all eight metrics, it is not easy to conclude which product is the best performing one. The CMORPH_BLD product had the values that are closest to the perfect values in four metrics (R^2, MAE, RMSE and FAR), and this product perhaps could be considered as the most favorable product compared to the others. The generation of CMORPH products, indeed, used daily rain gauge analysis and all bias correction and blending algorithms were conducted at the daily scale directly, while all other products conducted bias correction with rain gauge analysis at the coarser temporal scale (5-daily or monthly) if any. Similar better performance in daily precipitation of bias-corrected CMORPH product than other products (TRMM 3B42 and PERSIANN) was reported in China by Liu et al. (2015). It is worth noting that after recognizing the large portion of precipitation with intensity of 0-1 mm/d (shown later in Figure 2.4), all products were evaluated also by considering only precipitation greater than 1 mm/d. We found that the overall metrics were slightly worse than those presented in Table 2.3 but the relative performance of each product is the same as mentioned above (results not shown).

At the watershed scale, as expected, the statistical metrics for all precipitation products are closer to their corresponding perfect values than those at the grid scale, but the relative performances of all precipitation products are similar to those at the grid scale. The CMORPH_BLD product still had the values that are closest to the perfect values for four metrics, while the PGF product still had the worst performance with values that are most deviated from the optimal values for six metrics.

As already mentioned in Section 2.3, we also evaluated the CHIRPS and GSMaP products at their native finer spatial resolutions 0.05° and 0.10°, respectively. The scatterplots are included in Figure 2.2 for the evaluation at the grid scale and Figure 2.3 for the evaluation at watershed scale, the statistical metrics are also summarized in Table 2.3. For evaluation at both grid and watershed scales, there are very slight differences in the statistical metrics between the results at the aggregated spatial resolution 0.25° and their original finer resolutions, suggesting that the effect of spatial resolution on evaluation of products is very little.

Figure 2.2 Scatterplots of daily precipitation from all gridded products against the interpolated rain gauge data at the grid scale.

Figure 2.3 Scatterplots of daily precipitation from all gridded products against the interpolated rain gauge data at the watershed scale.

Table 2.3 Summary of statistical metrics for evaluation of gridded precipitation products at the daily temporal scale for both grid and watershed scales.

Scale	Product	R^2	ME (mm/d)	MAE (mm/d)	RMSE (mm/d)	BIAS	POD	FAR	CSI
Grid	TRMM	0.15	0.32	3.3	6.65	0.13	0.49	0.05	0.48
	CMORPH_RAW	0.06	-1.38	2.69	6.45	-0.55	0.43	0.05	0.42
	CMORPH_CRT	0.09	-0.77	2.97	6.8	-0.3	0.42	0.05	0.41
	CMORPH_BLD	0.48	-0.47	1.86	4.43	-0.19	0.54	0.02	0.53
	PCDR	0.08	0.69	3.64	6.9	0.27	0.61	0.07	0.58
	PGF	0.01	1.59	5.37	11.84	0.63	0.38	0.08	0.37
	CHIRPS025	0.15	0.09	3.31	7.24	0.04	0.34	0.05	0.33
	GSMaP_MVK025	0.07	-0.29	3.16	7.56	-0.11	0.67	0.06	0.64
	CHIRPS005	0.11	0	3.41	7.91	0	0.23	0.05	0.22
	GSMaP_MVK010	0.07	-0.37	3.16	7.83	-0.14	0.5	0.04	0.49
Watershed	TRMM	0.24	0.32	2.85	5.21	0.13	0.88	0.08	0.82
	CMORPH_RAW	0.1	-1.39	2.47	5.59	-0.55	0.82	0.07	0.77
	CMORPH_CRT	0.15	-0.77	2.62	5.44	-0.3	0.82	0.07	0.77
	CMORPH_BLD	0.64	-0.47	1.54	3.4	-0.19	0.92	0.06	0.87
	PCDR	0.11	0.69	3.39	6.14	0.27	0.89	0.09	0.82
	PGF	0.01	1.59	4.89	9.35	0.63	0.77	0.09	0.72
	CHIRPS025	0.21	0.09	2.95	5.9	0.04	0.63	0.06	0.61
	GSMaP_MVK025	0.12	-0.27	2.84	5.88	-0.11	0.95	0.09	0.88
	CHIRPS005	0.21	0	2.88	5.77	0	0.56	0.06	0.54
	GSMaP_MVK010	0.12	-0.32	2.76	5.84	-0.13	0.92	0.08	0.85

Figure 2.4 shows the occurrence frequency of daily precipitation with seven different intensity ranges for the interpolated rain gauge and evaluated eight gridded precipitation products at both the grid and watershed scale. The first class daily precipitation with intensity range (0-1 mm/d) is plotted separately in the left panel of Figure 2.4 for a more clear inter-comparison. At the grid scale, 69% of daily precipitation from the interpolated rain gauge data falls in the intensity of 0-1 mm/d. For this intensity range, the largest overestimation was from the CMORPH_RAW product (82%), and the overestimation was reduced for CMORPH_CRT (76%) and further for BLD (70%) product. The four products (BLD,

GSMaP_MVK, CHIRPS and PGF) had very close occurrence frequency to the interpolated rain gauge. The PCDR product (49%) considerably underestimated daily precipitation within this intensity range, and a relatively slight underestimation was observed from the TRMM product (59%). Similar large underestimation in the light precipitation (0-1 mm) by the PCDR product was also reported in Malaysia (Tan et al., 2015) and in the contiguous United States (Ashouri et al., 2015).

The CMORPH_CRT and CMORPH_RAW products almost consistently underestimated the daily precipitation beyond 1 mm/d, thereby resulting in quite large negative values for the metric BIAS for these two products (Table 2.3). The CMORPH_BLD product had very close occurrence frequency for daily precipitation with intensity of 1-10 mm/d to that from rain gauge data, but it showed underestimation for daily precipitation beyond 10 mm/d. The PCDR product considerably overestimated the precipitation within 1-20 mm/d intensity and underestimated the precipitation beyond 20 mm/d. Similar overestimation of daily precipitation within 1-20 mm/d intensity was also found for the TRMM product, but this product showed underestimation for precipitation between 20 and 50 mm/d. The PGF product underestimated precipitation with intensity of 1-5 mm/d and overestimated precipitation beyond 5 mm/d, thereby leading to the largest positive BIAS of 0.63 among all precipitation products. The largest underestimation in the light daily precipitation (1-2 mm/d) was found for the CHIRPS025 product, but this product showed quite close occurrence frequency for precipitation within 1-50 mm/d and showed moderate overestimation for precipitation beyond 50 mm/d. In general, the GSMaP_MVK025 showed similar result with the BLD except for the precipitation beyond 50 mm/d, overall the two products present the closest occurrence frequency to that of rain gauge data for the daily precipitation beyond 1 mm/d. At watershed scale, basically all products had similar relative performances in capturing the occurrence frequency with those at the grid scale. Some distinct exceptions are that the TRMM product became the one that showed the largest underestimation in the light precipitation, followed by the PCDR product.
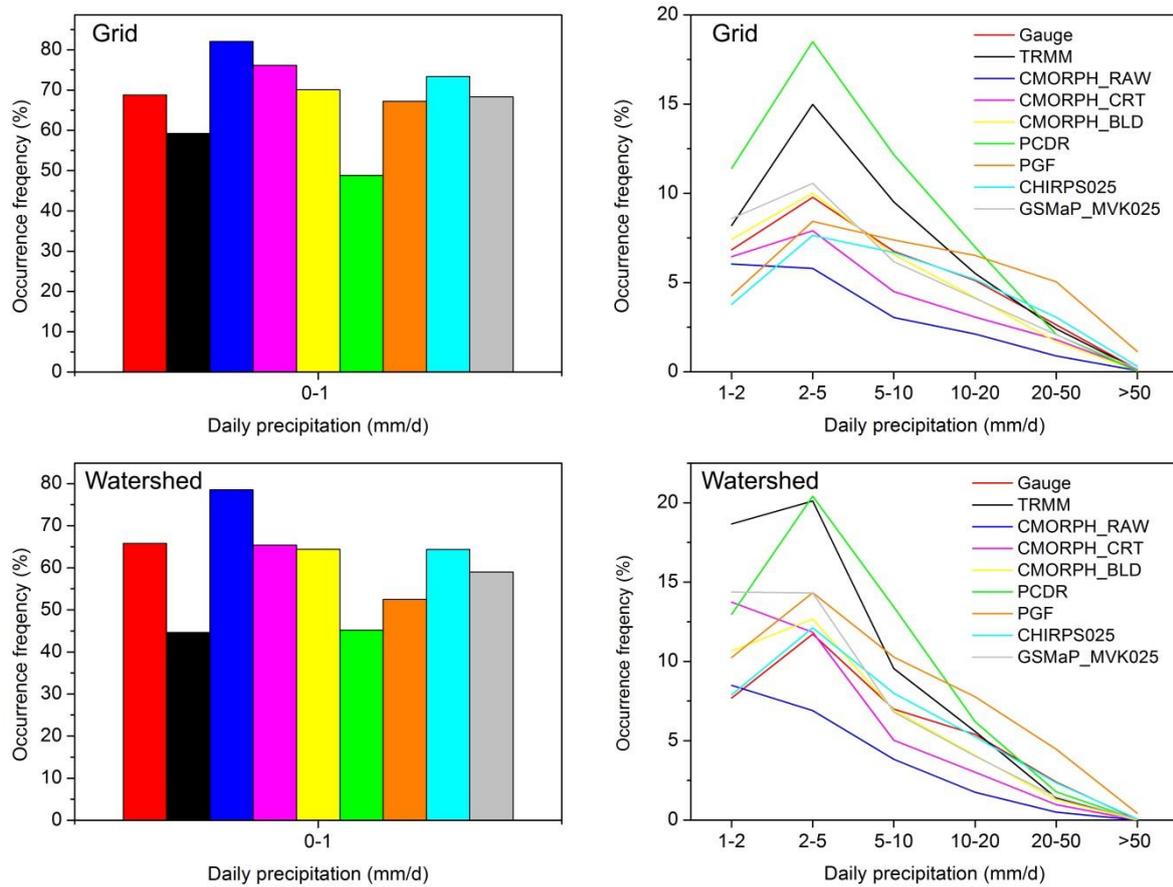
Figure 2.4 The occurrence frequency of daily precipitation with different intensity for interpolated rain gauge and eight evaluated gridded precipitation products at the grid and watershed scales.

We also computed the occurrence frequency of daily precipitation with different intensities for the product CHIRPS and GSMaP_MVK at their native finer spatial resolution 0.05° and 0.10°, respectively, and corresponding interpolated rain gauge data. For a clear inter-comparison, the occurrence frequencies of daily precipitation at all three spatial resolutions (the aggregated 0.25° and the original finer spatial resolution at the grid scale are presented in Table 2.4. As spatial resolutions aggregate from 0.05 to 0.25°, the rain gauge data showed reduced frequencies of the daily precipitation with intensity of 0-1 mm/d and beyond 20 mm/d while elevated frequencies of daily precipitation with intensity of 1-20 mm/d. When comparing the agreement in the computed occurrence frequency between the product and corresponding interpolated rain gauge data at different spatial resolutions, the product showed the same trend for underestimation or overestimation in the same daily precipitation intensity ranges. Identical to the results of CHIRPS025, the CHIRPS005 overestimated the precipitation with intensity of 0-1 mm/d and beyond 10 mm/d, and showed underestimation for precipitation with intensity of 1-10 mm/d. The same statement is also true for the GSMaP_MVK010 product except for the intensity 0-1 mm/d where GSMaP_MVK025 showed slight underestimation while GSMaP_MVK010 showed slight overestimation. Similar findings were also valid for results at the watershed scale. Therefore, taken

together it can be concluded that the spatial resolutions of products to which evaluation was conducted had almost no effect on their performance in terms of statistical metrics and precipitation occurrence frequency. The differences in the occurrence frequency of precipitation among the eight products would have significant implications for hydrological modeling and sediment and pollutant transport modeling due to the nonlinearity in the related processes (Tian et al. 2007). Our follow-up hydrological modeling study with the SWAT model found that different occurrence frequency of precipitation may lead to considerably different performances in streamflow simulation (Tuo et al., 2016). Thus, it is important to have a comprehensive understanding about the occurrence frequency of precipitation of these precipitation products before applying them for modeling-related applications.

Table 2.4 The occurrence frequency (%) of daily precipitation intensities for interpolated rain gauge and gridded precipitation product CHIRPS and GSMaP_MVK at the aggregated 0.25° and their original finer (0.05° and 0.10°) spatial resolution at the grid and watershed scale. "Gauge005" refers to interpolated rain gauge data at the same 0.05° grid with the CHIRPS at the finer resolution "CHIRPS005"; "Gauge010" refers to interpolated rain gauge data at the same 0.10° grid with the GSMaP_MVK at the finer resolution "GSMaP_MVK010"; "Gauge025" refers to interpolated rain gauge data at the same 0.25° grid with the CHIRPS and GSMaP_MVK at the common aggregated resolution "CHIRPS025" and "GSMaP_MVK025". Details can be found in Section 2.2.2 and 2.2.3.

Evaluation scale	Product	Daily precipitation intensity (mm/d)						
		0-1	1-2	2-5	5-10	10-20	20-50	≥50
Grid	Gauge005	72.23	5.57	8.18	6.16	4.92	2.72	0.21
	Gauge010	70.55	6.24	8.99	6.43	5.03	2.62	0.15
	Gauge025	68.77	6.83	9.78	6.76	5.11	2.63	0.13
	CHIRPS005	78.54	1.53	5.49	5.94	5.12	3.05	0.33
	CHIRPS025	73.37	3.78	7.64	6.68	5.17	3.05	0.3
	GSMaP_MVK010	72.01	6.55	9.19	5.78	4.17	2.12	0.19
	GSMaP_MVK025	68.32	8.59	10.56	6.16	4.12	2.09	0.17
Watershed	Gauge005	65.95	8.06	11.40	6.92	5.28	2.36	0.02
	Gauge010	65.88	8.01	11.57	6.92	5.25	2.34	0.02
	Gauge025	65.78	7.69	11.72	6.99	5.40	2.39	0.02
	CHIRPS005	66.08	7.52	11.50	7.52	5.00	2.31	0.07
	CHIRPS025	64.34	7.89	12.12	7.99	5.25	2.34	0.07
	GSMAP_MVK010	61.78	13.22	13.27	6.38	3.80	1.48	0.08
	GSMAP_MVK025	58.98	14.37	14.32	6.80	4.03	1.45	0.05

2.3.2   Evaluation at monthly scale

The daily precipitation data were accumulated to monthly total precipitation for the interpolated rain gauge data and all gridded precipitation products, and then similar evaluation were conducted at both the grid scale and the watershed scale. Figure 2.5 and Figure 2.6 show scatterplots of monthly total precipitation from all selected gridded precipitation products against interpolated rain gauge data at the grid scale and watershed scale, respectively. The statistical metrics are presented in Table 2.5.

Better agreements with the interpolated rain gauge data were observed for all products with improved metrics at both grid and watershed scales than those of daily precipitation. The PGF product was still the worst performing product with the largest errors and bias. At the grid scale, the CHIRPS025 had values that are closest to the perfect values for all five metrics except the R^2 (0.68) which is just slightly smaller than the best value (0.70) achieved by the TRMM product. Compared to the CHIRPS025, the TRMM and CMORPH_BLD products had higher or equal R^2 and just slightly larger MAE and RMSE values. At the watershed scale, CHIRPS025, TRMM and CMORPH_BLD showed quite similarly good performance with very slight difference in the statistical metrics. Therefore, these three products could be reasonably considered as the top three products for the monthly total precipitation. Similarly, there is very slight difference in the statistical metrics between evaluation results at the original finer spatial resolutions (CHIRPS005, GSMaP_MVK010) and the aggregated 0.25° spatial resolution (CHIRPS025, GSMaP_MVK025) for both CHIRPS and GSMaP_MVK products.
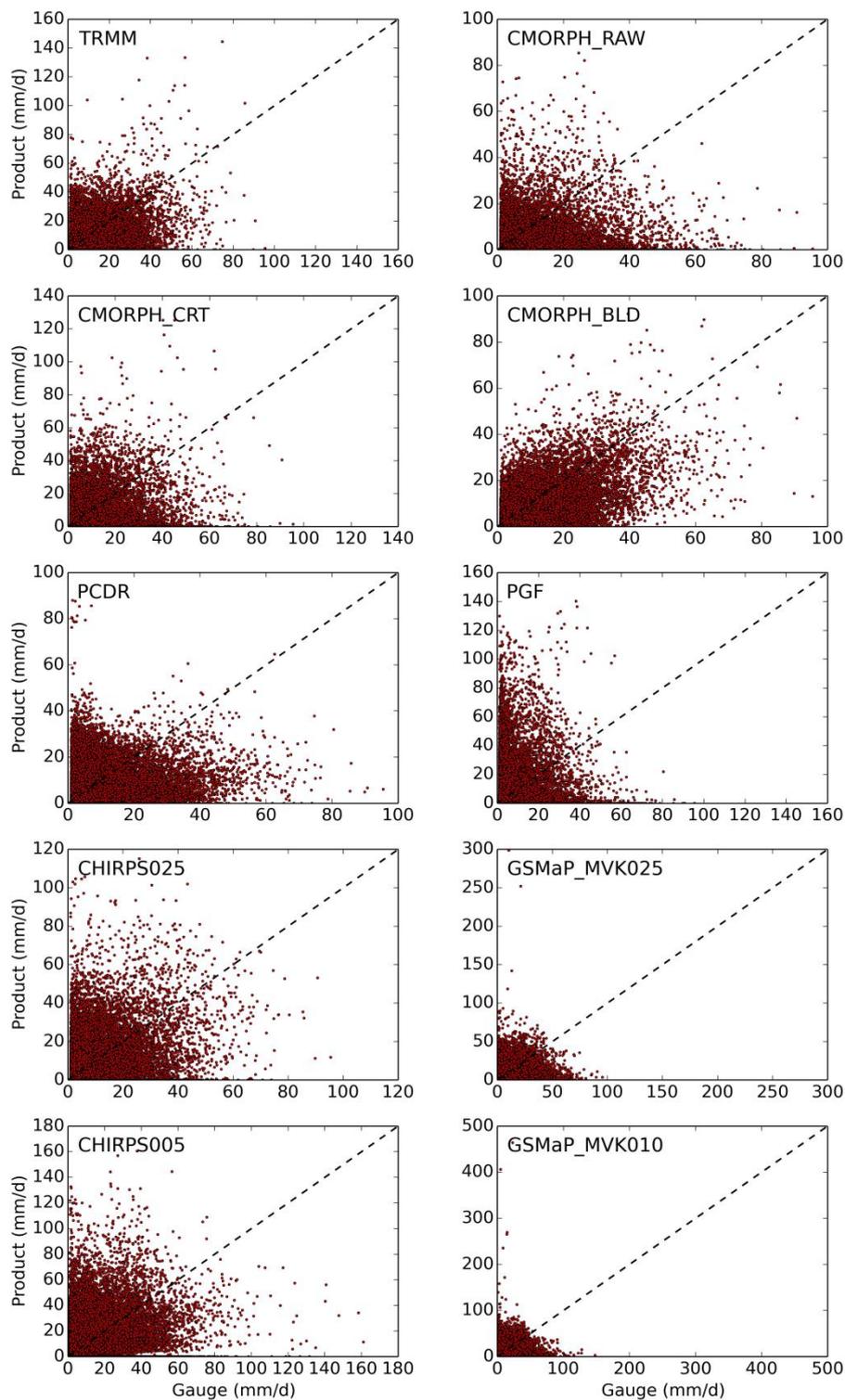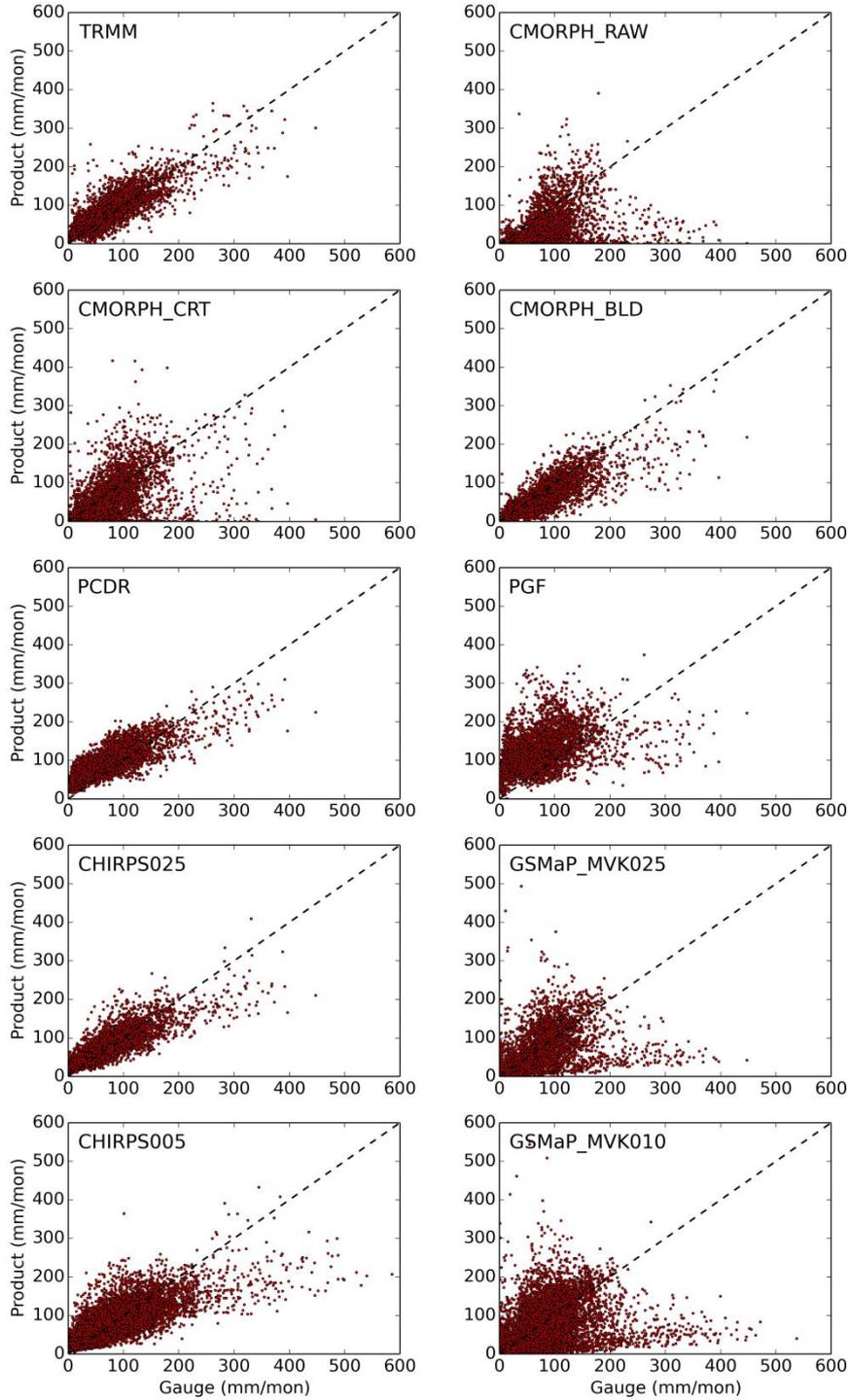
Figure 2.5 Scatterplots of monthly precipitation from all gridded products against the interpolated rain gauge data at the grid scale.
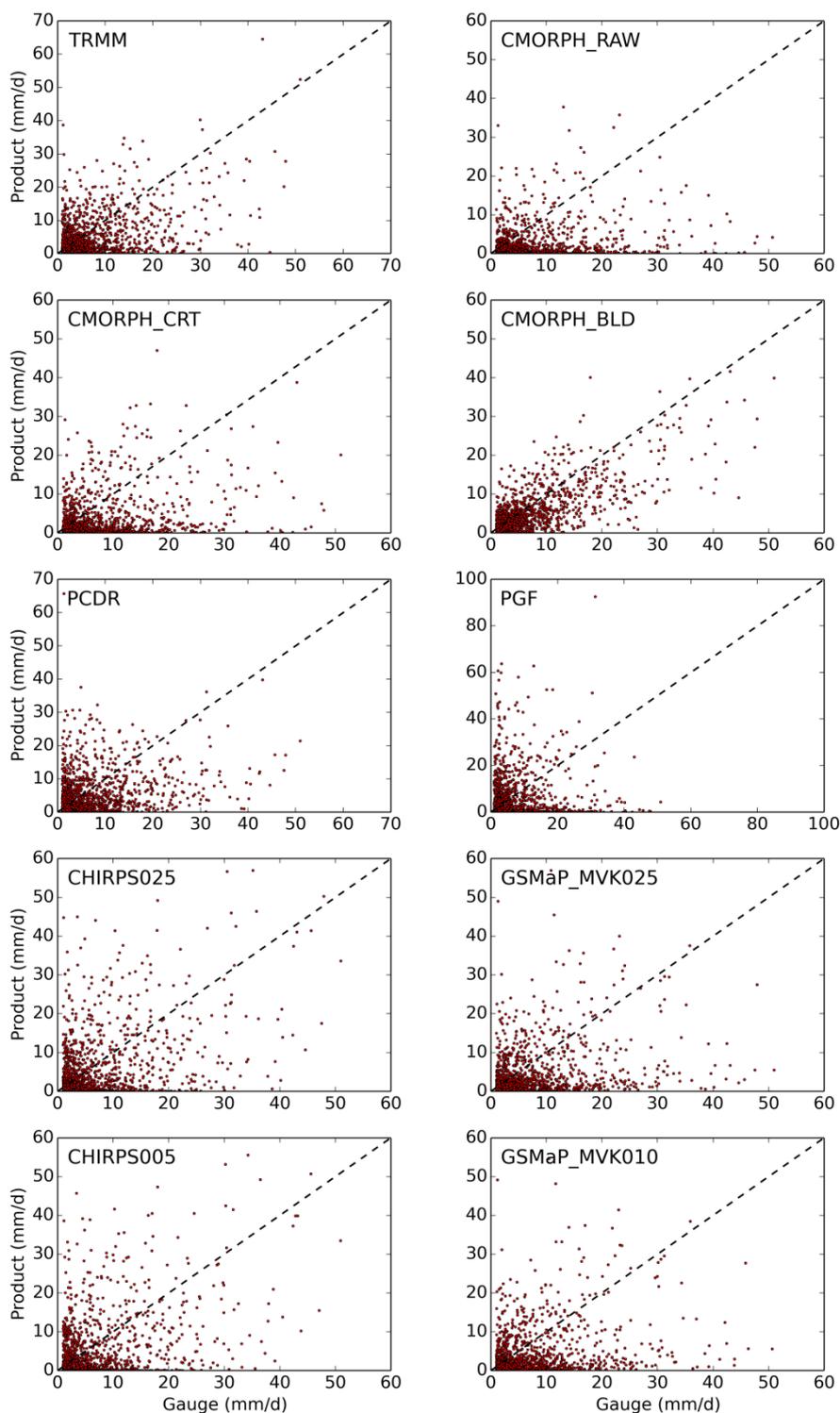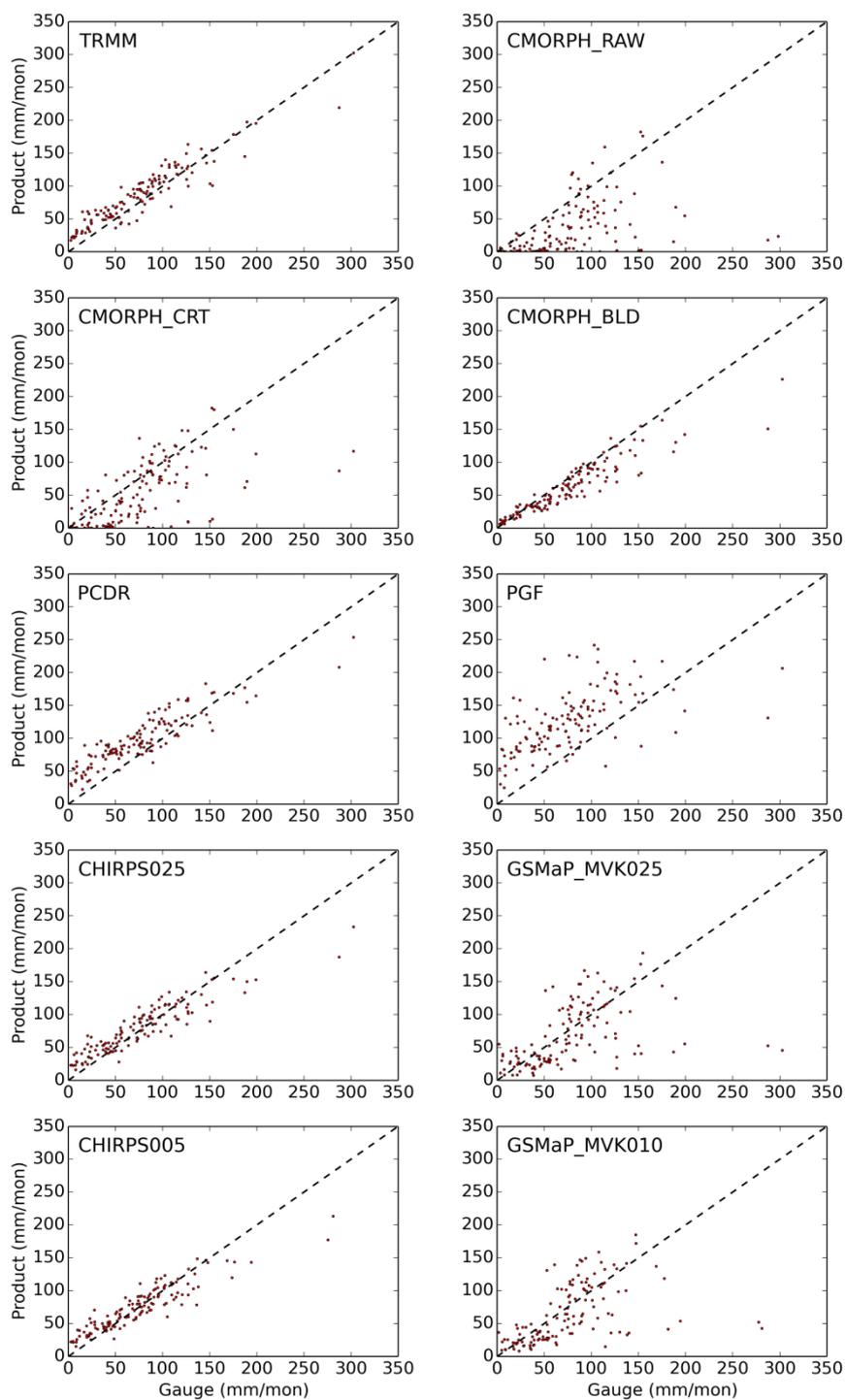
Figure 2.6 Scatterplots of monthly precipitation from all gridded products against the interpolated rain gauge data at the watershed scale.

Table 2.5 Summary of statistical metrics for evaluation of gridded precipitation products at the monthly temporal scale for both grid and watershed scales.

Scale	Product	R^2	ME (mm/mon)	MAE (mm/mon)	RMSE (mm/mon)	BIAS
Grid	TRMM	0.7	9.74	23.93	32.16	0.13
	CMORPH_RAW	0.17	-42.13	51.05	70.59	-0.55
	CMORPH_CRT	0.25	-23.44	44.38	62.8	-0.3
	CMORPH_BLD	0.68	-14.39	22.67	34.54	-0.19
	PCDR	0.64	21.06	32.57	39.46	0.27
	PGF	0.18	48.46	61.5	75.47	0.63
	CHIRPS025	0.68	2.83	22.3	31.19	0.04
	GSMaP_MVK025	0.16	-8.8	40.18	61.15	-0.11
	CHIRPS005	0.59	0.14	26.07	37.98	0
	GSMaP_MVK010	0.17	-10.85	40.16	62	-0.14
Watershed	TRMM	0.88	9.74	17.13	20.66	0.13
	CMORPH_RAW	0.21	-42.13	46.35	64.24	-0.55
	CMORPH_CRT	0.37	-23.44	33.7	48.69	-0.3
	CMORPH_BLD	0.86	-14.39	16.56	25.46	-0.19
	PCDR	0.81	21.06	27.3	31.44	0.27
	PGF	0.32	48.46	57.02	66.16	0.63
	CHIRPS025	0.84	2.83	16.92	22.35	0.04
	GSMaP_MVK025	0.22	-8.8	32.81	50.52	-0.11
	CHIRPS005	0.84	0.02	15.02	20.81	0
	GSMaP_MVK010	0.25	-11	30.35	47.37	-0.14
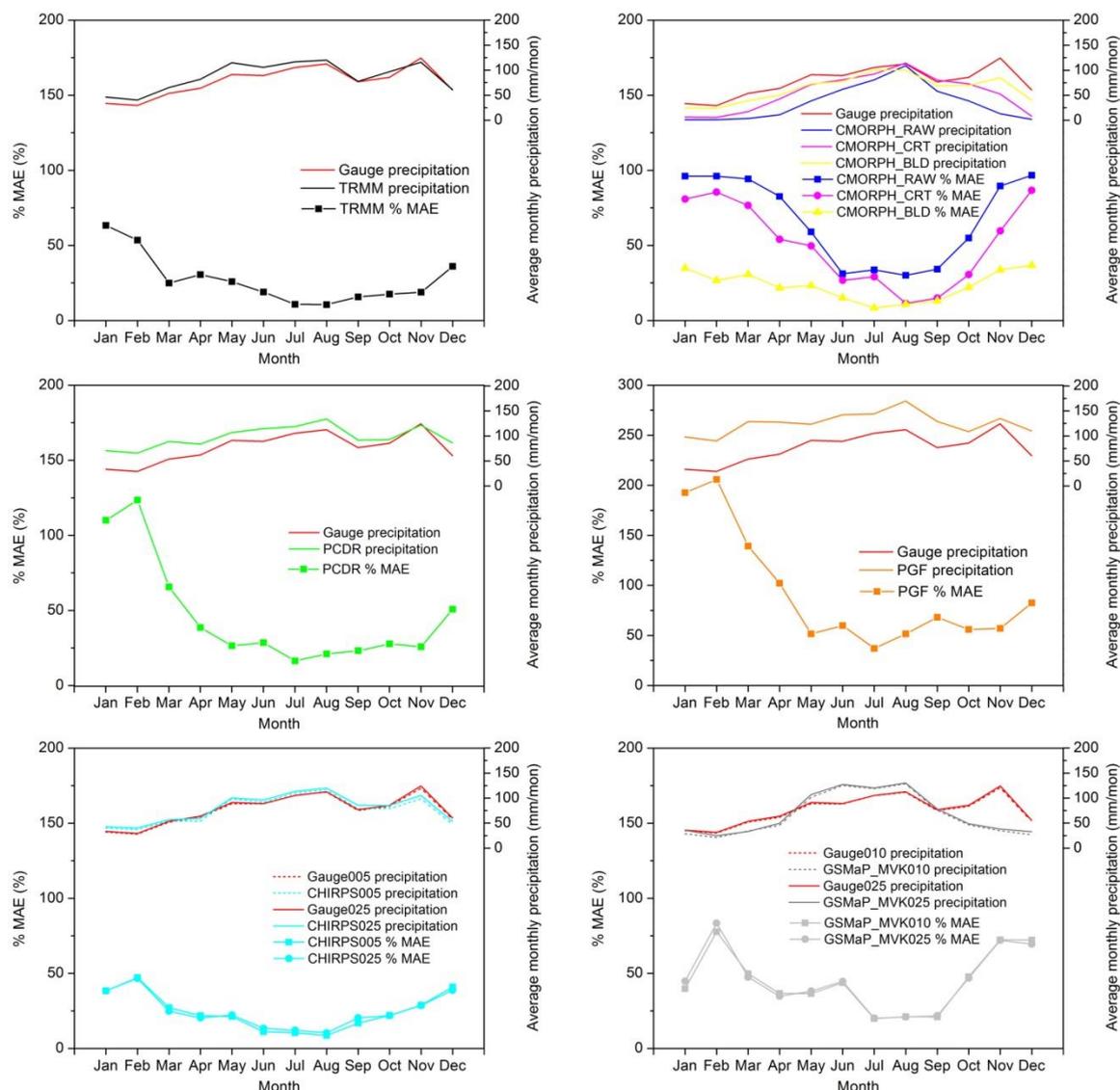
Figure 2.7 Seasonal variation of mean monthly precipitation during the period 2000-2010 and their relative mean absolute error (%MAE) for the Adige Basin at the watershed scale.

Figure 2.7 presents the mean temporal pattern of monthly precipitation (right Y-axis) from interpolated rain gauge data and all eight evaluated products at the watershed scale during the whole period 2000-2010, as well as temporal patterns of the mean absolute error (MAE) in relative percentage (%MAE, left Y-axis) for evaluated products. The %MAE was computed as the ratio of MAE value to the mean value from corresponding interpolated rain gauge data, and thus the %MAE enables us to make inter-comparisons among different months by taking into account the monthly variations of precipitation (Duan and Bastiaanssen, 2013). From interpolated rain gauge data, mean monthly precipitation in the Adige Basin is characterized by two peaks with a higher peak occurring in November and the other in August, and the winter months (December-February) received low precipitation with the lowest precipitation occurring in February. Snowfall will account for a larger portion of precipitation during the

35

October–April months (Mei et al., 2014), and rain gauge stations have higher possibility of under-sampling snowfall, resulting in underestimating of precipitation. The seasonal pattern with two peaks can be captured by most products except GSMaP_MVK, CMORPH_RAW and CMORPH_CRT which showed only one peak in August. For those products capable of capturing two peaks, the highest peak occurred in August rather than in November, suggesting that all eight products underestimated precipitation in November. Considering the %MAE, all eight products showed higher error in the winter months when precipitation is low while lower error in the summer and autumn months when precipitation is relatively high. Among all factors affecting the accuracy of evaluated precipitation products mentioned in Section 2.3.1, the consistent higher error in the winter months could be attributed to two factors: (1) winter precipitation in Adige Basin is often from non-convective systems and associated with shallower cloud, and such precipitation is difficult to be detected by satellite algorithms (Eber et al., 2007; Tian et al., 2007; Mei et al., 2014). On the other hand, satellite algorithms performed better in detecting precipitation in warm seasons that is often from convective systems; (2) more ice and snow would cover land surfaces in cold winter months, and the presence of ice and snow cover added more difficulties and errors in microwave-based precipitation retrievals from satellites because they produced strong disturbing signals to those of ice particles in the atmosphere (Eber et al., 2007; Tian et al., 2014). Comparatively, the CHIRPS product at both 0.05° and 0.25° spatial resolution performed best in reproducing the temporal pattern of monthly precipitation, followed by TRMM and CMORPH_BLD. All the three products CHIRPS, TRMM and CMORPH_BLD had comparably lower error in the 12 month with the mean value of %MAE being 25%, 23% and 27%, respectively. The PGF product had %MAE larger than 50% in 11 months and %MAE even higher than 100% in January-April months, resulting in the PGF having the largest error with the mean %MAE of 92%.

2.3.3   Evaluation at annual scale

The accumulated monthly total precipitation data were further accumulated to annual total precipitation. The annual precipitation data from all gridded precipitation products are plotted against those from the interpolated rain gauge data at the grid scale in Figure 2.8 and at the watershed scale in Figure 2.9. The statistical metrics are summarized in Table 2.6. The large deviation from the 1:1 line is clearly shown in Figure 2.8. At this grid scale, the agreement with interpolated rain gauge data were not improved but rather reduced with lower R^2 for most precipitation products as the temporal scale aggregates from monthly to annual. This suggests that the errors in daily and monthly precipitation were neither symmetrical nor random, yet they are characterized by a bias with dominant negative or positive errors. Thus temporal aggregation did not cancel each other and not improve the match between rain gauge data and gridded precipitation products. For example, large positive errors (overestimation) are clearly shown for the products PCDR and PGF while negative errors (underestimation) are clearly shown for the

CMORPH_BLD. The largest R^2 was just 0.57 which was achieved by the CMORPH_BLD, but this product showed consistent underestimation. The CHIRPS025 had the best values for four metrics ME, MAE, RMSE and BIAS. The TRMM product had second best values for MAE and RMSE, followed by the CMORPH_BLD. Other products showed relatively larger errors with the PGF product having the largest errors.

In contrast, at the watershed scale, the agreement with interpolated rain gauge data were improved with higher R^2 for most precipitation products as the temporal scale aggregates from monthly to annual. For all precipitation products, the errors were reduced compared to those at grid scale. The relatively performance of precipitation products are the same with those at the grid scale. The CHIRPS025 had the lowest MAE and RMSE, followed by the TRMM and CMORPH_BLD products. Similarly, very slight differences were observed in the statistical metrics between evaluation results of CHIRPS005 and GSMaP_MVK010 and the aggregated 0.25° spatial resolution (CHIRPS025, GSMaP_MVK025).

Figure 2.10 shows the spatial pattern of average annual precipitation over the Adige Basin during 2000-2010 from interpolated rain gauge data and eight evaluated products at the common spatial resolution of 0.25°. From the interpolated rain gauge data, the average annual precipitation ranged from 668 to 1152 mm/y and showed a general decreasing pattern from southeast to northwest. Overall, all evaluated products except PCDR, PGF and GSMaP_MVK captured the decreasing pattern but all products could not adequately reproduce the magnitude of average annual precipitation with overestimation and underestimation occurring in different places. The PGF product showed an opposite decreasing pattern. Comparatively, TRMM and CHIRPS025 had the spatial pattern that is closest to that of the interpolated rain gauge data. These results indicate that care should be taken when using these gridded precipitation products to study the spatial patterns of precipitation (Ji and Chen, 2012), and more researches on the spatial errors of precipitation should be conducted to further improve these products (Chen et al., 2013).
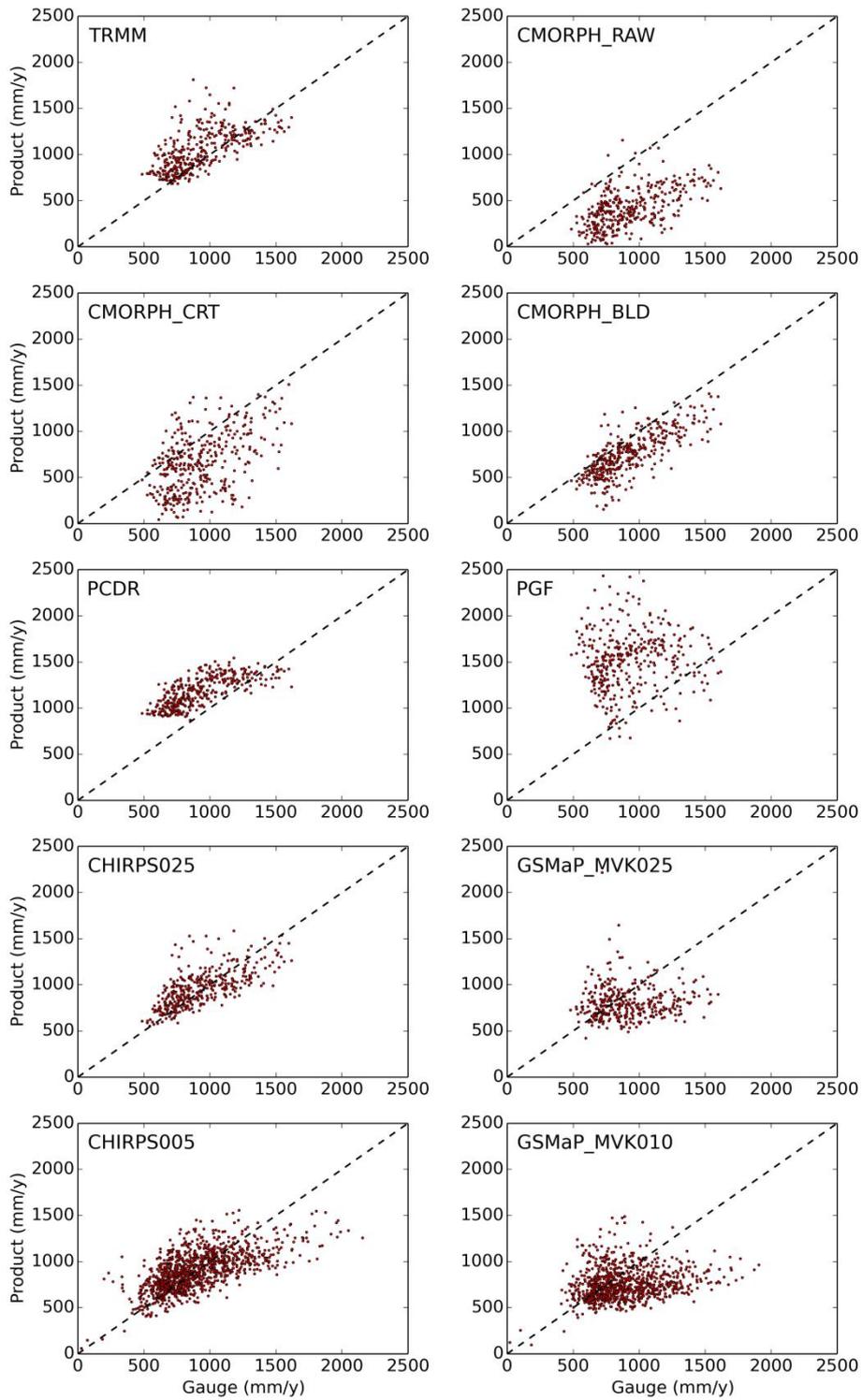
Figure 2.8 Scatterplots of annual precipitation from all gridded products against the interpolated rain gauge data at the grid scale.
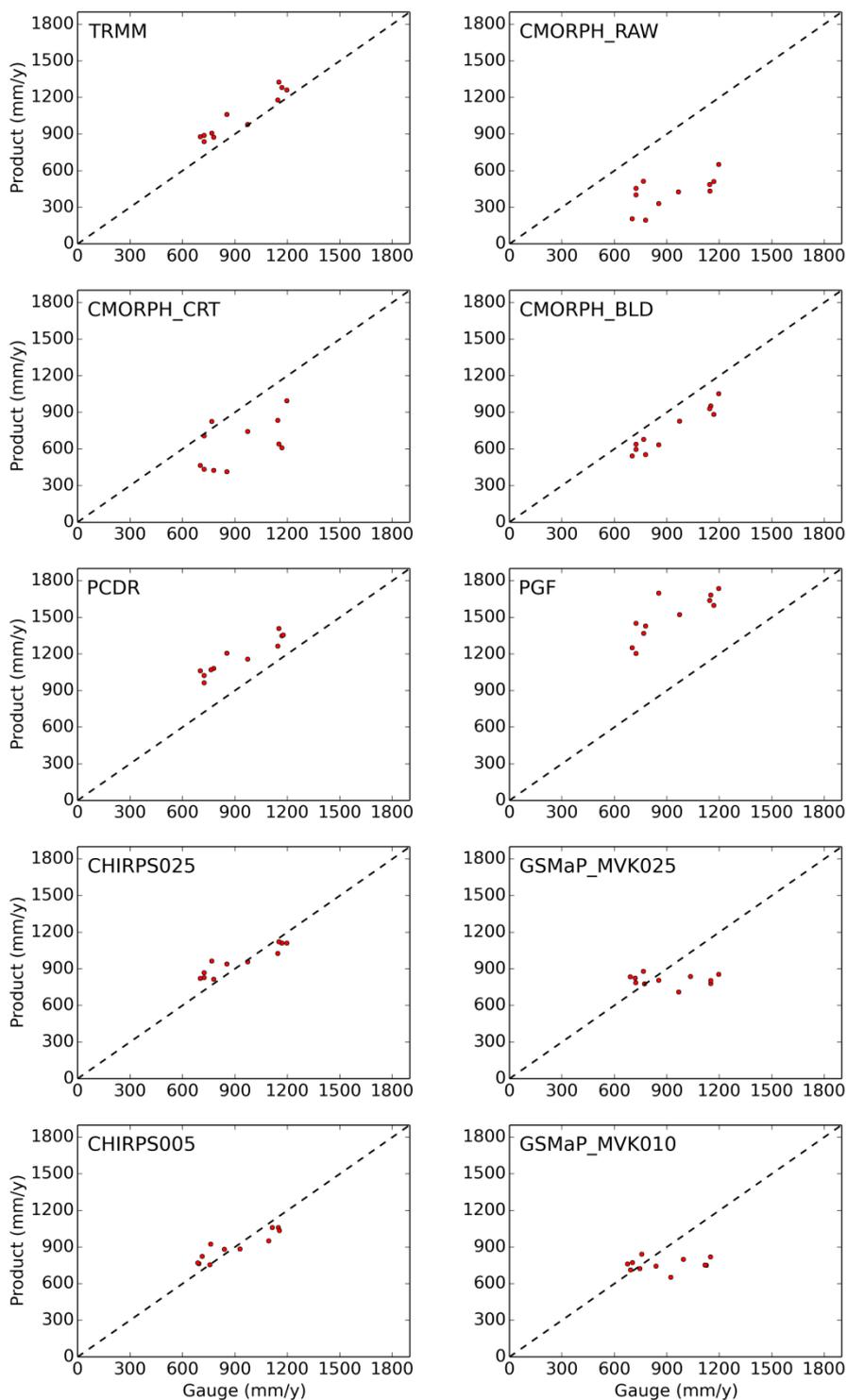
Figure 2.9 Scatterplots of annual precipitation from all gridded products against the interpolated rain gauge data at the watershed scale.

Table 2.6 Summary of statistical metrics for evaluation of gridded precipitation products at the annual temporal scale for both grid and watershed scales.

Scale	Product	R^2	ME (mm/y)	MAE (mm/y)	RMSE (mm/y)	BIAS
Grid	TRMM	0.42	116.91	171.57	229.83	0.13
	CMORPH_RAW	0.3	-505.59	509.02	550.36	-0.55
	CMORPH_CRT	0.24	-281.28	332.85	403.13	-0.3
	CMORPH_BLD	0.57	-172.69	194.69	240.2	-0.19
	PCDR	0.54	252.76	270.07	301.53	0.27
	PGF	0.01	581.5	606.07	694.72	0.63
	CHIRPS025	0.47	33.96	141.22	183.93	0.04
	GSMaP_MVK025	0.01	-103.2	233.06	304.48	-0.11
	CHIRPS005	0.42	1.66	172.48	222.88	0
	GSMaP_MVK010	0.05	-126.53	235.92	304.69	-0.14
Watershed	TRMM	0.90	116.91	116.91	131.73	0.13
	CMORPH_RAW	0.39	-505.59	505.59	527.95	-0.55
	CMORPH_CRT	0.31	-281.28	291.83	334.19	-0.3
	CMORPH_BLD	0.91	-172.69	172.69	182.67	-0.19
	PCDR	0.89	252.76	252.76	263.69	0.27
	PGF	0.66	581.5	581.5	592.6	0.63
	CHIRPS025	0.86	33.96	90.66	104.28	0.04
	GSMaP_MVK025	0.01	-103.2	181.51	220.1	-0.11
	CHIRPS005	0.84	1.53	83.3	94.92	0
	GSMaP_MVK010	0.02	-126.61	174.09	218.39	-0.14
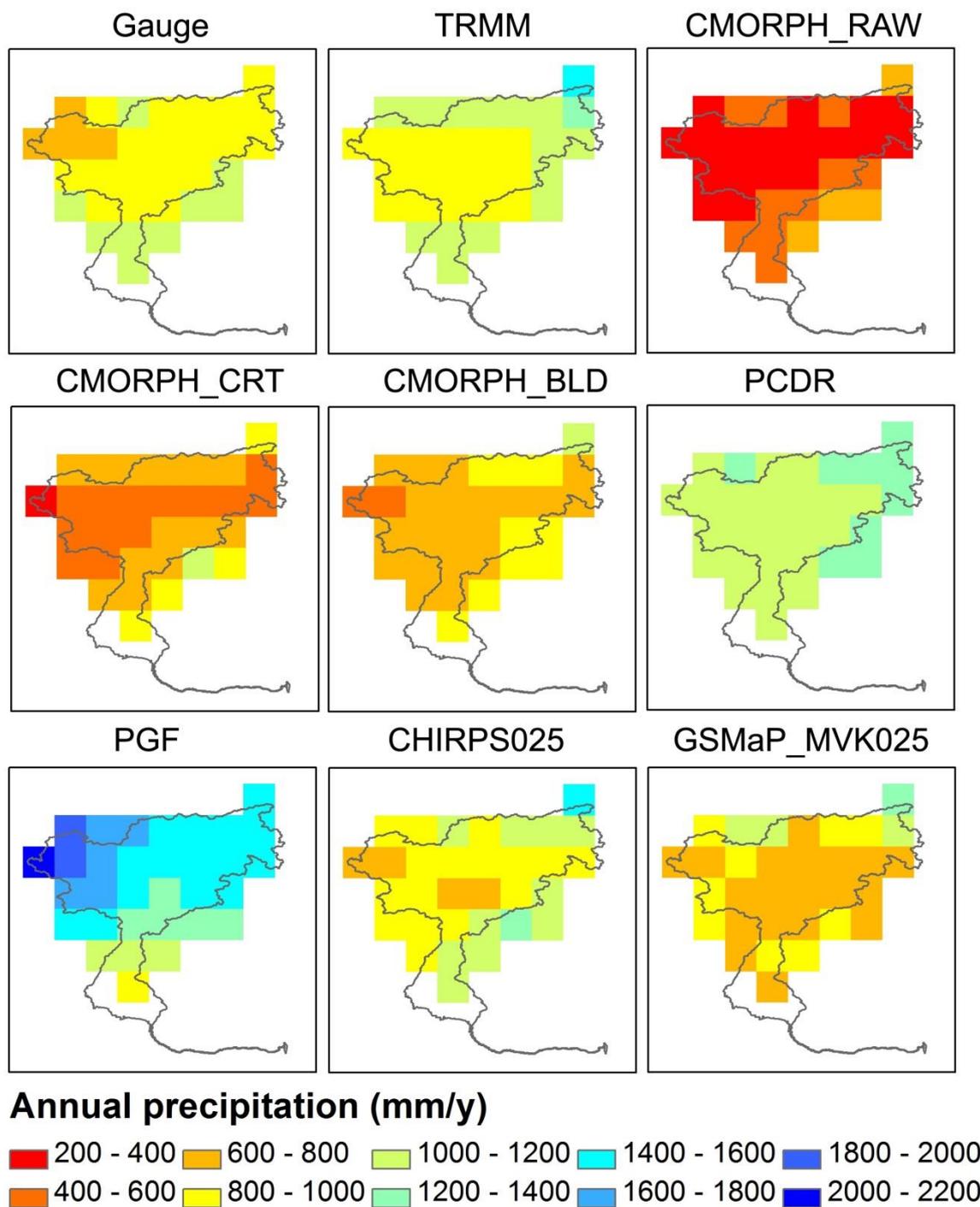
Figure 2.10 Spatial distribution of mean annual precipitation for the period 2000–2010 from (a) interpolated gauge station at the spatial resolution of 0.25 degree for grids which have at least one gauge station located inside covering Adige Basin.

2.4 Conclusions

In this study, we conducted a comprehensive evaluation of eight high resolution gridded precipitation products in the Adige Basin, Italy using measurements from 101 rain gauges during 2000-2010. Since most of satellite-based precipitation products have a quasi-global coverage (latitude bands of 50° N-S or

60° N-S), thus the nature of Adige Basin (located within 45-47.1° N latitude and characterized by complex topography) makes it a representative high-latitude site for assessing performances of various available products. This study presents a new contribution to the global view of the performance of various precipitation products. The evaluated eight products are the TRMM 3B42, CMORPH_RAW, CMORPH_CRT, CMORPH_BLD, PCDR, PGF, CHIRPS and GSMaP_MVK. All eight products were evaluated against interpolated rain gauge data at the common 0.25° spatial resolution, and additional evaluations at finer spatial resolution were also conducted for CHIRPS (0.05°) and GSMaP_MVK (0.10°). Evaluation was performed at various temporal (daily, monthly and annual) and spatial scales (grid and watershed).

Our evaluation found that considering all evaluation metrics and aspects, in general the CHIRPS (slight overall overestimation with BIAS < 0.05), TRMM (overall overestimation with BIAS of 0.13) and CMORPH_BLD (overall underestimation with BIAS of -0.19) can be considered to comparably perform better than others, while the PGF product performed worse with large overall overestimation with BIAS of 0.63. Among three different products from CMORPH, the CMROPH_RAW was consistently the worst performing product and improvements were achieved by CMROPH_CRT and further remarkably by CMROPH_ BLD. This confirms the effectiveness of the bias correction and blending algorithms used in CMORPH products. Consistent improved evaluation metrics are observed as the evaluation spatial scale increases from grid to watershed and the temporal scale increases from daily to monthly, but metrics for annual precipitation are worse than those at monthly scale. Evaluation metrics and aspects were only slightly different between them at the original finer spatial resolutions (0.05° and 0.10°, respectively) and the aggregated 0.25° spatial resolution for both CHIRPS and GSMaP_MVK products, suggesting the little effect of spatial resolution at which evaluation is performed on the evaluation results. All eight products presented different occurrence frequency of daily precipitation for some intensity ranges compared to rain gauge data, such differences are expected to have significant effects on hydrological modeling and sediment and pollutant transport modeling applications. As far as the performances in temporal and spatial distribution of precipitation are concerned, all eight products tend to show higher error in the winter months (December-February) when precipitation is low; all eight products except PCDR, PGF and GSMaP_MVK reproduced a general southeast-northwest decreasing spatial pattern with TRMM and CHIRPS performing better, while PGF captured an opposite spatial pattern.

This study has implications for precipitation product development and guidance to alternative choice of precipitation data for the local community. The findings clearly show that there is still room for improvement for all eight gridded precipitation products particularly at daily and finer temporal scales.

It is worth noting that the recent CHIRPS product available at the finest spatial resolution spatial resolution (0.05°) has shown the smallest bias and relatively better performance than other products, which favors its application in hydrological studies at small basin scales. Our follow-up study has shown that using the CHIRPS product as input to SWAT model resulted in satisfactory performance in simulating streamflow in all three tested headwater subbasins in the Adige Basin (Tuo et al., 2016). From the local perspective for the Adige Basin, one interesting future study could be to construct a further improved precipitation dataset at fine spatial resolution by merging the available rain gauge data with CHIRPS using similar effective blending algorithm used in CMORPH products, which will be beneficial to many studies identified within the GLOBAQUA project such as sediment and pollutant transport modeling and water resources assessment and management. Snowfall is of great significance during winter months in the Adige Basin and other basins of the Alps. However, accurate snowfall detection and quantification is still a challenging task for both ground and satellite remote sensing measurements (Lettenmaier et al., 2015), thus there is a need for improving the quantification of snowfall and its contribution to the total precipitation and further its effects on the seasonal hydrological balance in these study regions in future studies. The recently launched Global Precipitation Measurement (GPM) satellite mission (http://www.nasa.gov/mission_pages/GPM/overview/index.html) is expected to particularly improve the quantification of snow worldwide. The 0.25° spatial resolution of most available gridded products is still relatively coarse for applications in smaller basins. Spatial downscaling techniques can be performed to obtain precipitation at higher spatial resolution (e.g. 1 km). Various spatial downscaling techniques have been developed, but most of them are focused on monthly and annual scales (Duan and Bastiaanssen, 2013; Chen et al., 2015). An important focus for future work could thus be the development of robust spatial downscaling methods for precipitation at daily and finer time scales.

Chapter 3

Evaluation of precipitation input for SWAT modeling in Alpine catchment: A case study in the Adige River Basin (Italy)[2]

Precipitation is often the most important input data in hydrological models when simulating streamflow. The Soil and Water Assessment Tool (SWAT), a widely used hydrological model, only makes use of data from one precipitation gauge station that is nearest to the centroid of each subbasin, which is eventually corrected using the elevation band method. This leads in general to inaccurate representation of subbasin precipitation input data, particularly in catchments with complex topography. To investigate the impact of different precipitation inputs on the SWAT model simulations in Alpine catchments, 13 years (1998-2010) of daily precipitation data from four datasets including OP (Observed precipitation), IDW (Inverse Distance Weighting data), CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) and TRMM (Tropical Rainfall Measuring Mission) has been considered. Both model performances (comparing simulated and measured streamflow data at the catchment outlet) as well as parameter and prediction uncertainties have been quantified. For all three subbasins, the use of elevation bands is fundamental to match the water budget. Streamflow predictions obtained using IDW inputs are better than those obtained using the other datasets in terms of both model performance and prediction uncertainty. Models using the CHIRPS product as input provide satisfactory streamflow estimation, suggesting that this satellite product can be applied to this data-scarce Alpine region. Comparing the performance of SWAT models using different precipitation datasets is therefore important in data-scarce regions. This study has shown that, precipitation is the main source of uncertainty, and different precipitation datasets in SWAT models lead to different best estimate ranges for the calibrated parameters. This has important implications for the interpretation of the simulated hydrological processes.

3.1 Introduction

Quantitative hydrological models are useful tools to support the development of new water resource management policies and assess water quality issues (Abbaspour et al., 2015; Beven, 2011). Among these models, the Soil and Water Assessment Tool (SWAT) (Arnold and Fohrer, 2005; Arnold et al., 1998) has been widely used for river basins around the world (Guse et al., 2016; Malagò et al., 2016; Nerantzaki et al., 2015; Schmalz et al., 2015; Song et al., 2011). Over the last decades, this open source model has been continuously improved by integrating modules that evaluate the effects of various hydrological and chemical processes (Gassman et al., 2007). SWAT is frequently used by both the global academic community and the practitioners. Over 2500 peer-review papers have been published (Scopus, till 31 July 2016) using modeling results obtained applying SWAT. This model has been proven to be an effective tool for simulating hydrological processes, contaminant transport, soil erosion and for assessing the effects of climate change, land use change, and water management practices in diverse environmental conditions (Abbaspour et al., 2015; Abbaspour et al., 2007; Ayana et al., 2015; Dile et al., 2016; Guo et al., 2008; Rahman et al., 2013; Schuol and Abbaspour, 2006; Woznicki et al., 2016; Yang et al., 2016). It has also been used to support the implementation of environmental directives such as the Clean Water Act in the United States (Gabriel et al., 2014) or the European Water Framework directive (Volk et al., 2009). A new version (SWAT+) that is under development, will implement a landscape unit approach to improve the processes distribution and allocation of management operations in landscapes (Bonumá et al., 2014; Rathjens et al., 2015) and will be more flexible in watershed discretization and configuration so as to represent the physical characteristics of a watershed as realistically as possible (Bieger et al., 2017).

Precipitation is a major driving force of hydrological processes, sediment and chemical fluxes (Cho et al., 2009; Masih et al., 2011; Price et al., 2014), and therefore reliable precipitation data are important inputs for SWAT (Galván et al., 2014; Monteiro et al., 2016; Strauch et al., 2012) and other hydrological models (Andréassian et al., 2001; Bárdossy and Das, 2008; Mei et al., 2016a). Therefore, an accurate representation of the temporal and spatial variability of precipitation is of importance to achieve an accurate river basin model. In other words, physically based hydrological models such as SWAT cannot generate accurate predictions of hydrological processes without adequate representations of the regional precipitation distribution. Subsequently, without an accurate simulation of hydrological processes, reliable predictions of other relevant behaviors such as water quality and erosion cannot be achieved (Chaplot et al., 2005).

The sparse and heterogeneous spatial distribution of rain gauges often results in inaccurate precipitation inputs for SWAT, especially when modeling large river basin or basins with complex heterogeneous

terrains like mountainous regions (e.g. Alpine catchments) where the assumption of spatially uniform rainfall is not valid (Cho et al., 2009). Furthermore, the current method of representing precipitation in the SWAT model is simplistic, since it only uses data from one precipitation gauging station that is nearest to the centroid of each subbasin (Galván et al., 2014; Masih et al., 2011). Therefore, improved precipitation inputs which consider regional spatial variations are crucial for achieving reliable modeling results with SWAT (Tobin and Bennett, 2009).

Numerous methods are available for processing precipitation data in order to consider spatial effects. Among these methods, interpolation methods based on ground measurements (e.g., Inverse distance weighting, IDW) and satellites precipitation estimates (e.g., Tropical Rainfall Measuring Mission, TRMM; Climate Hazards Group InfraRed Precipitation with Station data, CHIRPS) have been considered in the present work. IDW-based interpolated precipitation and TRMM have been shown to be useful for obtaining satisfactory model performance in case of spatially varied precipitation patterns (Li et al., 2012; Ly et al., 2011; Shen et al., 2012; Wagner et al., 2012) in particular when used as input data for SWAT models (Galván et al., 2014; Masih et al., 2011; Shen et al., 2012; Strauch et al., 2012; Tobin and Bennett, 2009; Wagner et al., 2012). The CHIRPS dataset has a higher spatial resolution compared to TRMM (0.25°) and is expected to capture more representative precipitation features, because the recently released "satellite-gauge" type CHIRPS product has the finest spatial resolution of 0.05° (Funk et al., 2015).

Comparing with the previous works (Shen et al., 2012; Shope and Maharjan, 2015; Vu et al., 2012; Wagner et al., 2012; Zhang and Srinivasan, 2009), we can state that the choice of the best precipitation input data for the SWAT models is basin-specific. More importantly, few case studies have investigated catchments with high elevation gradients (Shope and Maharjan, 2015; Xu et al., 2010) and none of them refer to SWAT applications in the South-eastern Alps. Therefore, investigations about the optimal precipitation input dataset to model Alpine catchments with SWAT are limited. Furthermore, SWAT utilizes elevation bands to simulate precipitation variability in a subbasin due to orographic effects (Neitsch et al., 2011). Only a few studies have tried to assess the effectiveness of the elevation band approximation and the results may be case-specific (Grusson et al., 2015; Rahman et al., 2013; Strauch et al., 2012). This study aims at comparing different precipitation inputs for SWAT in Alpine headwater catchments belonging to the Adige river basin.

The specific objectives of this study are to: 1) analyze the impact of four different precipitation data sources on SWAT modeling without elevation correction; 2) evaluate the effect of the elevation band method on the performance of SWAT models in alpine catchments; 3) estimate parameter and prediction

uncertainties considering the four precipitation inputs in the three study areas; 4) investigate the exportability of the result of this study to other Alpine catchments.

3.2 Materials and methods

3.2.1   Study areas

The Adige river basin is located in North-East Italy and has a drainage area of 12,100 km^2 (Figure 3.1). It is the second longest river and third largest catchment in Italy and it was selected as one of the research catchments in the FP7 project GLOBAQUA (Navarro-Ortega et al., 2015). Streamflow generation in the catchment is expected to change due to climate change in the area, with important consequences for water management, in particular for its headwaters (Chiogna et al., 2016; Majone et al., 2016). Hence, climate change has been identified as one of the main stressors affecting the Adige river basin. Further environmental stressors in the basin are the occurrence of hydropeaking related to hydropower production, and pollution associated with touristic fluxes especially in mountain areas, and pollutants transported within the river basin due to agricultural activities. The role played by particle-facilitated transport of contaminants is also under investigation (Navarro-Ortega et al., 2015), and therefore erosion could also be listed among the most relevant environmental stressors. To propose adequate management solutions to alleviate the effect of these stressors, it is necessary to have an accurate representation of flow and transport processes at the catchment scale. SWAT can be a useful tool to achieve these goals.
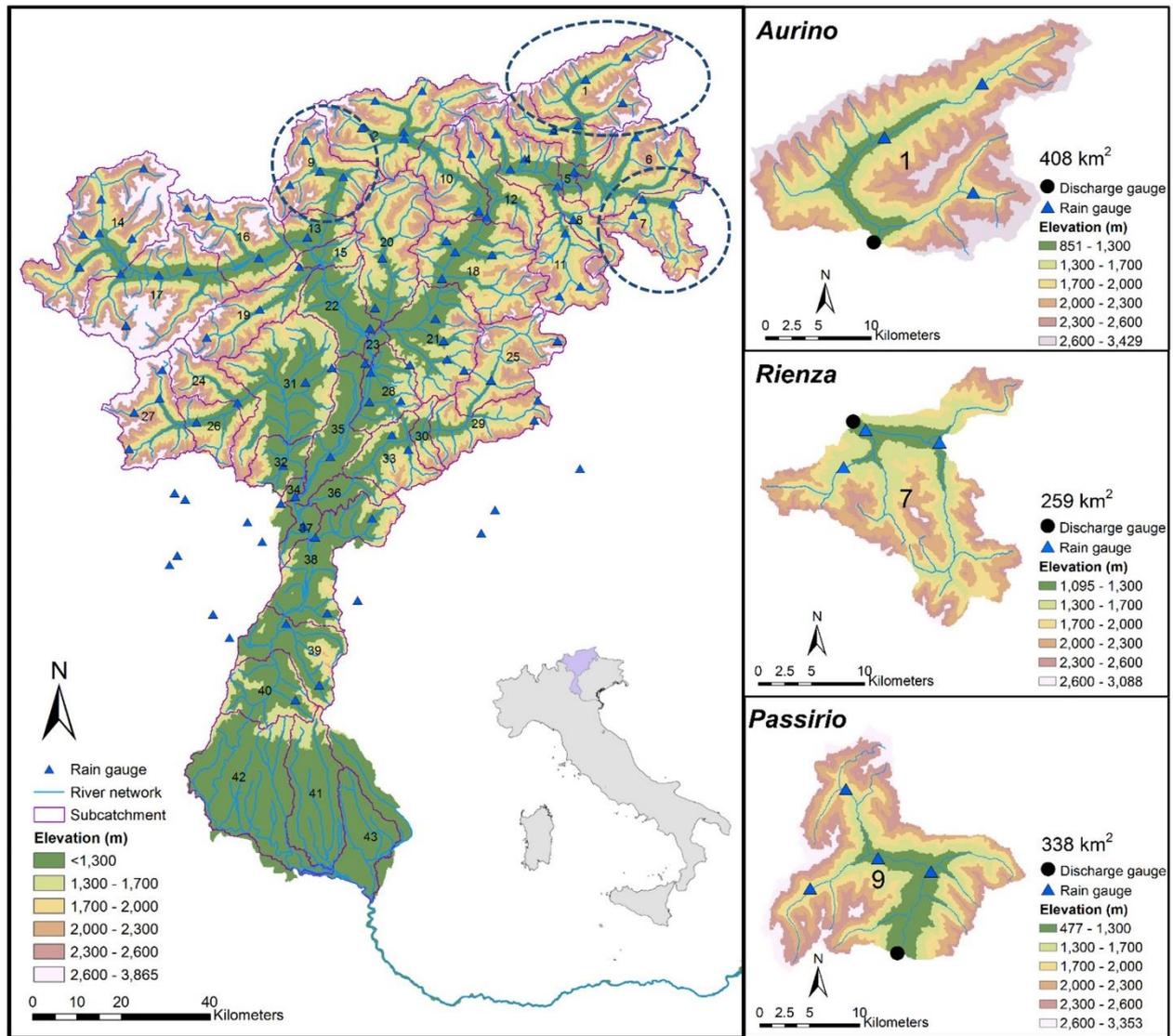
Figure 3.1 The Adige river basin and three studied mountainous alpine subbasins (Aurino, Rienza and Passirio).

The SWAT model has been set up for the entire Adige basin with 43 subbasins delineated in Figure 3.1. In this work, we focus on the results of three subbasins representing headwaters of the Adige river basin (Figure 3.1): Aurino (subbasin 1), Rienza (subbasin 7) and Passirio (subbasin 9). The three subcatchments are part of the larger SWAT model for the Adige river basin. Therefore, the three study areas have not been further subdivided into smaller subcatchments. These three subbasins were selected because they are headwaters of the basin and hydropower production does not exert a significant influence on their streamflow. This allows us to investigate the influence of different precipitation datasets on streamflow generation in mountainous catchments where high elevation gradients are present. The three subbasins are located near each other and belong to the same Alpine region. The climate of the region is characterized by snowmelt in spring, humid summers and autumns, and dry winters (Chiogna et al., 2016). According to the ground observations of the climate stations, the mean annual

precipitation is 909 mm, 767 mm and 1074 mm for the Aurino, Rienza and Passirio subcatchments, respectively; the mean daily average temperature is 3.6 °C, 6.0 °C and 8.5 °C for the Aurino, Rienza and Passirio subcatchments, respectively. According to the local soil investigation (Costantini et al., 2004), the major soils of three subbains are the same: the main types are Lithic Cryosols, Lithic, Mollic, Eutric, and Rendzic Leptosols, Eutric and Calcaric Cambisols, and Eutric Fluvisols. The main land use types of the three subbasins are forest, grassland, and barren land, which account of 49.2%, 14.6% and 22.2% in the Aurino subbasin, 64.6%, 9.3% and 14.1% in the Rienza subbasin, 49.0%, 23.9% and 12.6% in the Passirio subbasin, respectively. The land use map of the three subbasins is provided in the Supplementary material (Figure S3.1). The mean slope of the region is 51% (Costantini et al., 2004) and the slope map of each subbasin is offered in the Supplementary material (Figure S3.2). The elevation for Aurino, Rienza and Passirio ranges between 851-3429 m, 1095-3088 m, and 477-3353 m, respectively (Figure S3.3). The area of these three subbasins are 408 km^2, 259 km^2 and 338 km^2, respectively (Figure 3.1).

3.2.2   SWAT model

The SWAT model is a comprehensive, time-continuous, semi-distributed, process-based model (Arnold et al., 2012a). It was developed by the Agricultural Research Service of the United States Department of Agriculture (Arnold et al., 1998). SWAT can be used to model changes in hydrological processes, erosion, vegetation growth, and water quality in large river basins and evaluate the effects of climate change and water resources management (Abbaspour et al., 2015; Dile et al., 2016; Yang et al., 2016). It divides the river basin into subbasins and subsequently into Hydrologic Response Units (HRUs), characterized by different combinations of land use, soil characteristic, topography, and management schemes. The hydrological cycle is calculated based on water balance, which is controlled by climate inputs such as daily precipitation and maximum/minimum air temperature. Using daily input time series, SWAT simulates the daily, monthly and yearly fluxes of water and solutes in river basins. Simulations start by calculating the quantity of water, sediment and contaminants loading from land of each subbasin to the main channel. Then, these loads are transported and routed through the streams and reservoirs within the basin. More information about the model are provided in the literature (Arnold et al., 2012a; Arnold and Fohrer, 2005; Arnold et al., 1998; Gassman et al., 2007) and are available in the official model documentation (Neitsch et al., 2011).

3.2.3   Model setup

ArcSWAT 2012, with an interface in ArcGIS, was used to setup the model in this work. The datasets used in the model are listed in Table 3.1. Daily precipitation data come from 101 rain gauges (89 rain gauges inside the basin and 12 rain gauges within 25 km outside the entire Adige Basin boundary, see

Figure 3.1) for the period 1998-2010. This dataset has been used also for the IDW interpolation to generate an IDW-based precipitation dataset. Regarding the three subbasins of this study (Figure 3.1), three rain gauges with elevation of 1080 m, 1450 m and 1562 m are available at Aurino; three gauges with elevation of 1131 m, 1219 m and 1285 m are available at Rienza; four gauges with elevation of 644 m, 1147 m, 1618 m and 1716 m are available at Passirio. In this study, a total of twenty-four models have been set up for three Alpine subcatchments (Aurino, Rienza and Passirio) by inputting four different precipitation datasets into two types of SWAT programs: one with the elevation band method, the other without. The models considered a timeframe of thirteen years, from 1998 to 2010.

Table 3.1 Data source and description.

Data type	Scale	Data source
DEM	90 m × 90 m	Shuttle Radar Topography Mission (SRTM) produced by Consortium for Spatial Information (CGIAR-CSI)
Land use	100 m × 100 m	Corine Land Cover 2006(CLC2006) from European Environment Agency
Soil	1 : 1,500,000	Food and Agriculture Organization (FAO)
River network		EU-DEM product http://www.eea.europa.eu/data-and-maps/data/eu-dem.
Climate		Autonomous Province of Trento (http://www.meteotrentino.it) and Autonomous Province of Bolzano (http://www.provincia.bz.it/meteo/home.as)
River discharge		Autonomous Province of Bolzano

3.2.4 Precipitation datasets

In the present work, four precipitation datasets covering 1998-2010 have been used as SWAT input datasets. They can be classified into three categories: 1) Observed precipitation data from ground rain gauges; 2) IDW-based interpolated data, which were used as one representative of the datasets by interpolation techniques (Ly et al., 2011); 3) two satellite precipitation products: CHIRPS (Funk et al., 2015), which represents one of the latest products from remote sensing community and TRMM 3B42 (Huffman et al., 2007), which represents one of the frequently applied products from the remote sensing community. The two satellite precipitation products have been selected because they performed best in the Adige river basin among eight satellite precipitation products (Duan et al., 2016). The empirical cumulative density function (CDF) of the daily precipitation distribution during 1998-2010 has been

computed for the four datasets. The corresponding cumulative frequency has been evaluated based on the rainfall intensity classification of the World Meteorological Organization (WMO) standard (Tan et al., 2015): (1) rain < 1 mm (no/tiny rain), (2) 1 mm ≤ rain < 2 mm (light rain), (3) 2 mm ≤ rain < 5 mm (low moderate rain), (4) 5 mm ≤ rain < 10 mm (high moderate rain), (5) 10 mm ≤ rain < 20 mm (low heavy rain), (6) 20 mm ≤ rain < 50 mm (high heavy rain), and (7) rain ≥ 50 mm (violent rain).

3.2.4.1 Ground rain gauge dataset

All available precipitation data collected by pointed-based ground rain gauges have been directly used as input to SWAT. This represents the most widely used manner in which the rain gauge data are used in the SWAT model among all existing publications. However, SWAT only uses the data of the rain gauge closest to the centroid of each subbasin, disregarding all other stations (Galván et al., 2014; Masih et al., 2011). Therefore, only data from one single point has been utilized for the entire subbasin without considering any spatial heterogeneity. This precipitation dataset will hereafter be referred to as OP (observed precipitation) and accordingly the model result will be referred to as SWAT model using OP input (OP model).

3.2.4.2 IDW-based precipitation dataset

A huge variety of methods have been developed in the past to spatially interpolate precipitation datasets (Ly et al., 2011; Wagner et al., 2012; Zhang and Srinivasan, 2009). Babak and Deutsch (2009) state that "Variants of kriging are often proposed as statistical techniques with superior mathematical properties such as minimum error variance; however, the robustness and simplicity of IDW interpolation motivate its continued use in practice." The IDW method is probably one of the simplest interpolation methods and it does not rely on particular statistical assumptions. It is a widely used geometric interpolation method and is considered in most intercomparison studies focusing on the effect of different precipitation inputs on hydrological outputs (Chu et al., 2012; Di Luzio et al., 2008; Ly et al., 2011; Ly et al., 2013; Shen et al., 2012; Shope and Maharjan, 2015; Shope et al., 2014; Tuppad et al., 2010; van der Heijden and Haberlandt, 2010; Wagner et al., 2012; Yang et al., 2015; Zhang and Srinivasan, 2009). IDW estimates values at unknown points by the weighted average of observed data at neighboring points:

$$z(x_o) = \sum_{i=1}^{n} \lambda_i \cdot z(x_i) \tag{3.1}$$

$$\lambda_i = \frac{|d_{oi}|^{-d}}{\sum_{i=1}^{n} |d_{oi}|^{-d}} , d > 0 \tag{3.2}$$

where $z(x_o)$ represents the rain data of the unknown point, $z(x_i)$ is the rain data of the rain gauge i; λ_i is the weight as defined in Eq. (3.2), d_{oi} is the distance between the unknown point and the rainfall station i. The parameter d was set to 2, according to the recommendation of Ly et al. (2011).

IDW therefore depends only on the distance between stations and not on their elevation or on direction. This method is based on the assumption that points close to each other are more correlated than points at larger distances (Ly et al., 2011). This interpolation scheme has been chosen for its simplicity, without excluding the possibility that more complex interpolation schemes could improve the results obtained using IDW.

In this work, thirteen years (1998-2010) of daily measured data have been interpolated using the IDW tool in ArcGIS considering the 12 closest stations for each interpolated grid (100 m ×100 m). The number of stations was set as 12 by following the recommendations provided in previous works (Babak and Deutsch, 2009; Ly et al., 2011) and also considering the morphology of the catchments. Averaged IDW daily values for the subcatchment have been calculated and used as precipitation input at the centroid of the subbasin, which, in this way, consider the spatial heterogeneity in precipitation patterns. This IDW-based precipitation dataset will hereafter be referred to as IDW, and the corresponding SWAT model will be referred to as the IDW model.

3.2.4.3 CHIRPS dataset

The CHIRPS product provides daily precipitation data at spatial resolution of 0.05° for the quasi-global coverage of 50°N-50°S from 1981 to near present. The latest product is the Version 2.0 dataset that was completed and released in February 2015. The daily precipitation data from CHIRPS were downloaded at: http://chg.geog.ucsb.edu/data/chirps/. The CHIRPS product is based on integration of various datasets: the monthly precipitation climatology (CHPclim) that is created using rain gauge stations collected from FAO and GHCN, the Cold Cloud Duration (CCD) information based on thermal infrared data archived from CPC and NOAA National Climate Data Center (NCDC), the Version 7 TRMM 3B42 data, the Version 2 atmospheric model rainfall field from the NOAA Climate Forecast System (CFS), and the rain gauge stations data from multiple sources. Since rain gauge data are used for bias correction in the product, the CHIRPS product belongs to the "satellite-gauge" category. More detailed information on CHIRPS can be found in Funk et al. (2015). For each subbasin, daily averaged subbasin CHIRPS datasets for the period 1998-2010 have been calculated by averaging all effective 0.05° daily CHIRPS grids within the subbasin boundary, which were then used to force the SWAT model in this study (the CHIRPS model).

3.2.4.4 TRMM 3B42 dataset

The TRMM 3B42 product is one type of the TMPA (TRMM Multi-satellite Precipitation Analysis) products (Huffman et al., 2007). The TRMM 3B42 product provides 3-hourly precipitation at the spatial resolution of 0.25° for the quasi-global coverage of 50°N-50°S from 1998 to 2015. The latest product version is Version 7 and the applied algorithm is the TMPA algorithm that combines precipitation

estimates from microwave and infrared satellites, as well as the GPCC monthly gauge analysis. More details about TMPA algorithms can be found in Huffman et al. (2007). The TRMM 3B42 daily precipitation data were obtained from Goddard Earth Sciences Data and Information Services Center at http://mirador.gsfc.nasa.gov. The mean daily accumulated TRMM 3B42 data during 1998-2010 have been calculated for each subbasin and used as the SWAT inputs in this study. The TRMM 3B42 precipitation data will hereafter be referred to as TRMM, and the corresponding SWAT model will be referred to as the TRMM model.

3.2.5   Elevation bands

To consider the orographic effects on precipitation and temperature in mountainous areas, SWAT uses the elevation bands method which allows for up to ten elevation bands in each subbasin. In this work, five elevation bands have been applied. As shown in Eqs. (3.3) and (3.4), the method modifies the regional precipitation by weighting the elevation difference between the band of the rain gauge and the bands.

$$R_{band} = R_{day} + \left(EL_{band} - EL_{gauge}\right) \cdot \frac{plaps}{days_{pcp,yr} \cdot 1000}, R_{day} > 0.01 \qquad (3.3)$$

$$R_{day} = \sum_{bnd=1}^{b} R_{band} \cdot fr_{bnd} \qquad (3.4)$$

where R_{band} is the precipitation in the elevation band (mm), R_{day} is the precipitation recorded at the rain gauge (mm), EL_{band} is the mean elevation at the elevation band (m), EL_{gauge} is the elevation at the recording gauge (m), $plaps$ is the precipitation lapse rate (mm/km) and $days_{pcp,yr}$ is the average number of days of precipitation in the subbasin in a year, fr_{bnd} is the fraction of the subbasin area within the elevation band and b is the total number of elevation bands in the subbasin. Notice that in addition to Eqs. (3.3) and (3.4) SWAT imposes the following condition: if $R_{band} < 0$, then $R_{band} = 0$. This means that if $EL_{gauge} > EL_{band}$ and $abs\left(\left(EL_{band} - EL_{gauge}\right) \cdot \frac{plaps}{days_{pcp,yr} \cdot 1000}\right) > R_{day}$, precipitation of that band is set equal to 0. This condition prevents the elevation band method from being a constant adjustment of the input precipitation data. Galván et al. (2014) has pointed out that this method generally either underestimates or overestimates precipitation due to the difference between the altitude of the subbasin and the elevation of the rain gauge. In addition, they found that $plaps$ is erroneously introduced in mm/m instead of mm/km as indicated. Although it has some intrinsic limitations, this method has been proven to be useful in several Alpine catchments (Grusson et al., 2015; Rahman et al., 2013)

3.2.6   Model calibration, evaluation and uncertainty analysis

The first three years (1998-2000) have been used as the warm up period to mitigate the effect of initial conditions. For each SWAT model, monthly simulated stream flow of the three alpine subbasins have

been calibrated separately using the time period 2001-2005 and then validated in the period 2006-2010, based on the measured discharge records of stations Aurino-Caminata, Rienza-Monguelfo and Passirio-Saltusio for Aurino, Rienza and Passirio, respectively (Figure 3.1). The automatic calibration and validation were performed by using the Sequential Uncertainty Fitting algorithm version 2 (SUFI-2) (Abbaspour et al., 2004; Abbaspour et al., 2007) in the SWAT-CUP tool package (Abbaspour, 2015). Snow parameters can be very influential in SWAT (Grusson et al., 2015). Before calibration, we therefore applied snow parameters available in the literature for neighboring regions (Adler et al., 2015; Rahman et al., 2013; Zanotti et al., 2004) to avoid the potential interference of snow processes on the interpretation of the results obtained using different precipitation inputs. The sensitivity analysis has been performed by the one-at-a-time procedure of SWAT-CUP (Abbaspour, 2015) for several common sensitive hydrological parameters (Table 3.2) to select the sensitive hydrological parameters for each subbasin. This procedure tests the model sensitivity by changing one parameter while keeping all other parameters constant. Furthermore, starting with the initial ranges of parameters shown in Table 3.3, the models were calibrated with four iterations. The initial parameter ranges are qualified to physically reasonable intervals according to the literature (Grusson et al., 2015; Vu et al., 2012) and SWAT official documentation (Arnold et al., 2012b). For each iteration, 1500 simulations were run. After each iteration, the ranges of the parameters have been modified (normally narrowed down) according to both the new parameters suggested by the program (Abbaspour et al., 2004; Abbaspour et al., 2007) and their reasonable physical limitations. More details about the protocol to calibrate the model can be found in Abbaspour (2015) and Abbaspour et al. (2015).

Table 3.2 Hydrological parameters considered for sensitivity analysis (''a__'', ''v__'' and ''r__'' means an absolute increase, a replacement, and a relative change to the initial parameter values, respectively).

Parameters	Description	Range	Defualt
a_SOL_AWC().sol	Available water capacity of the soil layer[mm H_2O/mm soil]	0/0.9	Soil layer specific
a_SOL_K().sol	Saturated hydraulic conductivity[mm/hr]	-10/10	Soil layer specific
r_SOL_BD().sol	Moist bulk density[g/cm^3]	-0.5/0.5	Soil layer specific
a_CN2.mgt	SCS runoff curve number	-20/20	HRU specific
v_ESCO.hru	Soil evaporation compensation factor	0/1	0.95
v_EPCO.hru	Plant uptake compensation factor	0/1	1
a_HRU_SLP.hru	Average slope stepness [m/m]	-0.2/0.4	HRU specific
a_SLSUBBSN.hru	Average slope length [m]	-9/130	HRU specific
a_OV_N.hru	Manning's "n" value for overland flow	0.01/29	0.1

v_CH_K2.rte	Effective hydraulic conductivity [mm/hr]	0/400	0
v_CH_N2.rte	Manning's nvalue for main channel	0/0.3	0.014
a_GWQMN.gw	Threshold depth of water in the shallow aquifer required for return flow to occur [mm]	-500/500	1000
a_REVAPMN.gw	Threshold depth of water in the shallow aquifer for "revap" to occur [mm]	-500/500	750
v_GW_REVAP.gw	Groundwater "revap" coefficient	0.02/0.2	0.02
v_GW_DELAY.gw	Groundwater delay [days]	0/300	31
v_ALPHA_BF.gw	Baseflow alpha factor [days]	0/1	0.048

Several metrics are available in the literature to evaluate model performance (e.g., Bennet et al., 2013). In this work, we follow the approach suggested in Abbaspour et al. (2015) where model performance was evaluated taking into account the parameter uncertainties and 95% prediction uncertainty (95PPU) of the outputs by using SUFI-2 (Yang et al., 2008). The Nash–Sutcliffe coefficient (NS) (Nash and Sutcliffe, 1970) and the coefficient of determination (R^2) have been used as the goodness of fit indicators for the best simulation. NS measures the quantity difference between the predictions and the observed data, with NS = 1 being the optimal value. R^2 ranges from 0 to 1 and represents the trend similarity between the observed data and the simulated ones, with higher R^2 values indicating better model performance. The model performance has been classified using the NS value according to the work of Moriasi et al. (2007): unsatisfactory performance (NS≤ 0.50), satisfactory performance (0.50<NS≤0.65), good performance (0.65<NS≤0.75) and very good performance (0.75<NS≤1.00).

As described by Abbaspour (2015) and Abbaspour et al. (2007), uncertainty in parameters are expressed as the final ranges of the parameter sets with which the model reached the satisfactory result. Hence, these parameter uncertainties lead to the output uncertainty 95PPU which is calculated at the 2.5% and 97.5% levels of the cumulative distribution of an output variable generated through the propagation of the parameter uncertainties using Latin hypercube sampling. The uncertainty analysis of SUFI-2 is based on the theory that the model performance is represented by an envelope of good solutions expressed by the 95PPU, generated by certain parameter ranges, rather than a single signal.

To quantify the prediction uncertainties, two indices were introduced (Abbaspour et al., 2004): P-factor and R-factor. The P-factor is the fraction of measured data enveloped by the 95PPU band. It ranges from 0 to 1, in which 1 is optimal and indicates 100% bracketing of the observed data within model prediction uncertainty. The R-factor is the thickness of the 95PPU envelop, which means the ratio of the average width of the 95PPU band and the standard deviation of the measured variable. For discharge, P-

factor >0.7 and R-factor<1.5 are considered acceptable in terms of prediction uncertainty (Abbaspour, 2015).

Table 3.3 Initial parameter ranges for calibration. Final calibrated ranges for the different precipitation products are available in the Supplementary material (Figure S3.4-S3.6).

Parameters	Default	Calibration range
Aurino		
v__PLAPS.sub	0	0/0.15
a__SOL_AWC().sol	0.01─0.06	0/0.9
a__CN2.mgt	55─94	-20/4
v__ESCO.hru	0.95	0/1
a__SLSUBBSN.hru	9.146─15.244	-9/130
a__GWQMN.gw	1000	-300/300
a__REVAPMN.gw	750	-300/300
v__GW_REVAP.gw	0.02	0.02/0.2
v__GW_DELAY.gw	31	0/300
v__ALPHA_BF.gw	0.048	0/1
v__CH_K2.rte	0	0/400
Rienza		
v__PLAPS.sub	0	0/0.15
a__SOL_AWC().sol	0.06	0/0.9
a__CN2.mgt	55─60	-20/38
v__ESCO.hru	0.95	0/1
a__SLSUBBSN.hru	9.146─18.293	-9/130
a__HRU_SLP.hru	0.183─0.719	-0.2/0.4
a__GWQMN.gw	1000	-300/300
v__GW_REVAP.gw	0.02	0.02/0.2
v__GW_DELAY.gw	31	0/300
v__CH_K2.rte	0	0/400
Passirio		
v__PLAPS.sub	0	0/1
a__SOL_AWC().sol	0.01─0.06	0/0.9
a__CN2.mgt	55─84	-20/14
v__ESCO.hru	0.95	0/1
a__SLSUBBSN.hru	9.146─18.293	-9/130
a__GWQMN.gw	1000	-300/300
v__GW_REVAP.gw	0.02	0.02/0.2
v__GW_DELAY.gw	31	0/300
v__CH_K2.rte	0	0/400

3.3 Results and discussion

3.3.1   Comparison of the four different precipitation datasets

Considering the Aurino subbasin (Figure 3.2), the four products display different probability of occurrence of dry days (rain=0 mm/d), which are 57%, 31%, 67% and 45% for OP, IDW, CHIRPS and TRMM, respectively. The largest difference in the CDFs of the four products occurs for precipitation events smaller than moderate rain (< 2 mm/d): particularly, for the occurrence of tiny rain, the probabilities are 12%, 35%, 4% and 5% for OP, IDW, CHIRPS and TRMM, respectively. OP, IDW and CHIRPS show great similarity for low moderate rain (8%, 9%, and 7%, respectively), while TRMM displays a larger probability (18%). Results show negligible differences (less than 2%) of occurrence rates for rainfall larger than 5 mm/d between the four precipitation datasets. At Rienza (Figure 3.3), the occurrence of dry days shows the largest variability among the four datasets (62% for OP, 31% for IDW, 67% for CHIRPS, and 58% for TRMM). The occurrence of tiny rain presents the largest discrepancy and the probabilities are 10%, 31%, 4% and 3% for OP, IDW, CHIRPS and TRMM, respectively. For other precipitation events including light, moderate, heavy and violent rain, OP, IDW and CHIRPS datasets are quite similar, while TRMM has relative higher CDF values than the other three products for values up to 40 mm/d. In the Passirio subcatchment (Figure 3.4), the probabilities of dry day are 65%, 30%, 70% and 40% for OP, IDW, CHIRPS and TRMM, respectively. The largest difference of the four datasets occurs for the tiny rain (7% for OP, 38% for IDW, 4% for CHIRPS and 5% for TRMM). As shown in Figure 3.4, the CDF of the TRMM dataset evidently deviates from the others: higher cumulative rates in light, low moderate, and heavy rains.
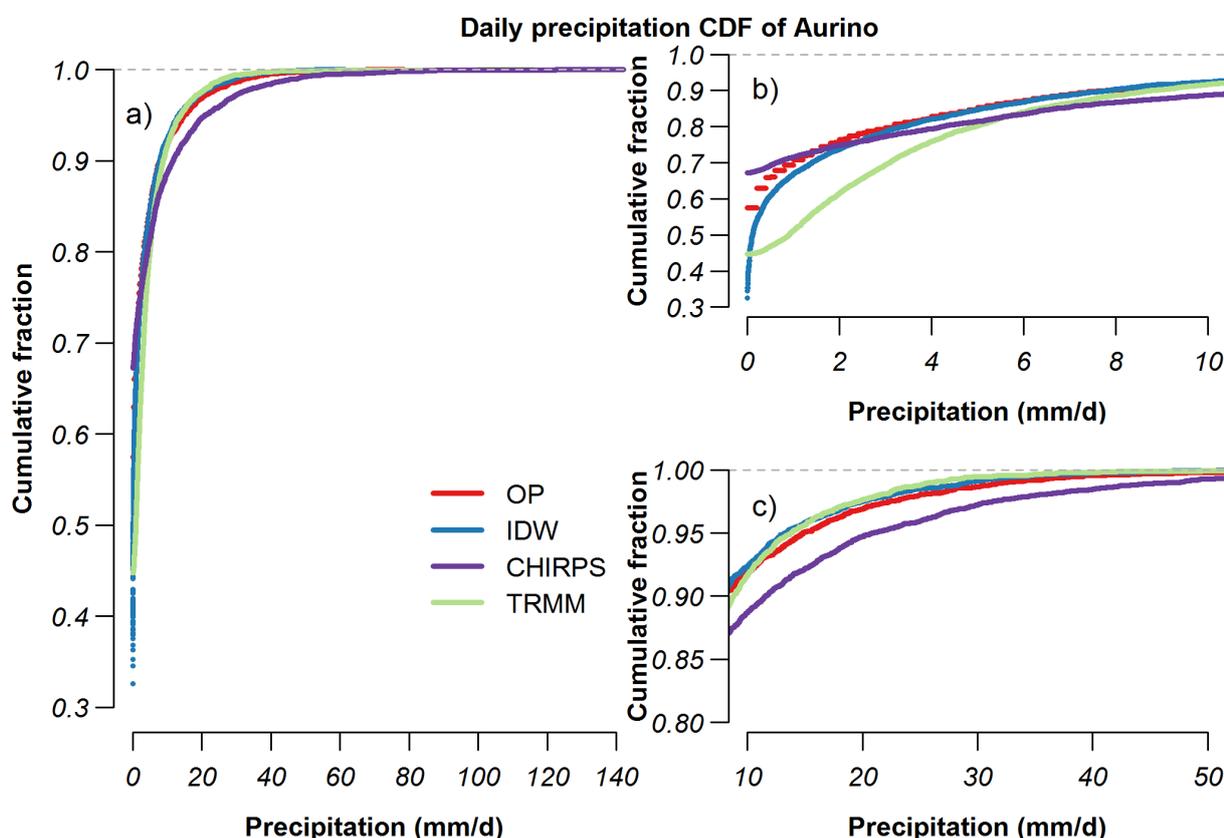
Figure 3.2 Distribution of daily precipitation values of the four precipitation inputs (OP, IDW, CHIRPS, TRMM) at Aurino: a) distribution of all precipitation values; b) distribution of precipitation <10 mm; c) distribution of 10 mm ≤ precipitation < 50 mm.
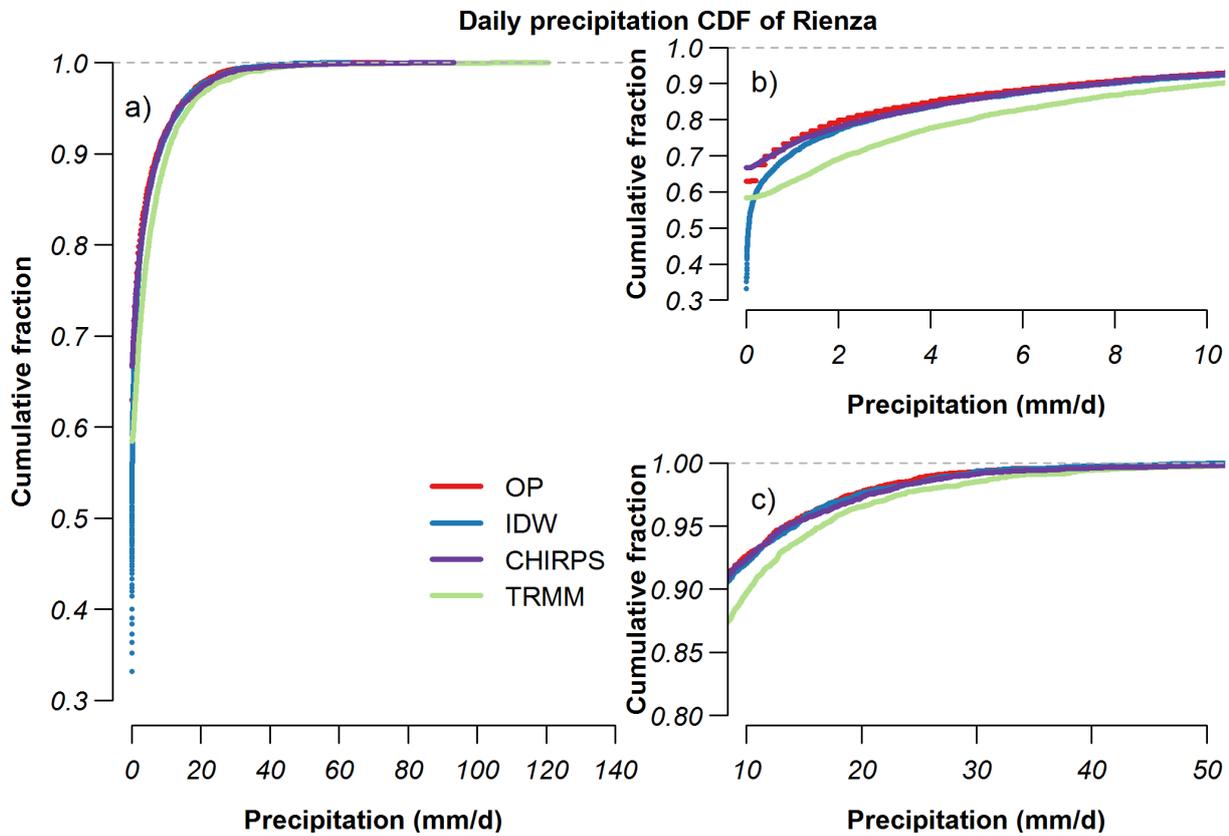
Figure 3.3 Distribution of daily precipitation values of the four precipitation inputs (OP, IDW, CHIRPS, TRMM) at Rienza: a) distribution of all precipitation values; b) distribution of precipitation <10 mm; c) distribution of 10 mm ≤ precipitation < 50 mm.
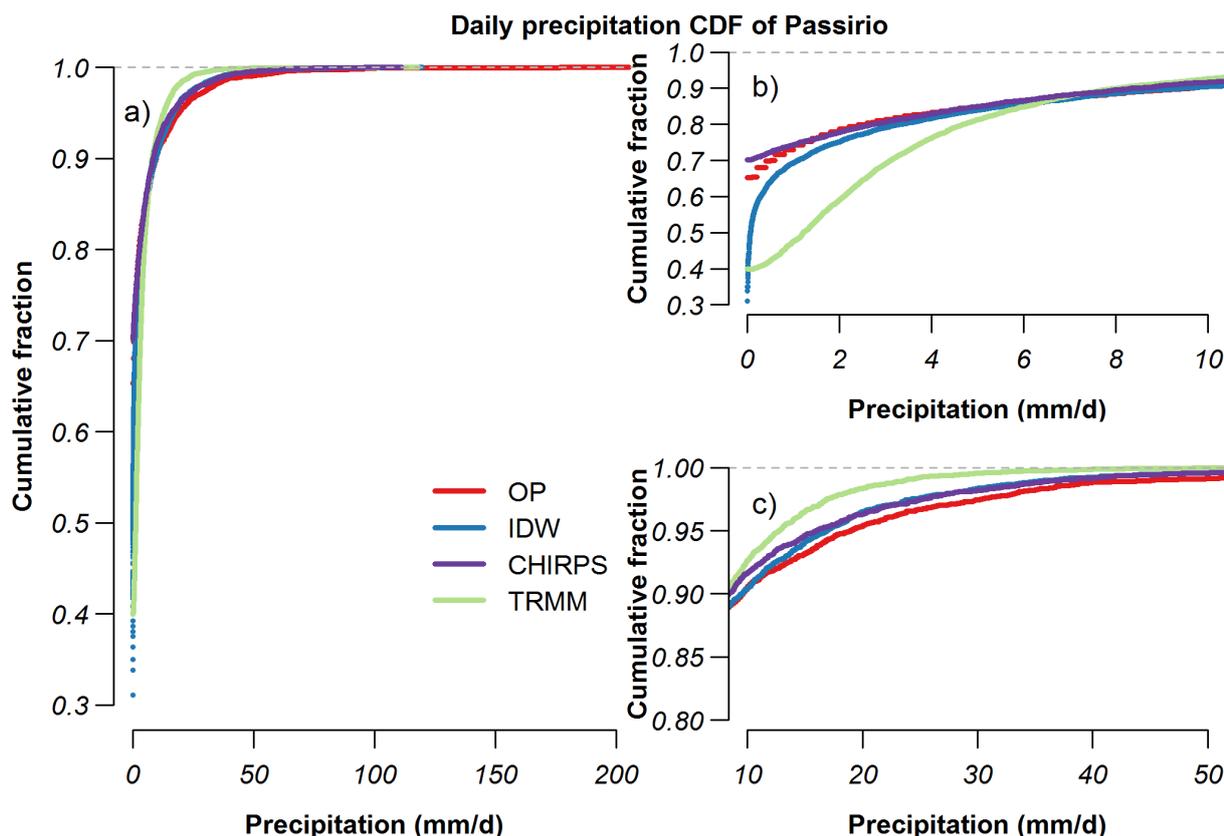
Figure 3.4 Distribution of daily precipitation values of the four precipitation inputs (OP, IDW, CHIRPS, TRMM) at Passirio: a) distribution of all precipitation values; b) distribution of precipitation <10 mm; c) distribution of 10 mm ≤ precipitation < 50 mm.

In general, IDW datasets have the largest occurrence probability of days with tiny rain (35%, 31% and 38% for the Aurino, Rienza and Passirio subcatchments, respectively), which is ascribed to the effect of the interpolation method. In fact, the precipitation value at a given point of the interpolated dataset is influenced by the nearest 12 observation stations (Figure 3.1). Hence, the IDW dataset of each subbasin is influenced also by stations outside the subbasin which may have recorded local precipitation events. The results of a set of two-sample Kolmogorov-Smirnov tests at the 5% significance level show that the four precipitation inputs have different statistical distributions in the three subcatchments. In particular, as shown in Figure 3.5, their monthly precipitation patterns throughout the period 2001-2010 are not identical, and hence they represent four different input datasets for the hydrological model. The mean standard deviation (i.e., standard deviation of the four precipitation datasets computed at the monthly time scale and averaged over the 10 years considered) is 31mm, 27 mm and 30 mm for the Aurino, Rienza and Passirio subcatchments, respectively.
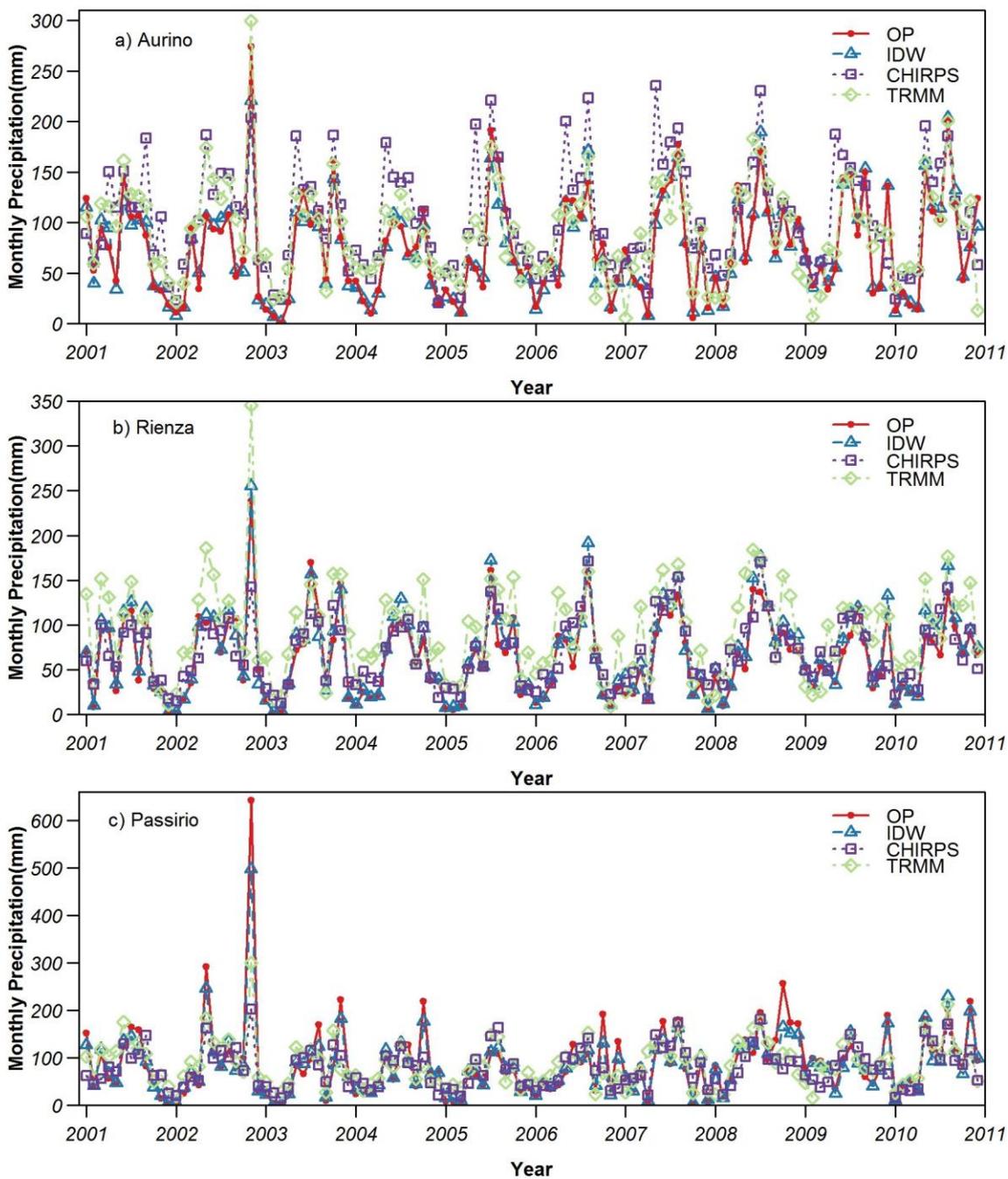
Figure 3.5 Monthly precipitation of the four precipitation datasets in the three Alpine catchments: a) Aurino, b) Rienza, and c) Passirio.

3.3.2 Model performance without elevation band

As shown in Figure 3.6, it is evident that the four precipitation inputs fail in promoting the model to reproduce the discharge records at all three subbasins. In particular, the peak flow in summer is underestimated by more than 50%. According to the model performance classification of Moriasi et al. (2007), the only model which achieves satisfactory performance($0.50<NS\leq0.65$) is the one for the

Aurino subcatchment using the CHIRPS precipitation input. All the other models achieve unsatisfactory results (NS ≤ 0.50).
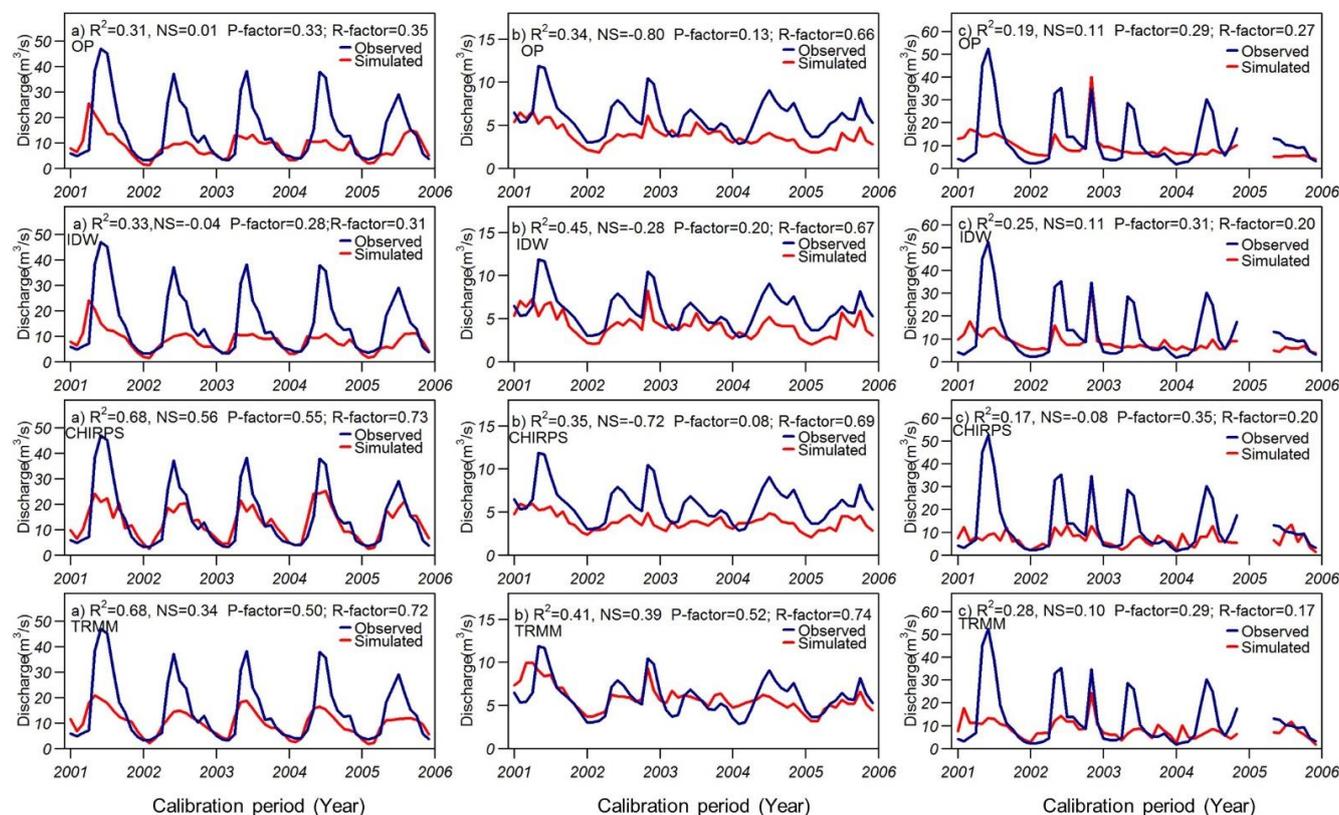
Figure 3.6 Calibration results of the three subbasin without elevation band correction: a) Aurino, b) Rienza, c) Passirio. The results represent monthly stream flow data of the best simulation, i.e., highest NS and R^2 values obtained in a set of 1500 simulations (red lines), and observed stream flow data (blue lines), along with the prediction uncertainty (P-factor and R-factor).

A rough calculation of the ratio of annual observed streamflow to annual precipitation leads to anomalously large values: mean values computed for the period 2001-2010 using OP dataset are 1.24±0.20 for Aurino, 0.91±0.16 for Rienza, and 1.02±0.26 for Passirio; mean values using IDW based dataset are 1.29±0.21 for Aurino, 0.84±0.14 for Rienza, and 1.07±0.25 for Passirio; mean values using CHIRPS dataset are 0.89±0.13 for Aurino, 0.84±0.15 for Rienza, and 1.11±0.26 for Passirio; mean values using TRMM dataset are 1.04±0.14 for Aurino, 0.62±0.15 for Rienza, and 1.00±0.20 for Passirio. All four datasets often underestimate the total precipitation of the studied subbasins. The underestimation is most likely due to the poor representation of the spatial variability of precipitation patterns in the region, thereby leading to the high ratio of streamflow-to-precipitation and results shown in Figure 3.6. Since the IDW model is driven by interpolated data, its results depend on the ground observations. Therefore, since the ground observations systematically underestimate the total

precipitation in the study area, the interpolated data will also underestimate it. Moreover, the underestimation of precipitation in the two remote sensing datasets CHIRPS and TRMM could be attributed to the fact that the topographic effects have not been considered in the bias correction using rain gauge analysis. Additionally, satellite precipitation estimates have their own uncertainties (Duan et al., 2016). Therefore, as a consequence of the underestimation of precipitation, it is not possible to close the water balance for the three subbasins considering the available products. We conclude that the four tested precipitation products considerably underestimate the amount of precipitation over a large part of the Adige headwaters. Such underestimation is common in Alpine catchments (Isotta et al., 2014). Indeed, the high variability in the morphology and orography of Alpine catchments, leads to a high heterogeneity in precipitation patterns and intensity (Panziera et al., 2015). Ground observations, used also for the validation of satellite products, are normally located at low elevations (below 2000 m a.s.l.) and are not representative for the entire catchment (Duan and Bastiaanssen, 2013; Javanmard et al., 2010). Only 3% of the rain gauges are located at elevations higher than 2000 m a.s.l. in the upper part of the Adige river basin, while more than 30% of the Adige river basin closed at Bronzolo is above 2000 m a.s.l. (Adler et al., 2015). Mei et al. (2016a) also found streamflow/precipitation ratios larger than 1 in several subbasins (Aurino, Passirio, Isarco, Rienza) of the Adige river basin. Furthermore, Mei et al. (2014; 2016b) also identified the critical role of precipitation input for this region in case of flood protection and assessed the quality of some available satellite products for this region. Our work therefore complements the current knowledge about these catchments involving different precipitation inputs, different spatial scales and different hydrological models.

3.3.3 Model performance with elevation band method

SWAT is designed to cope with the aforementioned underestimation problems related to precipitation input using elevation bands (Eq. (3.3) and (3.4)) which modulates the amount of precipitation depending on the orography of the catchment.Figure 3.7, Figure 3.8 and Figure 3.9 show evident improvements in the performances of the model for all three subbasins.

For the Aurino subbasin (Figure 3.7), all models driven by the four different precipitation input datasets well reproduce the measured streamflow. Using the performance classification of Moriasi et al. (2007), the model using OP as precipitation data achieves good performance in the calibration period and very good performance in the validation one, while the models utilizing IDW, CHIRPS and TRMM data attain very good performance. The IDW model reaches the highest NS (0.91) and R^2 (0.91) values in the validation period. For the Rienza subcatchment (Figure 3.8), during the calibration period, the use of OP and CHIRPS data lead to satisfactory model performance. The IDW model attains the level of good model performance, while TRMM model does not achieve satisfactory model performance. During the

validation period (Figure 3.8), NS and R^2 increase for all precipitation datasets. The IDW and OP model achieve very good model performance, while CHIRPS model does not exceed satisfactory performance. The best simulation obtained using the TRMM dataset improves to the level of good performance. In the Rienza subcatchment, considering both calibration and validation periods, the IDW model reaches the best performance; OP and CHIRPS model show a consistent satisfactory performance in reproducing streamflow. For the Passirio subcatchment (Figure 3.9), during the calibration period, the IDW model attains good model performance, the CHIRPS model obtains satisfactory performance, while the performances of OP model and TRMM model are unsatisfactory. During the validation period, all models obtain higher NS and R^2 values: the IDW model reaches very good performance, the CHIRPS model improves to good performance, while the OP model achieves satisfactory performance. The performance of TRMM model is still unsatisfactory.
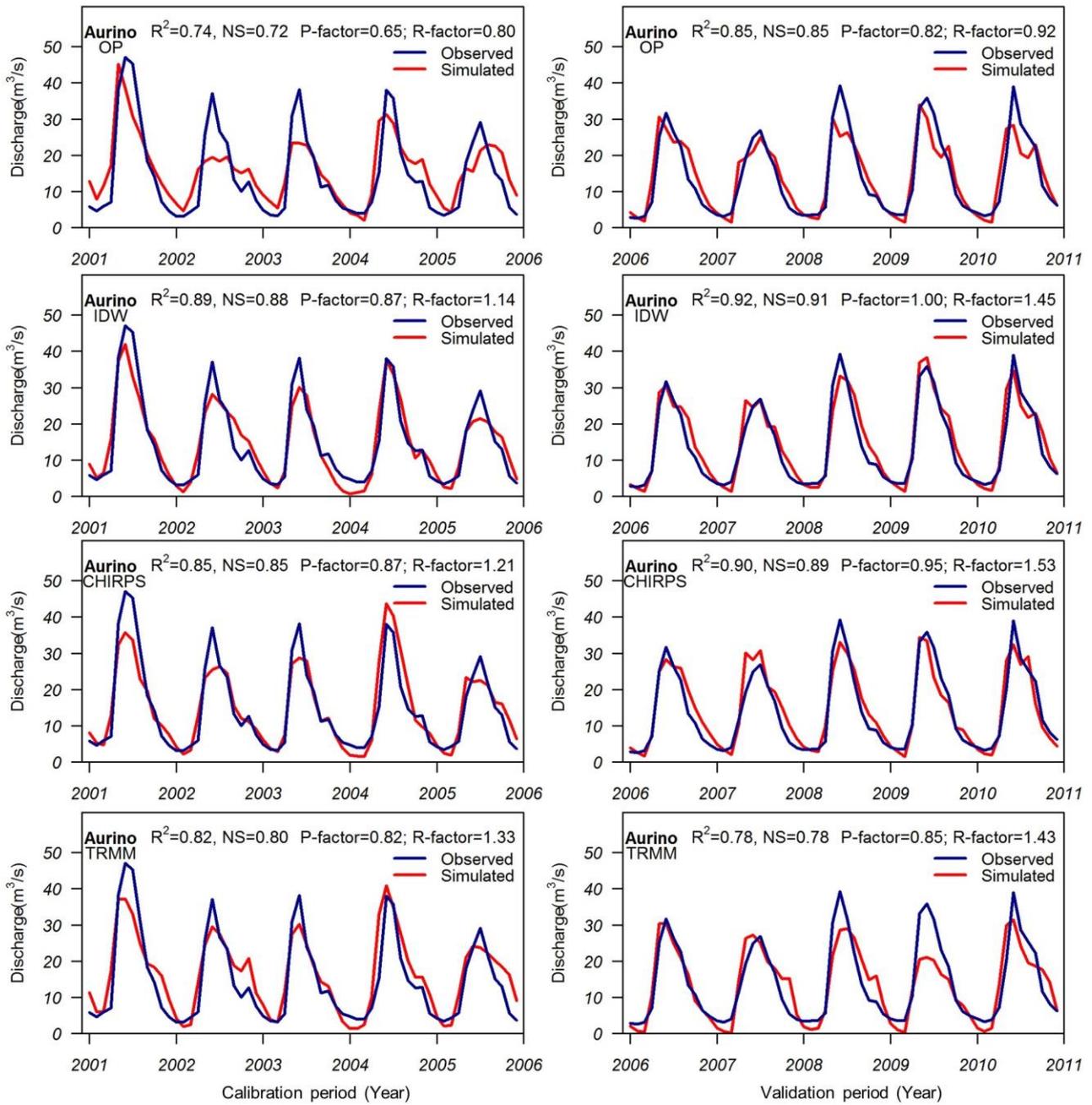
Figure 3.7 Calibration (2001-2005) and validation (2006-2010) results of subbasin Aurino with elevation band method. The results represent monthly stream flow data of the best simulation, i.e., highest NS and R^2 values obtained in a set of 1500 simulations (red lines), and observed stream flow data (blue lines), along with the prediction uncertainty (P-factor and R-factor).
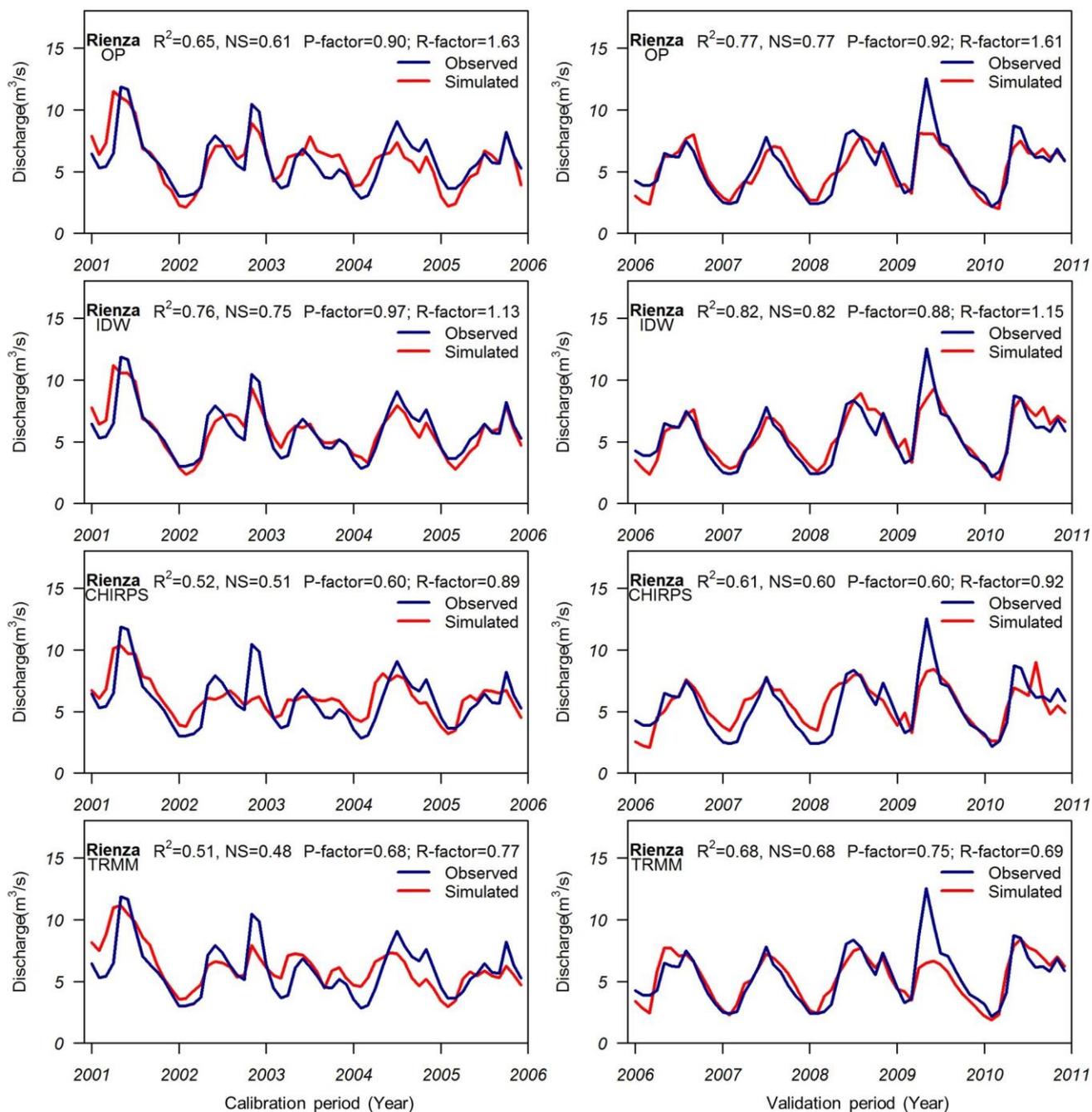
Figure 3.8 Calibration (2001-2005) and validation (2006-2010) results of subbasin Rienza with elevation band method. The results represent monthly stream flow data of the best simulation, i.e., highest NS and R^2 values obtained in a set of 1500 simulations (red lines), and observed stream flow data (blue lines), along with the prediction uncertainty (P-factor and R-factor).
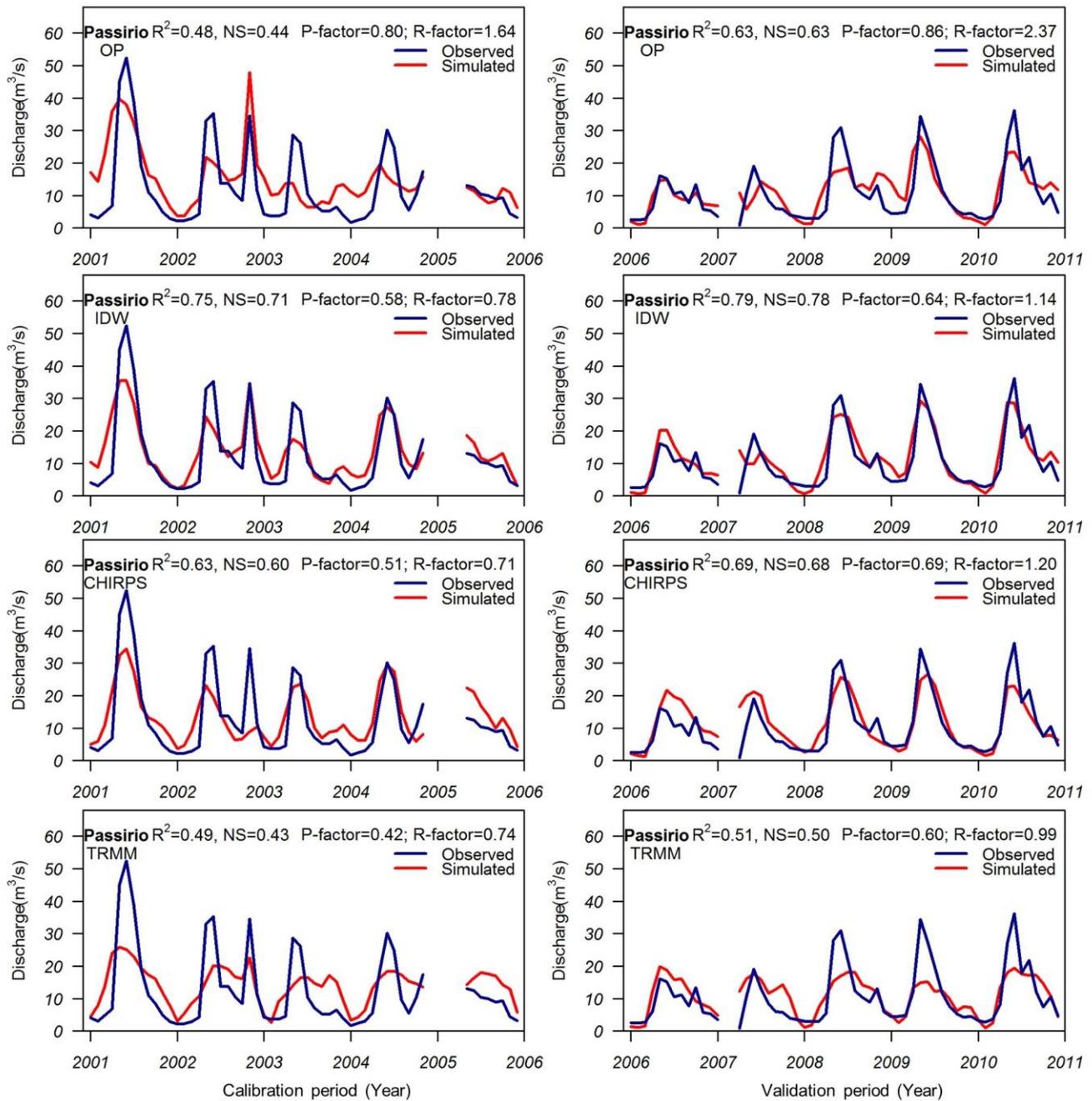
Figure 3.9 Calibration (2001-2005) and validation (2006-2010) results of subbasin Passirio with elevation band method. The results represent monthly stream flow data of the best simulation, i.e., highest NS and R^2 values obtained in a set of 1500 simulations (red lines), and observed stream flow data (blue lines), along with the prediction uncertainty (P-factor and R-factor).

In summary, the introduction of elevation bands, despite the limitations discussed in the method section, greatly improves model performance regardless of which precipitation input is used. We can see in Table 3.4 that elevation bands have a large effect (increase in precipitation between 10% and 103% by the elevation band method) on all precipitation products, although CHIRPS and TRMM consistently show the need for a lower correction in comparison to OP and IDW. However, we can observe differences in

66

the effect of elevation bands among different subcatchments, which prevents the generalization of the conclusions obtained for a specific catchment to other Alpine catchments. Despite the model performance improvements, clear differences are still present in the model results depending on the precipitation dataset applied for the different case studies. Models utilizing IDW data have the best performance in terms of NS and R^2. Concerning satellite products, the CHIRPS dataset is a feasible choice with at least satisfactory model performances at all subbasins.

Table 3.4 Average increase ratio (i.e., relative amount of precipitation added to the original value using elevation band method) of precipitation with elevation band method in order to obtain the best simulations over the calibration period for each subbasin shown in Figure 3.7-3.9.

	OP	IDW	CHIRPS	TRMM
Aurino	0.78±0.12	0.91±0.11	0.32±0.02	0.47±0.11
Rienza	0.49±0.07	0.45±0.05	0.39±0.03	0.10±0.02
Passirio	1.03±0.24	0.89±0.20	0.87±0.16	0.70±0.20

The best performance of models with IDW is probably caused by the distinct behavior of the IDW datasets that have a larger occurrence probability of rainfall days (Figure 3.2-3.5). As a result of interpolation, IDW precipitation time series have more rainfall days than the other datasets in form of tiny rain and more than 80% of them have intensity larger than 0.01 mm/d (Figure 3.2-3.5). According to the equation of the elevation band method (Eq. (3.3) and (3.4)), daily precipitation data will be modified only when they are larger than 0.01 mm/d. A larger amount of days corrected with the elevation band method increases the probability of obtaining a better model result for IDW models (Figure 3.10) as discussed in the following section in details. Besides, the similarities in the CDF of OP, IDW and CHIRPS products lead to the at least satisfactory performance of the best model simulations obtained using these three input datasets. The differences in the daily rainfall distribution of TRMM data in comparison to the other three datasets lead to unsatisfactory model performance (Figure 3.8-3.9). However, it is not possible to unequivocally identify the direct impact of rainfall distribution on the model performance of simulating discharge in the above study areas. In fact, the relation between streamflow and rainfall is strongly nonlinear and a change in the precipitation input leads to the definition of different best parameter sets involved in streamflow generation.

3.3.4 Evaluation of model performance

Beside the evaluation of the best simulation achieved with each precipitation dataset, as described above, we have also evaluated the ensemble of all available simulations. We have calculated the frequency of NS values obtained for both final calibration (1500 simulations) and validation (1500 simulations) model results. In total, 3000 simulation results are available for each precipitation dataset. Figure 3.10 shows

the NS distribution for the three subbasins. The performance classification of Moriasi et al. (2007) has been used as the reference for evaluation of the results.
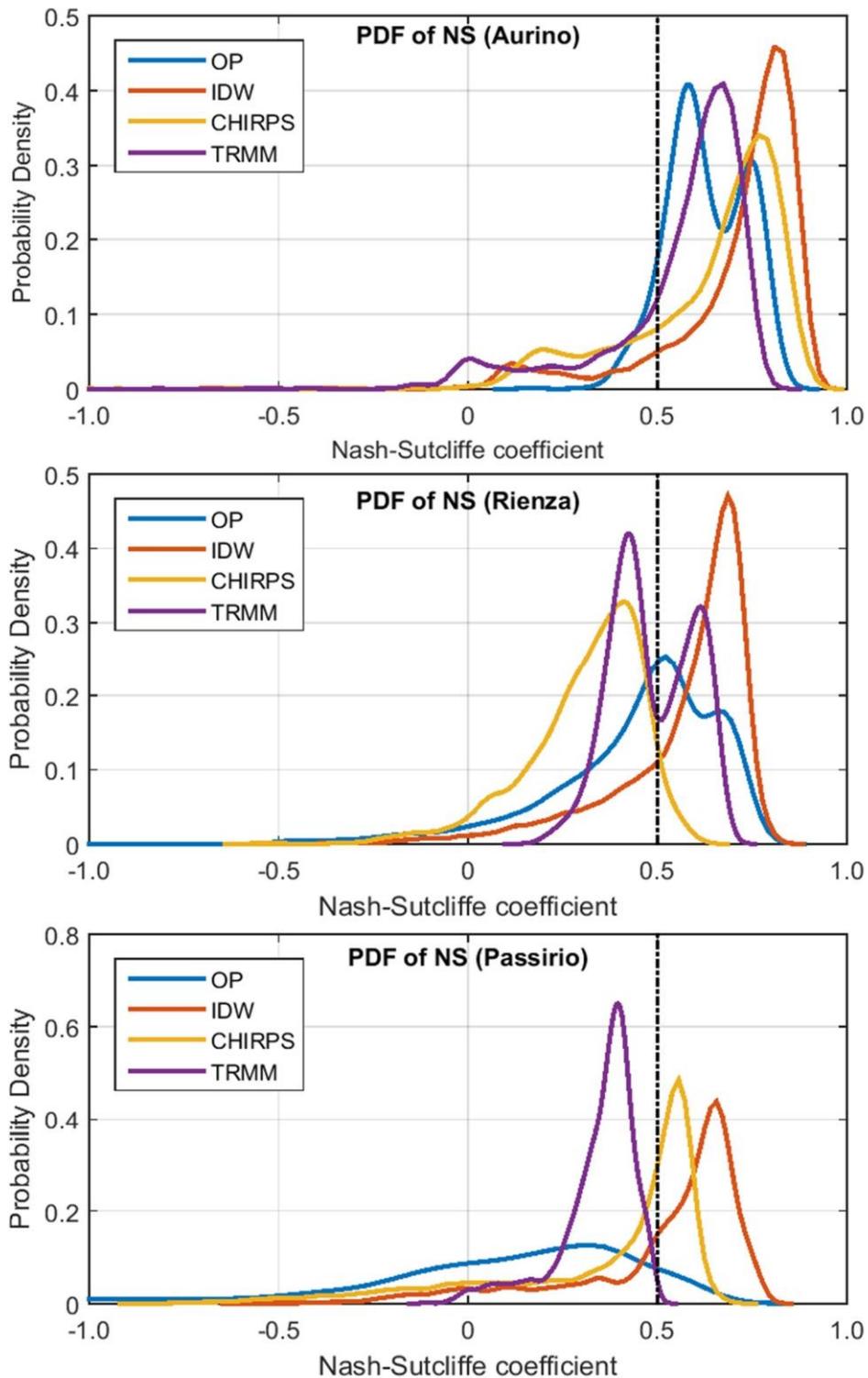
Figure 3.10 Nash-Sutcliffe coefficient distribution of the simulation results with elevation band method for three study catchments.

In the Aurino subbasin, all four models have good performance and most of model results reach at least a satisfactory level. Using IDW-based dataset, 77% of the model results reach at least a good level, while lower fractions are obtained using OP, CHIRPS, and TRMM (41%, 61% and 38%, respectively). Moreover, the IDW model displays the largest fraction of simulations having NS>0.75, which represents very good model performance. In the Rienza subbasin, the distribution of NS values is different in comparison to those of Aurino subbasin. The percentage of simulations with NS>0.5 are 75%, 49%, 44% and 5% for the IDW, OP, TRMM and CHIRPS models, respectively. Besides, the IDW model gets the highest percentage of NS larger than 0.75, which achieves very good model performance. In the Passirio subbasin, the IDW model consistently ranks first in occurrence rate of NS>0.5 (70%) in comparison to OP (9%), CHIRPS (47%) and TRMM (0.1%) models. It also has the highest frequency of NS>0.75.

Therefore, IDW is not only the precipitation dataset which leads to the best model run in terms of NS and R^2 values, but it is the dataset that generally leads to the best set of model simulations. As described above, this better performance of the IDW model in NS distribution can be ascribed to the most frequent application of an elevation correction factor in the case of the IDW-based dataset. Regarding the other products, it is not possible to identify a consistent ranking among the three subbasins. This could be caused by the joint effects of multiple calibrated parameters (Guse et al., 2016).

3.3.5   Parameter uncertainty

The application of the elevation band method to correct the precipitation inputs improves the fitting between model results and observed streamflow data. However, the models have been able to reproduce the streamflow by adjusting other relevant hydrological parameters and converging to different optimum intervals of calibrated parameters.

Among the selected parameters shown in Figure 3.11, CN2 is normally one of the most sensitive hydrological parameters and it indicates an influence of reducing the surface runoff caused by the precipitation (Strauch et al., 2012; Vu et al., 2012; Xu et al., 2010). Different precipitation inputs lead to different best CN2 values and ranges, among which the most apparent differences can be observed in the Passirio subbasin (the best fit values for a_CN2 range between -5.5 and 6.3). However, it is not possible to identify a common pattern for all basins and hence a correlation between the estimated CN2 range with a specific precipitation input cannot be derived. SOL_AWC, responsible for available water capacity of the soil layer, displays a smaller variability compared to CN2 (the coefficient of variation of the parameter range for a_SOL_AWC is 0.16 while for a_CN2 is 0.33) in both best values and ranges for different subbasins and for considering different precipitation inputs. Therefore, in these three subcatchments, SOL_AWC is less affected by a change in the precipitation input than CN2. ESCO is an important parameter related to soil evaporation. The values of ESCO span over the entire physical range

(0-1) not only for different precipitation inputs but also for different subbasins (Figure 3.11). GWQMN is responsible for base flow. Different precipitation inputs result in different best GWQMN values and ranges at each subbasin. The variability in the estimated parameter is again basin-specific (the coefficient of variation of the best fit value is 6.4, 2.0 and 32.1 for Aurino, Rienza and Passirio, respectively).
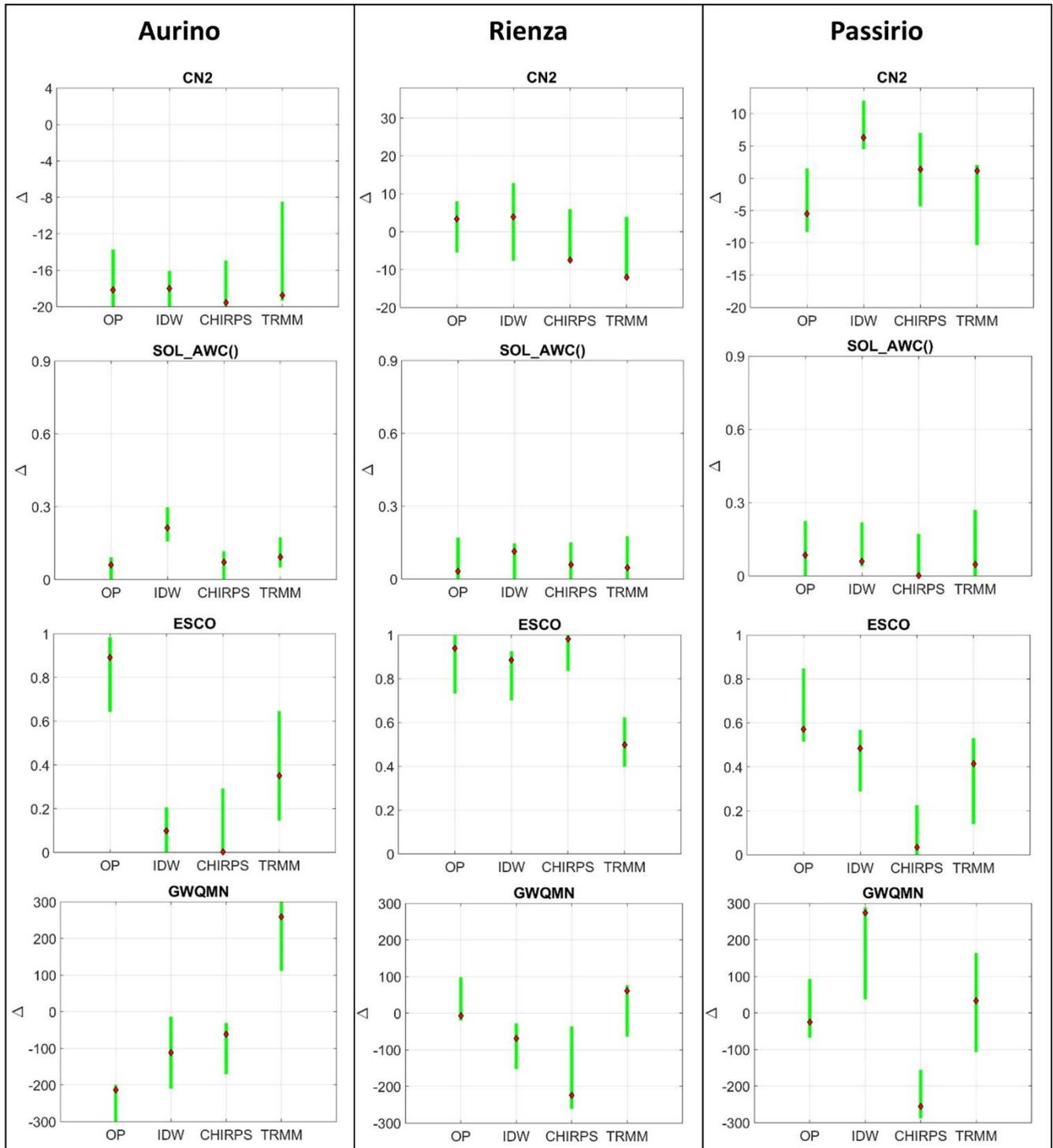
Figure 3.11 Calibrated parameter distributions of four global sensitive hydrological parameters for the four precipitation inputs within the initial parameter range (y-axis domain): green bars show the final

parameter ranges; red points represent the values of "best parameter". "Δ" means an absolute increase, without "Δ" means absolute value. The analysis of the other calibrated parameters is provided in the Supplementary material (Figure S3.4-S3.6).

In summary, no clear pattern emerges to correlate the ranges of the estimated parameters and their uncertainties with the precipitation products considered in this study. The results show that different precipitation inputs affect both the best estimate of a parameter as well as its uncertainty range. It is not possible to identify which precipitation datasets would generally have smaller or larger parameter uncertainties. Additionally, the sensitivity of a parameter towards a change in the precipitation input is catchment specific. Calibrated SWAT models with different precipitation datasets corrected using the elevation band method are able to reproduce the measured discharge. However, in order to fit the measured river discharge, SWAT-CUP adjusts the water volume of different hydrological components (e.g., surface runoff and groundwater contribution) by calibrating the parameters distinctly to cope with the different rainfall features of the four precipitation products (Figure 3.2-3.5). As shown in Table 3.5, despite the significant range of variability among the different fitted values and the different datasets, similar water volumes are assigned to the main hydrological compartments (evapotranspiration, percolation and base flow). Soil water and runoff display the largest variability depending on the applied precipitation input. This has a direct influence on erosion rates and contaminant transport (Neitsch et al., 2011). Hence, it is important to constrain the model using streamflow data, and to validate it using other datasets such as soil moisture, evapotranspiration and snow coverage measurements (Grusson et al., 2015).

As a result, even though all the models might fit streamflow data well (e.g. the model performances in of Aurino subbasin shown in Figure 3.7), the redistribution of the single discharge components across different hydrological compartments is different. For example, at Passirio, the model calibrated using IDW-based dataset has higher optimal CN2 ranges than the other three models calibrated with OP, CHIRPS and TRMM data. This suggests that facing similar rainfall, the IDW model tends to produce more surface runoff than the other models. As a consequence, the IDW model might predict different soil erosion rates in comparison to the other models (Neitsch et al., 2011). At Aurino, the TRMM calibrated model has higher GWQMN values than the other three models. Hence, the TRMM model reflects a different groundwater processes in comparison to the other models, which subsequently would affect groundwater management practice. In the Rienza subcatchment, the model calibrated using CHIRPS data converges towards a range of ESCO values much lower than the models calibrated using the other three datasets. Consequently, the CHIRPS model would probably present higher soil evaporation, which could result in a different implication for water resources management plans in

comparison to the other models. The parameter uncertainty caused by the use of different precipitation input therefore propagates to the uncertainties of predicted hydrological processes (Table 3.5) and then further propagates to subsequent processes controlled by hydrological drivers (Neitsch et al., 2011) and impact decision making processes (e.g., groundwater management and integrated water resources management).

Table 3.5 Water volume of major hydrological processes simulated by the best simulation. Mean annual data for calibration period.

	Evapotranspiration(mm)	Soil water(mm)	Percolation(mm)	Runoff(mm)	Baseflow(mm)
Aurino					
OP	320	96	872	137	827
IDW	367	201	707	83	611
CHIRPS	451	106	806	170	702
TRMM	333	134	927	96	851
Rienza					
OP	406	82	645	31	601
IDW	437	162	634	28	589
CHIRPS	387	112	633	10	631
TRMM	533	95	639	10	642
Passirio					
OP	621	137	991	234	833
IDW	506	105	896	258	748
CHIRPS	502	38	884	182	784
TRMM	422	91	730	78	602

3.3.6   Prediction uncertainties

Parameter uncertainties also contribute to the prediction uncertainties of the models (Abbaspour, 2015). The prediction uncertainties discussed here only refer to the prediction uncertainties of streamflow, which are reflected by the values of P-factor and R-factor of each model (Figure 3.7-3.9).

In the Aurino subbasin (Figure 3.7), in the calibration phase, all the other models except OP obtain acceptable uncertainties, capturing more than 70% of the observations and having an acceptable 95PPU envelope narrower than 1.5 (Abbaspour et al., 2015). Among the models, the IDW model presents the smallest prediction uncertainties with the highest P-factor and smallest R-factor. Considering the validation period, the values of the P-factor of the four models are all within the satisfactory range (Figure 3.7). The values of the R-factor are satisfactory for all products, except CHIRPS (unacceptable). As suggested by a rough balance between P and R, all the models generally present good prediction uncertainty in the Aurino subbasin.

In the Rienza subbasin, the four models perform distinctly for what concerns prediction uncertainties. The P-factor and R-factor also vary between the validation and the calibration periods (Figure 3.8). The IDW model performs consistently well in reaching desirable uncertainties with acceptable values of P-factor and R-factor (Figure 3.8). On the contrary, the other models have at least one factor with an unsatisfactory value either in the calibration or in the validation phase (Figure 3.8). Considering the entire period, a good prediction uncertainty is hence achieved by the SWAT model calibrated using the IDW datasets.

In the Passirio subbasin, only the model calibrated utilizing the OP dataset achieves an acceptable P-factor in both the calibration (0.80) and validation (0.86) periods. Considering the R-factor, only the model calibrated with the OP has a value above 1.5. It is evident that all models fail to achieve at a good balance between the P-factor and the R-factor, and therefore none of the models reach an acceptable prediction uncertainty (Figure 3.9).

Only IDW models obtain consistently acceptable or non-acceptable P-factors and R-factors in both calibration and validation periods; this is not the case for the models using other three precipitation datasets. Therefore, IDW based precipitation dataset leads to more consistent prediction uncertainties than the other three precipitation datasets. Different precipitation inputs generate distinct prediction uncertainties in modeling streamflow. Also in this case, the results are heterogeneous and it is not possible to generalize the outcomes obtained in one subcatchment to another one, i.e., the prediction uncertainty of each precipitation product is basin-specific.

3.3.7   Influence of the precipitation input on a model with fixed parameters

In the previous sections, we showed the impact of different precipitation inputs on the calibrated parameters of the SWAT model. We showed that, despite the catchment similarities in terms of land use, major soil types, and slopes (see also Figure S3.1-S3.2 in the Supplementary material), calibrating the model using different precipitation inputs leads to significantly different model parameters (i.e., the optimal parameter ranges identified using different precipitation inputs may not overlap). We have also shown that the best four simulations for each subbasin (i.e., the simulation that achieves a maximum NS value for a specific precipitation input) have NS values that may vary by more than 40%. This means that the crucial input parameter is the precipitation dataset used.

In this section, we define a fix set of parameters for each of the three subcatchments. The three sets of hydrological parameters have been defined for each subcatchment by averaging the parameter values obtained for the simulations that maximized NS using each precipitation dataset. This test is performed in order to focus on the effect of the precipitation input in a situation that is not biased by the auto-

calibration effect. The hydrological parameters used in this section have been provided in the Supplementary material (Table S3.1). The elevation band parameters for each subbasin and each precipitation input have been reported in Table 3.6.

Table 3.6 The elevation band parameters (PLAPS in Figure S3.4-S3.6 that maximized NS during calibration period) for each subbasin and each precipitation input.

	OP	IDW	CHIRPS	TRMM
Aurino	0.042	0.142	0.150	0.103
Rienza	0.041	0.103	0.141	0.042
Passirio	0.134	0.375	0.672	0.285

The results obtained using IDW-based precipitation data have the best performance in terms of NS and R^2 (the only exception being the NS value computed for Rienza subcatchment) (Table 3.7). In the three case studies considered in this work, a simple interpolation algorithm coupled with the elevation band correction provides the best model results in terms of NS and R^2 values. The result of the IDW model is particularly good for the Passirio subcatchment, where only the IDW model reaches NS and R^2 values larger than 0.75, while the other models do not reach NS and R^2 values larger than 0.55. CHIRPS models rank second in terms of NS and R^2 values in the Aurino and in the Passirio subcatchments, while the performance is poor in the Rienza subcatchment.

Table 3.7 NS and R^2 coefficients for the four precipitation datasets considering fixed hydrological parameters for each subbasin.

	OP		IDW		CHIRPS		TRMM	
	R^2	NS	R^2	NS	R^2	NS	R^2	NS
Aurino	0.81	0.80	0.84	0.83	0.84	0.81	0.75	0.73
Rienza	0.66	0.64	0.70	0.61	0.46	0.35	0.60	0.49
Passirio	0.51	0.38	0.79	0.78	0.55	0.48	0.40	0.36

Comparing the results presented in Table 3.7 with those shown in Figure 3.7 to 3.9, we can see that the Aurino subcatchment is the least sensitive to the applied parameters. In fact, NS and R^2 values obtained using parameter sets specifically calibrated for a particular precipitation dataset are comparable with the one obtained using the computed mean parameter set. Results obtained using the IDW model show less variability in terms of NS and R^2 values in comparison to the other precipitation input models. In fact, comparing the IDW results shown in Table 3.7 with those shown in Figure 3.7 to 3.9, we observe NS and R^2 values that vary by less than 25%.

3.4 Conclusions

This study investigates the impact of four different precipitation inputs on streamflow predicted using SWAT in three Alpine catchments. We have analyzed the rainfall features, model performances, parameter uncertainty, prediction uncertainty and the potential relationships among the above

components. High elevation Alpine catchments are generally data-scarce regions. In particular, the problem for hydrological modeling is that precipitation data are poorly resolved in space and do not capture heterogeneous orographic effects (e.g., Le Moine et al., 2015). Most available meteorological stations are located at low elevations, which often leads to an underestimation of precipitation input (Adler et al., 2015; Isotta et al., 2014).

SWAT uses the elevation band methods to consider the orographic effects on precipitation in mountainous areas. The elevation band method implemented in SWAT currently has some known limitations (Galván et al., 2014) and it is simplistic in comparison to other more physically based corrections (Bárdossy and Pegram, 2013). However, comparing the results obtained without any correction (Figure 3.6) with the results obtained using the elevation band method (Figure 3.7-3.9), there is an evident improvement in model performance (NS increases by at least 9%). In this study, the application of this method to four different precipitation datasets (OP, IDW, TRMM and CHIRPS) has been found to greatly improve the match between simulated streamflow and measurements in three high-elevation Alpine subcatchments of the Adige river basin (Aurino, Rienza and Passirio).

We have investigated the influence of the different datasets on model performance and on the calibrated range of the estimated parameters. The four different precipitation inputs are different in terms of the amount and the temporal distribution of precipitation. The models with the IDW dataset coupled with elevation band method reach the best NS and R^2 values in all three investigated catchments. This dataset is characterized by a greater number of rainy days in form of tiny rain (average 1646 days) resulting from the interpolation of 12 nearest rain gauges. The model with the CHIRPS dataset performs satisfactorily in simulating streamflow, and thus this satellite precipitation product can be a favorable choice for this Alpine region facing data scarcity. The TRMM dataset generally leads to unsatisfactory results when used as SWAT precipitation input for streamflow modeling in these Alpine catchments.

The uncertainty affecting the estimated parameters and their calibrated range of variability changes when applying different precipitation inputs and it is catchment specific, which prevents the generalization of the outcomes achieved for a single case study. This has important consequences on the hydrological interpretation of the results, on the computation of processes driven by hydrological forcing such as erosion and solute transport, and on water management.

We have also investigated the influence of the precipitation input on models with fixed parameters. The use of a precipitation input computed using a simple interpolation algorithm (IDW) allows us to obtain results which are less sensitive to the calibrated model parameters and have higher NS and R^2 values than results obtained using satellite products or single ground observations. This conclusion is only valid

for the study areas investigated in this work and should be verified in other Alpine catchments to be generalized.

In summary, selection of precipitation products has a crucial effect on model performance, model uncertainties and parameter uncertainties in streamflow simulations. Model simulations driven using different datasets can lead to different conclusions about the most relevant hydrological processes in a catchment. Moreover, this cascade of uncertainty then propagates towards processes such as erosion and contaminant transport, and in the end it would likely result in different water management strategies or policies. Nowadays, several precipitation datasets can be available for a single catchment. The uncertainty generated by the use of different precipitation inputs has then to be taken into consideration since it is at the base of the previously described cascade of uncertainty.

Chapter 4

Calibration of Snow Parameters in SWAT: Comparison of Three Approaches in the Upper Adige River Basin[3]

The Soil and Water Assessment Tool (SWAT) model is generally applied in alpine catchments using a unique set of snow parameters for the entire basin, and calibration was based on discharge records only. This technical note presents three calibration procedures for snow parameters of SWAT considering snow water equivalent (SWE) values computed using a dense network of snow depth measurement stations available in the Upper Adige River basin. The first two procedures calibrate snow parameters according to the average sub-basin SWE: the first one defines a unique set of parameters for the entire basin, while the second allows for sub-basin variability. The last approach includes the elevation band SWE output in the calibration for each sub-basin and qualitatively compares it to the SWE computed from the available snow depth monitoring stations. This last method provides the best agreement between SWAT model results and SWE data.

4.1 Introduction

Mountainous catchments are of vital importance for the freshwater supply all over the world, since they contribute to global discharge to a great extent (Viviroli et al. 2007, Viviroli et al. 2003). Due to the high elevation, snow is a dominant component of the hydrological cycle in alpine catchments, having a decisive impact on hydrology-related issues including water supply, erosion, hydropower management and flood control (Cunderlik and Ouarda 2009, Pradhanang et al. 2011, Rahman et al. 2013, Zampieri et al. 2013). Therefore, an accurate description of snow processes is of primary importance for hydrological research and water use in the alpine catchments.

The Soil and Water Assessment Tool (SWAT) (Arnold and Fohrer, 2005) has been widely used by both the global academic community and practitioners for hydrological modeling studies of river basins worldwide (Abbaspour et al. 2015, Francesconi et al. 2016, Golmohammadi et al. 2017, Liu et al. 2016, Malagò et al. 2016, Tuo et al. 2016). This model has been frequently applied in mountainous catchments to investigate the influence of snow processes on the hydrological cycle (Ahl et al. 2008, Debele et al. 2010, Grusson et al. 2015, Rostamian et al. 2008, Troin and Caya 2014). In particular, SWAT has been used to support regional water management practices like hydropower operation (Malagò et al. 2015, Rahman et al. 2013), flood protection (Yang et al. 2014) and policy formulation including climate change adaptation strategies (Pradhanang et al. 2011). The Snow module of SWAT is based on the temperature-index approach (Hock 2003, Neitsch et al. 2011, Walter et al. 2005, Zhang et al. 2008). This approach has been proven accurate enough for SWAT applications (Debele et al. 2010, Luo et al. 2013, Meng et al. 2015). More complex and physically based snow modules included for example in GEOTOP (Endrizzi et al. 2006, Zanotti et al. 2004), Wasim (Warscher et al. 2013) and AMUNDSEN (Strasser 2008) are preferable for hydrological distributed models, while the current structure of SWAT cannot take advantage of a more detailed description of snow processes (Debele et al. 2010, Meng et al. 2015).

Even if modeling snow dynamics is of fundamental importance for SWAT application in alpine catchments, a systematic investigation to identify an optimal procedure to calibrate snow parameters is missing. Most of the previous studies assigned a unique set of snow parameters to an entire basin (e.g., Fu et al. 2014, Fu et al. 2015, Rahman et al. 2013), although snow distribution can be highly heterogeneous (Egli and Jonas 2009). Only a few works recently tried to evaluate the model performance by considering the spatial variability of snow parameters at different sub-basins (Omani et al. 2017). Moreover, snow parameters are often calibrated using streamflow observations alone (e.g., Levesque et al. 2008, Wang and Melesse 2005), instead of including the information derived by snow observations in the calibration process. Discharge integrates the contribution of all the different hydrological components and can hardly allow for distinguishing their specific contributions leading to model

equifinality (Beven 2006). Besides, discharge-based snow parameter calibration is not possible in ungauged catchments and it can be extremely challenging in many alpine river basins strongly influenced by hydropower production and other anthropogenic activities (Majone et al. 2016).

In this technical note, we propose three methodologies to calibrate snow parameters in SWAT. Differently from the approaches used up to now in the literature, we explicitly consider snow observations in the calibration process. Snow observations are available in many alpine catchments mainly in terms of daily or subdaily snow depth values often with higher spatial resolution than river discharge gauging stations. Recently, simple regression equations have been proposed to convert snow depth data into snow water equivalent (SWE) (Grusson et al. 2015, Jonas et al. 2009, Pistocchi 2016, Qi et al. 2016, Zampieri et al. 2013). Besides, in SWAT all HRUs in a sub-basin receive the same SWE value. This value is derived from the weighted sum of the SWE computed for each elevation band in a sub-basin. We hence propose to apply these regression models to compute sub-basin scale SWE averages, and compare these results with the SWE values computed by SWAT on several spatial scales. The methodology has the advantage of using snow observations, by which the snow parameters can be quantified considering the SWAT snow module alone, without the need for discharge data. To the best of our knowledge no studies have taken this possibility into account, and only a few studies have tried to utilize the SWAT SWE output to evaluate the performance of the SWAT snow module independently of discharge records (Fu et al. 2014, Fu et al. 2015, Grusson et al. 2015, Pradhanang et al. 2011). Moreover, although a new version of SWAT, named SWAT+, is currently under development, it seems that the snow module has not been modified yet (Bieger et al. 2016). Our study can therefore provide the developers with an interesting approach that could be automated and implemented into SWAT+ in order to improve model calibration in alpine catchments (Arnold et al. 2015, Bieger et al. 2016).

To summarize, the specific objectives of this study are to: 1) propose a calibration procedure for the parameters used in the SWAT snow package based on sub-basin SWE values; 2) validate and assess the proposed procedure in the Upper Adige catchment, an important alpine basin in Northeast Italy; 3) test the spatial variability of SWAT snow parameters considering both basin-scale and sub-basin-scale calibration; 4) propose an improvement in SWAT to facilitate the use of snow measurements for model calibration.

4.2 Material and method

4.2.1 Study area

The Upper Adige River basin closed at Bronzolo (6875 km^2) is located in the Northeastern Italian Alps (Figure 4.1). It is a typical Alpine catchment characterized by steep slopes, sharp hydro-climatic

gradients and a large elevation variation ranging from 223 m a.s.l to 3865 m a.s.l. (Callegari et al. 2015).
This river basin was selected as a case-study in the FP7 project GLOBAQUA (Navarro-Ortega et al.
2015) to investigate a variety of stressors encountered in the Alpine region, in which snow processes are
major driving forces of the hydrological cycle (Chiogna et al. 2016). This area is typically dominated by
snowfall precipitation from late October to April, leading to an increase of the cumulative snow depth.
Streamflow in the following spring period is dominated by snow melting (Mei et al. 2014, Mei et al.
2016, Penna et al. 2014). The location of the precipitation and temperature stations are displayed in
Figure S4.1 of the supplemental material.
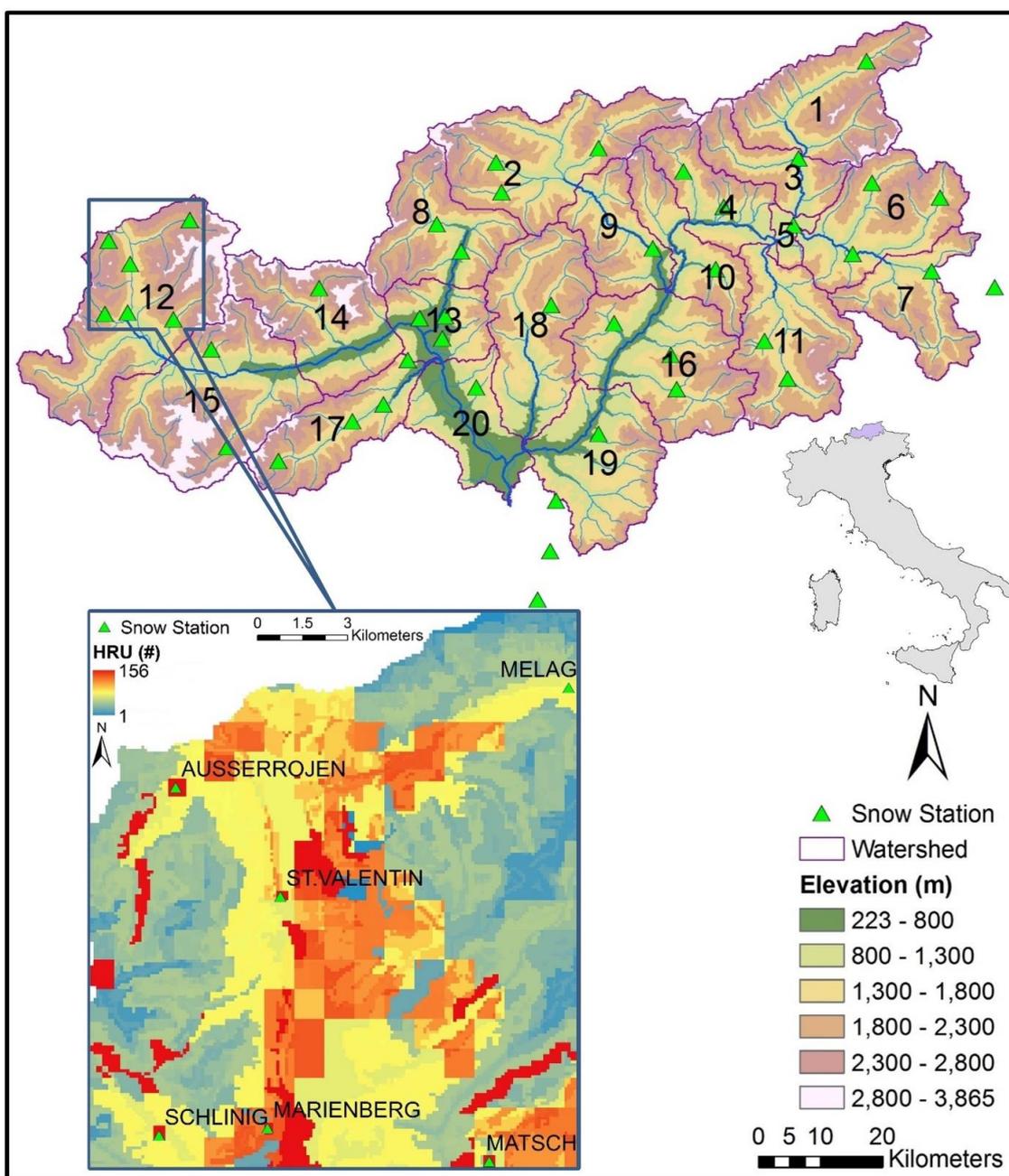
Figure 4.1 The Upper Adige River basin and the zoomed-in HRU map showing the locations of the snow
stations in sub-basin 12 (inset).

4.2.2   Computation of sub-basin snow water equivalent values

The dataset available for the Upper Adige catchment comprises 44 snow measurement stations (Figure 4.1 and Table S4.1) for which daily snow depth values are recorded. The data have been provided by the Autonomous Province of Bolzano and represent a subset of the data used by Adler et al. (2015) to analyze climatic changes occurring in this region. The available time series have been quality checked and their homogeneity has been tested using the standard normal homogeneity test (Marcolini et al. 2017). The stations used in this study were selected based on data availability for the period 2005-2010.

In order to make use of these data for SWAT model calibration, two steps have to be performed: first daily snow depth data (HS) are used to compute a mean sub-basin snow depth and then these values are converted in SWE data.

4.2.2.1 Computation of mean sub-basin snow depth

The HS measurements were used to calculate an average snow depth value for each of the 20 sub-basins of the catchment. López-Moreno and Nogués-Bravo (2006) provide an important intercomparison study between various local, geostatistical and global methods to interpolate snow depth data at the catchment scale. Their outcome is that global methods (e.g., linear regression and generalized additive models) have the best performance, and elevation is the most relevant parameter to be included in the model. The important correlation between snow depth and elevation at various spatial scales was also evidenced in Grünewald et al. (2014). In this study, we apply a simple power law model to establish a correlation between all available snow depth measurements in the catchment at a given day and elevation (h):

$$HS\_reg = a \cdot h^b \tag{4.1}$$

The parameters a and b are computed for each day of the period considered in this study. Such a model does not reproduce the effect observed by Grünewald et al. (2014) in which snow depth increases with elevation up to a certain altitude where it displays a distinct peak followed by a decrease at the highest elevations. This effect is most likely present in the Upper Adige catchment, but the highest available station is located at 1915 m a.s.l. (Table S4.1) and therefore it is not possible to observe it. Several tests were performed comparing the results of Eq. (4.1) with the results obtained using a more complex regression model reproducing the behavior observed by Grünewald et al. (2014). These tests did not lead to important differences (i.e., larger than 20%) when computing mean snow depth values at the subcatchment scale. Despite the simplicity of the model proposed in Eq. (4.1), we obtained an average correlation coefficient of 0.65±0.12 (computed between 2005 and 2010) between modelled and measured values, which is comparable with the results obtained by López-Moreno and Nogués-Bravo

(2006). Eq. (4.1) is then used to compute the mean sub-basin snow depth <Sub_HS>, depending on the topography of the sub-basin:

$$< Sub\_HS >= \int_0^{h_{max}} HS\_reg \cdot \tilde{h}(h) dh \qquad (4.2)$$

where $\tilde{h}$ is the probability density function of the elevation in each sub-basin and h_{max} is the maximum catchment elevation.

4.2.2.2 Conversion of snow depth measurements into snow water equivalent values

The conversion of observed HS [m] values for a single monitoring station into SWE [kg/m^2] is performed by multiplying HS by the snow bulk density ρ [kg/m^3]:

$$SWE = HS \cdot \rho \qquad (4.3)$$

For single monitoring stations, measured values of ρ can be available. However, in most cases continuous measurements of the snow bulk density are very rare (e.g., Pistocchi, 2016). Therefore, several empirical methods have been proposed to convert snow depth values into SWE values, developing regression equations representing the temporal variability of the snow bulk density at the basin or regional scale (e.g., Jonas et al. 2009, McCreight and Small 2014, Pistocchi 2016, Sturm et al. 2010). Such models generally imply a positive correlation between snow density and the number of days since the beginning of the winter period and are specific for a given geographic region. Despite the intrinsic limitations associated with the use of such models in particular at the daily scale (McCreight and Small 2014), this simple approach can be applied whenever direct density measurements are scarce or not available (McCreight and Small 2014, Pistocchi 2016). The uncertainty in estimating SWE by applying such empirical models is about 13% and it is comparable to the within-site spatial variability (Pistocchi, 2016).

For the Upper Adige catchment, Pistocchi (2016) derived the following empirical equation to estimate the temporal variability in snow bulk density, subsequently estimating SWE starting from snow depth observations:

$$\rho = 200 + (DOY + 61) \qquad (4.4)$$

where DOY is a counter of the day of the year which is set to 1 on January 1st, with October 1st being -92 to account for the winter season extending across two years in the northern hemisphere (Pistocchi 2016). In spite of its simplicity, Eq.(4.4) aims at describing the seasonal dynamics of snow density and in particular its increase during the winter season (Jonas et al., 2009; Pistocchi, 2016). As recommended by Pistocchi (2016), Eq. (4.4) represents a reasonable first estimate when no density time series are available with adequate spatial and temporal resolution, which is the case of the Upper Adige catchment.

Therefore, we computed the mean sub-basin snow water equivalent (<Sub_SWE>) for each day according to the following equation:

$$< Sub\_SWE > = < Sub\_HS > \cdot \rho \qquad\qquad (4.5)$$

The units of <Sub_SWE> were then converted in mm of H_2O so they could be compared to SWAT model results.

4.2.3 SWAT model

The SWAT model is a comprehensive, time-continuous, semi-distributed, process-based model (Arnold et al. 2012a). It divides river basins into sub-basins and subsequently into Hydrologic Response Units (HRUs), characterized by different combinations of land use, soil characteristics, topography, and management schemes. The hydrological cycle is calculated based on water balance, which is controlled by climate inputs such as daily precipitation and maximum/minimum air temperature. Using daily input time series, SWAT simulates the daily, monthly and yearly fluxes of water and solutes in river basins

4.2.4 Model setup

ArcSWAT 2012 (version 664), with an interface in ArcGIS, was used to setup the model in this work. The datasets used in the model are listed in Table 4.1. The SWAT model was set up for the Upper Adige River basin with 20 sub-basins delineated in Figure 4.1. In most sub-basins, snow depth measurements are available for stations located at similar elevations (Table S4.1). However, at sub-basin 12 and sub-basin 17, snow depth measurement stations are available for different elevations (Table 4.2). The elevations for sub-basin 12 and sub-basin 17 range between 883-3666 m and 261-3339 m, respectively (Figure 4.1). The areas of these two sub-basins are 654 km^2 and 285 km^2, respectively. We performed simulations at the daily scale.

Table 4.1 Source of input data used in building up the SWAT model. Despite in this work we focus only on the snow module of SWAT we provide all required input to construct a complete hydrological SWAT model.

Data type	Scale	Data source
DEM	90 m × 90 m	Shuttle Radar Topography Mission (SRTM) produced by Consortium for Spatial Information (CGIAR-CSI)
Land use	100 m × 100 m	Corine Land Cover 2006 (CLC2006) from European Environment Agency

Soil	1 : 1500,000	Data provided in the framework of Globaqua Project (Navarro-Ortega et al. 2015) following the procedure described in (Meyer et al. 2016)	
River network		EU-DEM product http://www.eea.europa.eu/data-and-maps/data/eu-dem.	
Climate input		Autonomous Province of Bolzano (http://www.provincia.bz.it/meteo/home.as)	
Snow measurements		Autonomous Province of Bolzano (Adler et al. 2015)	

Table 4.2 Elevation and Band distribution of snow stations.

Station	Elevation (m a.s.l)	Band	Elevation range (m a.s.l)
Sub-basin 12			
Marienberg	1310	1	883-1350
ST. Valentin	1499	3	1499-1650
Matsch	1570	3	1499-1650
Schlinig	1690	4	1650-1800
Ausserrojen	1833	5	1800-2050
Melag	1915	5	1800-2050
Sub-basin 17			
Stausee ST. Pankraz	810	1	261-950
Stausee Zoggl	1142	2	950-1500
Pawigl	1400	2	950-1500
Weissbrunn	1900	5	1800-2050

4.2.5 Snow package

The snow module of SWAT provides snow output in forms of snowmelt and SWE defined as the water contents of the snow pack with identical units to precipitation (Pradhanang et al. 2011). The mass balance of the snow pack is calculated as:

$$SWE_i = SWE_{i-1} + R_{day_i} - E_{sub_i} - SNO_{mlt_i} \tag{4.6}$$

where SWE_i is the snow water equivalent of the snow pack on day i (mm H$_2$O), SWE_{i-1} is the snow water equivalent of the snow pack on the previous day (mm H$_2$O), R_{day_i} is the amount of precipitation in terms of snowfall on day i (mm H$_2$O), E_{sub_i} is the snow sublimation on day i (mm H$_2$O), and SNO_{mlt_i} is the amount of snowmelt on day i (mm H$_2$O).

The snowfall temperature (SFTMP (°C)) is used as a threshold to classify the precipitation as snowfall or rainfall. When the mean daily air temperature is lower than SFTMP, the precipitation is classified as snowfall and is added to the snow pack. Snow sublimation is computed as a function of potential

evapotranspiration in the evapotranspiration module (Neitsch et al. 2011). The snow pack will not melt until the snow pack temperature exceeds the threshold value: snowmelt base temperature (SMTMP (°C)). The snow pack temperature is calculated as:

$$T_{snow_i} = T_{snow_{i-1}}(1 - TIMP) + \bar{T}_{av_i} \cdot TIMP \qquad (4.7)$$

where T_{snow_i} and $T_{snow_{i-1}}$ are the snow pack temperatures on the current day i and the previous day i-1 (°C), respectively, $\bar{T}_{av_i}$ is the mean air temperature on the current day i (°C), and $TIMP$ is the snow temperature lag factor.

When the snow pack reaches a higher temperature than SMTMP, it starts to melt. The amount of snowmelt is computed as:

$$SNO_{mlt_i} = b_{mlt_i} \cdot sno_{cov_i} \cdot \left[\frac{T_{snow_i}+T_{max_i}}{2} - SMTMP \right] \qquad (4.8)$$

$$b_{mlt_i} = \frac{(SMFMX+SMFMN)}{2} + \frac{(SMFMX-SMFMN)}{2} \cdot sin\left[\frac{2\pi}{365}(i - 81)\right] \qquad (4.9)$$

where SNO_{mlt_i} is the amount of snowmelt on day i (mm H_2O), b_{mlt_i} is the melt factor for day i (mm H_2O °$C^{-1}day^{-1}$), sno_{cov_i} is the fraction of the HRU area covered by snow, T_{max_i} is the maximum air temperature on the current day, $SMFMX$ is the melt factor for June 21(mm H_2O °$C^{-1}day^{-1}$), and $SMFMN$ is the melt factor for December 21 (mm H_2O °$C^{-1}day^{-1}$). As explained in the SWAT I/O files (Arnold et al. 2012b), SMFMX is the maximum melt factor in the Northern Hemisphere and the minimum melt factor in the Southern Hemisphere. Therefore, for our study area, SMFMX should be larger than SMFMN.

In SWAT, either a unique set of snow parameters can be assigned to the entire basin or a different set of snow parameters can be defined for each sub-basin. In the latter case, the source code has to be modified. We noticed that negative and zero values of SFTMP and SMTMP are not allowed at the sub-basin scale (see the code files named as "readsub.f" and "readsno.f"). However, as described in the SWAT official documentation (Arnold et al., 2012b), it should be possible to define values between -5 °C and 5 °C for SFTMP and SMTMP (e.g. Grusson et al. 2015). Therefore, we did not consider it appropriate to apply this "above 0" limitation to the range of SFTMP and SMTMP values.

In order to consider the orographic effects on precipitation and temperature in mountainous areas, SWAT uses the elevation bands method which permits defining up to ten elevation bands in each sub-basin (see supplemental material for a detailed description of the method). In this work, six elevation bands were applied and the bands were defined individually for each sub-basin according to the specific elevation distribution of the sub-basin. Using elevation bands has been proven useful and necessary in several

snow-dominated alpine catchments (Grusson et al. 2015, Luo et al. 2013, Malagò et al. 2015, Morán-Tejeda et al. 2014, Omani et al. 2017, Pradhanang et al. 2011, Rahman et al. 2013, Zhang et al. 2008).

SWAT output files provide SWE of the snow pack for each elevation band in the output.snw file. Snowmelt values and the SWE of the snow pack at the HRU scale are stored in the output.hru file. Since snowmelt measurements are very rare, we propose to calibrate snow parameters according to SWE data. In the following, we refer to the output available for each elevation band as EB-SWE. Since one elevation band covers an area with a range of elevations and since the SWE depends on the snow depth which is elevation dependent, the EB-SWE output cannot be calibrated to fit exactly with the data available for a single measurement point.

The following exercise exemplifies the reason why the measurement of a single monitoring station cannot be used to calibrate the model output obtained at the HRU scale. We constructed a specific model in which a single and unique HRU was defined for the location of each snow station. The HRU map of sub-basin 12 is provided as an example in Figure 4.1. In sub-basin 12, there are six snow stations located at different elevations (Table 4.2) and each of them recorded different snow data. However, the SWAT model output for each HRUs of the six stations was practically the same (Figure S4.2) and was equal to SUM_EB (mm) - the weighting sum of the SWE computed for each elevation band:

$$SUM\_EB = \sum_{bnd=1}^{b} SWE_{band} \cdot fr_{bnd} = Sub\text{-}SWE \qquad (4.10)$$

where SWE_{band} is the SWE value of each band (mm), fr_{bnd} is the fraction of the sub-basin area within the elevation band, and b is the total number of elevation bands in the sub-basin. Therefore, we can conclude that SWAT assigns the same SWE to all HRUs within one sub-basin, and this value represents the elevation weighted SWE value of the sub-basin. This sub-basin specific value can be used in calibration by comparing it to the sub-basin SWE computed using Eq. (4.5). This SWE output of SWAT is hereafter referred to as Sub-SWE.

4.2.6 Model calibration and evaluation

A time frame of 516 days (01/01/2005 - 05/31/2006) was used as a warm up period in order to mitigate the effects of the initial conditions for the 20 sub-basins. Calibration of snow parameters was performed during the time period 06/01/2006 - 05/31/2008, while the validation period starts on 06/01/2008 and lasts until 05/31/2010. The snow parameters to be calibrated are selected according to the sensitivity analysis of the one-at-a-time procedure in SWAT-CUP (Abbaspour 2015). In this work, all parameters used in the snow module (Eqs. (4.6) - (4.9)) except snow cover variables were identified as sensitive parameters. The snow parameters were calibrated such that the Sub-SWE output (Eq. (4.10)) fits the sub-basin average SWE as defined in Eq. (4.5).

Three calibration procedures were applied: I) the basin-scale snow parameters were estimated using an automatic calibration procedure to assign the same set of values to the entire basin; II) snow parameters were calibrated at the sub-basin scale to take the spatial variability of snow processes into account. Automatic calibration was first implemented, but an additional manual calibration was necessary when automatic calibration led to inconsistent parameter sets. For example, SMFMX values estimated using automatic calibration were often lower than SMFMN values, which is not correct for the Northern Hemisphere where the study basin is located; III) sub-basin-scale snow parameters were calibrated manually considering a qualitative comparison between EB-SWE and the SWE values computed for each monitoring station (the snow depth observations were converted into SWE values using Eq.(4.3) and Eq. (4.4)).

Starting with the initial ranges of parameters shown in Table 4.3, the automatic calibration was performed first with the Sequential Uncertainty Fitting algorithm version 2 (SUFI-2) (Abbaspour et al. 2004, Abbaspour et al. 2007) in the SWAT-CUP tool package (Abbaspour 2015) to narrow down and target the effective range of each parameter. SUFI-2 is a semi-automated inverse modeling procedure, which implements a combined calibration-uncertainty analysis. Detailed information and calibration procedures for the tool can be found in Abbaspour (2015). The initial parameter ranges are defined following the works of Grusson et al. (2015), Rahman et al. (2013) and the SWAT documentation (Arnold et al. 2012b). For each iteration, 1000 simulations were run. After each iteration, the ranges of the parameters were modified (normally narrowed down) according to both the new parameters suggested by the program (Abbaspour et al. 2004, Abbaspour et al. 2007) and their initial limitations.

The goodness of fit between Sub-SWE and the sub-basin average SWE defined in Eq. (4.5) for each sub-basin was first calculated during both the calibration and validation periods. Furthermore, the simulated EB-SWE was qualitatively compared to the data of each station belonging to a specific elevation band. Results are provided for sub-basins 12 and 17, because in these subcatchments the measurement stations are located in more than one elevation band. The Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe 1970) and the coefficient of determination (R^2) were used as goodness of fit indicators for the best simulation (Eqs. (4.11) and (4.12), respectively):

$$NSE = 1 - \frac{\sum_{t=1}^{T}(SWE_{s,t} - SWE_{o,t})^2}{\sum_{t=1}^{T}(SWE_{o,t} - \overline{SWE_o})^2} \tag{4.11}$$

$$R^2 = \frac{\left[\sum_{t=1}^{T}(SWE_{o,t} - \overline{SWE_o})(SWE_{s,t} - \overline{SWE_s})\right]^2}{\sum_{t=1}^{T}(SWE_{o,t} - \overline{SWE_o})^2 \sum_{t=1}^{T}(SWE_{s,t} - \overline{SWE_s})^2} \tag{4.12}$$

where $SWE_{o,t}$ is the observed SWE at time t, $SWE_{s,t}$ is the simulated SWE at time t, $\overline{SWE_o}$ is the mean of observed SWE values, and $\overline{SWE_s}$ is the mean simulated SWE.

NSE indicates whether the model performs better than the mean value of the data (NSE = 1 is the optimal value). R^2 ranges from 0 to 1 and represents the trend similarity between the observed and the simulated data: higher R^2 values indicate better model performance. A standard method to evaluate model performance for daily SWE modeling results is not available. In this work, we follow the recommendations of Moriasi et al. (2007) developed for monthly discharge values: unsatisfactory performance (NSE≤0.50), satisfactory performance (0.50<NSE≤0.65), good performance (0.65<NSE≤0.75) and very good performance (0.75<NSE≤1.00).

Table 4.3 Calibrated snow parameters and elevation band parameters.

Parameters	Description	Default	Range
SFTMP	Snowfall temperature [°C]	1	0/4
SMTMP	Snowmelt base temperature [°C]	0.5	-2/2
SMFMN	Melt factor for snow on December 21 [mm H_2O °C^{-1}day^{-1}]	4.5	1.5/7
SMFMX	Melt factor for snow on June 21 [mm H_2O °C^{-1}day^{-1}]	4.5	1.5/7
TIMP	Snow pack temperature lag factor	1	0/1
PLAPS	Precipitation lapse rate [mm H_2O/m]	0	0/0.15
TLAPS	Temperature lapse rate [°C/km]	0	-9/0

4.3 Results

We will focus on model results for sub-basins 12 and 17. We chose these two sub-basins, because there are multiple snow stations available at different elevations (Figure 4.1 and Table 4.2). This allowed us to consider the snow information available for different elevations when evaluating the model performance. Model performances for all sub-basins are summarized in a tabular form in the following sections.

4.3.1 Model performance using calibration procedure I

As shown in Table 4.4, the definition of a unique set of snow parameters for the entire basin (hereafter defined as basin-scale snow parameters) resulted in satisfactory model performance during both the calibration and validation periods in 90% of the sub-basins. Exceptions were represented by sub-basin 5 and sub-basin 20. During the calibration period, sub-basins 5 and 20 displayed much lower NSE values (-0.38 and 0.34, respectively) than the other 18 sub-basins (the lowest NSE is 0.65). During validation, the NSE of sub-basins 5 and 20 were also lower (-0.31 and 0.48, respectively) than the others (the lowest NSE is 0.52). This noteworthy difference in model performance could be ascribed to the topography of the sub-basins. Sub-basin 5 and 20 are characterized by lower elevations (about 90% of the area of the two sub-basins is below 1600 m a.s.l) than the other sub-basins (on average only 37% of the area of the
88

subcatchments is below 1600 m a.s.l). During calibration, the assignment of a unique set of parameters for the entire basin was biased by this disproportion between low and high elevation sub-basins and the optimal set of parameters was representative of the majority of the sub-basins. Therefore, the calibrated parameters better represented high elevation catchments characterized by high SWE values than sub-basins 5 and 20.

Table 4.4 Calibration and validation performance of basin-scale snow parameters (calibration procedure I).

Sub-basin	Calibration		Validation	
	R^2	NSE	R^2	NSE
1	0.77	0.74	0.96	0.96
2	0.82	0.81	0.94	0.87
3	0.93	0.92	0.72	0.65
4	0.77	0.76	0.68	0.67
5	0.09	-0.38	0.13	-0.31
6	0.93	0.90	0.94	0.94
7	0.73	0.73	0.86	0.85
8	0.85	0.71	0.96	0.81
9	0.90	0.88	0.92	0.85
10	0.80	0.67	0.86	0.76
11	0.76	0.65	0.96	0.92
12	0.81	0.81	0.93	0.93
13	0.81	0.72	0.80	0.52
14	0.87	0.87	0.82	0.81
15	0.71	0.71	0.87	0.84
16	0.92	0.87	0.91	0.72
17	0.81	0.73	0.95	0.76
18	0.88	0.88	0.96	0.61
19	0.77	0.67	0.88	0.87
20	0.75	0.34	0.85	0.48
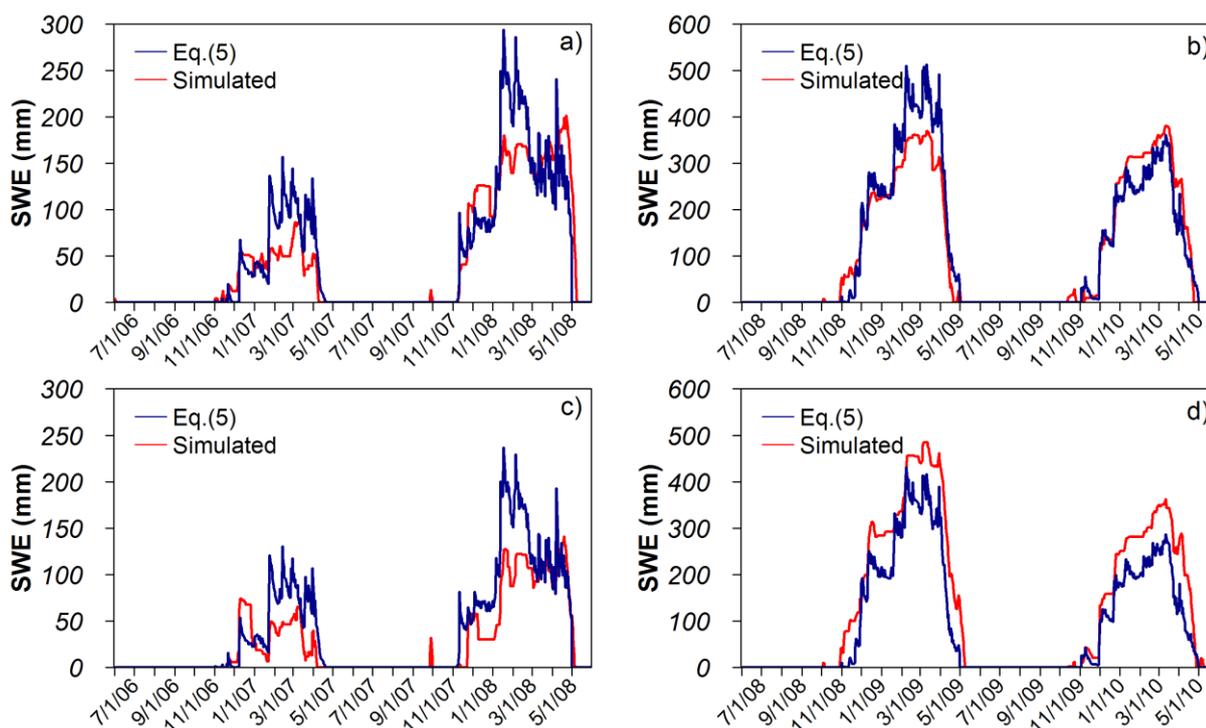
Figure 4.2 Model performance of basin-scale snow parameters for fitting sub-basin SWE during both the calibration and validation periods (calibration procedure I): (a) calibration and (b) validation results of sub-basin 12; (c) calibration and (d) validation results of sub-basin 17.

Figure 4.2 displays the sub-basin SWE of sub-basins 12 and 17. The snow covered periods were well captured at both sub-basins. Simulated Sub-SWE of sub-basin 12 fitted well with the SWE computed using Eq. (4.5) in both the calibration (R^2=0.81, NSE=0.81) and validation periods (R^2 =0.93 NSE=0.93). Generally, the simulation underestimated the maximum SWE except for the snow season of 2009-2010. Similarly, the same basin-scale snow parameters applied to sub-basin 17 led to an underestimation of the maximum SWE during the calibration period, but resulted in a slight overestimation during validation.

The calibrated basin-scale snow parameters generally did not lead to good model performance when comparing EB-SWE to the point SWE data of the snow stations (NSE values are provided in Figure 4.3 and 4.4). As shown in Figure 4.3 for sub-basin 12, the EB-SWE simulations of all considered elevation bands displayed clear underestimations in comparison to the station observations during calibration. In the validation period, the results were better in high elevation areas (bands 4 and 5), where the EB-SWE simulations matched the observation of three snow stations with acceptable NSE better than in low elevation areas. For sub-basin 17 (Figure 4.4), the EB-SWE values fitted better with station observations during calibration compared to sub-basin 12. Model performance was good in low elevation areas (bands 1 and 2) but underestimated the observed values in high elevation areas (band 5). During validation, performance was unsatisfactory in low elevation areas but improved to very good levels at high elevation

areas (band 5). Notice that the NSE values were used to support the qualitative comparison between EB-SWE and the observed values from snow monitoring stations. In fact, the EB-SWE is a representative value standing for areas of a range of elevations, whereas the station observation can only represent one point among the elevations. To summarize, in both sub-basins, the basin-scale snow parameters resulted in good performance in simulating sub-basin SWE, but normally failed to reach reasonable and acceptable fits between EB-SWE and the point observations.
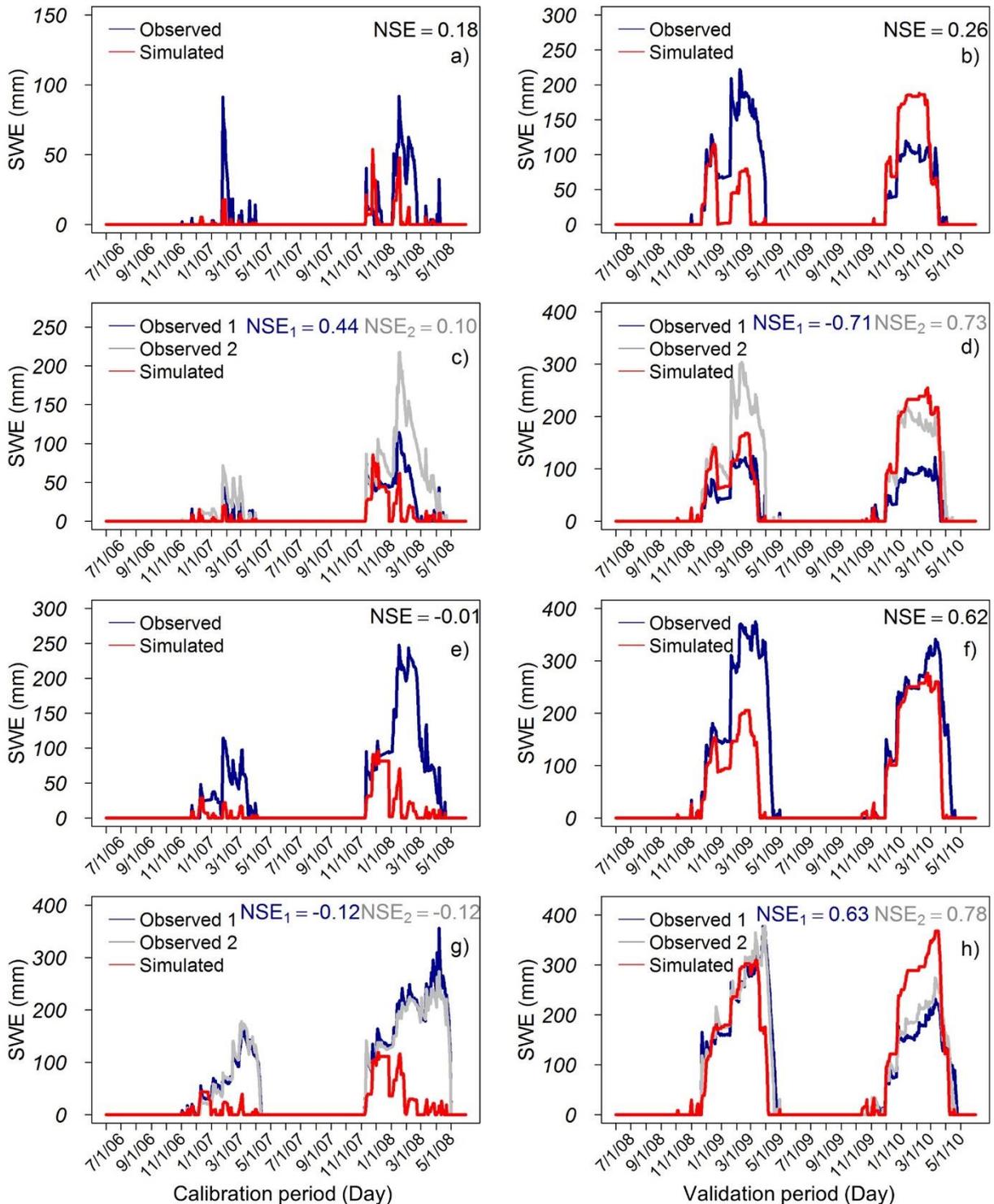
Figure 4.3 SWE comparison between values calculated from station observations and the simulated values of the corresponding elevation band at sub-basin 12 with basin-scale snow parameters (calibration procedure I): (a) calibration and (b) validation performances of band 1 and Marienberg station; (c) calibration and (d) validation performance of band 3: "Observed 1" stands for Matsch station, "Observed 2" represents ST. Valentin station; (e) calibration and (f) validation performance of band 4 and Schlinig station; (g) calibration and (h) validation performances of band 5: "Observed 1" stands for Ausserrojen station, and "Observed 2" represents Melag station.
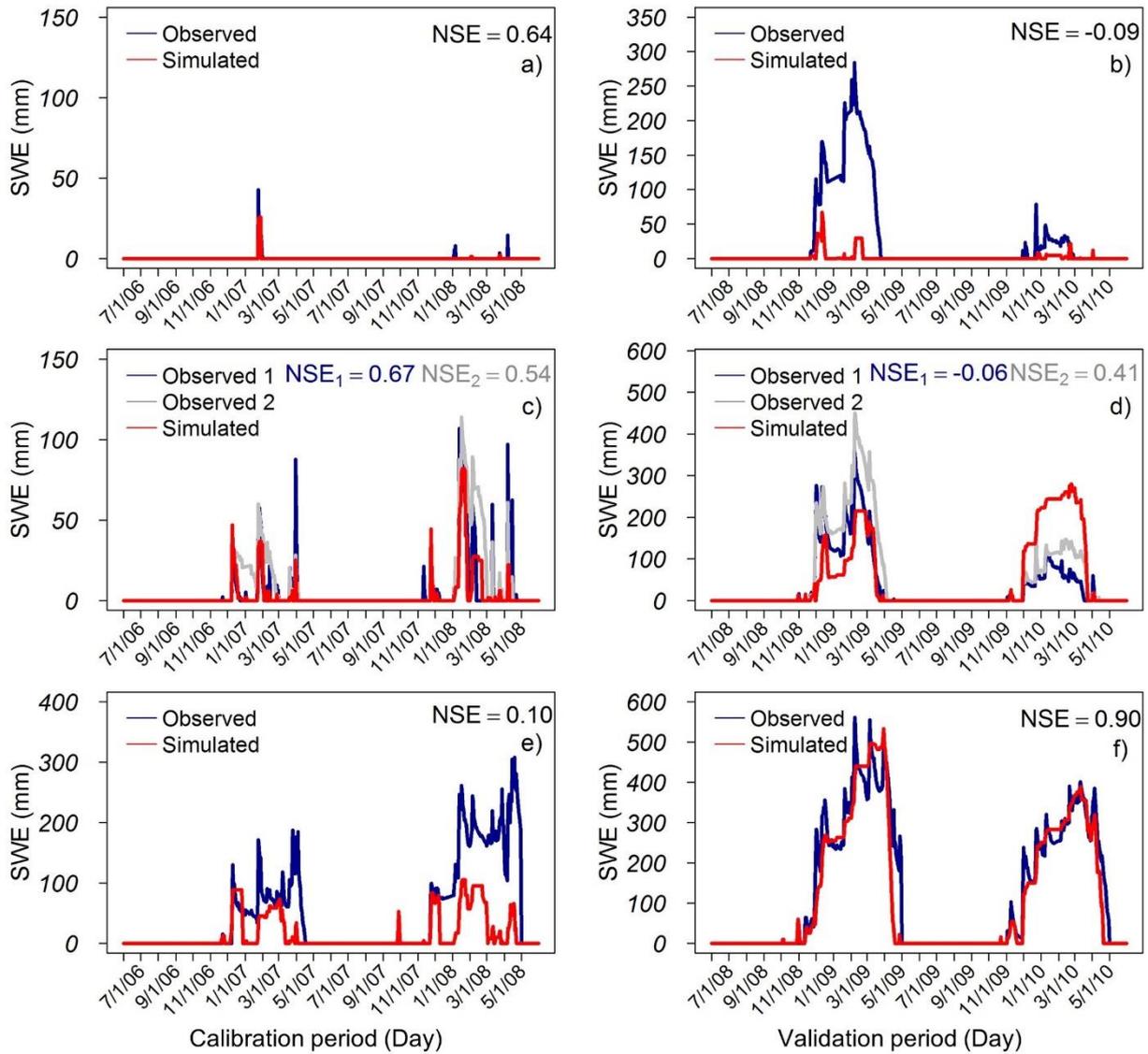
Figure 4.4 SWE comparison between values calculated from station observations and the simulated values of the corresponding elevation band at sub-basin 17 with basin-scale snow parameters (calibration procedure I): (a) calibration and (b) validation performance of band 1 and Stausee ST. Pankraz station; (c) calibration and (d) validation performance of band 2: "Observed 1" stands for Pawigl station, "Observed 2" represents Stausee Zoggl station; e and f are the calibration and validation performances of band 5 and Weissbrunn station.

4.3.2   Model performance using calibration procedure II

The second calibration procedure considers the definition of sub-basin specific snow parameters (hereafter defined as sub-basin-scale snow parameters). The lowest R^2 and NSE were 0.61 and 0.55 during calibration, and 0.65 and 0.56 during validation (Table 4.5). Compared to the performance with basin-scale snow parameters, sub-basin-scale parameterization led to higher R^2 and NSE values in 65% of the sub-basins (e.g. sub-basins 4, 7 and 8), but lower values were also observed in four sub-basins (sub-basins 3, 6, 14, and 18). The sub-basin SWE of the two low elevation-dominated sub-basins (5 and 20) were apparently better reproduced considering the improvement in the R^2 and NSE values. For the calibration and validation periods of sub-basin 5, R^2 increased by 0.74 and 0.61, and NSE increased by 1.20 and 1.04, respectively. For sub-basin 20, NSE increased to the acceptable level with 0.55 and 0.56 for each period. Except for sub-basins 5 and 20, however, the differences of the R^2 and NSE values (Table 4.4 and Table 4.5) between calibration procedure I and II was on average smaller than 0.2% and 3.5%, respectively, indicating comparable performances were reached between basin-scale and sub-basin-scale approaches.

Table 4.5 Calibration and validation performance of sub-basin-scale snow parameters (calibration procedure II).

Sub-basin	Calibration		Validation	
	R^2	NSE	R^2	NSE
1	0.75	0.74	0.96	0.94
2	0.81	0.81	0.95	0.94
3	0.90	0.88	0.69	0.61
4	0.88	0.86	0.65	0.63
5	0.83	0.82	0.74	0.73
6	0.87	0.85	0.91	0.90
7	0.87	0.84	0.95	0.95
8	0.81	0.79	0.97	0.90
9	0.88	0.87	0.91	0.87
10	0.86	0.82	0.83	0.77
11	0.70	0.67	0.94	0.94
12	0.82	0.81	0.93	0.93
13	0.83	0.73	0.79	0.66
14	0.86	0.84	0.68	0.66
15	0.82	0.82	0.94	0.92
16	0.89	0.89	0.87	0.81
17	0.83	0.75	0.95	0.84
18	0.88	0.88	0.91	0.57
19	0.80	0.72	0.93	0.91
20	0.61	0.55	0.83	0.56
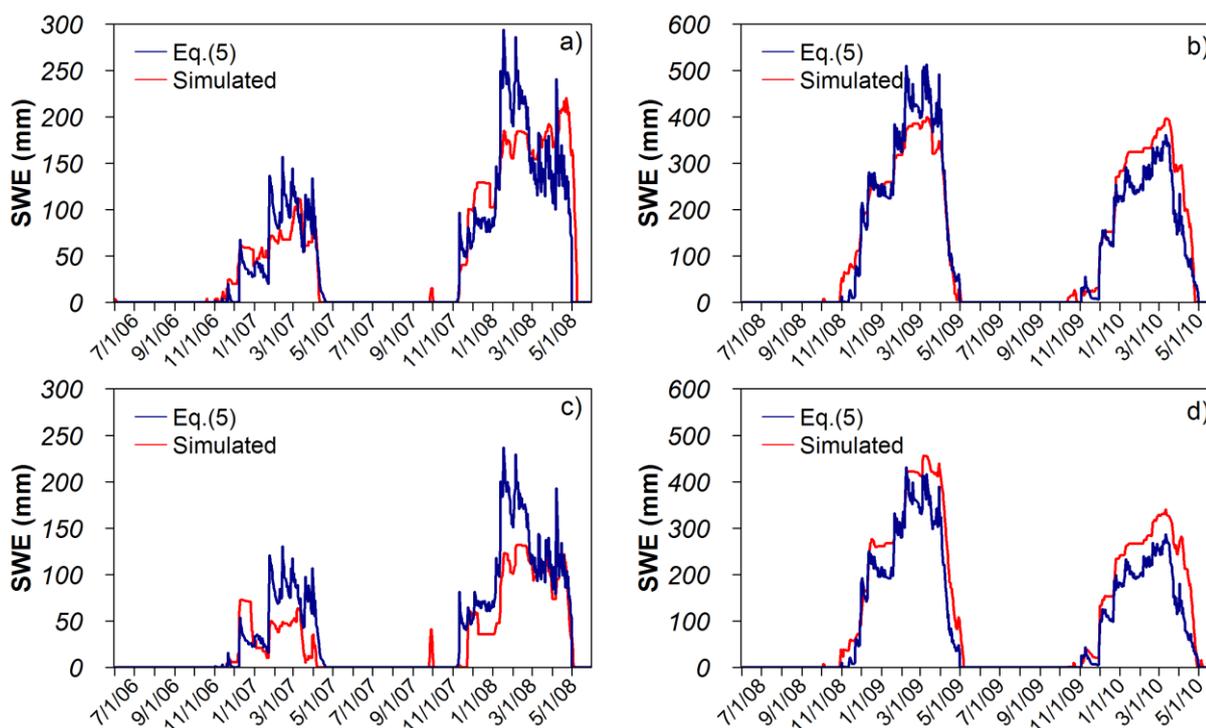
Figure 4.5 Model performance of sub-basin-scale snow parameters for fitting sub-basin SWE during both the calibration and validation periods (calibration procedure II): (a) calibration and (b) validation results of sub-basin 12; (c) calibration and (d) validation results of sub-basin 17.

With the sub-basin-scale snow parameters, the snow seasons of the two sub-basins were also captured well (Figure 4.5). In particular, the period of 2008-2009 was reproduced better in comparison to the performance of the basin-scale parameterization in Figure 4.2. At sub-basin 12 (Figure 4.5a and 4.5b), the sub-basin SWE values were reproduced well in both calibration (R^2=0.82, NSE=0.81) and validation (R^2=0.93, NSE=0.93). At sub-basin 17 (Figure 4.5c and 4.5d), the model results generally displayed underestimations in the maximum SWE during the calibration period, and presented overestimations during validation.

Even with the sub-basin-specific snow parameters, clear differences were observed at most elevation bands by comparing the EB-SWE with the point value from the corresponding snow stations at sub-basin 12 (Figure 4.6). However, in comparison to the basin-scale results (Figure 4.3), closer matching was observed at sub-basin 12 (Figure 4.6). For sub-basin 12 during the calibration period, only the EB-SWE of band 3 was close to the data of Matsch station with NSE of 0.65 (Figure 4.6c), while the EB-SWE of the other bands presented obvious underestimations (Figure 4.6). During validation, the EB-SWE of high elevation areas (Figure 4.6f and 4.6h) approached the observations with high values of NSE > 0.70, whereas the simulations of lower elevation areas (Figure 4.6b and 4.6d) generally showed underestimation except for the case of ST. Valentin station. For sub-basin 17, sub-basin-specific snow parameters generally led to improvements (Figure 4.7) in comparison to calibration approach I. During

94

the calibration period, close matching was reached between EB-SWE and the station observations of all elevation bands (Figure 4.7a, 4.7c and 4.7e). During validation, although higher NSE values (Figure 4.7) were generally obtained compared to the results of approach I (Figure 4.4), obvious differences were still observed at the low elevation bands with negative NSE values (Figure 4.7b and 4.7d). An exceptional case was displayed between the EB-SWE of band 2 and the observations from Stausee Zoggl station, which got a close match with NSE of 0.62. The EB-SWE of the high elevation band matched the observed value quite well with an NSE of 0.93 (Figure 4.7f).
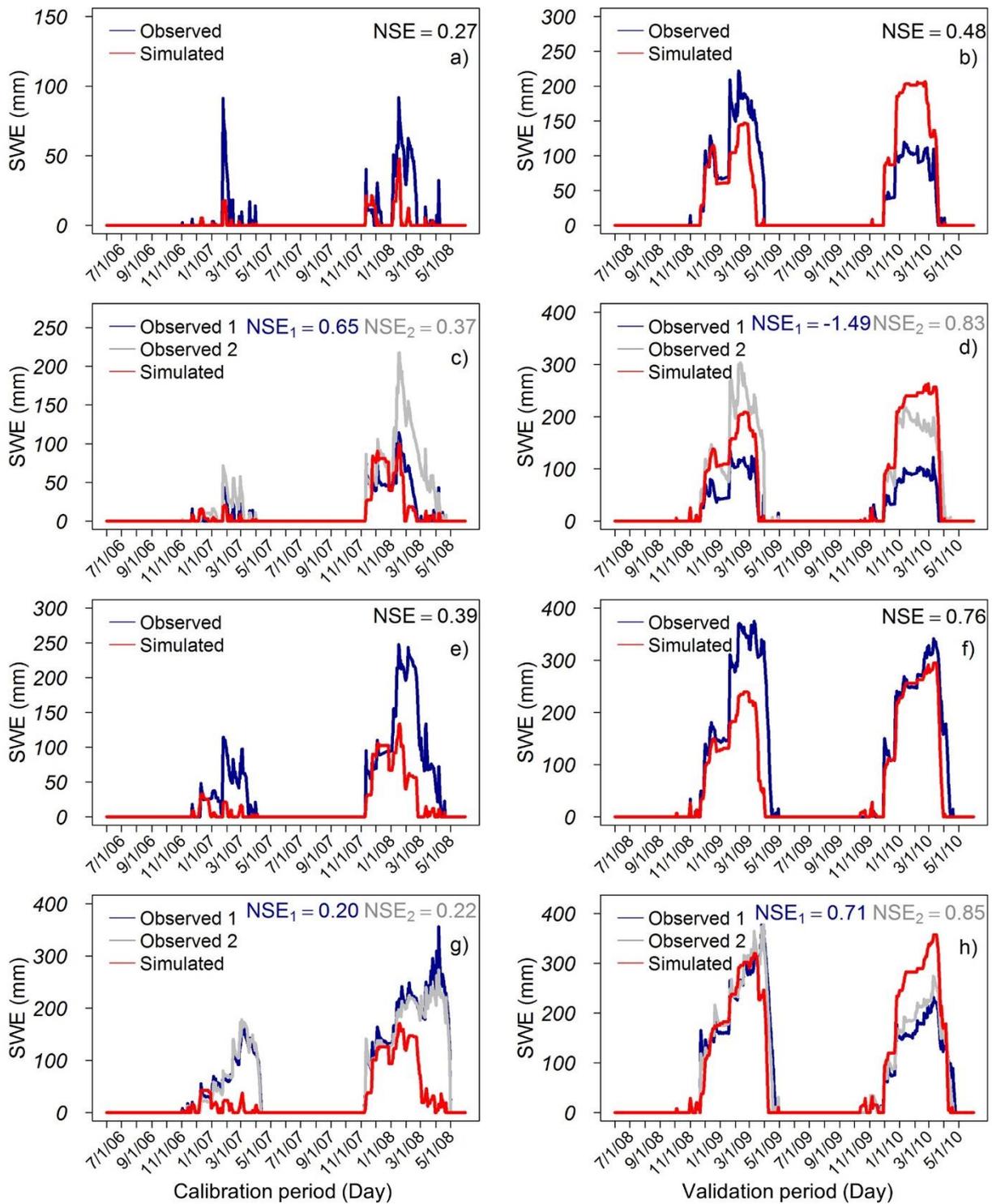
Figure 4.6 SWE comparison between values calculated from station observations and the simulated values of the corresponding elevation band at sub-basin 12 with sub-basin-scale snow parameters (calibration procedure II): (a) calibration and (b) validation performances of band 1 and Marienberg station; (c) calibration and (d) and validation performances of band 3: "Observed 1" stands for Matsch station, "Observed 2" represents ST. Valentin station; (e) calibration and (f) validation performances of

band 4 and Schlinig station; (g) calibration and (h) validation performances of band 5: "Observed 1" stands for Ausserrojen station, and "Observed 2" represents Melag station.
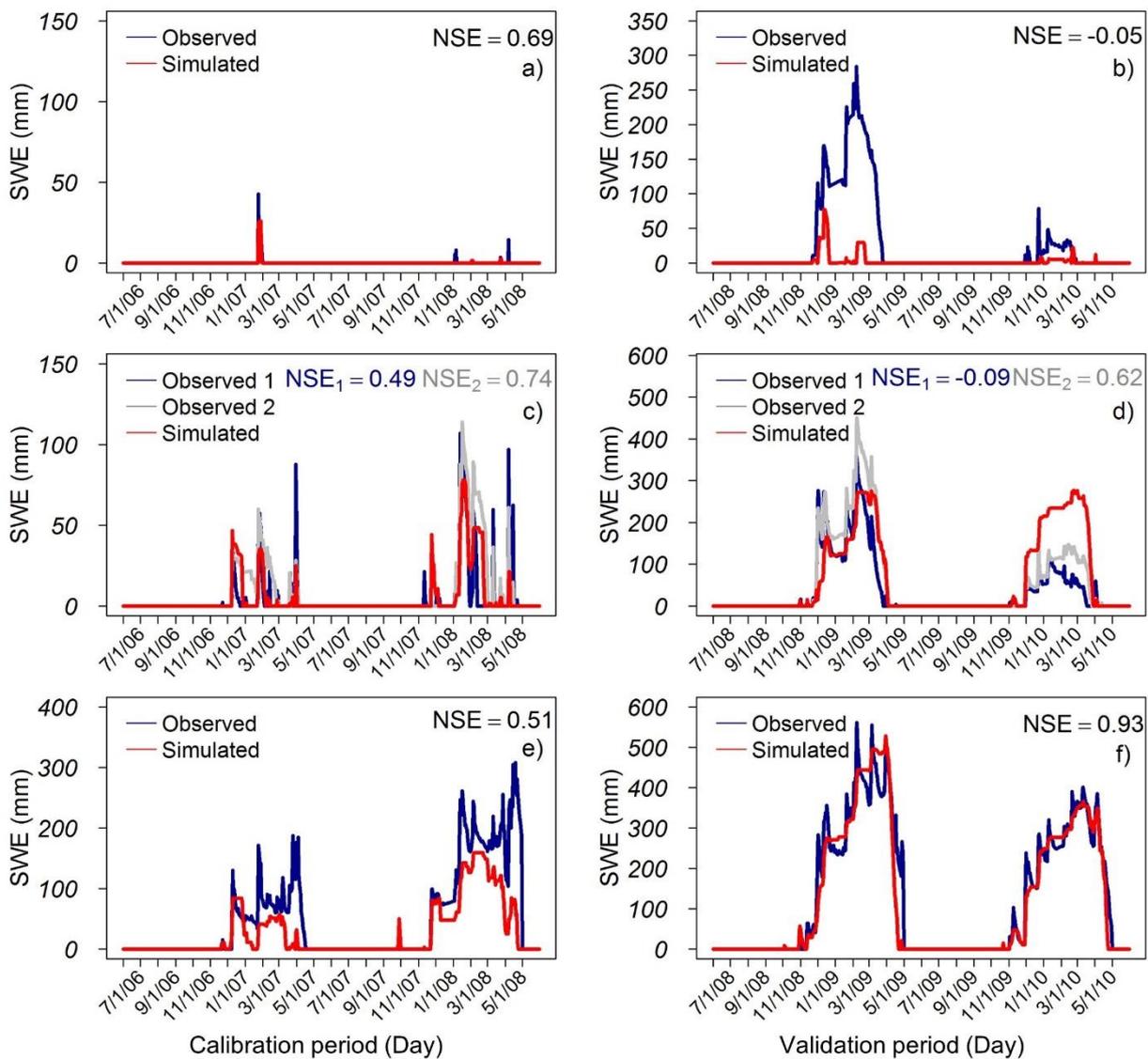
Figure 4.7 SWE comparison between values calculated from station observations and the simulated values of the corresponding elevation band at sub-basin 17 with sub-basin-scale snow parameters (calibration procedure II): (a) calibration and (b) validation performances of band 1 and Stausee ST. Pankraz station; (c) calibration and (d) validation performances of band 2: "Observed 1" stands for Pawigl station, "Observed 2" represents Stausee Zoggl station; (e) calibration and (f) validation performances of band 5 and Weissbrunn station.
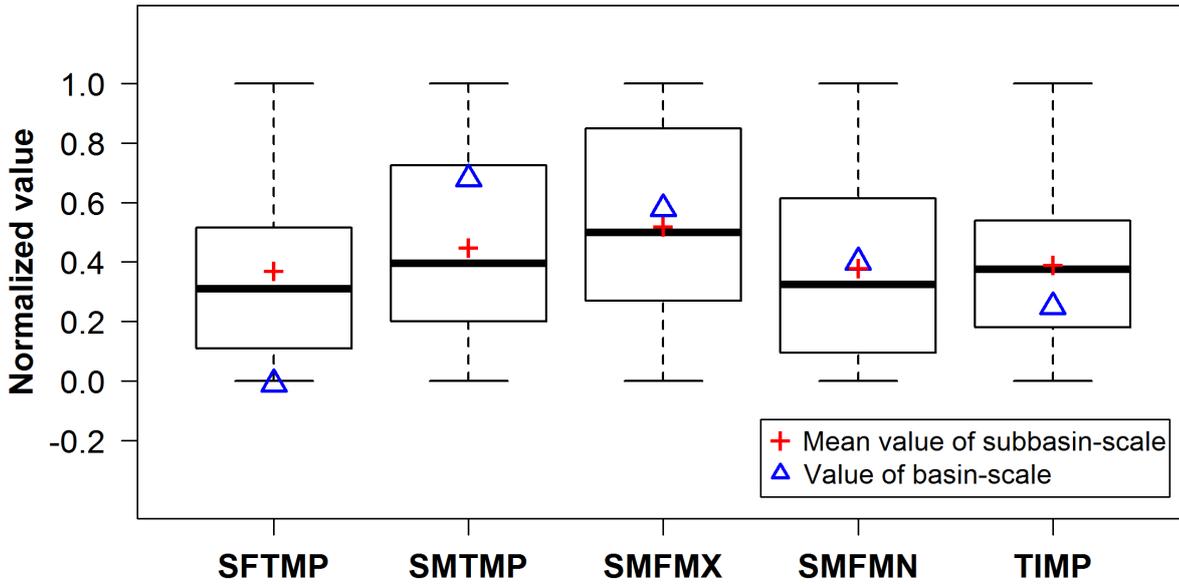
Figure 4.8 Comparison between the calibrated basin-scale snow parameter and sub-basin-scale parameters. Box plots display the distribution of sub-basin-scale snow parameters.

The calibrated values of each snow parameter for each sub-basin are provided in Figure S4.3 in the Supplementary material. In Figure 4.8, the normalized values of the snow parameters are displayed and they were calculated according to Eq. (4.13).

$$Normalized\ Value = \frac{Value - Min}{Max - Min} \qquad (4.13)$$

where *Value* represents the value of the calibrated parameter, *Max* and *Min* stand for the maximum and minimum values of each sub-basin-scale parameter.

As shown in Figure 4.8, great variability was observed for all five snow parameters among different sub-basins. SMFMX and SMFMN basin-scale parameters were very close to the mean value of the sub-basin-scale parameters. SMTMP and TIMP parameters were within the 25th and 75th percentile of the sub-basin-scale parameters. The largest discrepancy between basin-scale and sub-basin-scale parameters was observed for SFTMP.

4.3.3 Model performance using calibration procedure III
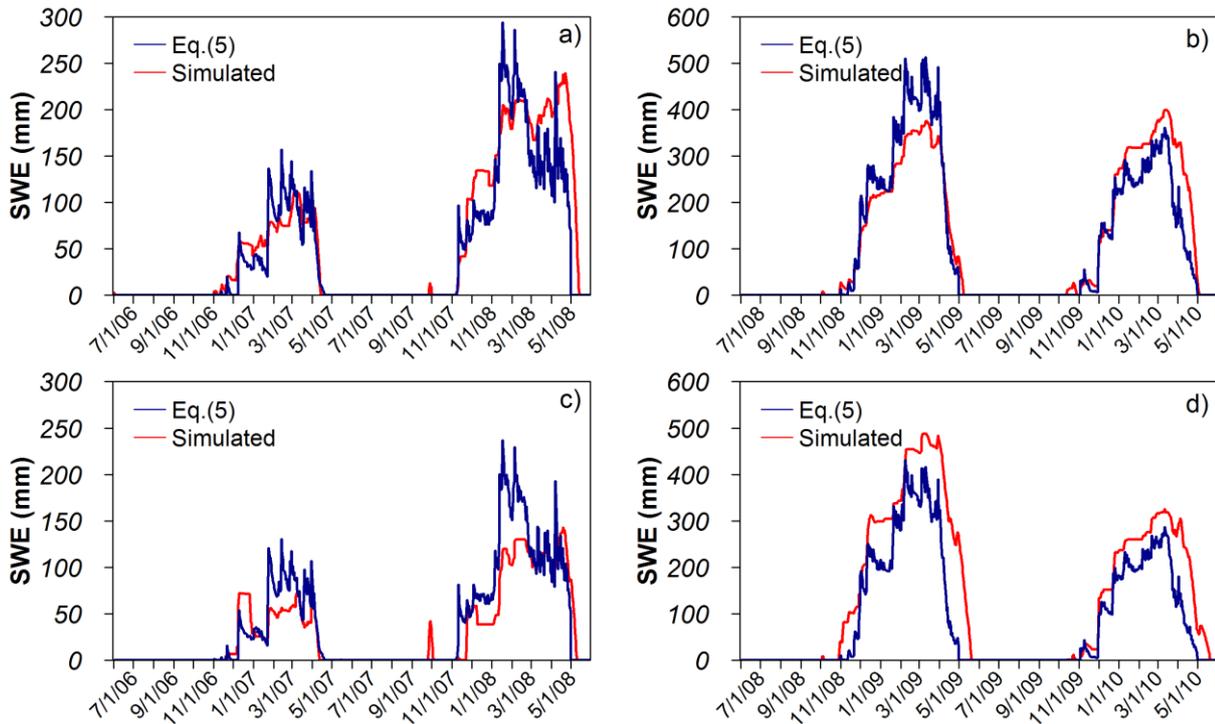
Figure 4.9 Model performance of manually modified snow parameters for fitting sub-basin SWE during both the calibration and validation periods (calibration procedure III): (a) calibration and (b) validation results of sub-basin 12; (c) calibration and (d) validation results of sub-basin 17.

In the third calibration procedure, after a first automatic calibration considering sub-basin specific snow data, the snow parameters were manually modified (Table 4.6) within the reasonable physical ranges to match the SWE values converted from the corresponding snow stations. The simulated Sub-SWE values of sub-basin 12 (Figure 4.9a and 4.9b) presented a very good estimation of reproducing both snow seasons and daily SWE behavior during both calibration (R^2=0.82 and NSE=0.78) and validation (R^2=0.89 and NSE=0.89). For sub-basin 17 (Figure 4.9c), the snow season of the calibration period was captured equally well as the results in Figure 4.2 and Figure 4.5. The simulated daily SWE displayed good performance (R^2=0.81 and NSE=0.78) with underestimations in the calibration period and also good performance (R^2=0.88 and NSE=0.61) with slight overestimations during validation.

Table 4.6 The calibrated values of parameters for sub-basins 12 and 17.

Parameters	Calibrated values		
	Approach I	Approach II	Approach III
Sub-basin 12			
PLAPS [mm H_2O/m]	0.056	0.051	0.055
TLAPS [°C/km]	-4.05	-4.11	-3.50
SFTMP [°C]	0.06	0.85	0.69

SMTMP [°C]	0.72	1.41	1.60
SMFMX [mm H_2O °$C^{-1}day^{-1}$]	5.04	6.41	3.49
SMFMN [mm H_2O °$C^{-1}day^{-1}$]	3.23	3.01	1.83
TIMP	0.25	0.42	0.30
Sub-basin 17			
PLAPS [mm H_2O/m]	0.045	0.032	0.025
TLAPS [°C/km]	-5.98	-5.38	-6.09
SFTMP [°C]	0.06	1.20	0.81
SMTMP [°C]	0.72	1.50	1.80
SMFMX [mm H_2O °$C^{-1}day^{-1}$]	5.04	5.23	3.14
SMFMN [mm H_2O °$C^{-1}day^{-1}$]	3.23	2.92	2.48
TIMP	0.25	0.25	0.43

With the manually modified snow parameters, most of the EB-SWE presented reasonable behaviors compared to snow observations. At sub-basin 12 during calibration (Figure 4.10a, 4.10c, 4.10e and 4.10g), the EB-SWE of all the highest elevation areas got evidently close to the station values with a minimal NSE of 0.57. During validation (Figure 4.10b, 4.10d, 4.10f and 4.10h), a big difference was observed between band 3 and the values from ST. Valentin station, while high NSE values were obtained for the other five stations. Compared to the results of approaches I and II (Figure 4.3, 4.4, 4.6, and 4.7), the improvements were very obvious at sub-basin 12. As for sub-basin 17, good estimations were presented at all the considered elevation bands during calibration, which had the lowest NSE of 0.51 (Figure 4.11a, 4.11c, and 4.11e). During validation, negative NSE values were observed between band 1 and Stausee ST. Pankraz station (Figure 4.11b), and also between band 2 and Pawigl station (Figure 4.11d). Located at the same elevation band, Stausee Zoggl station revealed values close to the EB-SWE of band 2, with an NSE of 0.61 (Figure 4.11d). For the high elevation areas (Figure 4.11f), the EB-SWE approached the observations from Weissbrunn station with a very high NSE=0.95. In comparison to the results of the other two approaches (Figure 4.3, 4.4, 4.6, and 4.7), the manually modified snow parameters led to improvements, particularly for the high elevation band. Overall, the modified snow parameters resulted in "all-acceptable-performance" for the majority of all the elevation bands where the snow stations were located at sub-basins 12 and 17. Meanwhile, they kept a good fit between the simulated Sub-SWE and the referred values calculated based on observations.
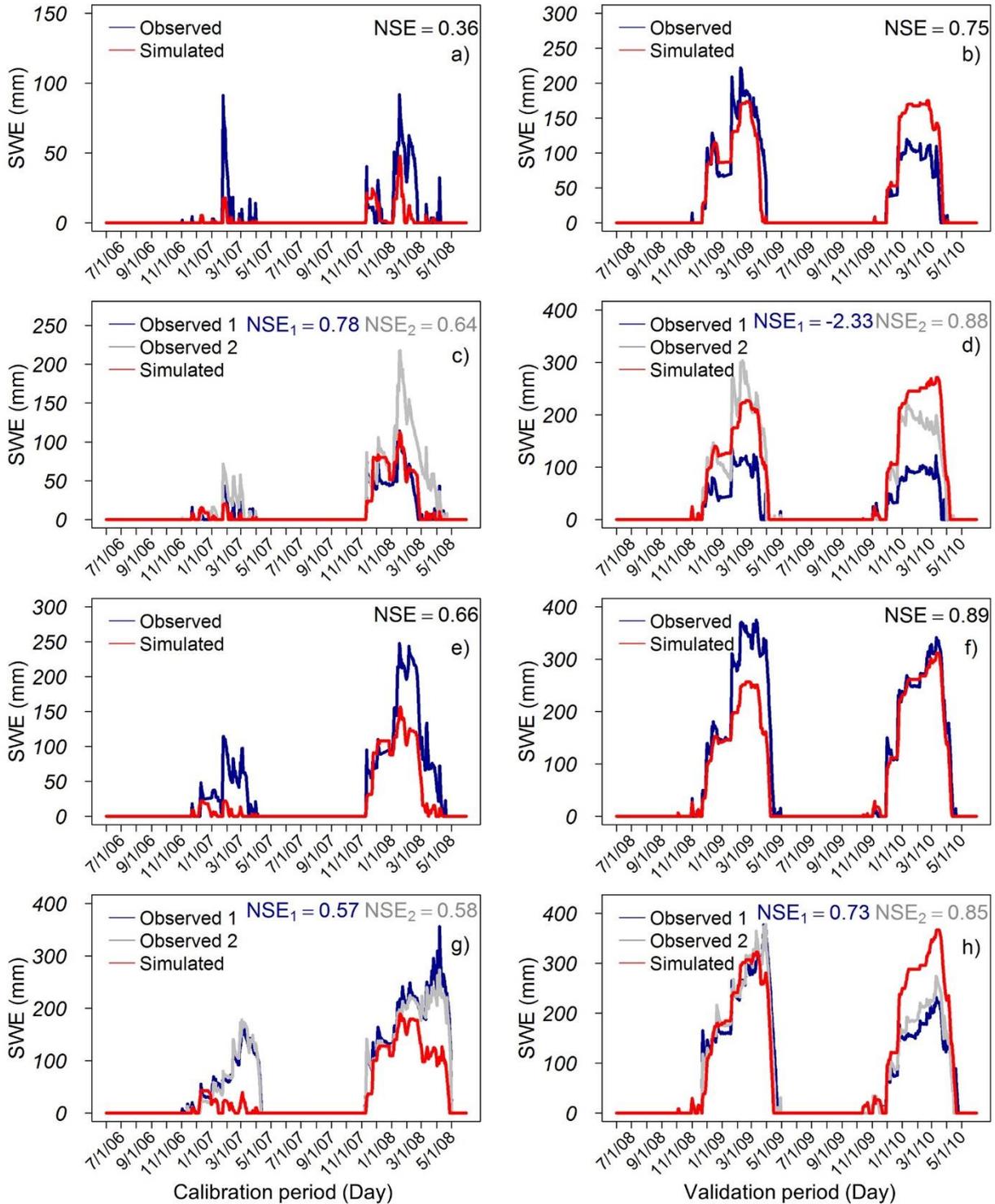
Figure 4.10 SWE comparison between values calculated from station observations and the simulated values of the corresponding elevation band at sub-basin 12 with manually modified snow parameters (calibration procedure III): (a) calibration and (b) validation performances of band 1 and Marienberg station; (c) calibration and (d) validation performances of band 3: "Observed 1" stands for Matsch station, "Observed 2" represents ST. Valentin station; (e) calibration and (f) validation performances of band 4

and Schlinig station; (g) calibration and (h) validation performances of band 5: "Observed 1" stands for Ausserrojen station, and "Observed 2" represents Melag station.
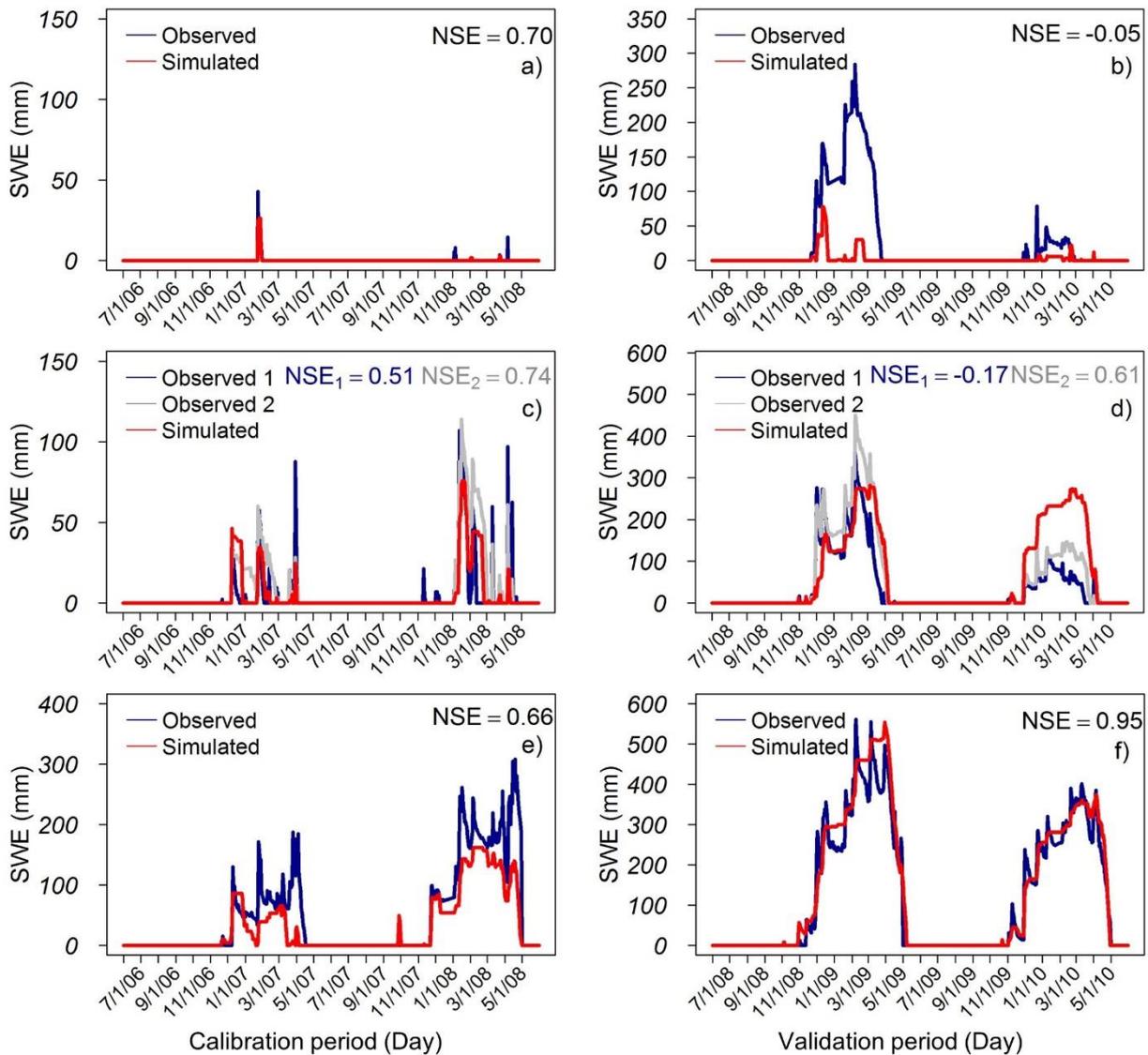
Figure 4.11 SWE comparison between values calculated from station observations and the simulated values of the corresponding elevation band at sub-basin 17 with manually modified snow parameters (calibration procedure III): (a) calibration and (b) validation performances of band 1 and Stausee ST. Pankraz station; (c) calibration and (d) validation performances of band 2: "Observed 1" stands for Pawigl station, "Observed 2" represents Stausee Zoggl station; (e) calibration and (f) validation performances of band 5 and Weissbrunn station.

4.3.4 Validation of SWAT model results against snow depth observations

As a final validation step of the applied calibration procedure we compare the snow depth observations available with the SWAT model results obtained using calibration procedure III. Since SWAT provides SWE time series, we converted them into snow depth values using Eqs. (4.3) - (4.5). We grouped the

available observations in six elevation classes and computed the mean of the observations (<HS>) and the mean of model results (HS_{SWAT}), for each elevation class, to simplify the comparison. Please notice that these elevation classes should not be confused with the elevation bands of the SWAT model. They have been defined only to provide a clearer representation of the results. As shown in Figure 4.12, observed snow depth values are highly variable also in the same elevation class, however, the model results capture very well the mean behavior of the observations, confirming that the calibrated model can reproduce observed snow dynamics, despite the strong assumptions introduced in the definition of <Sub_HS> and ρ. It is also interesting to observe that the model is able to capture well both snow accumulation as well as its melt, such that the snow cover duration is well described for all elevation classes.
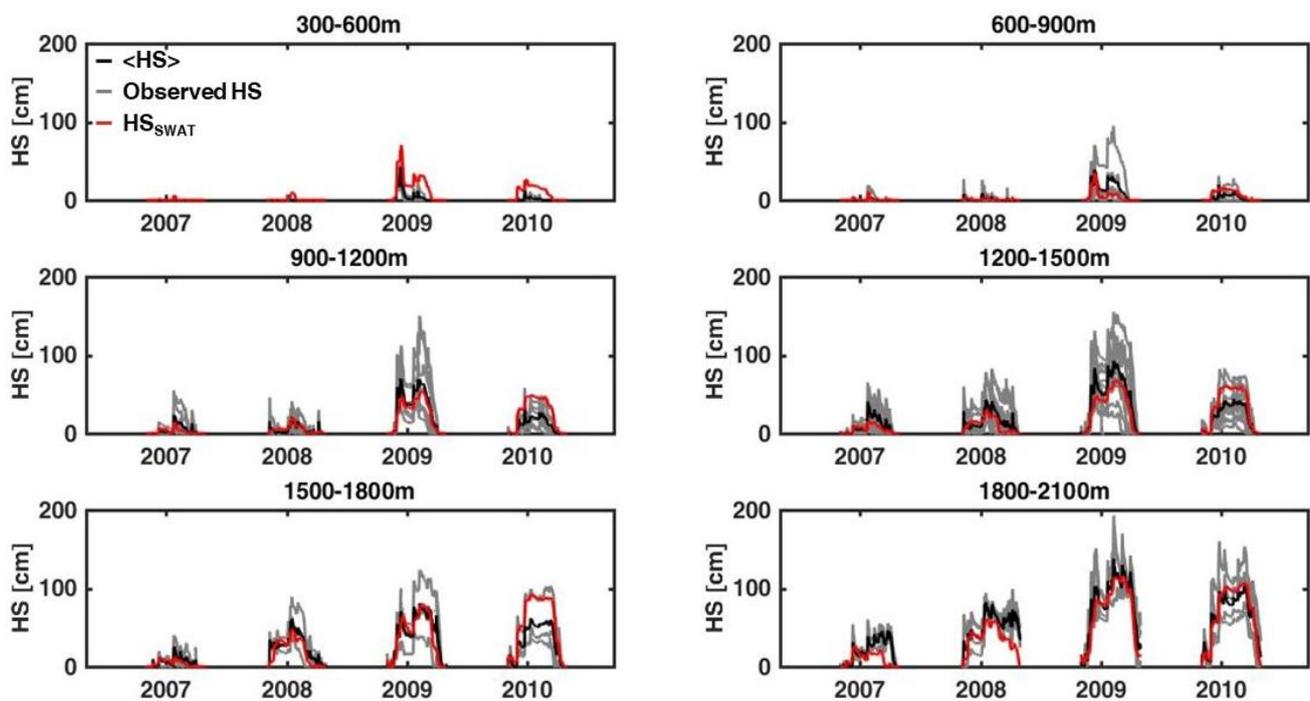
Figure 4.12 Comparison between observed snow depth for the available stations at various elevation ranges (Observed HS), their mean <HS> and the results of SWAT simulations (HS_{SWAT}) converted into snow depth using Eq. (4.5).

4.4 Discussion

4.4.1 Calibration of snow parameters considering Sub-SWE

Calibration of snow parameters according to the sub-basin averaged SWE can represent a useful approach for SWAT applications. The results of our study (Table 4.4 and Table 4.5) show that the definition of sub-basin specific snow parameter sets can be recommended when sub-basins have quite different elevation ranges, as is the case for sub-basins 5 and 20. Moreover, the sub-basin specific snow

parameterization is appropriate when snow monitoring stations are densely distributed in the catchment, i.e., enough data are available to perform model calibration at a finer spatial scale than the basin scale one.

It is well known that snow dynamics are highly dependent on several site specific conditions (Endrizzi et al. 2006, Hiemstra et al. 2002), such as exposure to wind, land cover, orientation and solar radiation. As described in the official documentation (Neitsch et al. 2011), the snow module of SWAT uses a non-distributed, temperature index-based approach. The model applies the same SWE value to each HRU. This value represents the sub-basin weighted average SWE considering the elevation band results. Moreover, snow processes in SWAT are controlled by precipitation and temperature inputs. SWAT assigns the same climate inputs to all HRUs of one sub-basin, independently of whether the elevation band method is applied or not (Neitsch et al. 2011, Tuo et al. 2016). This does not allow for more detailed snow parameterization at the HRU level, however, the definition of sub-basin-scale snow parameters can provide a rough description of the differences which are often observed in snow dynamics within a basin.

One possible way of improving the description of snow processes in SWAT could be the use of band specific climate data instead of the lumped climate data modified by elevation band. In addition, the elevation band could be used as a factor during HRU definition. This would allow calculation of the hydrological processes using different input data for different bands and differentiate between the processes occurring at high and low elevation HRUs. A dynamic HRU scheme based on snow cover change, along the lines of the dynamic glacier approach of Luo et al. (2013), may also improve the snow modeling performance of SWAT. Bieger et al. (2016) pointed out that a grid-based model setup option will be available in the new version SWAT+. Therefore, a grid-based approach could also be applied to the snow module in order to improve the spatially distributed simulation of snow processes. Meanwhile, this approach requires input data with higher spatial resolution and a longer computational time.

Considering model performance, the calibrated snow parameters were able to simulate the SWE of each sub-basin well (Table 4.4 and Table 4.5). During both the calibration and validation periods, acceptable R^2 and NSE values were obtained for 18 high elevation sub-basins with basin-scale snow parameters (Table 4.4) and all sub-basins with sub-basin-scale snow parameters (Table 4.5). Therefore, the performances of the two parameterization approaches could be considered as equally good at reproducing the overall behavior of snow process at the sub-basin scale for high elevation sub-basins. The length of the snow seasons was properly captured and daily variability of SWE was in general sufficiently well simulated (Figure 4.2 and 4.5 as examples). Although in some sub-basins (e.g. sub-basins 3,6, and 19) the basin scale parameters led to higher NSE or R^2 values than the sub-basin-scale

ones, a more specific parameterization for each sub-basin generally resulted in better performances (i.e., higher NSE and R^2 values) for the majority of the sub-basins. Considering the uncertainties related to data preprocessing described in section 4.2.2 as well as epistemic model uncertainty, it can be expected that the model is not able to perfectly reproduce SWE peaks and their fluctuations (Figure 4.2, 4.5 and 4.9). In particular, we consider the following to be the most important sources of uncertainty (Abbaspour et al. 2007): 1) the conversion process from snow depth to sub-basin averaged SWE; 2) the quality and uncertainty of the climate and temperature inputs, e.g. large uncertainty exists for precipitation in this mountainous region (Duan et al. 2016, Tuo et al. 2016); 3) the epistemic uncertainty of the model structure (e.g. lumped way of using climate input) and of the snow package itself (e.g. limitation of temperature index approach). These three sources of uncertainty mentioned above are to some extent known by the SWAT users' community (Abbaspour et al. 2007, Galván et al. 2014) and our recommendation is to avoid looking for a perfect match between model and computed SWE data, since it will most likely lead to model overfitting. Taking the issue of model and data uncertainty into proper consideration, the values of R^2 and NSE obtained in this study support the conclusion that the daily SWE time series can be reproduced using the SWAT snow module.

4.4.2 Calibration of snow parameters considering EB-SWE

Different combinations of snow parameter values have led to similar acceptable SWE simulations on the sub-basin level in both calibration and validation periods (Table 4.4-4.6, Figure 4.2, 4.5, and 4.9). Therefore, the use of Sub-SWE is not enough to identify the most suitable set of snow parameters. Specific snow outputs of each elevation band were used for further validation of the parameters and compared to SWE values computed from point observations of snow depth. As shown in Figure 4.3, 4.4, 4.6, 4.7, 4.10, and 4.11, the EB-SWE computed considering different sets of snow parameters led to different matching to the corresponding point observations. Calibrating snow parameters only referring to Sub-SWE, using both basin-scale or sub-basin-scale calibration, could fail to get acceptable EB-SWE values compared to the point observations (Figure 4.3, 4.4, 4.6, and 4.7). Especially for sub-basin 12 (Figure 4.3 and 4.6), the simulated SWE values were considerably different from the observed values, thus, for some cases these calibrated snow parameters were not able to reproduce the sub-basin snow conditions at the same time as well as the elevation band conditions. Unlike the sub-basin 12, the results of sub-basin 17 represented another situation in which the calibrated snow parameters of II may already lead to a close match with most point observations (Figure 4.7). As a consequence, the manual calibration of procedure III could not lead to large improvements (Figure 4.11). Under this situation, the calibration procedure III was suggested to be implemented as an additional attempt to optimize the model further.

Improvements (Figure 4.10-4.11) were observed after manual refinement of snow parameters by considering the performance of both Sub-SWE and EB-SWE. Calibration procedure III allowed us to obtain generally acceptable results at the sub-basin scale (Figure 4.9) and produced EB-SWE simulations closer to the values from station observations than calibration procedures I and II. For some bands, the model results were not satisfactory (Figure 4.10a, Figure 4.10d, Figure 4.11b, and Figure 4.11d) even after the calibration procedure III. It cannot be excluded that the poor performance of these cases could be ascribed to: 1) the simplified snow modeling structure of SWAT; 2) measurement errors; 3) the simplified approach used for the conversion from snow depth to snow water equivalent; 4) the specific local conditions influencing a single station. The local variability in snow observations within the same elevation band can be appreciated by comparing for example ST. Valentin (1499 m) and Matsch (1570 m) stations (for example, Figure 4.3c and 4.3d), in which the higher station even displayed lower SWE values than the lower station. In mountainous regions, in addition to elevation, local snow dynamics are also affected by the specific topographic factors of the location like valley or ridge, sunny side or shaded side due to radiation or wind effect (Endrizzi et al. 2006, Hiemstra et al. 2002). In SWAT, only elevation effects can be considered by applying the elevation band method to partially take into consideration this source of variability. The NSE of EB-SWE generally increased from the simulation of basin-scale parameters to that of the sub-basin-scale parameters, from simulation of the sub-basin-scale parameters to that of the manually modified parameters.

Independently of the three calibration procedures, some discrepancies between the sub-basin-averaged SWE and SWAT model results seem to be persistent. As shown in Figure 4.2, 4.5 and 4.9, we always observed an earlier snow peak simulated around late September of 2007 for sub-basin 17. Meanwhile, there was no station presenting a peak during the same period (Figure 4.4, 4.7, and 4.11). This disagreement could be ascribed to the earlier peak simulated in high elevation bands (Figure 4.4e, 4.7e and 4.11e). The high elevation areas accounted for more than 50% of sub-basin 17. As a result of the lumped approach of the elevation band method (Eq. (4.10)), the earlier peak was propagated to the simulated Sub-SWE.

Limited by climate data availability, computational time and snow module structure, it is also not feasible to delineate very small sub-basins for capturing the specific local condition of each station. In fact, it is not appropriate to calibrate the model by fitting EB-SWE to point observations, because one band covers a range of elevations that may contain several stations with clearly different snow records. The snow parameters of a lumped model calibrated based on one station cannot be considered representative of the situation in an entire region (Egli and Jonas, 2009). Therefore, we suggest using only the EB-SWE as an additional reference for model assessment by qualitatively judging whether the EB-SWE is reasonable or not in comparison to the station observation of each band. Furthermore, the

parameters could be modified accordingly by making a compromise between sub-basin averaged SWE outputs and the EB-SWE performance, which could actually provide reasonable snow simulations by systematically considering both model characteristics and the natural snow processes. Considering the recent coupling of SWAT with MODFLOW and RT3D performed by Bailey et al. (2017) and Bailey et al. (2016) to have a better spatial description of groundwater flow and transport processes, a combination of SWAT and a more detailed grid-based snow module could be envisioned. However, as snowmelt is considered a part of the effective rainfall, the precipitation method of SWAT should also be modified to fit the grid structure.

4.4.3   Spatial variability of snow parameters

Snow parameters obtained using calibration procedure I and II both led to good performance for the sub-basins located at high elevations, although they identified different combinations of the parameter set. The comparison of the basin-scale parameters with the mean values of sub-basin-scale parameters (Figure 4.8) presented a small difference for SMFMX, SMFMN and TIMP, indicating a similar melt process (Eq. (4.9)). Meanwhile, large discrepancies were observed for SFTMP and SMTMP. Lower SFTMP and higher SMTMP of basin scale parameters lead to less snow fall into the sub-basins and less snowmelt. On the other hand, the combination of higher SFTMP and lower SMTMP of the sub-basin parameter set indicated the situation of more snow fall and more snowmelt. As a consequence, the accumulated snow pack at the sub-basin scale was similar for both parameterization approaches. However, considering the EB-SWE results at sub-basins 12 and 17, sub-basin-scale parameters led to a closer match between the EB-SWE and the observations from snow monitoring stations than the basin-scale parameters. This result may suggest that by defining specific snow parameters individually for a sub-basin, the sub-basin-scale parameterization is capable of capturing snow processes better than the uniform assignment of snow values for the entire basin.

The great variability observed in the sub-basin-scale snow parameters (Figure 4.8) could be ascribed to the complex spatial variability of the natural snow processes in alpine catchments (Egli and Jonas 2009, Hiemstra et al. 2002). Moreover, the obvious differences between the sub-basin dependent snow parameter values and the uniform basin-scale values (Figure 4.8) also emphasized that a single set of snow parameters may not realistically reproduce the natural snow processes of every sub-basin. In addition, we can observe that different calibration approaches led to different optimal snow parameters sets (Table 4.6), which resulted in good performances in reproducing sub-basin SWE (Figure 4.2, 4.5, and 4.9) but resulted in the main difference in the results obtained by comparing the EB-SWE (Figure 4.3, 4.4, 4.6, 4.7, 4.10, and 4.11). What really matters is to choose the approach according to the topographic characteristics of the sub-basins. Therefore, it is concluded that for snow-dominated,

elevation-diverse river basins, sub-basin-scale snow parameters are preferable to help investigate the snow processes appropriately and accurately by better reflecting the complex spatial variability of snow. Basin-scale snow parameterization is suitable for river basins characterized by uniform topography, which display similar snow behavior over all sub-basins.

4.5 Conclusion

Modeling snow processes is of great importance in snow-dominated river basins like in the Alps. In view of the model structure and parameter complexity of SWAT, we recommend calibrating snow parameters using SWE information. Ground snow observations or remote sensing snow products can provide us with this information. In this technical note, we proposed three calibration approaches. First, a unique set of snow parameters for the entire basin was defined for model calibration using sub-basin averaged SWE values. The second approach aimed at considering the spatial variability of snow parameters, which defined a sub-basin specific set of snow parameters for calibrating the model with sub-basin averaged SWE. In the third and final procedure, in addition to sub-basin-averaged SWE values, we also utilized the station-specific information to calibrate the model. This last calibration approach is capable of reaching a good compromise that fits both SWE at the sub-basin level and the SWE specific to the stations. Currently, the third calibration procedure proposed in this work could be an easy and feasible option for the calibration of alpine snow processes. Moreover, a sub-basin-scale snow parameterization seems to be more suitable for models where the sub-basins display a large variability in elevation range. The use of elevation bands permits qualitatively comparing the SWAT EB-SWE results with the information available for each station. In this work, we also suggest possible modifications of the SWAT model structure to improve the representation of snow processes. In particular, a grid-based approach or the definition of elevation dependent HRUs and meteorological input can be effective solutions for an improved application of SWAT in alpine catchments. Moreover, the third calibration procedure could be jointly used with the traditional SWAT model calibration performed considering discharge. A comparison between multi-objective and single-objective calibration is outside the scope of this technical note, but it deserves further investigation (Hanzer et al. 2016, Rajib et al. 2016).

Chapter 5

A Multi-Objective Approach to Improve SWAT Model Calibration in Alpine Catchments[4]

Multi-objective hydrological model calibration can represent a valuable solution to reduce model equifinality and parameter uncertainty. The Soil and Water Assessment Tool (SWAT) model is widely applied to investigate water quality and water management issues in alpine catchments. However, the model calibration is generally based on discharge records only, and most of the previous studies have defined a unique set of snow parameters for an entire basin. Only a few studies have considered snow observations to validate model results or have taken into account the possible variability of snow parameters for different subbasins. This work presents and compares three possible calibration approaches. The first two procedures are single-objective calibration procedures, for which all parameters of the SWAT model were calibrated according to river discharge alone. Procedures I and II differ from each other by the assumption used to define snow parameters: The first approach assigned a unique set of snow parameters to the entire basin, whereas the second approach assigned different subbasin-specific sets of snow parameters to each subbasin. The third procedure is a multi-objective calibration, in which we considered snow water equivalent (SWE) information at two different spatial scales (i.e. subbasin and elevation band), in addition to discharge measurements. We tested these approaches in the Upper Adige river basin where a dense network of snow depth measurement stations is available. Only the set of parameters obtained with this multi-objective procedure provided an acceptable prediction of both river discharge and SWE. These findings offer the large community of SWAT users a strategy to improve SWAT modeling in alpine catchments.

5.1 Introduction

Multi-objective calibration of hydrological models has proven useful in producing reliable predictions of several hydrological components and reducing model equifinality and uncertainty (Bekele and Nicklow, 2007; Hanzer et al., 2016; Lu et al., 2014; Pfannerstill et al., 2017; Rajib et al., 2016). Snow observations represent an optimal dataset to calibrate hydrological models in conjunction with streamflow since they are available in many alpine catchments in terms of snow depth values and often have higher spatial resolution than river discharge gauging stations. Moreover, snow dynamics play a key role in determining the seasonal characteristics of river discharge, which further affects other hydrology-related processes like erosion, hydropower operation, water supply, as well as flood prevention (Cunderlik and Ouarda, 2009; Pradhanang et al., 2011; Rahman et al., 2013; Zampieri et al., 2013). Therefore, in order to conduct a comprehensive and reliable hydrological investigation in alpine catchments, it is not only important to focus on river discharge, but it is also necessary to accurately describe the snow processes (Endrizzi et al., 2006; Grusson et al., 2015; Hanzer et al., 2016; Troin and Caya, 2014; Warscher et al., 2013). Hence, a multi-objective calibration approach considering both discharge and snow measurements can be a feasible and useful method to improve the application of hydrological models, like the Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998), in alpine catchments.

SWAT is one of the most comprehensive and applied hydrological models (Abbaspour et al., 2015; Francesconi et al., 2016; Golmohammadi et al., 2017; Liu et al., 2016; Malagò et al., 2016; Shen et al., 2012; Tuo et al., 2016). As outlined in the present work, SWAT adopts a specific approach to describe snow processes: the same snow water equivalent (SWE) value is assigned to all hydrological response units (HRU) in a subbasin, and this value is derived from the weighted sum of the SWE computed for each elevation band in a subbasin. Therefore, output time series of SWE are available at two spatial scales: for the entire subbasin (the average subbasin SWE) and for each elevation band.

To the best of our knowledge, only a few studies have taken these two snow outputs into account to evaluate the performance of the SWAT snow module (Fu et al., 2014; Fu et al., 2015; Grusson et al., 2015; Pradhanang et al., 2011) and no study has included the snow information of subbasin and elevation band scales into model calibration processes. Previous studies mostly assigned a unique set of snow parameters to an entire basin (e.g., Fu et al., 2014; Fu et al., 2015; Rahman et al., 2013) without considering the heterogeneous morphology encountered in large (>1000 km^2) catchments (Egli, 2011; Egli and Jonas, 2009). Moreover, multi-objective model calibration has been attempted in SWAT, considering water quality, soil moisture, evapotranspiration, or glacier information along with streamflow (Abbaspour et al., 2007; Bekele and Nicklow, 2007; Cao et al., 2006; Lu et al., 2014;

Maringanti et al., 2011; Omani et al., 2017; Pfannerstill et al., 2017; Rajib et al., 2016; White and Chaubey, 2005), but snow data, which are very important information in snow-dominated catchments to represent snow dynamics, have rarely been used, due to the typical lack of SWE time series. In snow-dominated catchments, SWAT model parameters are generally estimated using discharge records alone (e.g., Levesque et al., 2008; Rahman et al., 2013; Wang and Melesse, 2005). Since river discharge is the combination of the contributions of different hydrological components, calibration based on discharge alone can easily lead to the problem of model equifinality (Beven, 2006; Hanzer et al., 2016).

In this work, for the first time in the framework of multi-objective SWAT model calibration, we take advantage of a simple empirical regression equation which has recently been proposed to convert snow depth data into snow water equivalent (Pistochhi, 2016). The conversion equation is useful since SWE is the typical variable modeled in hydrological models, while snow depth is the most common information about the snowpack. Daily SWE data of both subbasin and elevation band scales, are computed based on snow depth measurements and further applied in the multi-objective calibration procedure. We compare then three approaches to calibrate SWAT model parameters in alpine catchments. Two are classical single objective calibration procedures, while the third one is a novel multi-objective calibration. The novelties in the latter approach include the use of SWE data derived from snow depth measurements and the use of these SWE information at two spatial scales (subbasin and elevation band) to calibrate the model along with discharge. The possibility of modeling both discharge dynamics and changes in SWE allows us to verify whether it is possible to improve model performance in alpine catchments considering multiple datasets for calibration.

The specific objectives of this study are to: 1) investigate the effect of using the same or different snow model parameters for different subbasins on model performance considering a single-objective discharge-based calibration procedure; 2) propose a multi-objective calibration procedure involving both discharge and two spatial scales of snow output for SWAT modeling in snow-dominated alpine catchments; 3) assess the model performances of single-objective and multi-objective procedures in the Upper Adige catchment, an important alpine basin in Northeast Italy; 4) compare the behavior of major hydrological components between the single-objective and multi-objective procedures.

5.2 Material and method

5.2.1   Study area

The Upper Adige river basin closed at Bronzolo (6875 km^2) is a typical Alpine catchment located in the Northeastern Italian Alps (Figure 5.1). It has a large elevation range (223 m a.s.l. - 3865 m a.s.l.) and is characterized by steep slopes and sharp hydro-climatic gradients (Callegari et al., 2015). From late

October to April, snowfall is the dominant form of precipitation in this area, and the snow cover duration is elevation-dependent (Marcolini et al., 2017b). As a consequence, snowmelt water is the dominant contributor to streamflow in the spring (Mei et al., 2014; Mei et al., 2016; Penna et al., 2014). Land uses of the river basin are mainly forest (55.8 %), grassland and pasture (21.3 %), and barren land (11.5 %). The main soil types are loamy sand (45.5 %), silty clay (21.5 %), and sand (17.3 %). With 13 major hydropower plants present in the catchment, several subbasins are affected by hydropeaking (Figure 5.1). Further information about the catchment is provided in the recent review of Chiogna et al. (2016).
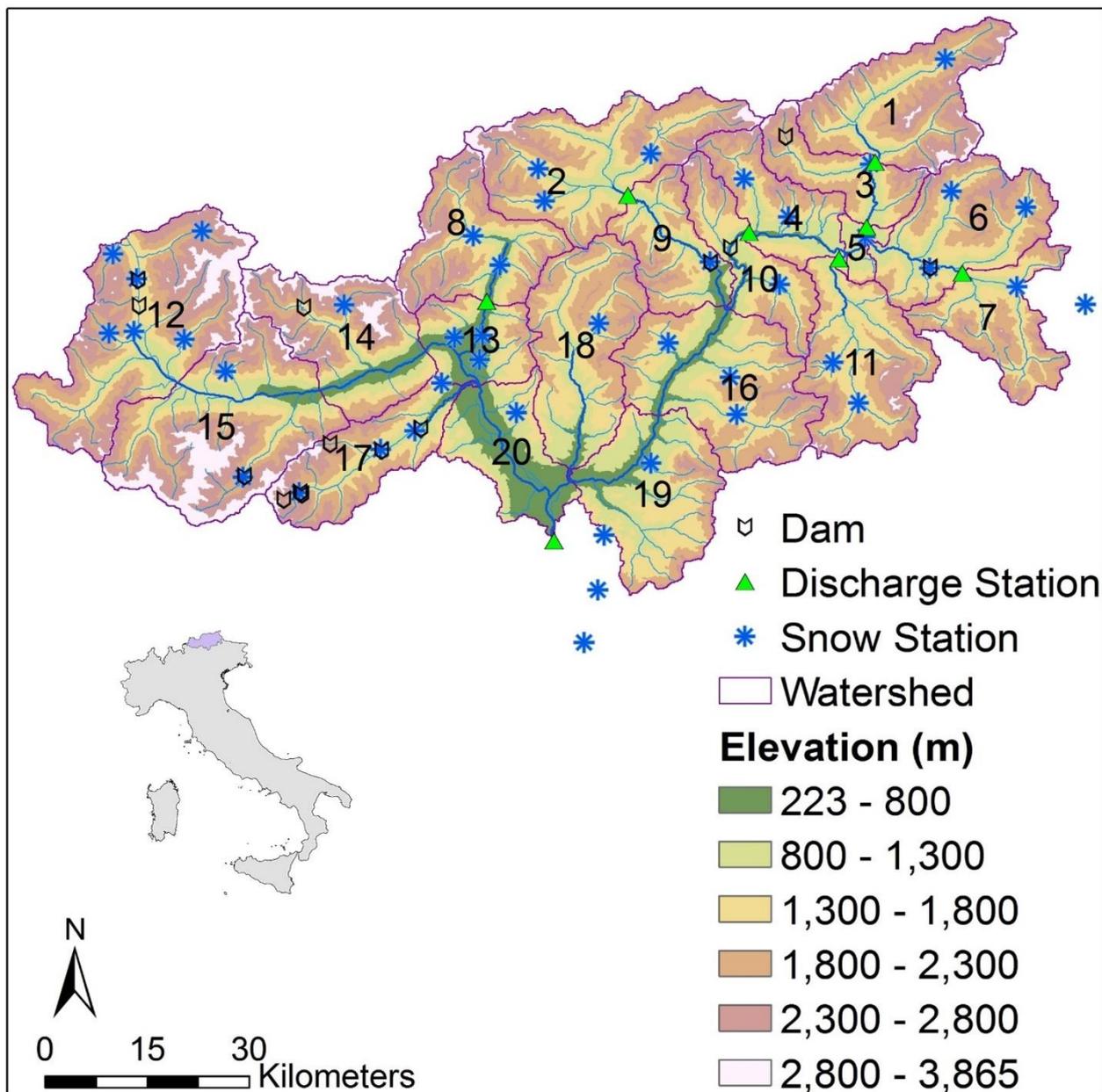
Figure 5.1 The Upper Adige river basin.

5.2.2   Computation of snow water equivalent values

Based on data availability for the period 2005-2010, we selected 44 snow measurement stations recording daily snow depth values in the Upper Adige river basin (Figure 5.1 and Table S5.1 in Supplementary material). These snow data represent a subset of the data used by Adler et al. (2015) and Marcolini et al. (2017b) to analyze climatic changes occurring in this region. The available time series have been quality checked and their homogeneity has been tested using the standard normal homogeneity test (Marcolini et al., 2017a).

In order to make use of these snow data for SWAT model calibration, two types of SWE data were computed, starting with the available snow depth observations: the mean subbasin scale SWE ($< Sub\_SWE >$) and the SWE of each station. A detailed description of the calculations is given in Appendix A; here, we illustrate the main steps of the procedure qualitatively. The subbasin scale SWE values were calculated from the mean snow depth data (HS) computed for each of the 20 subbasins (Figure 5.1) as follows. First, daily HS measurements were interpolated using a simple power law model to establish a correlation between HS and elevation. The interpolated values were then used to compute a mean subbasin snow depth, and finally these values were converted into SWE data using the equation suggested by Pistocchi (2016) for the temporal evolution of snow density. The SWE values for each snow station were computed by multiplying the snow depth observations by the temporal evolution of snow density according to the equation suggested by Pistocchi (2016). The computed $< Sub\_SWE >$ and the SWE of each station were then used to calibrate and validate the SWAT model. The $< Sub\_SWE >$ time series were compared to the subbasin SWE output of SWAT, while the SWE values for each snow station were compared qualitatively to the corresponding elevation band SWAT output.

5.2.3 SWAT model

5.2.3.1 Model description

The SWAT model, developed by the Agricultural Research Service of the United States Department of Agriculture, is a comprehensive, time-continuous, semi-distributed, process-based model (Arnold et al., 2012a; Arnold et al., 1998). It can be used to model the water cycle, erosion, vegetation growth, and water quality in large river basins and to assess the effects of climate change and water resources management (Abbaspour et al., 2015; Dile et al., 2016; Yang et al., 2016). SWAT divides river basins into subbasins composed of HRUs, which are characterized by different combinations of land use, soil characteristics and topography. The hydrological cycle is calculated based on water balance, which is highly affected by climate input such as daily precipitation and maximum/minimum air temperature. Using daily input time series, SWAT simulates the daily, monthly and yearly fluxes of water and solutes in river basins. The water yield is routed through the streams and reservoirs along the river network.

Even though SWAT was originally developed to quantify the impact of land management practices in large, complex catchments, its application in mountainous river basins is not rare (Ahl et al., 2008; Debele et al., 2010; Grusson et al., 2015; Rostamian et al., 2008; Troin and Caya, 2014). It has, for example, been used to support regional water management practices such as flood control (Yang et al., 2014), hydropower management (Malagò et al., 2015; Rahman et al., 2013) and assessment of climate change adaptation schemes (Pradhanang et al., 2011). SWAT uses the elevation band method to consider the orographic effects on precipitation and temperature in mountainous areas. Up to ten elevation bands can be defined in each subbasin. This method has been applied in several snow-dominated alpine catchments and proven useful and essential (Grusson et al., 2015; Malagò et al., 2015; Morán-Tejeda et al., 2014; Pradhanang et al., 2011; Rahman et al., 2013). In this work, six elevation bands were applied.

5.2.3.2 Snow package

The snow module of SWAT is based on the temperature-index approach (Hock, 2003; Neitsch et al., 2011; Walter et al., 2005) and it is able to simulate changes in SWE for several user-defined elevation bands. Although more complex snow modules have been tested in previous studies, no significant improvement has been observed (Debele et al., 2010; Meng et al., 2015).

SWAT provides snow output in the form of snowmelt and SWE with identical units of precipitation (Pradhanang et al., 2011). A detailed description and equations of the snow packages are available in the SWAT user's manual (Neitsch et al., 2011). Among the snow parameters (Table 5.1), *SMFMX* should be larger than *SMFMN* in our study area, because *SMFMX* is the maximum melt factor in the Northern Hemisphere and the minimum melt factor in the Southern Hemisphere, which is explained in the SWAT I/O files (Arnold et al., 2012b). In SWAT, the user can choose to assign either a unique set of snow parameters to the entire basin or define a different set of snow parameters for each subbasin. For the latter case, the source code has to be modified: The files "readsub.f" and "readsno.f" contain an inappropriate constraint: negative and zero values of *SFTMP* and *SMTMP* are not allowed at the subbasin scale. This is contradictory to the SWAT official documentation (Arnold et al., 2012b), which defines the physical range for *SFTMP* and *SMTMP* between -5 and 5 (e.g. Grusson et al., 2015).

SWAT output files provide SWE values for each elevation band in the output.snw file. In the following, we refer to this SWE output as EB-SWE (elevation band snow water equivalent). Each elevation band covers an area with a range of elevations, and the SWE depends on snow depth, which is elevation-dependent. Therefore, the EB-SWE output cannot be calibrated to fit exactly with the data available for a single monitoring station. Still, a qualitative comparison is feasible.

Snow outputs (snowmelt and SWE) are stored in the output.hru file at the HRU scale. SWAT assigns the elevation-weighted SWE value of the subbasin to all HRUs within a subbasin. Therefore, the same

SWE is assigned to each HRU. This SWE output of SWAT represents an average SWE value for the entire subbasin (detailed information is provided in the Supplementary material). This subbasin specific value can be used for model calibration by comparing it to the subbasin SWE that is computed using Eq. (A5). This SWE output in output.hru is hereafter referred to as Sub-SWE. Snowmelt values are not ideal for model calibration because their measurements are rarely available.

Table 5.1 Description and initial calibration range of the calibrated snow parameters, elevation band parameters and hydrological parameters. The designation "a_" means an absolute increase. The other parameters without "a_" mean a replacement.

Parameters	Description	Default	Range
SFTMP	Snowfall temperature [°C]	1	0/4
SMTMP	Snowmelt base temperature [°C]	0.5	-5/5
SMFMN	Melt factor for snow on December 21 [mm H_2O °C^{-1}day^{-1}]	4.5	1.5/7
SMFMX	Melt factor for snow on June 21 [mm H_2O °C^{-1}day^{-1}]	4.5	1.5/7
TIMP	Snow pack temperature lag factor	1	0/1
PLAPS	Precipitation lapse rate [mm H_2O /m]	0	0/0.15
TLAPS	Temperature lapse rate [°C/km]	0	-9/0
SOL_AWC	Available water capacity of the soil layer [mm H_2O /mm soil]	0.01-0.06	0/0.2 ("a_")
CN2	SCS runoff curve number	55-94	-20/4 ("a_")
ESCO	Soil evaporation compensation factor	0.95	0/1
GWQMN	Threshold depth of water in the shallow aquifer required for return flow to occur [mm]	1000	-300/300 ("a_")
REVAPMN	Threshold depth of water in the shallow aquifer for "revap" to occur [mm]	750	-300/300 ("a_")
GW_REVAP	Groundwater "revap" coefficient	0.02	0.02/0.2
GW_DELAY	Groundwater delay [days]	31	0/300
ALPHA_BF	Baseflow alpha factor [days]	0.048	0/1
SLSUBBSN	Average slope length [m]	9.15-15.24	-9/115 ("a_")
CH_N2	Manning's "n" value for the main channel	0.014	0/0.3
CH_K2	Effective hydraulic conductivity [mm/hr]	0	0/400

5.2.4   Model setup

ArcSWAT 2012 (version 664), with an interface in ArcGIS, was used to set up the model for the Upper Adige river basin with 20 subbasins (Figure 5.1). The datasets used in the model are listed in Table 5.2. Based on the availability of snow data and discharge records, the model considered the time frame from 2005 to 2010. All snow stations are located at a relatively low elevation, below 2000 m a.s.l. Due to the lack of high elevation stations, we cannot observe the effect described by Grünewald et al. (2014) in

which snow depth increases with elevation up to a certain altitude where it displays a distinct peak followed by a decrease at the highest elevations.

In most subbasins, snow depth measurements are available for stations located at similar elevations (Table S5.1). Eight gauging stations (Figure 5.1) provide effective discharge records for the investigated time frame. Among them, five stations record the natural streamflow of four headwater subbasins (1, 2, 7, 8, and 11), two (subbasins 3 and 4) are affected by the upstream subbasins, and one (subbasin 20) records the total streamflow behavior at the end of the entire basin. In this work, the simulations were performed at the daily scale.

Table 5.2 Source of input data used in setting up the SWAT model and data used for calibration.

Data type	Scale	Data source
DEM	90 m × 90 m	Shuttle Radar Topography Mission (SRTM) produced by Consortium for Spatial Information (CGIAR-CSI)
Land use map	100 m × 100 m	Corine Land Cover 2006 (CLC2006) from European Environment Agency
Soil map	1 : 1500,000	Data provided in the framework of Globaqua Project (Navarro-Ortega et al., 2015) following the procedure described in (Meyer et al. 2016)
River network		EU-DEM product http://www.eea.europa.eu/data-and-maps/data/eu-dem.
Climate input		Autonomous Province of Bolzano (http://www.provincia.bz.it/meteo/home.as)
Snow measurements		Autonomous Province of Bolzano (Adler et al., 2015)
River discharge		Autonomous Province of Bolzano (Adler et al., 2015)

5.2.5   Model calibration and evaluation

In order to mitigate the effects of the initial conditions, a time frame of 516 days (01/01/2005 - 05/31/2006) was used as a warm up period. Calibration began on 06/01/2006 and lasted until 05/31/2008, while validation was performed during the time period 06/01/2008 - 05/31/2010. The parameters to be calibrated (Table 5.1) were selected according to the one-at-a-time sensitivity analysis in SWAT-CUP (Abbaspour, 2015). Three calibration procedures were applied: I) All the parameters reported in Table 5.1 were calibrated according to discharge records using an automatic calibration procedure, and basin-scale snow parameters were estimated to assign the same set of values to the entire basin. II) All

parameters were calibrated automatically to reproduce the discharge measurements as in procedure I. Unlike procedure I, snow parameters were calibrated at the subbasin scale to consider the spatial variability of the snow processes. An additional manual calibration was required when automatic calibration led to unrealistic parameter sets. Such a step was necessary when *SMFMX* values that were estimated using automatic calibration were lower than *SMFMN* values, which is not correct for the Northern Hemisphere where the study area is located. III) Subbasin-scale snow parameters and the other parameters were calibrated manually so that the Sub-SWE output (Eq. (S5.1) in Supplementary material) fit the subbasin average SWE as defined in Eq. (A5) and the streamflow simulation fit the discharge records. Meanwhile, a qualitative comparison between EB-SWE and the SWE values computed for each monitoring station (Eq. (A3)) was also taken into account for parameter modification.

Starting with the initial ranges of parameters reported in Table 5.1, the automatic calibration was performed with the Sequential Uncertainty Fitting algorithm version 2 (SUFI-2) (Abbaspour et al., 2004; Abbaspour et al., 2007) in SWAT-CUP (Abbaspour, 2015) to narrow down and target the effective range of each parameter. SUFI-2 is based on a semi-automated inverse modeling procedure with the implementation of a combined calibration-uncertainty analysis (Abbaspour, 2015). The initial parameter ranges were defined following the SWAT documentation (Arnold et al., 2012b) and studies of neighboring Alpine catchments (Grusson et al., 2015; Rahman et al., 2013). For each iteration, 1000 simulations were run. After each iteration, the ranges of the parameters were modified and generally narrowed down according to both the suggested values of the program (Abbaspour et al., 2004; Abbaspour et al., 2007) and their initial limitations.

The goodness of fit between the modeled and the observed discharge was computed for each gauged subbasin during both the calibration and validation periods. The goodness of fit between Sub-SWE and the subbasin average SWE defined in Eq. (A5) for each subbasin was calculated during both periods. Furthermore, the simulated EB-SWE was qualitatively compared to the data of each station belonging to a specific elevation band. The Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970), the coefficient of determination (R^2), and the percent bias (PBIAS) (Gupta et al., 1999) were used as goodness of fit indicators for the best simulation (Eqs. (5.1) -(5.3), respectively):

$$NSE = 1 - \frac{\sum_{t=1}^{T}(OB_t - SIM_t)^2}{\sum_{t=1}^{T}(OB_t - \overline{OB})^2} \tag{5.1}$$

$$R^2 = \frac{\left[\sum_{t=1}^{T}(OB_t - \overline{OB})(SIM_t - \overline{SIM})\right]^2}{\sum_{t=1}^{T}(OB_t - \overline{OB})^2 \sum_{t=1}^{T}(SIM_t - \overline{SIM})^2} \tag{5.2}$$

$$PBIAS = \frac{\sum_{t=1}^{T}(OB_t - SIM_t)*(100)}{\sum_{t=1}^{T}(OB_t)} \tag{5.3}$$

where OB_t is the SWE computed by Eq. (A5) or the observed discharge at time t; SIM_t is the simulated SWE or the simulated discharge at time t; $\overline{OB}$ is the mean of SWE computed by Eq. (A5) or the mean of observed discharge values; and $\overline{SIM}$ is the mean simulated SWE or discharge.

NSE indicates the accuracy of the model output in comparison to the mean of the referred data (NSE = 1 is the optimal value). R^2 represents the trend similarity between the observed and the simulated data, and it ranges from 0 to 1. Higher values indicate better model performance. PBIAS indicates the average tendency of the simulations to be larger or smaller than the corresponding observations. The lower the PBIAS's absolute value, the more accurate the model simulation (PBIAS=0 is the optimal value). Positive values stand for underestimation bias, and negative values represent overestimation bias (Gupta et al., 1999). A standard method to evaluate model performance is not available for daily SWE or daily discharge modeling results. In this work, we followed the recommendations of Moriasi et al. (2007) developed for monthly discharge values: unsatisfactory performance (NSE≤0.50), satisfactory performance (0.50<NSE≤0.65), good performance (0.65<NSE≤0.75), and very good performance (0.75<NSE≤1.00). The values of R^2 and PBIAS were used to compare model performances.

In order to evaluate the model performance of reproducing snow pack characteristics, the mean absolute percentage error (MAPE) of snow cover duration (SCD, i.e., the number of days with HS>0 m), snow peak value (SPV, i.e., the maximum HS between June 1st and May 31st) and snow peak day (SPD, i.e., the day when SPV occurs) were calculated for the entire river basin as follows:

$$MAPE = \frac{100}{N}\sum_{t=1}^{N}\left|\frac{A_t - S_t}{A_t}\right| \tag{5.4}$$

where A_t is the SCD/SPV/SPD calculated based on the SWE value converted from Eq. (A5) of subbasin t, and S_t is the SCD/SPV/SPD calculated based on the simulated SWE value of subbasin t.

5.3 Results

In the following, we will focus on model results for subbasins 2 and 20. We chose these two subbasins, because the first is a headwater not affected by hydropeaking, and the second integrates the processes occurring in the entire river basin, and it is not heavily affected by hydropeaking. Graphic results are displayed for calibration procedures II and III in order to focus on the comparison between the single-objective and multi-objective calibration strategies. Model performances for all three calibration procedures and all subbasins are provided in tables.

5.3.1 Model performance using calibration procedure I

As shown in Table 5.3, calibration procedure I resulted in at least satisfactory model performances (NSE>0.5) in simulating river discharge during the calibration period. During validation, the discharge

of all gauged subbasins, except for subbasins 7 and 11, were also well reproduced with NSE>0.5. Comparing simulated and observed river discharge for all gauging stations and for the entire simulation period (Figure 5.2a), a close fit was observed with an R^2 value of 0.92. Underestimation of streamflow occurred for river discharge values larger than 250 m^3/s. The mean absolute value of PBIAS for streamflow was 8.42 % considering the entire period (Table 5.3). Despite the good performance in reproducing river discharge, the agreement between Sub-SWE and the subbasin averaged SWE was unsatisfactory (NSE<0.5) for most subbasins (70 % of the subbasins during the calibration period and 50 % during the validation period). In addition, the absolute values of PBIAS for Sub-SWE were higher than 50 % for 65 % of the subbasins during calibration and for 50 % of the subbasins during validation. Considering the entire period and all subbasins, poor agreement was shown between the Sub-SWE and the SWE computed according to Eq. (A5) (Figure 5.2d). Thus, the parameters of procedure I calibrated according to discharge were able to reproduce river discharge but failed to capture the behavior of SWE at the subbasin scale.

Table 5.3 Calibration and validation performance of both SWE and discharge with the calibrated parameters of procedure I.

Subbasin	Calibration period						Validation period					
	NSE_{DIS}	R^2_{DIS}	$PBIAS_{DIS}$	NSE_{SWE}	R^2_{SWE}	$PBIAS_{SWE}$	NSE_{DIS}	R^2_{DIS}	$PBIAS_{DIS}$	NSE_{SWE}	R^2_{SWE}	$PBIAS_{SWE}$
1	0.68	0.69	7.60	-2.11	0.40	-147.70	0.60	0.63	14.60	0.12	0.65	-73.50
2	0.67	0.72	9.80	-1.43	0.56	-147.00	0.67	0.71	11.20	0.33	0.77	-71.50
3	0.76	0.81	5.30	-1.17	0.48	-128.50	0.70	0.75	14.50	0.61	0.71	-34.80
4	0.83	0.86	-3.40	0.81	0.82	4.90	0.80	0.81	-2.20	0.61	0.63	23.00
5	—	—	—	0.80	0.86	-20.20	—	—	—	0.84	0.89	-18.00
6	—	—	—	-0.17	0.73	-92.90	—	—	—	0.80	0.85	-27.50
7	0.54	0.57	-2.50	-0.53	0.48	-108.40	0.31	0.37	10.40	0.19	0.70	-80.70
8	0.72	0.73	-5.90	-0.88	0.52	-104.60	0.65	0.68	-14.40	-0.80	0.64	-114.10
9	—	—	—	0.91	0.91	0.10	—	—	—	0.87	0.91	-17.80
10	—	—	—	-0.45	0.06	99.60	—	—	—	-0.41	0.09	93.90
11	0.66	0.68	0.60	0.66	0.68	-8.60	0.32	0.47	-14.20	0.95	0.95	-7.30
12	—	—	—	-0.42	0.43	-96.40	—	—	—	0.35	0.63	-59.30
13	—	—	—	-14.79	0.61	-440.30	—	—	—	-6.58	0.81	-292.90
14	—	—	—	-0.36	0.24	97.10	—	—	—	-0.16	0.41	84.20
15	—	—	—	-0.10	0.55	87.20	—	—	—	0.25	0.75	67.90
16	—	—	—	0.68	0.85	-41.20	—	—	—	0.64	0.88	-50.90
17	—	—	—	0.09	0.59	80.80	—	—	—	0.89	0.92	18.30
18	—	—	—	-0.02	0.34	81.00	—	—	—	0.80	0.87	32.30
19	—	—	—	0.58	0.78	38.80	—	—	—	0.92	0.93	7.40
20	0.73	0.82	11.50	-0.25	0.18	94.40	0.74	0.82	6.60	0.77	0.83	35.00
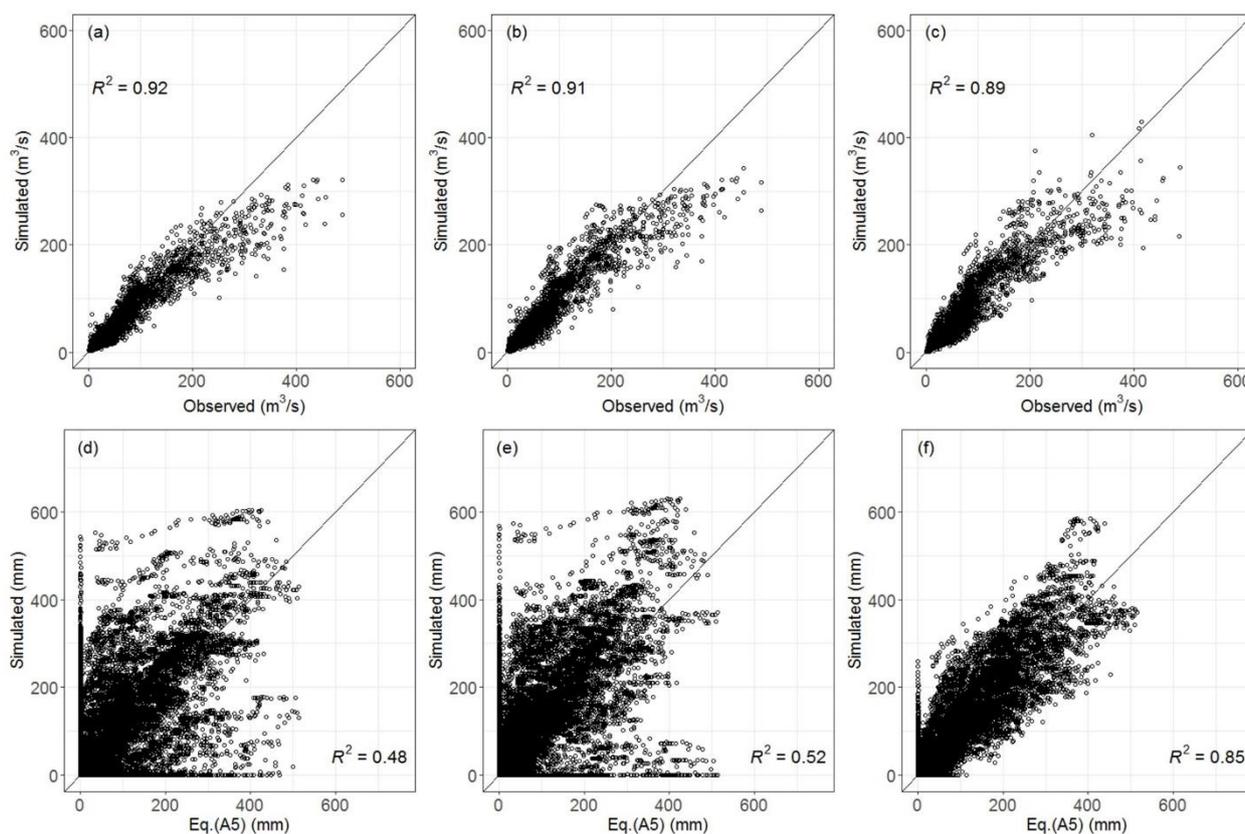
Figure 5.2 Scatterplots of the simulated discharge against the observed discharge of all the gauged subbasins (a-c: calibration procedures I-III) and the simulated Sub-SWE against the subbasin average SWE converted from Eq. (A5) of all the subbasins (d-f: calibration procedures I-III).

5.3.2   Model performance using calibration procedure II

Calibration procedure II differs from calibration procedure I in that it defines subbasin specific snow parameters. The calibrated parameters of procedure II showed performances as good as procedure I in simulating river discharge during both calibration and validation periods (Table 5.4). This procedure also obtained unsatisfactory performances during validation for subbasins 7 and 11. Still, the R^2 value between model simulation and all available streamflow observations (Figure 5.2b) was equal to 0.91. In this case, the model underestimated streamflow with values larger than 300 m^3/s. Considering the entire period, the mean absolute value of PBIAS was 7.75 % for streamflow simulations. With respect to the subbasin SWE, unsatisfactory results (NSE<0.5) were obtained for 70 % of the subbasins during calibration and 50 % of the subbasins during validation. The absolute value of PBIAS higher than 50 % occurred for 65 % of the subbasins during calibration and 50 % of the subbasins during validation. As shown in Figure 5.2e, poor agreement was also observed between all the Sub-SWE and the subbasin averaged SWE. Generally, no improvements were observed in simulating the subbasin averaged SWE even though different sets of snow parameters were assigned to each subbasin to consider local

differences in snow dynamics. Also, in this case, the model calibrated according to discharge was only able to reproduce streamflow dynamics adequately.

Table 5.4 Calibration and validation performance of both SWE and discharge with the calibrated parameters of procedure II.

	Calibration period						Validation period					
Subbasin	NSE_{DIS}	R^2_{DIS}	$PBIAS_{DIS}$	NSE_{SWE}	R^2_{SWE}	$PBIAS_{SWE}$	NSE_{DIS}	R^2_{DIS}	$PBIAS_{DIS}$	NSE_{SWE}	R^2_{SWE}	$PBIAS_{SWE}$
1	0.69	0.70	-7.70	-2.17	0.44	-147.80	0.62	0.64	-2.30	0.21	0.71	-72.10
2	0.58	0.64	15.90	-3.67	0.47	-220.40	0.68	0.73	16.50	-0.18	0.76	-105.80
3	0.73	0.77	-1.20	0.28	0.75	-66.20	0.66	0.70	7.30	0.74	0.75	12.20
4	0.78	0.81	-8.60	0.81	0.81	4.20	0.75	0.76	-7.20	0.57	0.62	15.50
5	—	—	—	0.82	0.83	-8.40	—	—	—	0.72	0.79	-15.90
6	—	—	—	0.02	0.75	-85.60	—	—	—	0.85	0.87	-21.00
7	0.57	0.59	1.90	-0.60	0.45	-117.70	0.23	0.34	13.20	0.19	0.67	-81.70
8	0.72	0.75	-5.70	-1.15	0.52	-114.80	0.68	0.76	-16.90	-1.02	0.65	-122.70
9	—	—	—	-4.29	0.39	-230.50	—	—	—	-0.52	0.61	-120.10
10	—	—	—	0.10	0.64	-78.60	—	—	—	0.64	0.82	-49.20
11	0.64	0.71	0.20	0.63	0.66	-12.60	0.34	0.49	-13.10	0.95	0.95	3.00
12	—	—	—	0.26	0.48	-48.90	—	—	—	0.55	0.65	-38.30
13	—	—	—	-5.27	0.80	-252.60	—	—	—	-2.99	0.90	-193.90
14	—	—	—	-0.19	0.40	90.10	—	—	—	-0.02	0.32	73.90
15	—	—	—	-0.37	0.19	97.60	—	—	—	-0.37	0.17	91.90
16	—	—	—	0.76	0.82	31.80	—	—	—	0.78	0.80	-1.10
17	—	—	—	0.68	0.69	15.50	—	—	—	0.17	0.87	-80.00
18	—	—	—	0.37	0.81	-63.90	—	—	—	0.34	0.92	-67.10
19	—	—	—	0.78	0.81	18.90	—	—	—	0.83	0.91	-21.80
20	0.77	0.79	3.60	0.24	0.57	68.50	0.77	0.79	-2.70	0.82	0.87	-15.40

Figure 5.3 and Figure 5.4 show the detailed model performances of subbasins 2 and 20, respectively. We focused on river discharge, subbasin scale SWE, and elevation band scale SWE compared to monitoring station values. The daily discharge behavior obtained using the calibrated parameters of procedure II was well captured at subbasin 2 (Figure 5.3a). Most of the peaks were closely reproduced, while base flow was slightly underestimated. In contrast, the subbasin averaged SWE values were overestimated (Figure 5.3c) with on average 59 % longer snow seasons. This result indicates that a much larger amount of snow accumulation was simulated at subbasin 2 in comparison to the subbasin averaged SWE reference. A larger amount of snow accumulated in the subbasin lead to a large amount of snowmelt water contributing to the discharge during the melting season. The EB-SWE of lower elevation areas (bands 1 and 2) well matched the SWE values (computed converting measured *HS* values using Eq. (A3)) of the three available snow stations in subbasin 2 (Figure 5.3e and 5.3g). As previously mentioned, EB-SWE is a representative value for a range of elevations, whereas the station observation can only represent one point among these elevations. The SWE time series of the snow stations at subbasin 2 were considered to be well reproduced by the model. For the higher elevation areas (band 3-6) of subbasin 2, no snow stations were available. The simulated EB-SWE values are displayed in Figure

121

5.3i. The highest elevation area (band 6), in particular, displayed a much larger SWE amount and longer snow cover duration than the other elevation bands. In fact, the Sub-SWE is the weighted mean of the EB-SWE of all the elevation bands (Eq. (S5.1)). Therefore, the overestimation of Sub-SWE observed in subbasin 2 (Figure 5.3c) could be explained by the overestimation of SWE at elevation bands 3 through 6 (Figure 5.3i).

Considering subbasin 20, the simulated discharge reproduced the general behavior of the observations well (Figure 5.4a). Daily fluctuations stemming from hydropeaking were not captured since the reservoir module was not used in the model due to the lack of information about reservoir management. Figure 5.4c displays the subbasin SWE of subbasin 20. The snow-covered periods were well captured. The simulation underestimated the subbasin scale SWE during the calibration period. During validation, simulated Sub-SWE fit quite well with the SWE of the 2008-2009 snow season, but overestimated the maximum SWE values for the snow season 2009-2010. Regarding the EB-SWE simulations, they matched with the SWE of the monitoring station "MOELTEN" during the first three snow seasons (2006-2009), but they were much higher than the station SWE values for the snow season 2009-2010.
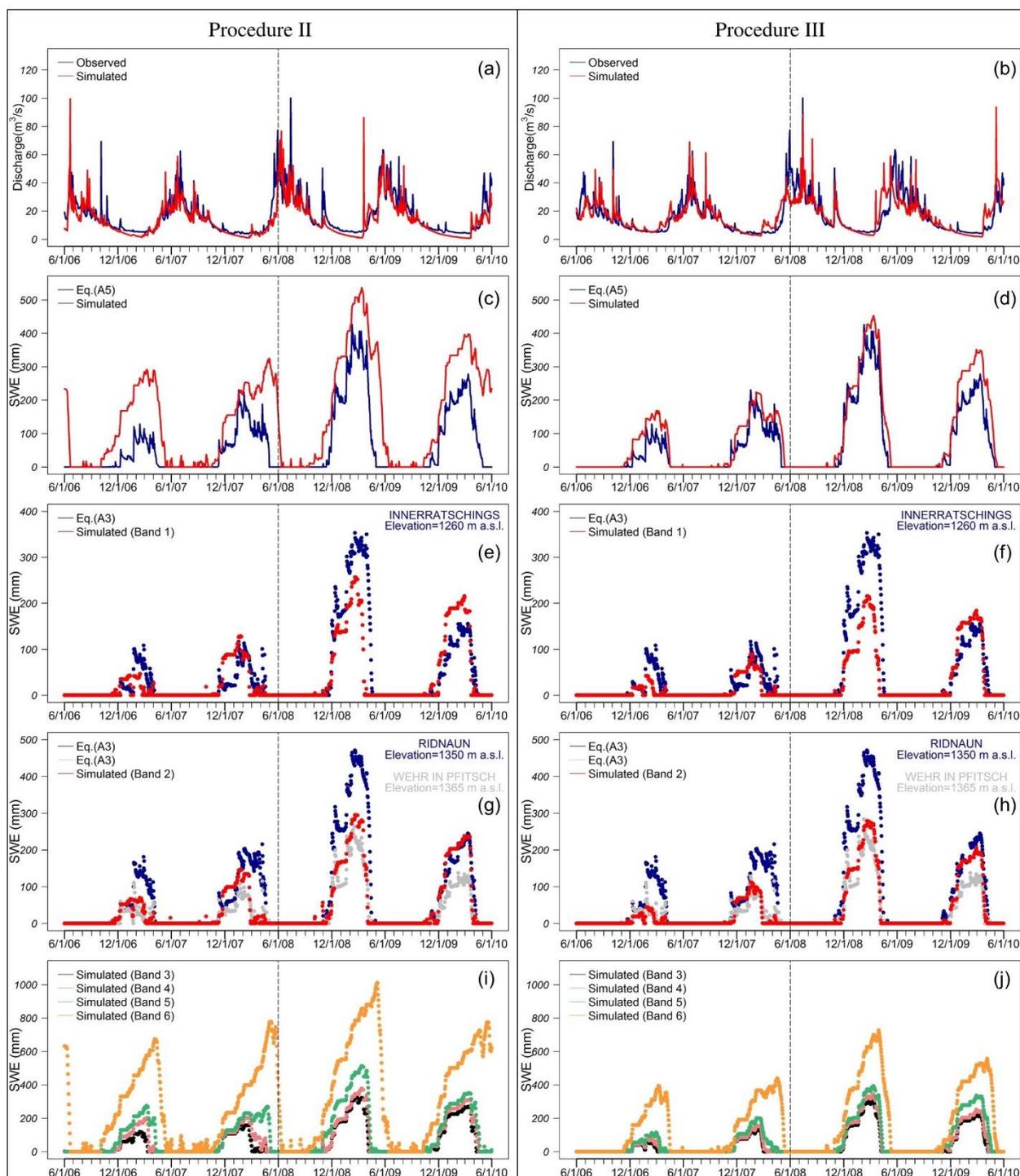
Figure 5.3 Model performance of the calibrated parameters for both calibration procedures II and III at subbasin 2. a-b: discharge; c-d: subbasin SWE; e-h: qualitative comparison between the SWE values of each snow station converted using Eq. (A3) and the EB-SWE output of the corresponding elevation bands; i-j: EB-SWE of the higher elevation areas where there are no stations (the dotted line divides the calibration and validation periods).
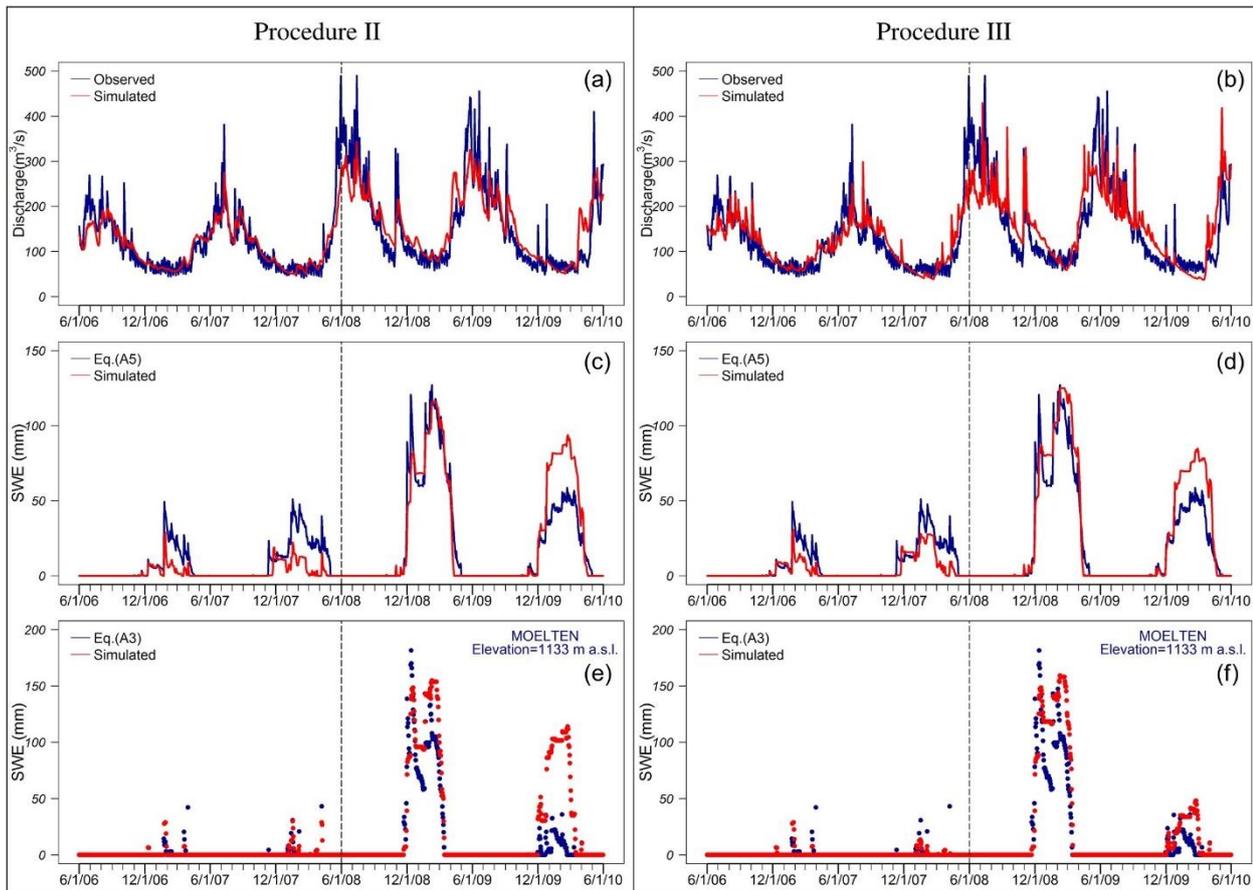
Figure 5.4 Model performance of the calibrated parameters of both calibration procedures II and III at subbasin 20. a-b: discharge; c-d: subbasin SWE; e-h: qualitative comparison between the SWE values of snow station converted using Eq. (A3) and the EB-SWE output of the corresponding elevation band (the dotted line divides the calibration and validation periods).

5.3.3 Model performance using calibration procedure III

A third calibration procedure was developed to meet both good performances in subbasin SWE and river discharge simulations, as well as suitable matches between EB-SWE and SWE values obtained converting the *HS* observations of the corresponding snow stations. This procedure was able to reproduce both subbasin SWE and river discharge properly (Table 5.5). The simulated river discharge values fit well with the observations, as shown by the at least satisfactory model performances (NSE>0.5) for all the gauged subbasins during both the calibration and validation periods (Table 5.5). Procedure III appeared to lead to a more accurate set of model parameters than procedures I and II, since subbasins 7 and 11 also displayed satisfactory results during the validation period. Such a conclusion, however, should be supported by a systematic comparison to be performed in a larger set of alpine river basins. In addition, the scatterplot displayed in Figure 5.2c shows good agreement between observed and simulated values (R^2=0.89). Underestimation of river discharge was observed for values larger than 300 m^3/s

(Figure 5.2c). The mean absolute value of PBIAS for streamflow simulations was 4.63 %. The agreement between Sub-SWE and the subbasin averaged SWE was at least satisfactory (NSE>0.5) for all 20 subbasins during both the calibration and validation periods (Table 5.5). Good model performance (0.65<NSE≤0.75) was achieved for 75 % of the subbasins during calibration and 75 % of the subbasins reached very good model performance (0.75<NSE≤1.00) during validation. In addition, absolute PBIAS values were smaller than 50 % during the entire simulation period for 95 % of SWE. Considering the entire period and all subbasins, close agreement was reached between Sub-SWE and the subbasin averaged SWE with an R^2 of 0.85 (Figure 5.2f). The multi-objective based calibration procedure was able to capture the behavior of both river discharge and averaged SWE of each subbasin well.

Table 5.5 Calibration and validation performance of both SWE and discharge with the calibrated parameters of procedure III.

	Calibration period						Validation period					
Subbasin	NSE_{DIS}	R^2_{DIS}	$PBIAS_{DIS}$	NSE_{SWE}	R^2_{SWE}	$PBIAS_{SWE}$	NSE_{DIS}	R^2_{DIS}	$PBIAS_{DIS}$	NSE_{SWE}	R^2_{SWE}	$PBIAS_{SWE}$
1	0.65	0.66	-2.60	0.50	0.69	-34.60	0.53	0.54	6.10	0.86	0.87	-12.40
2	0.66	0.67	-3.60	0.54	0.86	-59.20	0.61	0.63	3.60	0.82	0.94	-34.10
3	0.63	0.63	-2.80	0.70	0.88	-40.20	0.50	0.52	9.60	0.75	0.76	14.60
4	0.70	0.73	-9.80	0.59	0.76	-37.10	0.57	0.59	-1.80	0.67	0.69	-0.70
5	—	—	—	0.84	0.84	4.30	—	—	—	0.78	0.78	9.70
6	—	—	—	0.86	0.89	-8.90	—	—	—	0.91	0.92	9.10
7	0.76	0.84	-6.90	0.75	0.82	-29.50	0.63	0.70	11.40	0.77	0.86	-33.10
8	0.63	0.63	-2.70	0.74	0.85	-30.80	0.50	0.52	-7.10	0.51	0.97	-62.10
9	—	—	—	0.87	0.88	9.30	—	—	—	0.88	0.91	-13.50
10	—	—	—	0.73	0.78	27.70	—	—	—	0.82	0.83	5.00
11	0.73	0.77	-2.50	0.64	0.70	-19.10	0.56	0.59	-0.70	0.92	0.93	-9.90
12	—	—	—	0.78	0.82	-17.50	—	—	—	0.89	0.89	-5.10
13	—	—	—	0.82	0.90	26.50	—	—	—	0.84	0.90	-17.30
14	—	—	—	0.88	0.91	9.70	—	—	—	0.77	0.81	-11.00
15	—	—	—	0.80	0.80	-9.30	—	—	—	0.87	0.88	-9.60
16	—	—	—	0.78	0.85	32.80	—	—	—	0.83	0.86	-9.10
17	—	—	—	0.78	0.81	19.90	—	—	—	0.61	0.88	-51.80
18	—	—	—	0.88	0.88	3.50	—	—	—	0.62	0.96	-47.00
19	—	—	—	0.75	0.83	27.00	—	—	—	0.93	0.94	-8.50
20	0.71	0.71	0.20	0.60	0.74	44.80	0.70	0.70	-2.70	0.84	0.92	-23.40

The model performance of procedure III for subbasins 2 and 20 is also displayed in Figure 5.3 and Figure 5.4. For subbasin 2, the multi-objective calibrated parameters led to a good fit between simulated and observed discharge values (Figure 5.3b). Most peaks and base flow were well estimated. The simulated Sub-SWE fit quite well with the subbasin average SWE, capturing the length of the snow seasons and providing a satisfactory estimate of the SWE amount (Figure 5.3d). EB-SWE simulation results for bands 1 and 2 (Figure 5.3f and 5.3h) were similar to those obtained in procedure II. A noticeable difference between model results obtained using procedure II and procedure III was observed for elevation bands from 3 to 6. Using calibration procedure III, snow cover duration (Figure 5.3j) did not

increase as much as when using calibration procedure II, leading to an improved agreement between Sub-SWE and subbasin average SWE.

For subbasin 20, the simulated discharge fit well with the observations (Figure 5.4b) except for the fluctuations during the winter season caused by hydropeaking. Although the subbasin mean SWE values displayed a large variability between calibration and validation periods, the model calibrated using procedure III closely followed the observations (Figure 5.4d). The simulated snow seasons matched well the values computed using Eq. (A5). Figure 5.4f shows a close match between simulated EB-SWE and the SWE of the snow station converted using Eq. (A3). Calibration procedure III allows us to reproduce SWE and streamflow appropriately, both in low snow accumulation years and in high snow accumulation years.

5.3.4 Validation of SWAT model results against snow depth observations

A final validation step was implemented to compare the *HS* observations of the 40 snow gauges within the basin (Figure 5.1) to the model results obtained using calibration procedure II and III. The SWE time series of the model output were converted into *HS* values using Eqs. (A3) and (A4). To simplify the comparison, we classified the observations and the corresponding model results into six classes according to 300 m elevation intervals (i.e., class 1 stations from 300 m a.s.l. to 600 m a.s.l., class 2 from 600 m a.s.l. to 900 m a.s.l. and so on). We then computed the mean snow depth between November 1st and April 30th for the entire simulation period (<HS>) for each elevation class. As shown in Figure 5.5, observed <HS> values presented a positive correlation with the elevations, which can be reproduced with a power law as given in Eq. (A1). The model results of procedure III captured very well the mean behavior of the observations of the elevation areas for classes 2 through 5, where 82.5 % of the snow stations were located (Table S5.1), while slight deviations were observed for classes 1 and 6. Results of procedure II captured well the mean behavior of the observations for classes 3 and 4, corresponding to 57.5 % of the snow stations (Table S5.1). It is particularly relevant to observe that calibration procedure II performed poorly for class 5 (1500 m a.s.l. - 1800 m a.s.l.). This is the most sensitive elevation for snow variability in the Adige catchment, as evidenced by Marcolini et al. (2017b). Therefore, procedure III performed better in reproducing the observed snow dynamics than procedure II.
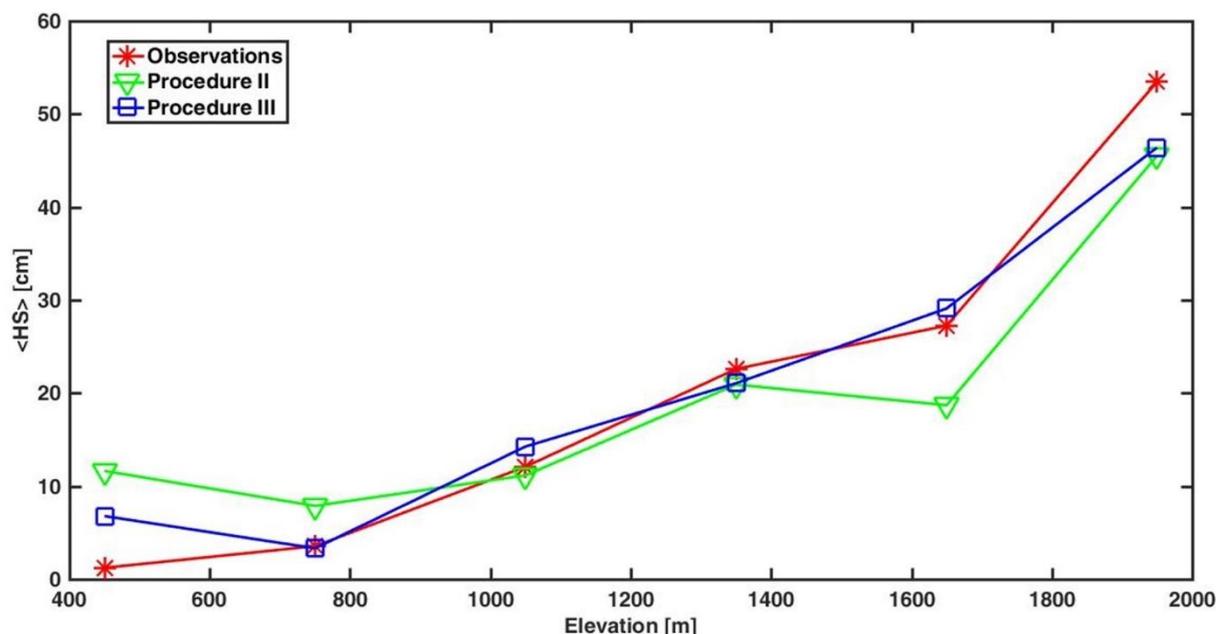
Figure 5.5 Comparison between the mean values of observed *HS* for the available stations at various elevation ranges and the mean values of SWAT simulation results converted into *HS* using Eqs. (A3) and (A4).

As shown in Table 5.6, evident differences were observed between the multi-objective and single-objective calibration procedures in reproducing snow pack characteristics, such as snow cover duration (SCD), snow peak value (SPV) and snow peak day (SPD). The two single-objective procedures (I and II) led to similar errors for all three snow characteristics during both the calibration and validation periods. Their values were higher than those obtained by the multi-objective approach (III). Among them, MAPE values of SCD and SPV obtained by procedure I and II were three and two times higher, respectively, than the corresponding values obtained using procedure III. The MAPE values of SPD were close among the results of the three procedures, and procedure III still led to the smallest values. Generally, procedure III performed better in reproducing the snow pack characteristics than procedures I and II. The absolute percentage error of SCD, SPV and SPD for each subbasin are provided in the Supplementary material (Table S5.2-S5.4) for all calibration procedures.

Table 5.6 Mean absolute percentage error of snow pack characteristics: snow cover duration (SCD), snow peak value (SPV) and snow peak day (SPD) for the three calibration procedures.

	Calibration			Validation		
	I	II	III	I	II	III
MAPE$_{SCD}$ (%)	44.0	37.7	8.4	31.5	30.4	8.8
MAPE$_{SPV}$ (%)	51.5	43.3	19.3	30.9	31.2	12.0
MAPE$_{SPD}$ (%)	9.3	10.6	7.9	3.1	3.5	1.0

5.4   Discussion

5.4.1   Effect of different snow parameterizations on discharge-based, single-objective calibration

Table 5.3, Table 5.4 and Figure 5.2 show that the set of parameters obtained using calibration procedures I and II both led to equally good performances in simulating the river discharge but failed to reproduce the subbasin average SWE in most subbasins. The definition of subbasin specific snow parameters in procedure II led to a higher R^2 (Figure 5.2d and 5.2e) and smaller MAPE of SCD and SPV (Table 5.6) than procedure I. However, the improvements were minor. Therefore, we can conclude that assigning subbasin specific snow parameters does not significantly improve the model performance of snow simulations when calibrating all model parameters according to river discharge records. Considering a single-objective calibration procedure, the defined set of subbasin specific snow parameters could still not fit the specific snow conditions of each subbasin. Moreover, river discharge in subbasins 3, 6, 9, 10, 12, 14, 15, and 17 was highly affected by hydropeaking, which cannot be reproduced by our hydrological model due to the lack of input data. Using discharge-based single-objective calibration, the identified optimal set of snow parameters was also influenced by this source of uncertainty. Thus, the model calibrated according to discharge only could not properly capture the snow observations in the Upper Adige catchment. Therefore, the SWAT model calibrated using a single-objective approach was not able to reproduce both snow and discharge dynamics, independently of the scale at which snow parameters were defined.

5.4.2   Importance of applying a multi-objective calibration procedure in alpine catchments

Calibration procedures I and II displayed quite similar model performances. We will now focus only on procedure II in comparison to procedure III, since they both define subbasin specific snow parameters. The multi-objective calibration approach offered using procedure III resulted in much better model performances than the single-objective procedure, especially in reproducing snow dynamics.

As shown in Table 5.3-5.5, the calibrated parameters of procedure III led to acceptable simulations for streamflow in all gauged subbasins during the entire period. In contrast, procedure II could not reproduce the discharge of subbasins 7 and 11 during validation. More importantly, the use of procedure III allowed us to simulate the subbasin SWE of each subbasin properly with at least satisfactory performance, while the use of a single-objective procedure failed to reproduce SWE adequately in most subbasins. Moreover, including snow data during calibration allowed us to reproduce important properties better, such as snow cover duration in the catchment and peak snow accumulation (Table 5.6).

Considering the overall performance of the models, both calibration procedures were able to reproduce the discharge records with similar R^2 (Figure 5.2b and 5.2c). The underestimation of high streamflow

discharge values could be ascribed to the impact of river damming. Without data to implement the reservoir module, all these models were not able to simulate the high discharge and hydropeaking effects properly.

The comparisons of model performances in subbasins 2 and 20 further illustrated the positive effect of using a multi-objective calibration approach. As shown in section 5.3, subbasin 2 represented a headwater subbasin with natural flow and subbasin 20 represented the outlet of the river basin. In addition, subbasin 2 was a snow-dominated alpine subbasin with high elevation areas, and subbasin 20 was dominated by low elevation areas (< 800 m). Based on the hydrographs, procedures II and III presented an equally good ability in simulating the discharge time series for both subbasins (Figure 5.3a and 5.3b, and Figure 5.4a and 5.4b). However, SWE simulations were quite different between the two procedures at both the high elevation and the low elevation subbasins. This can lead to distinct interpretations about the hydrological modeling behaviors caused by different calibration approaches. In subbasin 2 (Figure 5.3), procedures II and III performed equally well in qualitatively matching the SWE of snow stations converted using Eq. (A3), but different subbasin average SWE were obtained. Procedure II led to a large overestimation of the subbasin SWE and much longer snow seasons due to the overestimation of SWE for high elevations (Figure 5.3i). Procedure III led to a close fit between Sub-SWE and subbasin averaged SWE references, although the amounts of subbasin SWE were very different between calibration and validation periods (the values of SWE during validation were almost twice as high as the calibration values). As a consequence, the contribution of snowmelt to streamflow was very different in the model that was calibrated using the single-objective calibration procedure compared to the multi-objective procedure. This was shown by comparing the snowmelt results of the two procedures (Figure 5.6a and Figure 5.7). The cumulated snowmelt simulated by calibration procedure II reached 2333 mm, which was almost twice as much as the amount simulated by procedure III during the entire period (Figure 5.7). The simulated and observed cumulative discharge in the case of procedure III fit closely (Figure 5.7). Despite the larger snowmelt contribution to streamflow, the results of procedure II still underestimated the cumulative behaviors of discharge records (Figure 5.7a). These results show that although the simulated discharge values may look reasonable, they can be inaccurate in terms of the relative contribution of runoff components. In other words, the multi-objective calibration approach reduced model equifinality (Beven, 2006; Hanzer et al., 2016). Considering subbasin 20 (Figure 5.4), there was no significant difference between the simulated Sub-SWE of procedure II and procedure III, although improvements were observed during the calibration period of procedure III. The predicted snowmelt was similar between the two procedures as observed in Figure 5.6b. In this low elevation subbasin, however, the contribution of snowmelt was not as relevant as for subbasin 2. In general, the snow and elevation band parameters obtained using procedure III allowed us

to reproduce snow dynamics at both subbasin and elevation band spatial scales better than those of procedure II, since snow data were considered during model calibration. Moreover, a multi-objective model calibration allowed us to obtain a more robust estimate of snow parameters than a single-objective calibration in subbasins affected by hydropeaking. The lack of data to implement the reservoir module of SWAT led to inaccurate streamflow predictions in subbasins affected by hydropower production, which consequently affected the snow parameters that were calibrated according to river discharge. Lastly, the final validation of section 5.3.4 from another perspective further confirmed the better snow dynamic simulation ability of calibration procedure III compared to procedure II, which closely matched the snow measurements of the river basin.
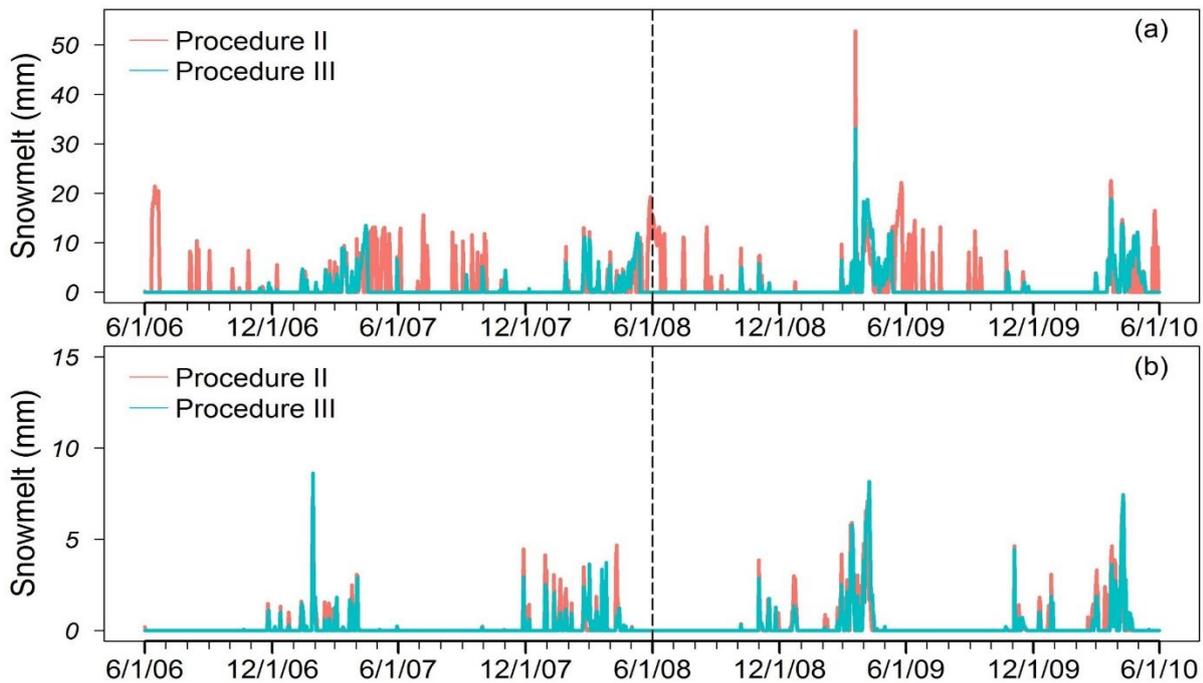
Figure 5.6 Comparison of the snowmelt behaviors of procedures II and III at subbasin 2 (a) and subbasin 20 (b) (the dotted line divides the calibration and validation periods).
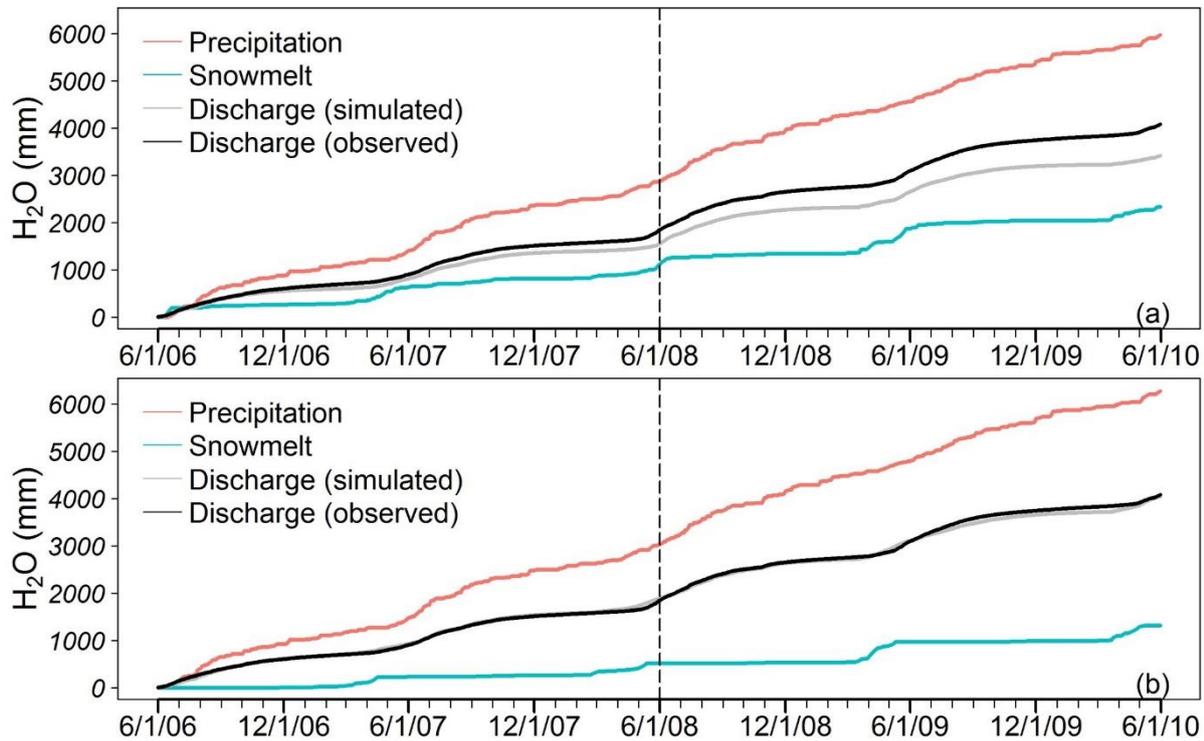
Figure 5.7 Cumulative curve of precipitation, snowmelt water and discharge at subbasin 2 of calibration procedures II (a) and III (b) (the dotted line divides the calibration and validation periods).

The way SWAT models snow processes, the computation of subbasin mean SWE and the conversion from *HS* to SWE present some limitations, which to some extent justify the discrepancy between SWE model results and observations. First, the snow module of SWAT uses a non-distributed, temperature index-based approach (Neitsch et al., 2011), whereas in reality, snow dynamics are highly dependent on several site-specific conditions such as exposure to wind, land cover, and solar radiation (Endrizzi et al., 2006; Hiemstra et al., 2002). Moreover, the uncertainty in estimating SWE by the equation proposed by Pistocchi (2016) is about 13 %. We should also consider the uncertainty of the climate and temperature input in mountainous regions (Duan et al., 2016; Tuo et al., 2016), the epistemic uncertainty of the model structure (Abbaspour et al., 2007) and of the snow package itself, and the uncertainty propagated from hydropower operation (Majone et al., 2016).

5.5 Conclusion

Modeling snow processes is of great importance in snow-dominated alpine river basins like the Upper Adige river basin. A discharge-based, single-objective calibration approach is not capable of simulating snow dynamics reasonably well, regardless of whether the spatial variability of snow parameters is considered in the parameterization or not. By referring to river discharge alone, the model has a high chance of obtaining streamflow simulations that are in good agreement with the observations but are

composed of unreliable contributions of different hydrological processes. In the end, this would result in inappropriate considerations for water resources management plans and policies.

Due to the rapid development of measurement devices and remote sensing techniques, snow data are available or become available for many snow-dominated alpine basins and therefore represent an optimal dataset for model calibration to integrate streamflow observations. The multi-objective calibration procedure proposed in this work was based on the inclusion of SWE time series in the calibration procedure of SWAT. Measurements of SWE values are rare in alpine catchments and continuous SWE time series are generally not available. To overcome this important limitation, in the multi-objective calibration approach we used the simple empirical equation proposed by Pistocchi (2016) to convert snow depth observations into SWE values. Moreover, SWE time series at both the subbasin and the elevation band scales were computed, which allowed us to take advantage of the SWAT model structure to reproduce the spatial variability of snow dynamics properly. One disadvantage of the multi-objective calibration procedure compared to the single-objective approach is the time required for model calibration. As a payoff, however, model predictions are much more robust and reliable, which in the end is the most important goal to be achieved by a hydrological model.

The multi-objective calibration procedure was superior to the single-objective procedures. We therefore conclude that it is crucial to apply a multi-objective calibration procedure in hydrological modeling of alpine catchments, since: 1) When snow data of different spatial scales were included in the calibration, the multi-objective procedure was able to capture snow dynamics very well. As a major hydrological component, an improved simulation of snow dynamics is fundamental for alpine hydrology. 2) Multi-objective calibration also reproduced river discharge properly, not only by fitting the observations very well in terms of statistical metrics and hydrographs, but also by simulating discharge with a reasonable snowmelt contribution. Multi-objective calibration hence allowed us to reduce model equifinality and uncertainty. 3) We believe that a multi-objective calibration procedure could reduce the parameter uncertainty that is propagated from hydropeaking effects by providing more reasonable and reliable snow and elevation band parameters according to snow data of different spatial scales. To better explore this point, however, a specific uncertainty analysis should be performed in the future with the support of hydropower operation data, but it is beyond the scope of this work. Finally, to facilitate multi-objective modeling applications in snow-dominated catchments, the automatic calibration software SWAT-CUP (Abbaspour, 2015) could be further developed. In particular, an automatic rejection criterion could be introduced when the parameter set converges to a maximum snowmelt factor smaller than the minimum snowmelt factor. Moreover, it would be useful to allow the user to select the elevation band scale SWE result as target variable for calibration.

Appendix A

Computation of mean subbasin snow depth

López-Moreno and Nogués-Bravo (2006) showed that among various local, geostatistical and global methods used to interpolate snow depth data at the catchment scale, global methods (e.g., linear regression and generalized additive models) have the best performance. Moreover, several studies have indicated that snow depth shows the strongest correlation with elevation (López-Moreno and Nogués-Bravo (2006), Grünewald et al. (2014)). In this study, we applied a simple power law model to establish a correlation between all available snow depth measurements in the catchment at a given day and elevation (h):

$$HS = a \cdot h^b \tag{A1}$$

The parameters a and b were estimated at the daily scale. This simple interpolation scheme allowed us to obtain results that were comparable to those obtained by López-Moreno and Nogués-Bravo (2006). The mean subbasin snow depth <Sub_HS> was then computed as a weighted mean for each subbasin:

$$< Sub\_HS > = \int_0^{h_{max}} HS \cdot \tilde{h}(h)dh \tag{A2}$$

where $\tilde{h}$ is the probability density function of the elevation in each subbasin and h_{max} is the maximum subbasin elevation.

Conversion of snow depth measurements into snow water equivalent values

Snow depth values can be converted into SWE [kg/m^2] values by multiplying them by the snow bulk density ρ [kg/m^3]:

$$SWE = HS \cdot \rho \tag{A3}$$

Several empirical methods, specific for a given geographic region, have been proposed to convert HS values into SWE values, estimating the temporal variability of the snow bulk density (e.g., Jonas et al., 2009; McCreight and Small, 2014; Pistocchi, 2016; Sturm et al., 2010). For the Upper Adige catchment, the following equation was proposed by Pistocchi (2016):

$$\rho = 200 + (DOY + 61) \tag{A4}$$

where DOY is a counter of the day of the year, which is set to 1 on January 1st, with October 1st being -92 to account for the winter season extending across two years in the Northern Hemisphere (Pistocchi, 2016).

Eq. (A4) aims at describing the seasonal dynamics of snow density and assumes a positive correlation between snow density and the number of days since the beginning of the winter period (Jonas et al., 2009; Pistocchi, 2016). Despite the intrinsic limitations associated with the use of such models, in particular at the daily scale, this simple approach can be applied whenever snow density time series with adequate spatial and temporal resolution are not available, which is the case for the Upper Adige catchment (McCreight and Small, 2014; Pistocchi, 2016). The uncertainty in estimating SWE by applying Eq. (A4) is comparable to the within-site spatial variability (Pistocchi, 2016).

The mean subbasin snow water equivalent (<Sub_SWE>) for each day was computed according to the following equation:

$$< Sub\_SWE > = < Sub\_HS > \cdot \rho \qquad\qquad\qquad (A5)$$

The units of <Sub_SWE> were then converted into mm of H_2O so they could be compared to SWAT model results.

Chapter 6

Coupling Hydrological Modeling and Support Vector Regression to Model Hydropeaking in Alpine Catchments[5]

Water management in the alpine region has an important impact on streamflow. In particular, hydropower production is known to cause hydropeaking i.e., sudden fluctuations in river stage caused by the release or storage of water in artificial reservoirs. Modeling hydropeaking with hydrological models, such as the Soil Water Assessment Tool (SWAT), requires the knowledge of the management rules of the reservoirs. These data are often not available since they are sensitive information for hydropower production companies. In this work, we propose to couple the results of a calibrated hydrological model with a machine learning method to reproduce hydropeaking without the knowledge of the actual management operation of the reservoirs. In particular, we trained a support vector machine (SVM) with the output of a SWAT model, the day of the week and the energy price. We tested the model for the Upper Adige river basin in North-East Italy. A wavelet analysis showed that energy price has a significant influence on river discharge, and a wavelet coherence analysis indicated the improved performance of the SVM model in comparison to the SWAT model alone. The SVM model was also able to capture the fluctuations in streamflow caused by hydropeaking when both energy price and river discharge displayed a complex temporal dynamic.

[5] Chiogna, G., Marcolini, G., Liu W.Y., Ciria T.P., Tuo, Y., 2018. Coupling Hydrological Modeling and Support Vector Regression to Model Hydropeaking in Alpine Catchments. Science of the Total Environmental, 633: 220-229. DOI: 10.1016/j.scitotenv.2018.03.162

6.1 Introduction

Alpine catchments play a pivotal role in Europe for water provisioning as well as for hydropower production (Wagner et al., 2015, Viviroli and Weingartner, 2004; Hastik et al., 2015) and fluctuations of river discharge occur at multiple temporal scales due to natural and anthropogenic driving forces (Poff et al., 2007; Rahman et al., 2013). Therefore, water management in these catchments strongly influences the hydrological cycle and vice versa (Beniston, 2012; Basso and Botter, 2012). Dams and hydropower plants greatly impact alpine rivers since they generate highly variable river stage fluctuations. Hydropeaking i.e., the artificial increase and decrease of discharge and corresponding water levels in rivers (Hauer et al., 2017b) is particularly relevant, since it threatens ecosystem integrity (Zolezzi et al., 2009; Botter et al., 2010; García et al., 2011; FitzHugh and Vogel, 2011; Hauer et al., 2017a; Holzapfel et al., 2017).

Modeling hydrological processes in alpine catchments requires that the coupled interaction between natural processes and economic drivers, which control hydropower production, are taken into account (Gaudard et al., 2014; Tonolla et al., 2017). Several models exist in the literature to reproduce reservoir operation and have been successfully applied in several case studies (e.g., Majone et al., 2016; Rahman et al., 2013; Finger et al., 2012; Rahman et al., 2014). However, data availability is one of the main challenges in reproducing streamflow when considering reservoir operation for hydropower production (Gaudard et al., 2014). In fact, in many countries, the energy market is an open market and reservoir management rules are considered sensitive information by hydropower companies. Hence reliable management data may not be publically available, hindering the possibility of using reservoir operation modules in hydrological models. Moreover, the operation of the major hydropower plants now generally depends on energy price fluctuations, while in the past it was a regulated sector (Gaudard et al., 2014; Massarutto and Pontoni, 2015).

This study aims at showing how machine learning methods can be used to improve hydrological models in alpine catchments highly affected by hydropeaking at the daily time scale. Machine learning methods have been widely applied in hydrology as a surrogate for distributed and semi-distributed models for various purposes (e.g., Deka, 2014; Nourani et al., 2014; Rasouli et al., 2012; Solomatine and Shrestha, 2009). A peculiar characteristic of learning machines is their ability to deduce the dynamic response of the system depending on available measured data. Several works highlight the use of machine learning methods as a valuable and accurate tool for modeling complex river basin systems in support of water management information needs (Karamouz et al., 2009; Khalil et al., 2005; Ticlavilca and McKee, 2011).

In this work, we first performed a wavelet analysis of streamflow and energy price time series for the Upper Adige catchment (North-East Italy), since the ecological relevance of hydropeaking in this river

basin has been highlighted in several studies (Bruno et al., 2013; Zolezzi et al., 2009). The wavelet analysis is able to reveal the main mode of the frequency of a signal and it can be used to identify when any change in these modes occurred (Daubechies, 1990; Labat, 2005; Schaefli et al., 2007; Torrence and Compo, 1998). Moreover, using a wavelet coherence analysis (Grinsted et al., 2004), we identified the correlation between energy price and streamflow to investigate the impact of energy market fluctuations on river discharge. Using a support vector machine (SVM), we coupled hydrological model results obtained with the Soil Water Assessment Tool (SWAT) with energy price time series and the day of the week to improve model prediction of hydropeaking on a daily time scale. Previous works (e.g., Torres-Rua et al., 2012) coupled physical models and machine learning to correct, for example, hydraulic simulation models. However, this combination has only rarely been attempted in hydrology. In order to estimate the performance of our model, beside using traditional metrics like mean absolute error, mean absolute percentage error, root mean square error, ratio of root mean squared error to the standard deviation of observations and Nash-Sutcliffe efficiency, we propose to use the wavelet coherence between measured and modeled streamflow time series (Rathinasamy et al., 2014). This analysis allowed us to highlight at which temporal scales the coupled model outperformed the SWAT model calibrated without the reservoir management tool.

6.2 Methodology

6.2.1   Study area

The Upper Adige river basin is located in the North-East Italian Alps (Figure 6.1) and has a drainage area of 6875 km^2 closed at Bronzolo. It is a typical alpine catchment with a large elevation range from 223 m a.s.l. to 3865 m a.s.l. The river basin is characterized by steep slopes and sharp hydroclimatic gradients (Callegari et al., 2015). The Upper Adige catchment is affected by a variety of stressors that affect its ecological status, and hydropeaking is the most important one (Navarro-Ortega et al., 2015). Although several studies have focused on Upper Adige subcatchments (Mei et al., 2014; Mei et al., 2016b; Penna et al., 2014; Penna et al., 2017a; Penna et al., 2017b; Tuo et al., 2016), few of them have tried to reproduce the hydrological behavior of the system including reservoir operation. Water discharge for hydropower production in the Upper Adige is managed with 15 artificial reservoirs and their operational rules influence the measured discharge at the gauging station of Bronzolo. Reservoir management operational rules are not available and that is what motivated this research.
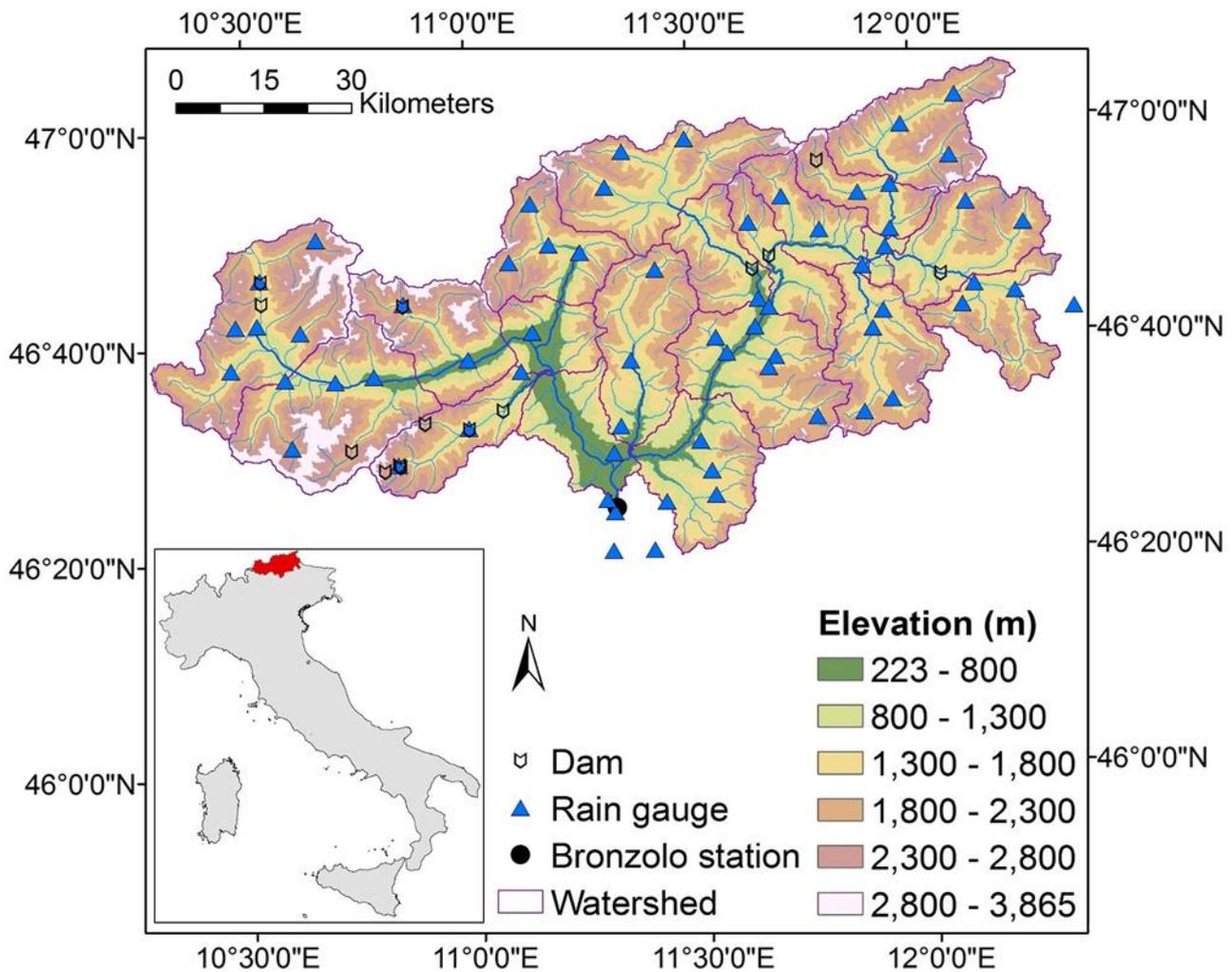
Figure 6.1 Upper Adige river basin

Land uses of the river basin are mainly forest, grassland and pasture, and barren land. The main soil types are loamy sand, silty clay, and sand. The data sets used to setup the SWAT model are the following:

- Digital elevation map (DEM) from Shuttle Radar Topography Mission (SRTM) produced by Consortium for Spatial Information (CGIAR-CSI), spatial resolution: 90 m x 90 m.

- Land use map: Corine Land Cover 2006 (CLC2006) from the European Environment Agency, spatial resolution: 100 m x 100 m.

- Soil map, developed by the Food and Agriculture Organization (FAO) of the United Nations.

- Digital stream network, EU-DEM product available at http://www.eea.europa.eu/data-and-maps/data/eu-dem.

- Daily precipitation and temperature data from a total of 65 rain gauges (60 rain gauges inside the basin and 5 rain gauges close to the basin boundary, see Figure 6.1) were collected from the meteorological surveys of the Autonomous Province of Bolzano (data available at http://weather.provinz.bz.it/default.asp). Previous studies have shown that the sensitivity of

hydrological models of the Upper Adige river basin depends on precipitation input (Duan et al., 2016; Mei et al., 2016a; Tuo et al., 2016) and this aspect will not be investigated further in this work.

- Daily measured streamflow data for the period at the Bronzolo gauging station (Figure 6.1) were collected in the GLOBAQUA project (Chiogna et al., 2016). In this work, we considered the period November 1, 2004 - October 31, 2010.

The river discharge time series measured at the Bronzolo gauging station is shown in Figure 6.2A. Streamflow displays the typical features of an alpine catchment, with low winter flows and high flows during the melting period in spring and summer. Figure 6.2B shows in more detail that river discharge fluctuates greatly during the winter period due to hydropeaking. Such fluctuations appear to be less pronounced during the high flow period.
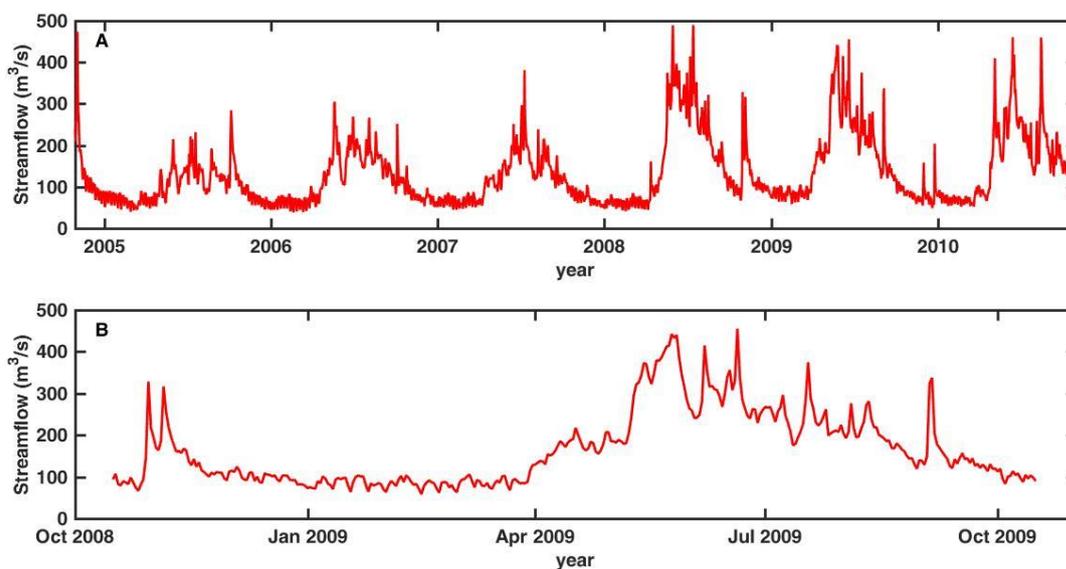
Figure 6.2 A) Streamflow at the Bronzolo gauging station for the entire period considered in this study. B) Streamflow at the Bronzolo gauging station from November 1, 2008 to October 31, 2009 shows the different impacts of hydropeaking during low flow and high flow conditions.

6.2.2   Hydrological model

SWAT was developed by the Agricultural Research Service of the United States Department of Agriculture (Arnold et al., 2012; Neitsch et al., 2011). It can be utilized to simulate daily water cycles, crop growth, sediment, nutrient and pesticide transport in large river basins and assess the effects of climate change and water resources management. The hydrological cycle simulated by SWAT is based on the water balance of various components including precipitation, infiltration, evapotranspiration, surface runoff, lateral flow, and percolation to shallow and/or deep aquifers. A SWAT model was set up

for the Upper Adige basin with 20 subbasins delineated in Figure 6.1. A two-year warm up period was used to mitigate the effect of the initial conditions as described by Tuo et al. (2016) and Vigiak et al. (2018). Sensitive parameters were identified with the one-at-a-time sensitivity analysis approach of SWAT-CUP (Abbaspour, 2015). Automatic model calibration was performed using the Sequential Uncertainty Fitting algorithm version 2 (SUFI-2) in the SWAT-CUP tool package (Abbaspour et al., 2015). Starting with the initial parameter ranges that were bounded to physically reasonable intervals according to literature (Arnold et al., 2012; Grusson et al., 2015), the model was calibrated with four iterations. After each iteration, parameter ranges were narrowed down considering both their reasonable physical limitations and the suggested ranges from the iteration results. The Nash–Sutcliffe efficiency (NSE) was used to assess goodness of fit of the simulations, following the guidelines of Moriasi et al. (2007).

6.2.3 Support Vector Machine

A brief mathematical introduction to support vector machines (SVM) is provided in Appendix A. For a more detailed description, we refer to e.g., the works of Schölkopf et al. (1999), Schölkopf and Smola (2002) and Smola and Schölkopf (2004). The basic idea behind SVM is that of using the maximum margin algorithm. This algorithm searches for the hyperplane (a line in the two dimensional case) with the largest separating margin between the observed data, to find the optimal function fitting the observations, paying attention to avoid overfitting to maintain the generality of the model.

Finding the optimal hyperplane is often a non-linear optimization problem and hence it is of difficult solution. The support vector algorithm applies the so-called kernel trick to solve this complex problem. The kernel trick aims to solve the non-linear problem in a space, called feature space, whose dimension is higher than that of the input space, i.e. the space of the observed data. In the feature space a linear representation of the dependency among variables is possible. The application of a kernel function allows us to solve the problem in the feature space, hence, solving a linear problem. This leads to much lower computational costs in comparison to other solvers.

6.2.3.1 Calibration of hyperparameters

In this work, we followed Hsu et al. (2003) and chose a Gaussian RBF kernel, and the calibration was executed by adjusting three hyperparameters - the kernel width 'rbf', the regularization constant 'C' and the ε-insensitive loss function parameter 'ε' - using a grid search method (Bergstra and Bengio, 2012). In the grid search method, we defined an upper and a lower bound for each hyperparameter. Then, one hyperparameter was kept fixed, while combinations of the other two hyperparameters were tested.

Each model run (i.e., each combination of hyperparameters) was evaluated by K-fold cross validation (Anguita et al., 2009). This method requires that you subdivide the data into k subsets of roughly equal size (four subsets of equal length were used in our case). In each step, the k-1 subsets were used to train the model, while the remaining subset was used for validation. Each subset was used exactly once for validation. Finally, the averaged error of all k trials was computed. Using the K-fold cross validation allowed us to find a set of parameters, which was not only suited for a particular subset. The advantage was that the optimal set of parameters identified should be less sensitive to the calibration period than using other calibration approaches. On the other hand, one drawback of this K-fold cross validation was that optimal parameters obtained for a subset and for the whole data set might not coincide (Schölkopf and Smola, 2002). Moreover, the smaller the data set, the stronger the regularization tended to be. In our work, kernel width was 6.5, the value of C was 20 and the ε value was 0.0288.

6.2.3.2 Selection of the variables used for training the SVM

The SVM was trained according to the river discharge modeled at Bronzolo with the SWAT model, the energy price and the day of the week starting from November 1, 2004. The last two pieces of information are in fact complementary: Energy price fluctuations are subject to market drivers, which depend on, for example, the cost of other energy sources (both fossil and renewable); the weekly fluctuations depend on the fact that energy demands are lower on weekends than on work days. Energy price and week days are to some extent correlated (less energy use in the weekend could mean a structural lower price), however, the variability of the energy price for different seasons and years, leads to a very small negative Kendall's tau (-0.06) with a significance level smaller than 95%. Several energy price data sets are available at http://www.mercatoelettrico.org/it/. Figure 6.3A represents the daily mean energy price for North Italy, which is the one considered in this study, while Figure 6.3B shows more clearly the daily variability of the signal. We observed that the time series displayed both weekly and seasonal fluctuations. Tests were also performed using other publicly available energy prices but no significant difference was identified for the purpose of this study.

Considering the constraints due to data availability, we used the period from November 1, 2004 to October 31, 2008 to train the SVM, while we used the following two years for model validation.
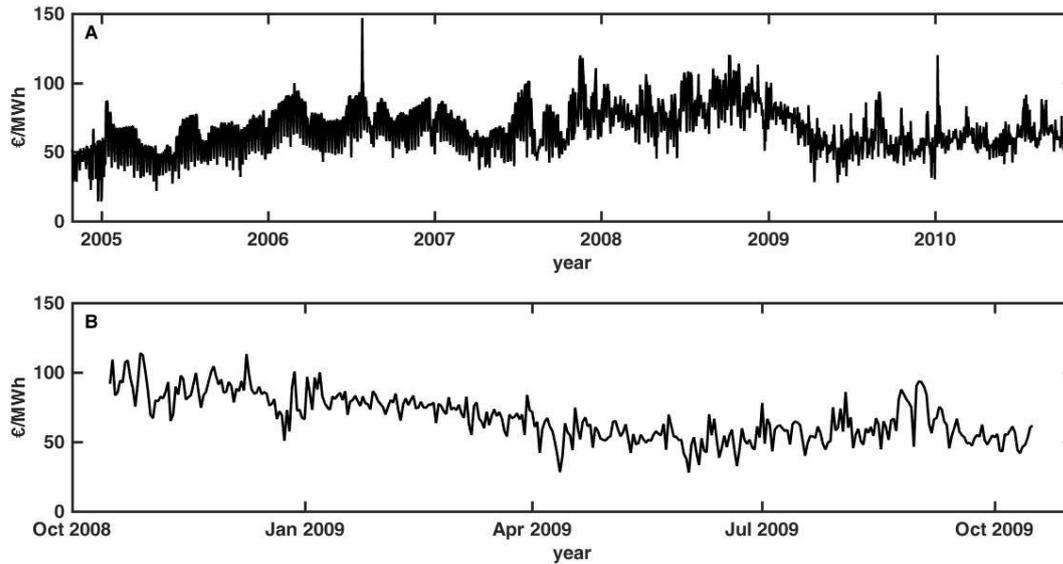
Figure 6.3 A) Mean daily energy price for North Italy for the entire simulation period. B) Mean daily energy price for North Italy from November 1, 2008 to October 31, 2009 showed more clearly the high frequency variability in the time series.

6.2.4   Continuous Wavelet Transform

Wavelet analysis has been shown to be a useful tool for detecting localized variations in the modes of variability within time series (Carey et al., 2013; Coulibaly and Burn, 2004; Guan et al., 2011; Lau and Weng, 1999; Marcolini et al., 2017; Zolezzi et al., 2009). Wavelet analysis is performed by decomposing the time series into a transformed variable in time-frequency space to determine both the dominant modes of variability and how these modes vary in time. We performed wavelet analysis using the Matlab toolbox developed by Torrence and Compo (1998) .

The continuous wavelet transform looks for similarity between a signal and a well-known mathematical function (Labat, 2005). The major difference compared to the Fourier transform, which is also the strength of this kind of analysis, is that the well-known mathematical function, called the wavelet function, is applied several times to the analyzed time series at different scales and at different temporal positions. Wavelet transform allowed us to determine not only the frequency content of a signal, which Fourier analysis is also able to do, but also the frequency time-dependence (Labat, 2005). Similar to other studies analyzing variability in climate and hydrological signals, the Morlet function was used as a wavelet function for its ability to evidence fluctuations in time series (Labat et al., 2004; Labat et al., 2000).

The wavelet power spectrum (WPS) represents the energy of the scale *s* and is useful for identifying fluctuations for scales with the largest influence on the signal. In particular, a larger positive amplitude

implies a higher positive correlation between the signal and the wavelet, and a large negative amplitude implies a high negative correlation.

The wavelet transform assumes that the time series is periodic. Since hydrological time series are not periodic and only a fraction of the entire time series is available, a bias is introduced at the beginning and at the end of the time series. To remove this bias, Torrence and Compo (1998) suggested padding the end of the time series with zeros. The introduction of the zero padding (not shown in the pictures) leads to the definition of a cone of influence, where edge effects are significant and hence the results are uncertain (Torrence and Compo, 1998). The cone of influence, indicated by shadowing the area of the contour plot, shows the wavelet transform influenced by edge effects.

Wavelet coherence (WTC) can be used to identify the scales where two time series focus on the power and how this may change in time (Torrence and Compo, 1998; Grinsted et al., 2004). An intuitive interpretation of wavelet coherence is as a generalization in the scale-time space of the squared cross-correlation coefficient r^2 between the signal content of the two time series. Similar to the classical correlation coefficient, high coherence values (close to 1) are displayed for the time windows and scales in which model and observations are highly correlated, while values close to 0 indicate low coherence. The possibility of identifying how model performance varies depending on the signal scales and time intervals is one of the main advantages of wavelet coherence analysis compared to classical correlation analysis. For more details about wavelet transform and wavelet coherence, the reader is referred to Torrence and Compo (1998), Grinsted et al. (2004) and Labat (2005).

6.2.5   Evaluation of model performance

The performance of the two models was assessed by global model performance indicators: MAE (mean absolute error), MAPE (mean absolute percentage error), RMSE (root mean square error), RSR (ratio of RMSE to the standard deviation of observations) and NSE (Nash-Sutcliffe Efficiency) (Table 6.1). MAE and RMSE measure the absolute deviation between model results and observations, while MAPE and RSR provide a relative error estimate which is not influenced by the different magnitude in the data. NSE evaluates whether the model output outweighs a simulator just taking the mean of the observations. Note that, due to the square operation, NSE puts high weight on large discharge values.

Table 6.1 Global model performance indicators. Y_j^{obs} indicates the data of the observed time series, Y_j^{sim} are the simulation results, the over bar represents the mean value of the time series.

Name	Equation	Description
Mean absolute error	$\text{MAE} = \dfrac{1}{n}\sum_{j=1}^{n}\left|Y_j^{obs} - Y_j^{sim}\right|$	$[0,+\infty)$, 0 is the optimum value (o.v.)

Mean absolute percentage error	$\text{MAPE} = \frac{1}{n}\sum_{j=1}^{n} \frac{\left|Y_j^{obs} - Y_j^{sim}\right|}{Y_j^{obs}} * 100$	$[0,+\infty)$, 0 is the o.v.
Root mean square error	$\text{RMSE} = \sqrt{\frac{1}{n}\Sigma_{j=1}^{n}(Y_j^{obs} - Y_j^{sim})^2}$	$[0,+\infty)$, 0 is the o.v.
RSR	$\text{RSR} = \frac{\text{RMSE}}{\text{STDEV}_{obs}} = \frac{\sqrt{\Sigma_{j=1}^{n}\left(Y_j^{obs}-Y_j^{sim}\right)^2}}{\sqrt{\Sigma_{j=1}^{n}\left(Y_j^{obs}-\overline{Y^{obs}}\right)^2}}$	$[0,+\infty)$, 0 is the o.v.
Nash-Sutcliffe Efficiency	$\text{NSE} = 1 - \frac{\Sigma_{j=1}^{n}\left(Y_j^{obs} - Y_j^{sim}\right)^2}{\Sigma_{j=1}^{n}\left(Y_j^{obs} - \overline{Y^{obs}}\right)^2}$	$(-\infty,1]$, 1 is the o.v.

In addition to these traditional metrics, we also applied the WTC between the measured and the simulated discharge. This analysis aimed at identifying the scales at which observed data and model results focused on the wavelet power spectrum, with particular attention to the evolution of this behavior in time (Torrence and Compo, 1998; Grinsted et al., 2004). High coherence values mean that the modeled and the observed time series share the same frequency content in a specific time window.

6.3 Results and discussion

Figure 6.4 shows the wavelet transform of the logarithm of the river discharge and energy price signals. The analysis was performed on the log transformed streamflow data to highlight the effect of hydropeaking. Both signals show a very interesting behavior: A high power is present for the periods of 7 days and 3 days until 2009. After this date, the power of the signal becomes weaker, in particular the energy price signal. The decrease in the power signal indicates that the energy price time series involves at least one clear change in 2009. This change may be attributed to various reasons: It could be caused by, for example, random energy price fluctuations at scales smaller than one day and/or by a more extensive use of other renewable energies. For the purpose of this study, however, it is not relevant to identify the exact cause of this change. The change in the streamflow time series is hard to explain considering purely hydrological reasons, since no abrupt change in meteorological or land use data was identified in 2009 in previous studies (Chiogna et al., 2016).

Both streamflow and energy price signals display high power during longer periods (at about 180 and 365 days). However, this seasonal and yearly dynamic behavior is not of high relevance for this study and will not be discussed further.
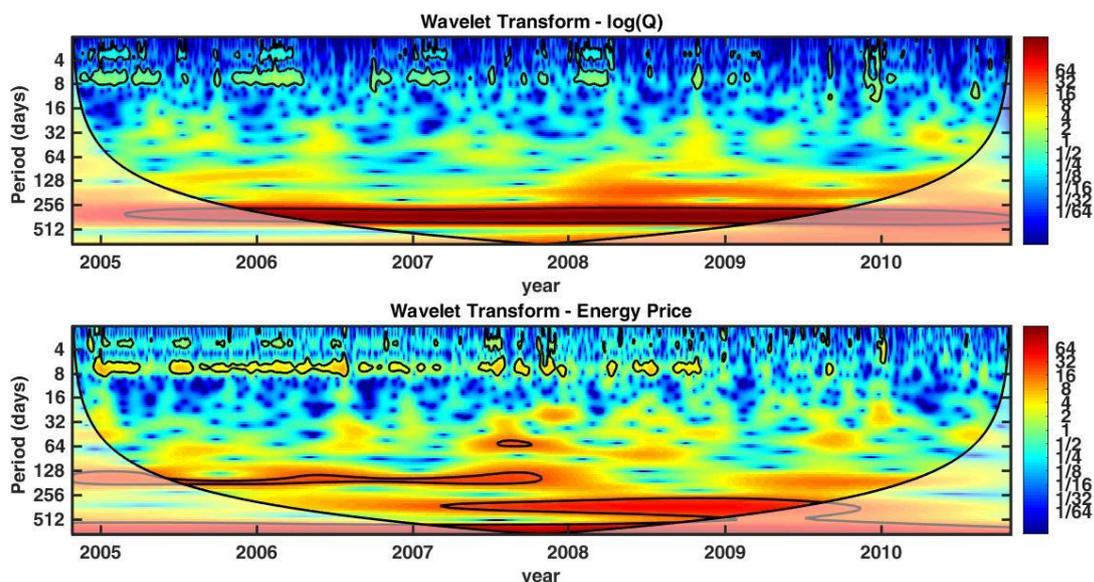
Figure 6.4 A) Wavelet transforms of log(Q) and B) of the energy price. Red indicates high power while blue indicates low power.

Figure 6.5A shows the observed streamflow data at Bronzolo, the model results obtained using SWAT and the SVM model trained using the energy price, the day of the week and SWAT model output results. Figure 6.5B shows in more detail that the SWAT model was not able to capture the fluctuations in river discharge due to hydropeaking, since no data related to hydropower management were available for this catchment in this study. However, the SVM model results clearly showed that the model was capable of including the impact of energy market fluctuations into the predicted river discharge. The SVM was trained in a period when the energy price displayed a strong weekly and sub-weekly periodicity (Figure 6.4). The test of the SVM was performed in the following period when this periodicity was much weaker (Figure 6.4). A visual evaluation of Figure 6.5A allows us to conclude that the SVM model performance was not affected by this change in the energy price behavior, showing the robustness of the approach towards changes in the water management operation of the reservoirs.
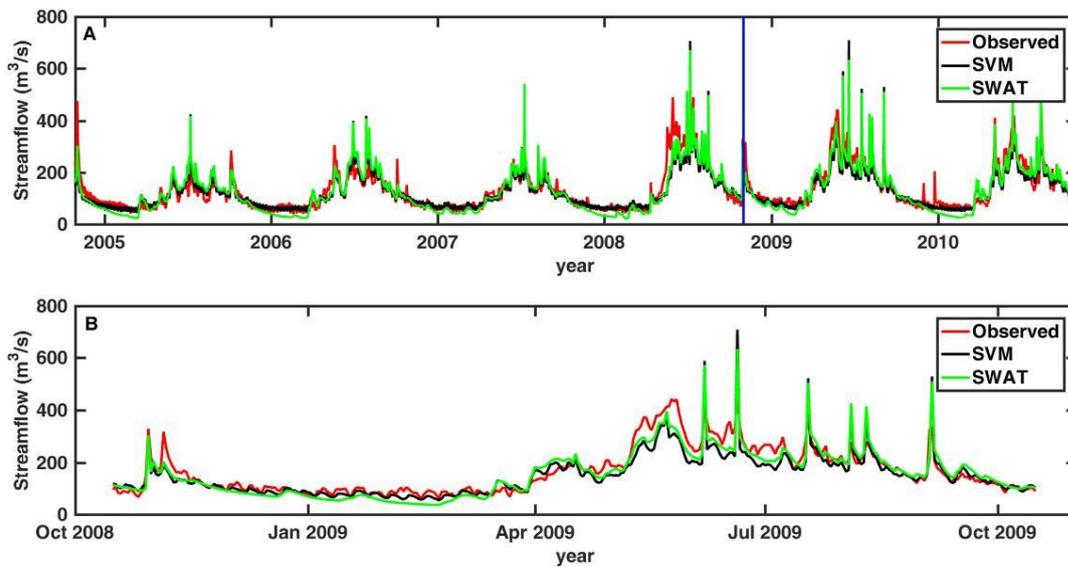
Figure 6.5 A) Streamflow observations in Bronzolo (red line), SWAT model results (green line) and SVM model results (black line) for the entire simulation period. The vertical blue line separates the SVM training period from the test period. B) Streamflow observations in Bronzolo (red line), SWAT model results (green line) and SVM model results (black line) from November 1, 2008 to October 31, 2009 show more clearly the different impacts of hydropeaking during low flow and high flow conditions.

This strength of the proposed approach becomes even more evident from the analysis of Figure 6.6, which shows the coherence between energy prices and both the measured and the modeled river discharge time series. Coherence between observed streamflow and energy price is almost equal to 1 for all winters before 2009 for the period of 3-7 days and the two signals are in phase (the arrows in the first panel of Figure 6.6 point to the right). During high flow conditions, however, coherence decreased significantly. The high coherence between the two signals can be considered an indication of causality, since the significant region of Figure 6.6 is so extensive that it is very unlikely that this is simply by chance (Grinsted et al., 2004). According to the coherence analysis, the fluctuations in the energy price drive therefore the fluctuations in the streamflow. Hence, we can also attribute the change in streamflow variability for the period of 3-7 days observed after 2009 to the change in the energy price variability. Indeed, it was shown that energy price variability affects reservoir management operations and hence water release (storage) from (in) the reservoirs (Gaudard et al., 2016; Russell and Campbell, 1996).

Coherence between energy price and SVM model results show a very similar behavior: High coherence during the winters before 2009 at the period of 3-7 days and low coherence elsewhere. This indicates that the SVM was able to learn the behavior of the impact of the energy market fluctuations on river discharge and it was able to reproduce such behavior properly in the case of a complex temporal dynamic as well. As expected, the SWAT model could not reproduce such behavior and it did not display high

coherence with the energy price signal anywhere. The results of Person (2013) showed that hydropeaking has a strong impact on fish habitat mainly during winter where suitable habitat is strongly displaced under unstable flow conditions. Therefore, it is important from an ecological perspective, to properly capture hydropeaking in river discharge during the winter season.

Moreover, a hypothetical SWAT model with implemented reservoirs should be updated in case of changes in reservoir management operations, and it would therefore require strong collaboration in terms of data exchange between hydropower companies and water authorities. Such cooperation is always desired, however, it is often impracticable because their interests differ. We should also consider that the energy market is becoming more and more variable due to, for example, the increased introduction of renewable energy sources, which depend on local atmospheric conditions (e.g., cloudiness, presence of wind) and consequently reservoir management rules have to frequently adapt to this variability.
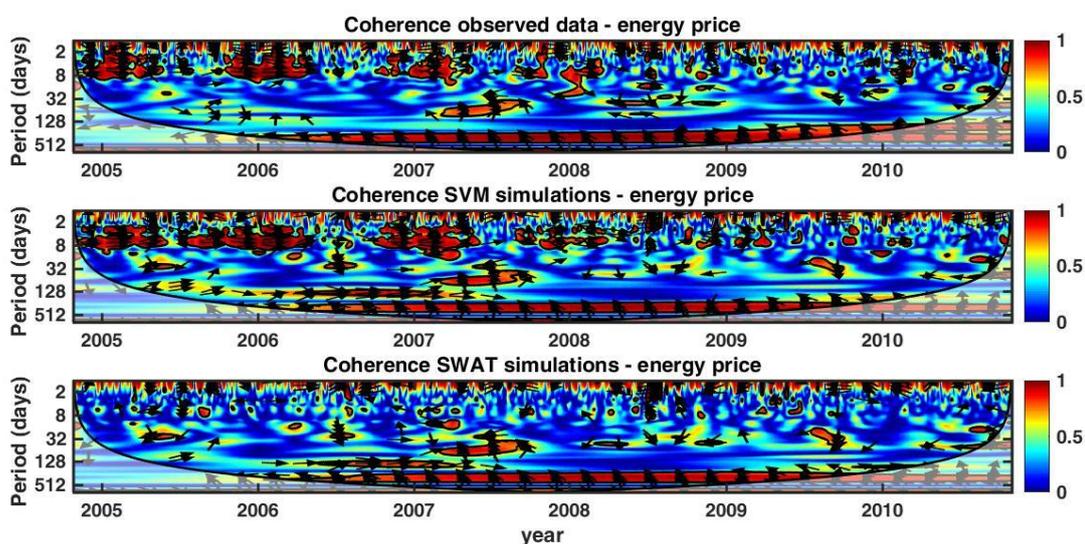
Figure 6.6 Wavelet coherence between energy price and observed streamflow data, SVM simulations and SWAT simulations. Red indicates high coherence while blue indicates low coherence. The direction of the arrows indicates a phase shift between the signals: arrows pointing to the right indicate that the two signals are in phase, while arrows pointing to the left indicate a phase shift of π.

In Figure 6.7, we compare the cumulative distribution functions of the observed streamflow with the SWAT and the SVM models considering the entire time series (i.e., validation and calibration periods are considered together). The main improvement from using the SVM model was achieved for low discharges, where hydropeaking fluctuations were more evident. On the other hand, for a range of high flows ($200\,\mathrm{m^3/s} < Q < 300\,\mathrm{m^3/s}$) the SWAT model slightly outperformed the SVM model. This behavior can be explained by observing Figure 6.5: During the training period of the SVM, events with a discharge larger than $300\,\mathrm{m^3/s}$ were very rare, and the SVM was trained considering high flow events that were

generally underestimated by the SWAT model. Therefore, during the training period, the SVM model did not learn to optimally reproduce the high flow conditions observed during the testing period. In order to solve this problem, the SVM could be trained with additional parameters controlling the hydrologic response of the catchment, such as temperature, precipitation or groundwater levels. However, such an analysis is beyond the scope of this work, which focuses on improving model performance in reproducing hydropeaking.
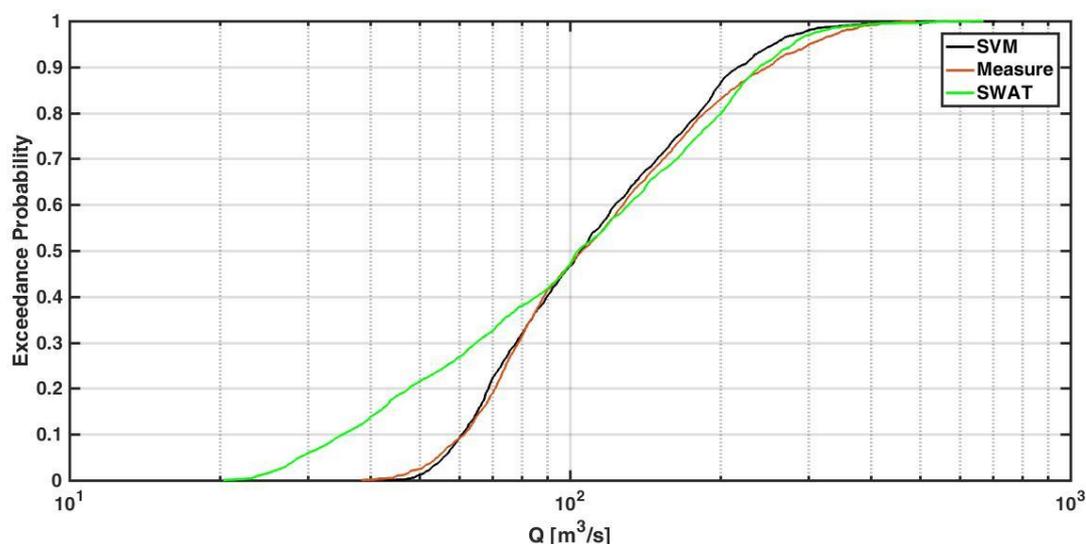
Figure 6.7 Cumulative distribution functions of the observed river discharge (red line) of the SWAT model (green line) and of the SVM model (black line).

In Table 6.2, we summarize the global model performance indicators computed for the SWAT and SVM models. In all cases, the SVM model outperformed the SWAT model and reached a closer prediction of observed streamflow. However, the difference between the two models according to these indicators is relatively small considering that the SVM model has three additional calibration parameters (the kernel width, the regularization constant and the ε-insensitive loss function parameter) compared to the SWAT model. Therefore, in our opinion, this kind of comparison is not sufficient to judge whether the SVM model outperforms the SWAT model.

Table 6.2 Model performance indicators.

	MAE [m^3/s]	MAPE [-]	RMSE [m^3/s]	RSR [-]	NSE [-]
SWAT	26.6	0.23	36.3	0.46	0.79
SVM	21.0	0.14	34.8	0.44	0.81

Figure 6.8 allows us to better explain where the coupled model performed better than the SWAT model. We evaluated the hydrological model results considering their coherence with the observations for several periods from one day to one year. Both SWAT and SVM models displayed high coherence (on

average higher than 0.8) with the observations at the period of one day. However, the difference in model performance for longer periods between 3 and 30 days was large. The SWAT model displayed an average coherence (computed over the entire simulation period) smaller than 0.5 with the observations at 3 days and 7 days, while for the SVM model, coherence ranged between 0.7 and 0.8 for the same periods. Both models were capable of reproducing the observed discharge at the monthly (i.e., about 30 days) and yearly periods showing an average coherence of 0.8 and 1, respectively.
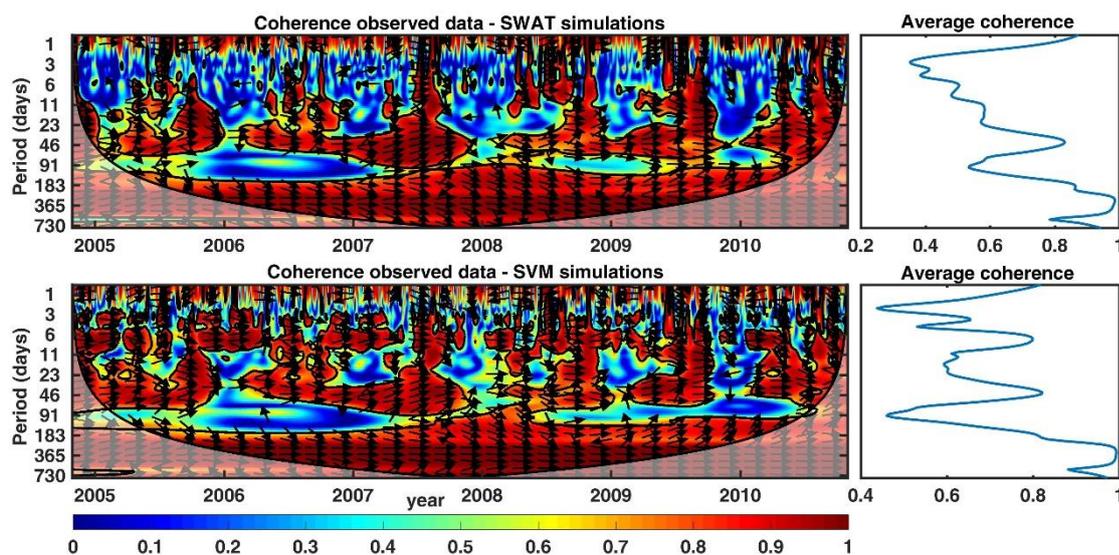
Figure 6.8 Coherence between streamflow observations and SWAT model simulations as well as streamflow observations and SVM model simulations. Red indicates high coherence while blue indicates low coherence. The direction of the arrows indicates a phase shift between the signals: arrows pointing to the right indicate that the two signals are in phase, while arrows pointing to the left indicate a phase shift of π.

6.4 Conclusion

In this work, our aim was to improve the model results obtained using a hydrological model such as SWAT for catchments affected by hydropeaking, where no reservoir operation data are available. We chose to study the Upper Adige River basin in North-East Italy. In order to improve hydrological model simulations, we proposed to reproduce streamflow using an SVM trained with the output of the hydrological model, the energy price and the day of the week. Our results showed that the SVM outcomes were able to capture hydropeaking behavior and led to better model performances. In particular, we proposed to use the wavelet coherence between the measured and the modeled data to identify at which scales and for which time windows the model performed the best. Wavelet coherence analysis indicated that the SVM model was particularly effective in reproducing hydropeaking during the winter period when low flow conditions are highly affected by weekly fluctuations in river discharge.

On the other hand, we observed that the SVM model was not able to outperform the SWAT model during high flow conditions when daily streamflow was less affected by hydropeaking than during low flow conditions.

This work represents an attempt to couple SVM and more traditional hydrological models such as SWAT. The obtained results were highly satisfactory but there is still room for improvements. In particular, modeling hydropeaking at the hourly scale should be explored in the future. This would provide important information for ecological studies. Moreover, the development of an SVM-based reservoir management module directly implemented within the hydrological model, allowing for streamflow routing and for a more realistic reproduction of water storage and release, would be greatly beneficial to model hydrological processes in catchments highly affected by human impact. Another test to be performed in future studies is to investigate the exportability of the trained SVM to other gauging stations in the catchment and to investigate the spatial variability of the SVM hyperparameters.

Appendix A

In this section, we will give a brief introduction to the mathematical description of the support vector regression. Suppose we have some observed data $\{(x_1, y_1), \dots, (x_l, y_l)\} \subset X \times \mathbb{R}$, where X is the input space, which in our case is $\mathbb{R}^d$. In the linear case, we look for a function $f = \langle w, x \rangle + b$ that solves the following convex optimization problem

$$\text{minimize} \quad \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{l} (\xi_i + \xi_i^*) \tag{1}$$

$$\text{subject to} \quad \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

where ε is a parameter indicating the maximum allowed error between observations and model predictions. In order to deal with otherwise unfeasible constraints, two slack variables ξ_i and ξ_i^* are introduced. The introduction of these variables allows for some flexibility in the definition of the error range. The positive constant C determines the trade-off between which deviations larger than ε are tolerated and the flatness of the function f. Summarizing, Eq. (1) aims at finding a function f as flat as possible that has at most ε deviation from the observations.

In order to solve the problem, we deal with its dual form. We first construct a Lagrange function from the objective function

$$L := \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) - \sum_{i=1}^{l}(\eta_i\xi_i + \eta_i^*\xi_i^*)$$

$$-\sum_{i=1}^{l}\alpha_i(\varepsilon + \xi_i - y_i + \langle w, x_i\rangle + b)$$

$$-\sum_{i=1}^{l}\alpha_i^*(\varepsilon + \xi_i^* + y_i - \langle w, x_i\rangle - b) \qquad (2)$$

where η_i, η_i^*, α_i and α_i^* are Lagrangian multipliers. It can be shown that the solution of Eq. (2) represents a saddle point with respect to the primal and dual variables for the Lagrange function L. Hence, to solve Eq. (2), the partial derivative of L with respect to the primal variables has to vanish.

$$\partial_b L = \sum_{i=1}^{l}(\alpha_i^* - \alpha_i) = 0 \qquad (3)$$

$$\partial_w L = w - \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)x_i = 0 \qquad (4)$$

$$\partial_{\xi_i^{(*)}} L = C - \alpha_i^{(*)} - \eta_i^{(*)} = 0 \qquad (5)$$

If we substitute Eqs (3), (4) and (5) in the Lagrange function, we obtain the dual optimization problem

$$\text{minimize} \quad \begin{cases} -\dfrac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)\langle x_i, x_j\rangle \\ -\varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{l}y_i(\alpha_i - \alpha_i^*) \end{cases} \qquad (6)$$

$$\text{subject to} \quad \begin{cases} \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, C] \end{cases}$$

Notice that in the dual optimization problem, the variables η_i and η_i^* are not present anymore.

From Eq. (4) we have

$$w = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)x_i \qquad (7)$$

Notice that according to Eq. (7), w can be described completely in terms of a linear combination of the training patterns x_i.

This leads to the so-called Support Vector expansion

$$f(x) = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*)\langle x_i, x \rangle + b \tag{8}$$

For the computation of $f(x)$, it is therefore not necessary to explicitly compute w. Moreover, the complete algorithm can be described in terms of dot products. In order to compute b, we apply the Karush-Kuhn-Tucker conditions, according to which we have

$$\alpha_i(\varepsilon + \xi_i - y_i + \langle w, x_i \rangle + b) = 0$$

$$\alpha_i^*(\varepsilon + \xi_i^* + y_i - \langle w, x_i \rangle - b) = 0$$

$$(C - \alpha_i)\xi_i = 0$$

$$(C - \alpha_i^*)\xi_i^* = 0 \tag{9}$$

From Eq. (9), we have that for all the samples with a distance to f smaller than ε, α_i and α_i^* vanish. Therefore, not all x_i are needed to compute w. Since, as we observed before, w can be described as a linear combination of the training patterns, this leads to a reduction in the complexity of the problem. The observations for which the coefficients are non-vanishing (and hence contributing to w) are called Support Vectors.

So far, we have considered a linear problem. If the problem is not linear, a solution could be to map the input data in some feature space F, where the problem can be linearly solved, with some map $\Phi: X \longrightarrow F$. Of course, this could lead to a large increase in the complexity of the problem and make it computationally unfeasible. Since the Support Vector algorithm can be described in terms of dot products of the patterns x_i, we can avoid increasing computational costs by applying a kernel function $k(x, x') := \langle \Phi(x), \Phi(x') \rangle$. The new representation of the function f would then be

$$f(x) = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*)k(x_i, x) + b \tag{11}$$

Therefore, in the nonlinear case, with the application of the so-called kernel trick, we can linearly solve the problem in a feature space, performing our computations in the input space, and hence limit the computational costs. As described by Schölkopf et al.(1999); Schölkopf and Smola (2002); Smola and

Schölkopf (2004), the map Φ has to satisfy some specific conditions to be applied for the computation of the dot products in the feature space.

Chapter 7

Conclusion and outlook

Mountainous catchments play a vital role in affecting water resources around the world. Therefore, reasonable understandings and reliable representations of hydrological behaviors in mountainous regions are important. In this work, we investigate the solutions for improving SWAT applications, regarding precipitation, snow processes, and the effect of hydropower operation, in Alpine catchments of Adige River Basin, Italy. Addressing how to evaluate and select the optimal precipitation input, how to utilize snow data, and how to take use of data related to hydropower plant operation effectively is meaningful for improving SWAT modeling applications in the Alpine catchments. The rational utilization of the available data considering model structure is of primary importance. Based on this concept, this work aims to improve SWAT application in Alpine catchments and hence increase the reliability of model results by utilizing the data related to precipitation, snow processes, and hydropower plant operation, respectively, while considering the corresponding SWAT model structures. In the following, the main outcomes of this study are summarized.

- Optimal precipitation inputs for SWAT modeling in Alpine catchments

  In chapter 2, we conduct a comprehensive evaluation of eight high resolution remote sensed precipitation products in the Adige River Basin, Italy by comparing to measurements of rain gauges at various temporal (daily, monthly and annual) and spatial scales (grid and watershed). All products present different occurrence frequency of daily precipitation for some intensity ranges compared to rain gauge data, such differences are expected to have significant effects on hydrological modeling applications. And all products also tend to show higher error in the winter months (December-February) when precipitation is low. Consistent improved evaluation metrics are observed as the evaluation spatial scale increases from grid to watershed and the temporal scale increases from daily to monthly. Considering all evaluation metrics and aspects, in general the CHIRPS (satellite-gauge based, 0.05°) and TRMM (satellite-gauge based, 0.25°) can be ranked as the top two best products, with relatively lower bias and error than others and reproducing a general southeast-northwest decreasing spatial pattern better. Therefore, CHIRPS and TRMM are selected as two representative remote sensed datasets to be evaluated as precipitation input for the hydrological modeling application in Chapter 3.

  In Chapter 3, the impact of four different precipitation inputs (observed, IDW, CHIRPS, and TRMM datasets) on streamflow predicted using SWAT in three Alpine catchments of the Adige River Basin is investigated, by analyzing the rainfall features, model performances, parameter uncertainty, prediction uncertainty and the potential relationships among the above components. Since all these four precipitations do not consider properly the orographic effects and the precipitation information at high elevation (Adler et al., 2015; Isotta et al., 2014), the elevation band method of SWAT presents to be necessary for estimating the orographic effects in the Alpine catchments, in order to reproduce

the hydrograph properly and close the water balance as well. The four different precipitation inputs are different in terms of the amount and the temporal distribution of precipitation. Using elevation band method, all the calibrated SWAT models with different precipitation datasets are able to reproduce the measured discharge. However, in order to fit the measured river discharge, the parameters are calibrated distinctly to adjust the water volume of different hydrological components for coping with the different rainfall features of the four precipitation products. As a results, although the discharge is reproduced well, different precipitation inputs result in different best parameter values and optimal parameter ranges, which indicates distinct hydrological components and information to the water management practices. Besides, no pattern is observed to identify which precipitation would generally lead to smaller/larger parameter uncertainty for a catchment nor for a specific parameter. In other words, parameter uncertainty is both precipitation-varied and catchment specific. Reflected by P-factor (fraction of measured data enveloped by the 95PPU band) and R-factor (thickness of the 95PPU envelop) values, precipitation-varied and catchment specific output uncertainty is also observed, although IDW based precipitation dataset leads to more consistent output uncertainties than the other three precipitation datasets. Considering the values of "goodness of fit" for calibrated model results and fixed parameter test and the uncertainty in general, IDW based precipitation outperforms the other studied precipitation inputs when coupled with elevation band method in this Alpine region.

In summary, selection of precipitation products has decisive influences on model performance, model parameter uncertainties and output uncertainties during hydrological modeling application in the Alpine catchments. Due to the lack of information in high elevation areas, the available precipitation datasets are of large uncertainty for the studied catchments. The precipitation uncertainty propagates to model parameter uncertainty and output uncertainty, which can lead to distinct conclusions about the most relevant hydrological processes in a catchment and result in different water management strategies or policies. Therefore, in order to select the optimal precipitation input from a specific mountainous catchment application, the uncertainty propagated from different precipitation datasets has to be taken into consideration. In addition to discharge, information of other hydrological components such as snow could be helpful to validate the model and constrain the parameters to reduce the model uncertainty and thus increase the reliability of model results.

- Constrain calibration of model parameters considering snow information

  Snow processes are of great importance in Alpine catchments. Snow water equivalent (SWE) of the snow packs is actually the difference between accumulated snow and snowmelt water. And it is the common form for describing snow in hydrological models. Although snowmelt output is also

available in SWAT, owing to the scarcity of snowmelt data in large scale Alpine catchments, the snowmelt parameters cannot be calibrated accordingly and therefore the snowmelt contributions to streamflow cannot be validated or verified directly. Calibration of SWAT snow parameters using SWE information of different spatial scales is of significant implications for reducing the uncertainty of the modeled snow dynamics. The performance of this calibration approach is restricted by the model structure of SWAT and the practical condition that sparse data cannot represent the complex nature of snow dynamics.

In Chapter 4, the SWE based calibration strategy is proposed and the feasibility of calibrating SWAT snow parameters according to SWE data of subbasin and elevation band scales is proven. As shown in Chapter 4, SWAT estimates the snow dynamics at relative larger spatial scales (i.e. subbasin and elevation band), which matches the sparse availability and incomplete spatial distribution of the snow data. The model reproduces well the general snow behavior of large scale, considering the uncertainties of model structure, original snow data and SWE conversion processes. Regarding the model performance of subbasin-scale SWE, basin scale and subbasin scale snow parameters lead to equally good estimations of the overall behavior of subbasin scale snow dynamics for high elevation subbasins. Subbasin scale parameterization is more suitable when the subbasins display a large variability in elevation range. However, it is difficult to clarify that which snow parameterization approach can lead to a smaller model uncertainty. Based on the same sources of uncertainty, the basin scale parameters manage to represent the common characteristics throughout the entire basin, while the subbasin scales parameters consider and respect the heterogeneous nature between the subbasins. Furthermore, owing to that different combinations of snow parameter values have led to similar acceptable simulations of the subbasin-scale SWE, the use of Sub-SWE is not enough to identify the most suitable set of snow parameters. Therefore, the elevation band scale SWE (EB-SWE), another snow output of SWAT, is introduced as an additional reference for parameter assessment by qualitatively judging whether the EB-SWE estimation is reasonable or not in comparison to the station observation of each band. Furthermore, the parameters are modified accordingly by making a compromise between Sub-SWE and EB-SWE performances. As a result, the model can both reproduce the subbasin scale snow dynamics well and respect the observations of the elevation band level. In this way, as evidently supported by the outcomes of Chapter 4, the uncertainty of snow simulations is reduced by constraining the snow parameters considering both model structure and the natural snow dynamics of different spatial scales. However, since SWAT only consider the orographic effect by elevation band method, the local climate conditions in mountainous region (Endrizzi et al. 2006, Hiemstra et al. 2002) are not represented by SWAT model.

Therefore, some unusual snow dynamics, such as the higher station obtaining lower SWE values than the lower station, cannot be reproduced well in the model results.

Furthermore, in Chapter 5, the SWE information is introduced together with discharge observations in the multi-objective calibration procedure to compare with the traditional discharge based, single objective calibration applications in Alpine catchments. When involving the discharge estimations, snow information also played crucial roles not only in addressing more reliable snow dynamics, but also in reproducing discharge observations with more reliable contributions from different hydrological components. Referring to river discharge alone, the calibrated model is not capable of represent the snow dynamics well, no matter whether the spatial variability of snow parameters is considered or not. What is worse, the simulated streamflow time series look in good agreement with the observations, but they are highly uncertain and are very likely composed of unreliable contributions of different hydrological processes in a lumped way, without the constrain from other important information such as snow. This is evident in Chapter 5 supported by the results of snowmelt simulations and the cumulative curves. This uncertainty definitely will lead to inappropriate instruction for local water resource management practices. The multi-objective calibration approach shows evident improvements in obtain a win-win situation for snow and streamflow estimations in Alpine catchments. Following the strategy verified in Chapter 4 to constrain the snow parameters, the snow dynamics of both subbasin and elevation band scales are represented properly while keeping the streamflow simulations fitting the observations well. Based on the water balance frame, as snow processes are major components of the water cycle in Alpine catchments, more reliable estimations of snow processes supported by snow information can both reduce the uncertainty of snow modeling and indirectly constrain the uncertainty of the other hydrological components in a lumped way by restraining their total quantity within a smaller variation range. However, the impacts of hydropower plant operation on the river discharge are aware but cannot been handled properly due to the lack of operation data, which further complicates and enlarges the model uncertainty of simulated discharge that is already affected by the uncertain hydrological components. Therefore, from another prospective, involving snow information into model calibration is extraordinarily necessary when the river discharge is interrupted by hydropower plant operation. The use of snow information as calibration reference could keep snow parameters independent from the impacts of the uncertainty propagated from hydropower plant operation and the other hydrological processes.

- Improvements of coupling Support Vector Machine (SVM) method with hydrological model on hydropeaking simulation

In Chapter 6, SVM method is utilized to improve the SWAT model results for Alpine catchments affected by hydropower plant operation. With no operation data, the SVM is trained with simulated discharge of SWAT, the energy price and the day of the week, in order to reproduce the effect of hydropeaking on streamflow. As shown in the outcome of Chapter 6, without data to implement the reservoir operation module, the SWAT model fails to capture the fluctuations in river discharge because of hydropeaking effect. Conversely, able to involve the energy price data, SVM incorporates the energy market fluctuations into the river discharge, which captures the behavior of hydropeaking. As a result, improved streamflow simulation is obtained with higher values of metrics. In fact, energy price variability has been proven to affect hydropower plant operations and hence impact the water release (storage) from (in) the reservoirs (Gaudard et al., 2016; Russell and Campbell, 1996). In other words, energy price functions highly affect and drive the river discharge fluctuations due to hydropeaking, which is proven by the high coherence between the observed streamflow and the energy price data. Relevance is also supported by the similar behaviors of wavelet transform signals between the logarithm of river discharge and energy price. The coherence analysis between SVM results and the observed streamflow further indicates that SVM is capable of reproducing the impact of the energy market fluctuations on river discharge. In particular, SVM performs effectively in reproducing hydropeaking of the low flow conditions when hydropeaking fluctuations are more evident, which is a significant improvement in comparison to SWAT. In fact, SWAT model cannot reproduce the impact of hydropower plant operation on river discharge, because it does not have the operation data and it cannot use the operation related energy price data. The simulated discharge of SWAT not only fails to match the fluctuated observations, but also contains large uncertainties because the model tries to imitate the hydropeaking fluctuations by forcing the hydrological parameters incorrectly. Utilization of SVM helps to reduce this uncertainty and improve the model performance in simulating the river discharge impacted by hydropeaking by learning from the energy price data that drive the hydropower operation. This coupling use of SVM and the SWAT model has important implications to fill the gap/disability of the model structure in simulating the hydropower plant affected river discharge by taking use of alternative available data.

Based on the results of the work, in order to improve the SWAT modeling applications in mountainous regions, future researches could better explore the typical regional hydrological influencing factors by addressing the following problems:

- Improvements of SWAT model structure

  Data availability of mountainous catchments is not only limited by the problem of data scarcity in comparison to the relative large scale of the catchments, but also is restrained by the serious lack of

information from high elevation areas. This problem is difficult to improve in a short time and especially impossible to solve for the historical period. Hydrological model relies on historical data to simulate the catchment characteristics and build up a calibrated model. The current model structure of SWAT makes it a suitable tool matching the situation of large scale mountainous catchments, because SWAT is capable of using sparse data effectively by utilizing precipitation in subbasin scale and producing snow outputs at relative large spatial scale (subbasin and elevation band scales). What we have done so far is about how to take use of the available resources effectively and efficiently based on the current model structure, however, a proper modification of the model structure would bring fundamental improvements. Since orographic effect is important for both precipitation and snow, structure modification could be implemented to perform hydrological calculations based on elevation classified areas. From a technical point of view, the modification is possible by introducing elevation band as the fourth factor to classify the HRU. In this way, the HRU still keeps the unique physical information such as soil, land use and slope. More importantly, a selection protocol could be set up to let users choose either the subbasin structure or elevation band structure, depending on the specific condition of the study area. From the perspective of data availability, this modification is also feasible: Elevation band precipitation inputs could be produced by the elevation band method and calibrated by constraining the precipitation lapse rate according to hydrological references and validated by comparing to the precipitation observations of certain elevation areas; Elevation band and subbasin-scale SWE output will also be helpful to constrain the parameters. This modified elevation band based structure could improve the model descriptions of the hydrological dynamics in mountainous region and differentiate between the processes occurring at high and low elevation HRUs, while still keeping the practicality of the model. Besides, although from a theoretical point of view, grid based model structure modification would bring improvements with finer spatial distributed hydrological simulations, which could benefit the snow dynamics and ground water estimations, it is not recommended and not practical at the large scale mountainous regions where accessible data are limited.

- Application of datasets of different sources

  Applying data of different hydrological processes to constrain the parameters and model uncertainty is always beneficial. One possible way is to install measurement devices in the catchments to get the information of high elevation or sensitive areas, such as precipitation, snow depth, snowmelt, soil moisture, groundwater and so on. These datasets could be used to calibrate or validate the model, and they could be reliable but have limited spatial distributions. Another possible way is to take advantage of innovative satellite products to further obtain data for the catchments and/or constrain the hydrological model. For example, the remote sensed soil moisture data of Soil Moisture Active

Passive (SMAP) satellite, available since 2015, could be utilized to validate and improve the hydrological model (McNairn et al., 2015). The Surface Water and Ocean Topography Mission (SWOT) offers observations of the temporal and spatial variations in water volumes stored in rivers, lakes, and wetlands (Solander et al., 2016), which is useful and could be applied to constrain the hydrological model. Compared to field measurements, remote sensed data have less accuracy but have better spatial distributions. A combined use of the datasets from different sources could be useful to improve the credibility of the model results.

- Improve the use of effective tools such as SWAT-CUP and SVM

  Utilizing useful tools effectively facilitates the hydrological modeling applications. Based on our work, we propose two relevant improvements for the future research. First, in order to perform the automatic calibration based on snow information for SWAT, the calibration software SWAT-CUP has to be further developed to allow the user to select the elevation band scale SWE results as target variable for calibration. Moreover, detailed investigations about the objective function for snow calibration could be performed with SWAT-CUP to identify the most suitable metrics for evaluating SWE simulation. Besides, an automatic rejection criterion could be introduced when the parameter set converges to a maximum snowmelt factor smaller than the minimum snowmelt factor. Second, compiling a SVM-based hydropower management module in to SWAT could outperform the current coupled use of SVM with SWAT. This direct coupling could realize the dynamic interactions between the SVM and SWAT, which could ensure the streamflow routing in-between and produce a more realistic reproduction of water storage and release. A dynamic and synchronous parameter calibration could be implemented further to optimize the model performances. These improvements would be greatly beneficial to the modeling applications in mountainous catchments.

- Analysis of the regional total precipitation reversely

  Limited by the data scarcity partially in high elevation areas, the average precipitation value is normally underestimated in mountainous region, which lead to the lack of a reasonable understanding about the total water contribution of precipitation for the area (Bárdossy and Pegram, 2009). Therefore, based on the water balance frame, it is possible to estimate the average total precipitation of a catchment reversely by taking use of the subbasin-scale precipitation input mechanism of SWAT model. This idea is based on the hypothesis that precipitation, snow processes and river discharge are the most important components of the water cycle during snow seasons in natural Alpine catchments, so that the magnitude of the other hydrological processes could be small enough in the water balance frame. During the model calibration, the snow and discharge information are used together to constrain the parameters including the precipitation lapse rate, namely the total precipitation of a subbasin is modified to reproduce the snow and discharge behaviors. Therefore,

once the snow and river discharge dynamics are reproduced properly, the estimated total precipitation could be considered representative for the subbasin during snow seasons. The uncertainty could be further validated under the water balance frame. This reverse reanalysis takes use of the SWAT model structure and could provide a relative reasonable estimation about the total snow season precipitation for a subbasin. This new precipitation dataset could be used to evaluate the grid based remote sensed products in Alpine catchments. The method could also be applied in the warm period if evapotranspiration data are available. Besides, it is not recommended for catchments affected by hydropower plant operation. Additional analyses should be performed to discuss about the rationality and feasibility of this idea in future study.

Effective utilization of available data based on the corresponding model structure is important for increasing the reliability of hydrological model results in alpine catchments. Based on this concept, this work expands new visions on the improvements of SWAT applications in alpine catchments by rationally using available data relevant to precipitation, snow, and hydropower plant operation. Moreover, new concepts and potential modifications are proposed in order to improve future SWAT applications and further benefit the hydrology researches in alpine catchments.

References

Abbaspour, K., 2015. SWAT-CUP 2012: SWAT calibration and uncertainty programs: A user manual. Department of Systems Analysis. Integrated Assessment and Modelling (SIAM), Eawag, Swiss Federal Institute of Aquatic Science and Technology, Duebendorf, Switzerland.

Abbaspour, K.C., Johnson, C.A., van Genuchten, M.T., 2004. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. Vadose Zone Journal, 3(4): 1340-1352.

Abbaspour, K.C. et al., 2015. A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. Journal of Hydrology, 524: 733-752. DOI:10.1016/j.jhydrol.2015.03.027

Abbaspour, K.C. et al., 2007. Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. Journal of Hydrology, 333(2-4): 413-430. DOI:10.1016/j.jhydrol.2006.09.014

Adler, R.F. et al., 2003. The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present). Journal of Hydrometeorology, 4(6): 1147-1167. DOI:10.1175/1525-7541(2003)004<1147:Tvgpcp>2.0.Co;2

Adler, S. et al., 2015. Il clima del Tirolo-Alto Adige-Bellunese, Zentralanstalt für Meteorologie und Geodynamik (ZAMG), Ripartizione Protezione antincendi e civile – Provincia Autonoma di Bolzano, Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto (ARPAV).

AghaKouchak, A., Mehran, A., 2013. Extended contingency table: Performance metrics for satellite observations and climate model simulations. Water Resources Research, 49(10): 7144-7149. DOI:10.1002/wrcr.20498

Ahl, R.S., Woods, S.W., Zuuring, H.R., 2008. Hydrologic Calibration and Validation of SWAT in a Snow-Dominated Rocky Mountain Watershed, Montana, USA. Journal of the American Water Resources Association, 44(6): 1411-1430. DOI:10.1111/j.1752-1688.2008.00233.x

Andreassian, V., Perrin, C., Michel, C., Usart-Sanchez, I., Lavabre, J., 2001. Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. Journal of Hydrology, 250(1-4): 206-223. DOI:10.1016/S0022-1694(01)00437-1

Anguita, D., Ghio, A., Ridella, S., Sterpi, D., 2009. K-Fold Cross Validation for Error Rate Estimate in Support Vector Machines, DMIN, pp. 291-297.

Arkin, P.A., Meisner, B.N., 1987. The Relationship between Large-Scale Convective Rainfall and Cold Cloud over the Western-Hemisphere during 1982-84. Monthly Weather Review, 115(1): 51-74. DOI:10.1175/1520-0493(1987)115<0051:Trblsc>2.0.Co;2

Arnold, J.G., Fohrer, N., 2005. SWAT2000: current capabilities and research opportunities in applied watershed modelling. Hydrological Processes, 19(3): 563-572. DOI:10.1002/hyp.5611

Arnold, J.G. et al., 2012a. Soil and Water Assessment Tool Input/Output Documentation: Version 2012. Texas Water Resources Institute.

Arnold, J.G. et al., 2012b. Swat: Model Use, Calibration, and Validation. Transactions of the ASABE, 55(4): 1491-1508.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large Area Hydrologic Modeling and Assessment Part I: Model Development. Journal of the American Water Resources Association, 34(1): 73-89. DOI:10.1111/j.1752-1688.1998.tb05961.x

Arnold, J.G. et al., 2015. Hydrological Processes and Model Representation: Impact of Soft Data on Calibration. Transactions of the ASABE, 58(6): 1637-1660.

Ashouri, H. et al., 2015. PERSIANN-CDR Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. Bulletin of the American Meteorological society, 96(1): 69-+. DOI:10.1175/Bams-D-13-00068.1

References

Awange, J.L. et al., 2016. Uncertainties in remotely sensed precipitation data over Africa. International Journal of Climatology, 36(1): 303-323. DOI:10.1002/joc.4346

Ayana, E.K., Worqlul, A.W., Steenhuis, T.S., 2015. Evaluation of stream water quality data generated from MODIS images in modeling total suspended solid emission to a freshwater lake. Science of The Total Environment, 523: 170-7. DOI:10.1016/j.scitotenv.2015.03.132

Babak, O., Deutsch, C.V., 2009. Statistical approach to inverse distance interpolation. Stochastic Environmental Research and Risk Assessment, 23(5): 543-553. DOI:10.1007/s00477-008-0226-6

Bailey, R., Rathjens, H., Bieger, K., Chaubey, I., Arnold, J., 2017. Swatmod-Prep: Graphical User Interface for Preparing Coupled Swat-Modflow Simulations. Journal of the American Water Resources Association, 53(2): 400-410. DOI:10.1111/1752-1688.12502

Bailey, R.T., Wible, T.C., Arabi, M., Records, R.M., Ditty, J., 2016. Assessing regional-scale spatio-temporal patterns of groundwater-surface water interactions using a coupled SWAT-MODFLOW model. Hydrological Processes, 30(23): 4420-4433. DOI:10.1002/hyp.10933

Balsamo, G. et al., 2015. ERA-Interim/Land: a global land surface reanalysis data set. Hydrology and Earth System Sciences, 19(1): 389-407. DOI:10.5194/hess-19-389-2015

Bárdossy, A., Das, T., 2008. Influence of rainfall observation network on model calibration and application. Hydrology and Earth System Sciences, 12(1): 77-89. DOI:10.5194/hess-12-77-2008

Bárdossy, A., Pegram, G., 2013. Interpolation of precipitation under topographic influence at different time scales. Water Resources Research, 49(8): 4545-4565. DOI:10.1002/wrcr.20307

Bárdossy, A., Pegram, G., 2009. Copula based multisite model for daily precipitation simulation. Hydrology and Earth System Sciences, 13(12): 2299-2314. DOI:10.5194/hess-13-2299-2009

Basso, S., Botter, G., 2012. Streamflow variability and optimal capacity of run-of-river hydropower plants. Water Resources Research, 48(10).

Bekele, E.G., Nicklow, J.W., 2007. Multi-objective automatic calibration of SWAT using NSGA-II. Journal of Hydrology, 341(3-4): 165-176. DOI:10.1016/j.jhydrol.2007.05.014

Beniston, M., 2012. Impacts of climatic change on water and associated economic activities in the Swiss Alps. Journal of Hydrology, 412: 291-296. DOI:10.1016/j.jhydrol.2010.06.046

Bennett, N.D. et al., 2013. Characterising performance of environmental models. Environmental Modelling and Software, 40: 1-20. DOI:10.1016/j.envsoft.2012.09.011

Bergstra, J., Bengio, Y., 2012. Random Search for Hyper-Parameter Optimization. Journal of Machine Learning Research, 13(Feb): 281-305.

Bertoldi, G. et al., 2010. Topographical and ecohydrological controls on land surface temperature in an alpine catchment. Ecohydrology, 3(2): 189-204. DOI:10.1002/eco.129

Beven, K., 2006. A manifesto for the equifinality thesis. Journal of Hydrology, 320(1-2): 18-36. DOI:10.1016/j.jhydrol.2005.07.007

Beven, K., Binley, A., 1992. The Future of Distributed Models - Model Calibration and Uncertainty Prediction. Hydrological Processes, 6(3): 279-298. DOI:10.1002/hyp.3360060305

Beven, K.J., 2011. Rainfall-runoff modelling: The primer, 2nd edition. John Wiley & Sons, 488 pp.

Bieger, K. et al., 2017. Introduction to SWAT+ , a completely restructured version of the soil and water assessment tool. Journal of the American Water Resources Association, 53(1): 115-130. DOI:10.1111/1752-1688.12482

Bloschl, G. et al., 2017. Changing climate shifts timing of European floods. Science, 357(6351): 588-590. DOI:10.1126/science.aan2506

Bohm, O., Jacobeit, J., Glaser, R., Wetzel, K.F., 2015. Flood sensitivity of the Bavarian Alpine Foreland since the late Middle Ages in the context of internal and external climate forcing factors. Hydrology and Earth System Sciences, 19(12): 4721-4734. DOI:10.5194/hess-19-4721-2015

164

Bonuma, N.B. et al., 2014. Simulating Landscape Sediment Transport Capacity by Using a Modified SWAT Model. Journal of Environmental Quality, 43(1): 55-66. DOI:10.2134/jeq2012.0217

Botter, G., Basso, S., Porporato, A., Rodriguez-Iturbe, I., Rinaldo, A., 2010. Natural streamflow regime alterations: Damming of the Piave river basin (Italy). Water Resources Research, 46(6). DOI:10.1029/2009wr008523

Brunetti, M., Maugeri, M., Monti, F., Nanni, T., 2006. Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series. International Journal of Climatology, 26(3): 345-381. DOI:10.1002/joc.1251

Bruno, M.C., Siviglia, A., Carolli, M., Maiolini, B., 2013. Multiple drift responses of benthic invertebrates to interacting hydropeaking and thermopeaking waves. Ecohydrology, 6(4): 511-522. DOI:10.1002/eco.1275

Callegari, M. et al., 2015. Seasonal River Discharge Forecasting Using Support Vector Regression: A Case Study in the Italian Alps. Water, 7(5): 2494-2515. DOI:10.3390/w7052494

Cao, W.Z., Bowden, W.B., Davie, T., Fenemor, A., 2006. Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. Hydrological Processes, 20(5): 1057-1073. DOI:10.1002/hyp.5933

Carey, S.K. et al., 2013. Use of color maps and wavelet coherence to discern seasonal and interannual climate influences on streamflow variability in northern catchments. Water Resources Research, 49(10): 6194-6207. DOI:10.1002/wrcr.20469

Chaplot, V., Saleh, A., Jaynes, D.B., 2005. Effect of the accuracy of spatial rainfall information on the modeling of water, sediment, and NO3–N loads at the watershed level. Journal of Hydrology, 312(1-4): 223-234. DOI:10.1016/j.jhydrol.2005.02.019

Chauvin, G.M. et al., 2011. Long-term water balance and conceptual model of a semi-arid mountainous catchment. Journal of Hydrology, 400(1-2): 133-143. DOI:10.1016/j.jhydrol.2011.01.031

Chen, C., Zhao, S., Duan, Z., Qin, Z., 2015. An Improved Spatial Downscaling Procedure for TRMM 3B43 Precipitation Product Using Geographically Weighted Regression. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(9): 4592-4604. DOI:10.1109/JSTARS.2015.2441734

Chen, M. et al., 2008. Assessing objective techniques for gauge-based analyses of global daily precipitation. Journal of Geophysical Research Atmospheres, 113(4). DOI:10.1029/2007JD009132

Chen, M., Xie, P., 2008. CPC unified gauge-based analysis of global daily precipitation. CPC Unified Gauge-based Analysis of Global Daily Precipiation.

Chen, S. et al., 2013. Evaluation of Spatial Errors of Precipitation Rates and Types from TRMM Spaceborne Radar over the Southern CONUS. Journal of Hydrometeorology, 14(6): 1884-1896. DOI:10.1175/Jhm-D-13-027.1

Chiogna, G. et al., 2016. A review of hydrological and chemical stressors in the Adige catchment and its ecological status. Science of The Total Environment, 540: 429-43. DOI:10.1016/j.scitotenv.2015.06.149

Cho, J., Bosch, D., Lowrance, R., Strickland, T., Vellidis, G., 2009. Effect of Spatial Distribution of Rainfall on Temporal and Spatial Uncertainty of Swat Output. Transactions of the ASABE, 52(5): 1545-1555.

Chu, J.G., Zhang, C., Wang, Y.L., Zhou, H.C., Shoemaker, C.A., 2012. A watershed rainfall data recovery approach with application to distributed hydrological models. Hydrological Processes, 26(13): 1937-1948. DOI:10.1002/hyp.8242

Costantini, E., L'Abate, G., Urbano, F., 2004. Soil regions of Italy. CRA-ISSDS, Firenze.

References

Coulibaly, P., Burn, D.H., 2004. Wavelet analysis of variability in annual Canadian streamflows. Water Resources Research, 40(3): 1-14. DOI:10.1029/2003wr002667

Cunderlik, J.M., Ouarda, T.B.M.J., 2009. Trends in the timing and magnitude of floods in Canada. Journal of Hydrology, 375(3-4): 471-480. DOI:10.1016/j.jhydrol.2009.06.050

Daubechies, I., 1990. The Wavelet Transform, Time-Frequency Localization and Signal Analysis. IEEE transactions on information theory, 36(5): 961-1005. DOI:10.1109/18.57199

Debele, B., Srinivasan, R., Gosain, A.K., 2010. Comparison of Process-Based and Temperature-Index Snowmelt Modeling in SWAT. Water Resources Management, 24(6): 1065-1088. DOI:10.1007/s11269-009-9486-2

Di Luzio, M., Johnson, G.L., Daly, C., Eischeid, J.K., Arnold, J.G., 2008. Constructing retrospective gridded daily precipitation and temperature datasets for the conterminous United States. Journal of Applied Meteorology and Climatology, 47(2): 475-497.

Dile, Y.T. et al., 2016. Assessing the implications of water harvesting intensification on upstream-downstream ecosystem services: A case study in the Lake Tana basin. Science of The Total Environment, 542(Pt A): 22-35. DOI:10.1016/j.scitotenv.2015.10.065

Duan, Z., Bastiaanssen, W.G.M., 2013. First results from Version 7 TRMM 3B43 precipitation product in combination with a new downscaling-calibration procedure. Remote Sensing of Environment, 131: 1-13. DOI:10.1016/j.rse.2012.12.002

Duan, Z., Bastiaanssen, W.G.M., Liu, J., 2012. Monthly and annual validation of TRMM Mulitisatellite Precipitation Analysis (TMPA) products in the Caspian Sea Region for the period 1999-2003, International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3696-3699. DOI:10.1109/IGARSS.2012.6350613

Duan, Z., Liu, J., Tuo, Y., Chiogna, G., Disse, M., 2016. Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. Science of The Total Environment, 573: 1536-1553. DOI:10.1016/j.scitotenv.2016.08.213

E., P., 2013. Impact of hydropeaking on fish and their habitat. EPFL-LCH.

Ebert, E.E., Janowiak, J.E., Kidd, C., 2007. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. Bulletin of the American Meteorological society, 88(1): 47-64. DOI:10.1175/BAMS-88-1-47

Egli, L., 2011. Spatial variability of seasonal snow cover at different scales in the Swiss Alps. Dissertation No. 19658, ETH ZURICH, Switzerland.

Egli, L., Jonas, T., 2009. Hysteretic dynamics of seasonal snow depth distribution in the Swiss Alps. Geophysical Research Letters, 36(2): L02501. DOI:10.1029/2008gl035545

Endrizzi, S., Bertoldi, G., Neteler, M., 2006. Snow cover patterns and evolution at basin scale: GEOtop model simulations and remote sensing observations, 63rd EASTERN SNOW CONFERENCE, Newark, Delaware USA

Farinotti, D., Usselmann, S., Huss, M., Bauder, A., Funk, M., 2012. Runoff evolution in the Swiss Alps: projections for selected high-alpine catchments based on ENSEMBLES scenarios. Hydrological Processes, 26(13): 1909-1924. DOI:10.1002/hyp.8276

Fatichi, S., Rimkus, S., Burlando, P., Bordoy, R., Molnar, P., 2015. High-resolution distributed analysis of climate and anthropogenic changes on the hydrology of an Alpine catchment. Journal of Hydrology, 525: 362-382. DOI:10.1016/j.jhydrol.2015.03.036

Ferraro, R.R., 1997. Special sensor microwave imager derived global rainfall estimates for climatological applications. Journal of Geophysical Research Atmospheres, 102(14): 16715-16735.

Finger, D., Heinrich, G., Gobiet, A., Bauder, A., 2012. Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century. Water Resources Research, 48(2).

FitzHugh, T.W., Vogel, R.M., 2011. The impact of dams on flood flows in the United States. River Research and Applications, 27(10): 1192-1215. DOI:10.1002/rra.1417

Francesconi, W., Srinivasan, R., Perez-Minana, E., Willcock, S.P., Quintero, M., 2016. Using the Soil and Water Assessment Tool (SWAT) to model ecosystem services: A systematic review. Journal of Hydrology, 535: 625-636. DOI:10.1016/j.jhydrol.2016.01.034

Fu, C.S., James, A.L., Yao, H.X., 2014. SWAT-CS: Revision and testing of SWAT for Canadian Shield catchments. Journal of Hydrology, 511: 719-735. DOI:10.1016/j.jhydrol.2014.02.023

Fu, C.S., James, A.L., Yao, H.X., 2015. Investigations of uncertainty in SWAT hydrologic simulations: a case study of a Canadian Shield catchment. Hydrological Processes, 29(18): 4000-4017. DOI:10.1002/hyp.10477

Funk, C. et al., 2015. The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes. Scientific data, 2. DOI:10.1038/sdata.2015.66

Gabriel, M., Knightes, C., Dennis, R., Cooter, E., 2014. Potential Impact of Clean Air Act Regulations on Nitrogen Fate and Transport in the Neuse River Basin: a Modeling Investigation Using CMAQ and SWAT. Environmental Modeling and Assessment, 19(6): 451-465. DOI:10.1007/s10666-014-9410-x

Galván, L., Olías, M., Izquierdo, T., Cerón, J.C., Fernández de Villarán, R., 2014. Rainfall estimation in SWAT: An alternative method to simulate orographic precipitation. Journal of Hydrology, 509: 257-265. DOI:10.1016/j.jhydrol.2013.11.044

Garcia, A., Jorde, K., Habit, E., Caamano, D., Parra, O., 2011. Downstream Environmental Effects of Dam Operations: Changes in Habitat Quality for Native Fish Species. River Research and Applications, 27(3): 312-327. DOI:10.1002/rra.1358

Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., 2007. The soil and water assessment tool: Historical development, applications, and future research directions. Transactions of the ASABE, 50(4): 1211-1250.

Gaudard, L., Gabbi, J., Bauder, A., Romerio, F., 2016. Long-term Uncertainty of Hydropower Revenue Due to Climate Change and Electricity Prices. Water Resources Management, 30(4): 1325-1343. DOI:10.1007/s11269-015-1216-3

Gaudard, L. et al., 2014. Climate change impacts on hydropower in the Swiss and Italian Alps. Science of The Total Environment, 493: 1211-21. DOI:10.1016/j.scitotenv.2013.10.012

Golmohammadi, G., Rudra, R., Dickinson, T., Goel, P., Veliz, M., 2017. Predicting the temporal variation of flow contributing areas using SWAT. Journal of Hydrology, 547: 375-386. DOI:10.1016/j.jhydrol.2017.02.008

Grinsted, A., Moore, J.C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear processes in geophysics, 11(5-6): 561-566. DOI:10.5194/npg-11-561-2004

Grunewald, T., Buhler, Y., Lehning, M., 2014. Elevation dependency of mountain snow depth. Cryosphere, 8(6): 2381-2394. DOI:10.5194/tc-8-2381-2014

Grusson, Y. et al., 2015. Assessing the capability of the SWAT model to simulate snow, snow melt and streamflow dynamics over an alpine watershed. Journal of Hydrology, 531: 574-588. DOI:10.1016/j.jhydrol.2015.10.070

Guan, K. et al., 2011. Spatiotemporal scaling of hydrological and agrochemical export dynamics in a tile-drained Midwestern watershed. Water Resources Research, 47(10).

Guo, H., Hu, Q., Jiang, T., 2008. Annual and seasonal streamflow responses to climate and land-cover changes in the Poyang Lake basin, China. Journal of Hydrology, 355(1-4): 106-122. DOI:10.1016/j.jhydrol.2008.03.020

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1999. Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. Journal of Hydrologic Engineering, 4(2): 135-143. DOI:10.1061/(Asce)1084-0699(1999)4:2(135)

Gurtz, J. et al., 2003. A comparative study in modelling runoff and its components in two mountainous catchments. Hydrological Processes, 17(2): 297-311. DOI:10.1002/hyp.1125

Guse, B. et al., 2016. On characterizing the temporal dominance patterns of model parameters and processes. Hydrological Processes, 30(13): 2255-2270. DOI:10.1002/hyp.10764

Hanzer, F., Helfricht, K., Marke, T., Strasser, U., 2016. Multilevel spatiotemporal validation of snow/ice mass balance and runoff modeling in glacierized catchments. Cryosphere, 10(4): 1859-1881. DOI:10.5194/tc-10-1859-2016

Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H., 2014. Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. International Journal of Climatology, 34(3): 623-642. DOI:10.1002/joc.3711

Hastik, R. et al., 2015. Renewable energies and ecosystem service impacts. Renewable & Sustainable Energy Reviews, 48: 608-623. DOI:10.1016/j.rser.2015.04.004

Hauer, C., Holzapfel, P., Leitner, P., Graf, W., 2017a. Longitudinal assessment of hydropeaking impacts on various scales for an improved process understanding and the design of mitigation measures. Science of The Total Environment, 575: 1503-1514. DOI:10.1016/j.scitotenv.2016.10.031

Hauer, C., Siviglia, A., Zolezzi, G., 2017b. Hydropeaking in regulated rivers–From process understanding to design of mitigation measures. Elsevier.

Hiemstra, C.A., Liston, G.E., Reiners, W.A., 2002. Snow Redistribution by Wind and Interactions with Vegetation at Upper Treeline in the Medicine Bow Mountains, Wyoming, U.S.A. Arctic, Antarctic, and Alpine Research, 34(3): 262-273. DOI:10.2307/1552483

Hock, R., 2003. Temperature index melt modelling in mountain areas. Journal of Hydrology, 282(1-4): 104-115. DOI:10.1016/S0022-1694(03)00257-9

Hofstra, N., Haylock, M., New, M., Jones, P., Frei, C., 2008. Comparison of six methods for the interpolation of daily, European climate data. Journal of Geophysical Research Atmospheres, 113(21). DOI:10.1029/2008JD010100

Holzapfel, P., Leitner, P., Habersack, H., Graf, W., Hauer, C., 2017. Evaluation of hydropeaking impacts on the food web in alpine streams based on modelling of fish-and macroinvertebrate habitats. Science of The Total Environment, 575: 1489-1502.

Horton, P., Schaefli, B., Mezghani, A., Hingray, B., Musy, A., 2006. Assessment of climate-change impacts on alpine discharge regimes with climate model uncertainty. Hydrological Processes, 20(10): 2091-2109. DOI:10.1002/hyp.6197

Hsu, C.-W., Chang, C.-C., Lin, C.-J., 2003. A practical guide to support vector classification.

Hsu, K.L., Gao, X.G., Sorooshian, S., Gupta, H.V., 1997. Precipitation estimation from remotely sensed information using artificial neural networks. Journal of Applied Meteorology, 36(9): 1176-1190. DOI:10.1175/1520-0450(1997)036<1176:Pefrsi>2.0.Co;2

Hu, Q. et al., 2014. Multi-scale evaluation of six high-resolution satellite monthly rainfall estimates over a humid region in China with dense rain gauges. International Journal of Remote Sensing, 35(4): 1272-1294. DOI:10.1080/01431161.2013.876118

Huffman, G.J. et al., 2007. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. Journal of Hydrometeorology, 8(1): 38-55. DOI:10.1175/Jhm560.1

Huffman, G.J., Bolvin, D.T., 2013. TRMM and other data precipitation data set documentation. TRMM and Other Data Precipitation Data Set Documentation.

Iida, T., Kajihara, A., Okubo, H., Okajima, K., 2012. Effect of seasonal snow cover on suspended sediment runoff in a mountainous catchment. Journal of Hydrology, 428: 116-128. DOI:10.1016/j.jhydrol.2012.01.029

Isotta, F.A. et al., 2014. The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. International Journal of Climatology, 34(5): 1657-1675. DOI:10.1002/joc.3794

Javanmard, S., Yatagai, A., Nodzu, M.I., BodaghJamali, J., Kawamoto, H., 2010. Comparing high-resolution gridded precipitation data with satellite rainfall estimates of TRMM_3B42 over Iran. Advances in Geosciences, 25: 119-125. DOI:10.5194/adgeo-25-119-2010

Ji, X., Chen, Y., 2012. Characterizing spatial patterns of precipitation based on corrected TRMM 3B 43 data over the mid Tianshan Mountains of China. Journal of Mountain Science, 9(5): 628-645. DOI:10.1007/s11629-012-2283-z

Jonas, T., Marty, C., Magnusson, J., 2009. Estimating the snow water equivalent from snow depth measurements in the Swiss Alps. Journal of Hydrology, 378(1–2): 161-167. DOI:10.1016/j.jhydrol.2009.09.021

Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. Journal of Hydrometeorology, 5(3): 487-503. DOI:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

Kalnay, E. et al., 1996. The NCEP/NCAR 40-year reanalysis project. Bulletin of the American Meteorological society, 77(3): 437-471. DOI:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Karamouz, M., Ahmadi, A., Moridi, A., 2009. Probabilistic reservoir operation using Bayesian stochastic model and support vector machine. Advances in water resources, 32(11): 1588-1600. DOI:10.1016/j.advwatres.2009.08.003

Katsanos, D., Retalis, A., Michaelides, S., 2016. Validation of a high-resolution precipitation database (CHIRPS) over Cyprus for a 30-year period. Atmospheric Research, 169: 459-464. DOI:10.1016/j.atmosres.2015.05.015

Khalil, A., McKee, M., Kemblowski, M., Asefa, T., 2005. Sparse Bayesian learning machine for real-time management of reservoir releases. Water Resources Research, 41(11).

Khandu, Awange, J.L., Forootan, E., 2016. An evaluation of high-resolution gridded precipitation products over Bhutan (1998-2012). International Journal of Climatology, 36(3): 1067-1087. DOI:10.1002/joc.4402

Kidd, C., Huffman, G., 2011. Global precipitation measurement. Meteorological Applications, 18(3): 334-353. DOI:10.1002/met.284

Kummerow, C. et al., 2001. The evolution of the Goddard profiling algorithm (GPROF) for rainfall estimation from passive microwave sensors. Journal of Applied Meteorology, 40(11): 1801-1820. DOI:10.1175/1520-0450(2001)040<1801:TEOTGP>2.0.CO;2

Labat, D., 2005. Recent advances in wavelet analyses: Part 1. A review of concepts. Journal of Hydrology, 314(1): 275-288.

Labat, D., Ababou, R., Mangin, A., 2000. Rainfall–runoff relations for karstic springs. Part II: continuous wavelet and discrete orthogonal multiresolution analyses. Journal of Hydrology, 238(3): 149-178.

Labat, D. et al., 2004. Wavelet analysis of Amazon hydrological regime variability. Geophysical Research Letters, 31(2). DOI:10.1029/2003gl018741

Laghari, A.N., Vanham, D., Rauch, W., 2012. To what extent does climate change result in a shift in Alpine hydrology? A case study in the Austrian Alps. Hydrological Sciences Journal, 57(1): 103-117.

Lau, K., Weng, H., 1999. Interannual, decadal–interdecadal, and global warming signals in sea surface temperature during 1955–97. Journal of Climate, 12(5): 1257-1267.

Le Moine, N., Hendrickx, F., Gailhard, J., Garçon, R., Gottardi, F., 2015. Hydrologically Aided Interpolation of Daily Precipitation and Temperature Fields in a Mesoscale Alpine Catchment. Journal of Hydrometeorology, 16(6): 2595-2618.

Lettenmaier, D.P. et al., 2015. Inroads of remote sensing into hydrologic science during the WRR era. Water Resources Research, 51(9): 7309-7342. DOI:10.1002/2015WR017616

Levesque, E., Anctil, F., van Griensven, A., Beauchamp, N., 2008. Evaluation of streamflow simulation by SWAT model for two small watersheds under snowmelt and rainfall. Hydrological Sciences Journal, 53(5): 961-976. DOI:10.1623/hysj.53.5.961

Li, X.H., Zhang, Q., Xu, C.Y., 2012. Suitability of the TRMM satellite rainfalls in driving a distributed hydrological model for water balance computations in Xinjiang catchment, Poyang lake basin. Journal of Hydrology, 426: 28-38. DOI:10.1016/j.jhydrol.2012.01.013

Liu, J.Z., Duan, Z., Jiang, J.C., Zhu, A.X., 2015. Evaluation of Three Satellite Precipitation Products TRMM 3B42, CMORPH, and PERSIANN over a Subtropical Watershed in China. Advances in Meteorology, 2015. DOI:10.1155/2015/151239

Liu, R.M., Xu, F., Zhang, P.P., Yu, W.W., Men, C., 2016. Identifying non-point source critical source areas based on multi-factors at a basin scale with SWAT. Journal of Hydrology, 533: 379-388. DOI:10.1016/j.jhydrol.2015.12.024

Lo Conti, F., Hsu, K.L., Noto, L.V., Sorooshian, S., 2014. Evaluation and comparison of satellite precipitation estimates with reference to a local area in the Mediterranean Sea. Atmospheric Research, 138: 189-204. DOI:10.1016/j.atmosres.2013.11.011

López-Moreno, J.I., Nogués-Bravo, D., 2006. Interpolating local snow depth data: an evaluation of methods. Hydrological Processes, 20(10): 2217-2232. DOI:10.1002/hyp.6199

Lu, S., Kayastha, N., Thodsen, H., van Griensven, A., Andersen, H.E., 2014. Multiobjective Calibration for Comparing Channel Sediment Routing Models in the Soil and Water Assessment Tool. Journal of Environmental Quality, 43(1): 110-120. DOI:10.2134/jeq2011.0364

Luo, Y., Arnold, J., Liu, S.Y., Wang, X.Y., Chen, X., 2013. Inclusion of glacier processes for distributed hydrological modeling at basin scale with application to a watershed in Tianshan Mountains, northwest China. Journal of Hydrology, 477: 72-85. DOI:10.1016/j.jhydrol.2012.11.005

Ly, S., Charles, C., Degré, A., 2011. Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium. Hydrology and Earth System Sciences, 15(7): 2259-2274. DOI:10.5194/hess-15-2259-2011

Ly, S., Charles, C., Degré, A., 2013. Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale. A review. Biotechnologie, Agronomie, Société et Environnement, 17(2): 392.

Maggioni, V., Meyers, P.C., Robinson, M.D., 2016. A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM) era. Journal of Hydrometeorology, 17(4): 1101-1117. DOI:10.1175/JHM-D-15-0190.1

Majone, B., Villa, F., Deidda, R., Bellin, A., 2016. Impact of climate change and water use policies on hydropower potential in the south-eastern Alpine region. Science of The Total Environment, 543, Part B: 965-980. DOI:10.1016/j.scitotenv.2015.05.009

Malago, A. et al., 2016. Regional scale hydrologic modeling of a karst-dominant geomorphology: The case study of the Island of Crete. Journal of Hydrology, 540: 64-81. DOI:10.1016/j.jhydrol.2016.05.061

Malagò, A., Pagliero, L., Bouraoui, F., Franchini, M., 2015. Comparing calibrated parameter sets of the SWAT model for the Scandinavian and Iberian peninsulas. Hydrological Sciences Journal, 60(5): 949-967. DOI:10.1080/02626667.2014.978332

Marcolini, G., Bellin, A., Chiogna, G., 2017a. Performance of the Standard Normal Homogeneity Test for the homogenization of mean seasonal snow depth time series. International Journal of Climatology, 37: 1267-1277. DOI:10.1002/joc.4977

Marcolini, G., Bellin, A., Disse, M., Chiogna, G., 2017b. Variability in snow depth time series in the Adige catchment. Journal of Hydrology: Regional Studies, 13: 240-254. DOI:10.1016/j.ejrh.2017.08.007

Maringanti, C., Chaubey, I., Arabi, M., Engel, B., 2011. Application of a Multi-Objective Optimization Method to Provide Least Cost Alternatives for NPS Pollution Control. Environmental Management, 48(3): 448-461. DOI:10.1007/s00267-011-9696-2

Masih, I., Maskey, S., Uhlenbrook, S., Smakhtin, V., 2011. Assessing the impact of areal precipitation input on streamflow simulations using the SWAT model. Journal of the American Water Resources Association, 47(1): 179-195. DOI:10.1111/j.1752-1688.2010.00502.x

Massarutto, A., Pontoni, F., 2015. Rent seizing and environmental concerns: A parametric valuation of the Italian hydropower sector. Energy Policy, 78: 31-40. DOI:10.1016/j.enpol.2014.12.016

McCreight, J.L., Small, E.E., 2014. Modeling bulk density and snow water equivalent using daily snow depth observations. Cryosphere, 8(2): 521-536. DOI:10.5194/tc-8-521-2014

McNairn, H. et al., 2015. The Soil Moisture Active Passive Validation Experiment 2012 (SMAPVEX12): Prelaunch Calibration and Validation of the SMAP Soil Moisture Algorithms. IEEE Transactions on Geoscience and Remote Sensing, 53(5): 2784-2801. DOI:10.1109/Tgrs.2014.2364913

Mei, Y., Anagnostou, E.N., Nikolopoulos, E.I., Borga, M., 2014. Error analysis of satellite precipitation products in mountainous basins. Journal of Hydrometeorology, 15(5): 1778-1793. DOI:10.1175/JHM-D-13-0194.1

Mei, Y.W., Nikolopoulos, E.I., Anagnostou, E.N., Borga, M., 2016a. Evaluating Satellite Precipitation Error Propagation in Runoff Simulations of Mountainous Basins. Journal of Hydrometeorology, 17(5): 1407-1423. DOI:10.1175/Jhm-D-15-0081.1

Mei, Y.W., Nikolopoulos, E.I., Anagnostou, E.N., Zoccatelli, D., Borga, M., 2016b. Error Analysis of Satellite Precipitation-Driven Modeling of Flood Events in Complex Alpine Terrain. Remote Sensing, 8(4): 293. DOI:10.3390/rs8040293

Meißl, G. et al., 2017. Climate change effects on hydrological system conditions influencing generation of storm runoff in small Alpine catchments. Hydrological Processes, 31(6): 1314-1330. DOI:10.1002/hyp.11104

Meng, X., Yu, D., Liu, Z., 2015. Energy balance-based SWAT model to simulate the mountain snowmelt and runoff — taking the application in Juntanghu watershed (China) as an example. Journal of Mountain Science, 12(2): 368-381. DOI:10.1007/s11629-014-3081-6

Meyer, S., Blaschek, M., Duttmann, R., Ludwig, R., 2016. Improved hydrological model parametrization for climate change impact assessment under data scarcity — The potential of field monitoring techniques and geostatistics. Science of The Total Environment, 543: 906-923. DOI:10.1016/j.scitotenv.2015.07.116

Monteiro, J.A.F., Strauch, M., Srinivasan, R., Abbaspour, K., Gucker, B., 2016. Accuracy of grid precipitation data for Brazil: application in river discharge modelling of the Tocantins catchment. Hydrological Processes, 30(9): 1419-1430. DOI:10.1002/hyp.10708

Moran-Tejeda, E., Lorenzo-Lacruz, J., Lopez-Moreno, J.I., Rahman, K., Beniston, M., 2014. Streamflow timing of mountain rivers in Spain: Recent changes and future projections. Journal of Hydrology, 517: 1114-1127. DOI:10.1016/j.jhydrol.2014.06.053

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

References

Moussa, R., Chahinian, N., Bocquillon, C., 2007. Distributed hydrological modelling of a Mediterranean mountainous catchment - Model construction and multi-site validation. Journal of Hydrology, 337(1-2): 35-51.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. Journal of Hydrology, 10(3): 282-290. DOI:10.1016/0022-1694(70)90255-6

Navarro-Ortega, A. et al., 2015. Managing the effects of multiple stressors on aquatic ecosystems under water scarcity. The GLOBAQUA project. Science of The Total Environment, 503-504: 3-9. DOI:10.1016/j.scitotenv.2014.06.081

Negrón Juárez, R.I., Li, W., Fernandes, K., de Oliveira Cardoso, A., 2009. Comparison of precipitation data sets over the tropical South American and African continents. Journal of Hydrometeorology, 10(1): 289-299. DOI:10.1175/2008JHM1023.1

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and water assessment tool theoretical documentation version 2009, Texas Water Resources Institute.

Nepal, S., Krause, P., Flugel, W.A., Fink, M., Fischer, C., 2014. Understanding the hydrological system dynamics of a glaciated alpine catchment in the Himalayan region using the J2000 hydrological model. Hydrological Processes, 28(3): 1329-1344. DOI:10.1002/hyp.9627

Nerantzaki, S.D. et al., 2015. Modeling suspended sediment transport and assessing the impacts of climate change in a karstic Mediterranean watershed. Science of The Total Environment, 538: 288-297. DOI:10.1016/j.scitotenv.2015.07.092

Nikolopoulos, E.I., Anagnostou, E.N., Borga, M., 2013. Using high-resolution satellite rainfall products to simulate a major flash flood event in northern Italy. Journal of Hydrometeorology, 14(1): 171-185. DOI:10.1175/JHM-D-12-09.1

Nourani, V., Baghanam, A.H., Adamowski, J., Kisi, O., 2014. Applications of hybrid wavelet–Artificial Intelligence models in hydrology: A review. Journal of Hydrology, 514: 358-377.

Omani, N., Srinivasan, R., Smith, P.K., Karthikeyan, R., 2017. Glacier mass balance simulation using SWAT distributed snow algorithm. Hydrological Sciences Journal, 62(4): 546-560. DOI:10.1080/02626667.2016.1162907

Panziera, L., Giovannini, L., Laiti, L., Zardi, D., 2015. The relation between circulation types and regional Alpine climate. Part I: synoptic climatology of Trentino. International Journal of Climatology, 35(15): 4655-4672. DOI:10.1002/joc.4314

Penna, D., Engel, M., Bertoldi, G., Comiti, F., 2017a. Towards a tracer-based conceptualization of meltwater dynamics and streamflow response in a glacierized catchment. Hydrology and Earth System Sciences, 21(1): 23-41. DOI:10.5194/hess-21-23-2017

Penna, D. et al., 2014. Tracer-based analysis of spatial and temporal variations of water sources in a glacierized catchment. Hydrology and Earth System Sciences, 18(12): 5271-5288. DOI:10.5194/hess-18-5271-2014

Penna, D. et al., 2017b. Response time and water origin in a steep nested catchment in the Italian Dolomites. Hydrological Processes, 31(4): 768-782. DOI:10.1002/hyp.11050

Pfannerstill, M. et al., 2017. How to Constrain Multi-Objective Calibrations of the SWAT Model Using Water Balance Components. Journal of the American Water Resources Association, 53(3): 532-546.

Pistocchi, A., 2016. Simple estimation of snow density in an Alpine region. Journal of Hydrology: Regional Studies, 6: 82-89. DOI:10.1016/j.ejrh.2016.03.004

Poff, N.L., Olden, J.D., Merritt, D.M., Pepin, D.M., 2007. Homogenization of regional river dynamics by dams and global biodiversity implications. Proceedings of the National Academy of Sciences, 104(14): 5732-5737.

Poulenard, J. et al., 2012. Tracing sediment sources during floods using Diffuse Reflectance Infrared Fourier Transform Spectrometry (DRIFTS): A case study in a highly erosive mountainous

catchment (Southern French Alps). Journal of Hydrology, 414: 452-462. DOI:10.1016/j.jhydrol.2011.11.022

Pradhanang, S.M. et al., 2011. Application of SWAT model to assess snowpack development and streamflow in the Cannonsville watershed, New York, USA. Hydrological Processes, 25(21): 3268-3277. DOI:10.1002/hyp.8171

Prakash, S., Mitra, A.K., Rajagopal, E.N., Pai, D.S., 2015. Assessment of TRMM-based TMPA-3B42 and GSMaP precipitation products over India for the peak southwest monsoon season. Int J Climatol.

Price, K., Purucker, S.T., Kraemer, S.R., Babendreier, J.E., Knightes, C.D., 2014. Comparison of radar and gauge precipitation data in watershed models across varying spatial and temporal scales. Hydrological Processes, 28(9): 3505-3520. DOI:10.1002/hyp.9890

Qi, J.Y. et al., 2016. A new soil-temperature module for SWAT application in regions with seasonal snow cover. Journal of Hydrology, 538: 863-877. DOI:10.1016/j.jhydrol.2016.05.003

Raghavendra, N.S., Deka, P.C., 2014. Support vector machine applications in the field of hydrology: A review. Applied Soft Computing, 19: 372-386. DOI:10.1016/j.asoc.2014.02.002

Rahman, K. et al., 2014. Streamflow response to regional climate model output in the mountainous watershed: a case study from the Swiss Alps. Environmental earth sciences, 72(11): 4357-4369. DOI:10.1007/s12665-014-3336-0

Rahman, K. et al., 2013. Streamflow Modeling in a Highly Managed Mountainous Glacier Watershed Using SWAT: The Upper Rhone River Watershed Case in Switzerland. Water Resources Management, 27(2): 323-339. DOI:10.1007/s11269-012-0188-9

Rajib, M.A., Merwade, V., Yu, Z.Q., 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

Rasouli, K., Hsieh, W.W., Cannon, A.J., 2012. Daily streamflow forecasting by machine learning methods with weather and climate inputs. Journal of Hydrology, 414: 284-293. DOI:10.1016/j.jhydrol.2011.10.039

Rathinasamy, M. et al., 2014. Wavelet-based multiscale performance analysis: An approach to assess and improve hydrological models. Water Resources Research, 50(12): 9721-9737.

Rathjens, H., Oppelt, N., Bosch, D.D., Arnold, J.G., Volk, M., 2015. Development of a grid-based version of the SWAT landscape model. Hydrological Processes, 29(6): 900-914. DOI:10.1002/hyp.10197

Rostamian, R. et al., 2008. Application of a SWAT model for estimating runoff and sediment in two mountainous basins in central Iran. Hydrological Sciences Journal, 53(5): 977-988. DOI:10.1623/hysj.53.5.977

Russell, S.O., Campbell, P.F., 1996. Reservoir operating rules with fuzzy programming. Journal of Water Resources Planning and Management-Asce, 122(3): 165-170. DOI:10.1061/(Asce)0733-9496(1996)122:3(165)

Schaefli, B., Maraun, D., Holschneider, M., 2007. What drives high flow events in the Swiss Alps? Recent developments in wavelet spectral analysis and their application to hydrology. Advances in water resources, 30(12): 2511-2525. DOI:10.1016/j.advwatres.2007.06.004

Schmalz, B. et al., 2015. Impacts of land use changes on hydrological components and macroinvertebrate distributions in the Poyang lake area. Ecohydrology, 8(6): 1119-1136. DOI:10.1002/eco.1569

Schneider, U. et al., 2015. GPCC Full Data Reanalysis Version 7.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. GPCC full data reanalysis version 7.0 at 0.5°: Monthly land-surface precipitation from rain-gauges built on GTS-based and historic data.

References

Schneider, U. et al., 2014. GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theoretical and Applied Climatology, 115(1-2): 15-40. DOI:10.1007/s00704-013-0860-x

Schölkopf, B., Burges, C.J., Smola, A.J., 1999. Advances in kernel methods: support vector learning. MIT press.

Schölkopf, B., Smola, A.J., 2002. Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press.

Schuol, J., Abbaspour, K., 2006. Calibration and uncertainty issues of a hydrological model (SWAT) applied to West Africa. Advances in Geosciences, 9.

Sheffield, J., Goteti, G., Wood, E.F., 2006. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. Journal of Climate, 19(13): 3088-3111. DOI:10.1175/JCLI3790.1

Shen, Y., Xiong, A., Wang, Y., Xie, P., 2010. Performance of high-resolution satellite precipitation products over China. Journal of Geophysical Research Atmospheres, 115(D2). DOI:10.1029/2009JD012097

Shen, Z.Y., Chen, L., Liao, Q., Liu, R.M., Hong, Q., 2012. Impact of spatial rainfall variability on hydrology and nonpoint source pollution modeling. Journal of Hydrology, 472: 205-215. DOI:10.1016/j.jhydrol.2012.09.019

Shinohara, Y., Kumagai, T., Otsuki, K., Kume, A., Wada, N., 2009. Impact of climate change on runoff from a mid-latitude mountainous catchment in central Japan. Hydrological Processes, 23(10): 1418-1429. DOI:10.1002/hyp.7264

Shope, C.L., Maharjan, G.R., 2015. Modeling Spatiotemporal Precipitation: Effects of Density, Interpolation, and Land Use Distribution. Advances in Meteorology, 2015: 16. DOI:10.1155/2015/174196

Shope, C.L. et al., 2014. Using the SWAT model to improve process descriptions and define hydrologic partitioning in South Korea. Hydrology and Earth System Sciences, 18(2): 539-557. DOI:10.5194/hess-18-539-2014

Shrestha, R., Takara, K., Tachikawa, Y., Jha, R.N., 2004. Water resources assessment in a poorly gauged mountainous catchment using a geographical information system and remote sensing. Hydrological Processes, 18(16): 3061-3079. DOI:10.1002/hyp.5749

Smola, A.J., Scholkopf, B., 2004. A tutorial on support vector regression. Statistics and computing, 14(3): 199-222. DOI:10.1023/B:Stco.0000035301.49549.88

Solander, K.C., Reager, J.T., Famiglietti, J.S., 2016. How well will the Surface Water and Ocean Topography (SWOT) mission observe global reservoirs? Water Resources Research, 52(3): 2123-2140. DOI:10.1002/2015wr017952

Solomatine, D.P., Shrestha, D.L., 2009. A novel method to estimate model uncertainty using machine learning techniques. Water Resources Research, 45(12). DOI:10.1029/2008wr006839

Song, X.F., Duan, Z., Kono, Y., Wang, M.Y., 2011. Integration of remotely sensed C factor into SWAT for modelling sediment yield. Hydrological Processes, 25(22): 3387-3398. DOI:10.1002/hyp.8066

Strasser, U., 2008. Modelling of the mountain snow cover in the Berchtesgaden National Park. Technical Report Berchtesgaden, 1(55).

Strauch, M. et al., 2012. Using precipitation data ensemble for uncertainty analysis in SWAT streamflow simulation. Journal of Hydrology, 414: 413-424. DOI:10.1016/j.jhydrol.2011.11.014

Sturm, M. et al., 2010. Estimating Snow Water Equivalent Using Snow Depth Data and Climate Classes. Journal of Hydrometeorology, 11(6): 1380-1394. DOI:10.1175/2010jhm1202.1

Tan, M.L., Ibrahim, A., Duan, Z., Cracknell, A.P., Chaplot, V., 2015. Evaluation of Six High-Resolution Satellite and Ground-Based Precipitation Products over Malaysia. Remote Sensing, 7(2): 1504-1528. DOI:10.3390/rs70201504

Tapiador, F.J. et al., 2012. Global precipitation measurement: Methods, datasets and applications. Atmospheric Research, 104-105: 70-97. DOI:10.1016/j.atmosres.2011.10.021

Tian, Y., Peters-Lidard, C.D., Adler, R.F., Kubota, T., Ushio, T., 2010. Evaluation of GSMaP precipitation estimates over the contiguous United States. Journal of Hydrometeorology, 11(2): 566-574. DOI:10.1175/2009JHM1190.1

Tian, Y., Peters-Lidard, C.D., Choudhury, B.J., Garcia, M., 2007. Multitemporal analysis of TRMM-based satellite precipitation products for land data assimilation applications. Journal of Hydrometeorology, 8(6): 1165-1183. DOI:10.1175/2007JHM859.1

Tian, Y.D., Liu, Y.Q., Arsenault, K.R., Behrangi, A., 2014. A new approach to satellite-based estimation of precipitation over snow cover. International Journal of Remote Sensing, 35(13): 4940-4951. DOI:10.1080/01431161.2014.930208

Ticlavilca, A.M., Mckee, M., 2011. Multivariate Bayesian Regression Approach to Forecast Releases from a System of Multiple Reservoirs. Water Resources Management, 25(2): 523-543. DOI:10.1007/s11269-010-9712-y

Tobin, C., Nicotina, L., Parlange, M.B., Berne, A., Rinaldo, A., 2011. Improved interpolation of meteorological forcings for hydrologic applications in a Swiss Alpine region. Journal of Hydrology, 401(1-2): 77-89. DOI:10.1016/j.jhydrol.2011.02.010

Tobin, K.J., Bennett, M.E., 2009. Using Swat to Model Streamflow in Two River Basins with Ground and Satellite Precipitation Data. Journal of the American Water Resources Association, 45(1): 253-271. DOI:10.1111/j.1752-1688.2008.00276.x

Tonolla, D., Bruder, A., Schweizer, S., 2017. Evaluation of mitigation measures to reduce hydropeaking impacts on river ecosystems – a case study from the Swiss Alps. Science of The Total Environment, 574: 594–604.

Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bulletin of the American Meteorological society, 79(1): 61-78. DOI:10.1175/1520-0477(1998)079<0061: Apgtwa>2.0.Co;2

Torres-Rua, A.F., Ticlavilca, A.M., Walker, W.R., McKee, M., 2012. Machine learning approaches for error correction of hydraulic simulation models for canal flow schemes. Journal of Irrigation and Drainage Engineering, 138(11): 999-1010.

Troin, M., Caya, D., 2014. Evaluating the SWAT's snow hydrology over a Northern Quebec watershed. Hydrological Processes, 28(4): 1858-1873. DOI:10.1002/hyp.9730

Tuo, Y., Duan, Z., Disse, M., Chiogna, G., 2016. Evaluation of precipitation input for SWAT modeling in Alpine catchment: A case study in the Adige river basin (Italy). Science of The Total Environment, 573: 66-82. DOI:10.1016/j.scitotenv.2016.08.034

Tuo, Y., Marcolini, G., Disse, M., Chiogna, G., 2018a. Calibration of snow parameters in SWAT: comparison of three approaches in the Upper Adige River basin (Italy). Hydrological Sciences Journal: 1-22. DOI:10.1080/02626667.2018.1439172

Tuo, Y., Marcolini, G., Disse, M., Chiogna, G., 2018b. A multi-objective approach to improve SWAT model calibration in alpine catchments. Journal of Hydrology, 559: 347-360. DOI:10.1016/j.jhydrol.2018.02.055

Tuppad, P., Douglas-Mankin, K.R., Koelliker, J.K., Hutchinson, J.M.S., 2010. Swat Discharge Response to Spatial Rainfall Variability in a Kansas Watershed. Transactions of the ASABE, 53(1): 65-74.

Ushio, T. et al., 2009. A kalman filter approach to the global satellite mapping of precipitation (GSMaP) from combined passive microwave and infrared radiometric data. Journal of the Meteorological Society of Japan, 87 A: 137-151. DOI:10.2151/jmsj.87A.137

van der Heijden, S., Haberlandt, U., 2010. Influence of spatial interpolation methods for climate variables on the simulation of discharge and nitrate fate with SWAT. Advances in Geosciences, 27: 91-98. DOI:10.5194/adgeo-27-91-2010

Van Loon, A.F., Laaha, G., 2015. Hydrological drought severity explained by climate and catchment characteristics. Journal of Hydrology, 526: 3-14. DOI:10.1016/j.jhydrol.2014.10.059

Verbunt, M. et al., 2003. The hydrological role of snow and glaciers in alpine river basins and their distributed modeling. Journal of Hydrology, 282(1-4): 36-55. DOI:10.1016/S0022-1694(03)00251-8

Vigiak, O. et al., 2018. Uncertainty of modelled flow regime for flow-ecological assessment in Southern Europe. Science of The Total Environment, 615(Supplement C): 1028-1047. DOI:10.1016/j.scitotenv.2017.09.295

Viviroli, D., Durr, H.H., Messerli, B., Meybeck, M., Weingartner, R., 2007. Mountains of the world, water towers for humanity: Typology, mapping, and global significance. Water Resources Research, 43(7): W07447. DOI:10.1029/2006wr005653

Viviroli, D., Weingartner, R., 2004. The hydrological significance of mountains: from regional to global scale. Hydrology and Earth System Sciences, 8(6): 1016-1029.

Viviroli, D., Weingartner, R., Messerli, B., 2003. Assessing the hydrological significance of the world's mountains. Mountain Research and Development, 23(1): 32-40. DOI:10.1659/0276-4741(2003)023[0032:Athsot]2.0.Co;2

Volk, M., Liersch, S., Schmidt, G., 2009. Towards the implementation of the European Water Framework Directive? Land Use Policy, 26(3): 580-588. DOI:10.1016/j.landusepol.2008.08.005

Vu, M.T., Raghavan, S.V., Liong, S.Y., 2012. SWAT use of gridded observations for simulating runoff – a Vietnam river basin study. Hydrology and Earth System Sciences, 16(8): 2801-2811. DOI:10.5194/hess-16-2801-2012

Wagner, B., Hauer, C., Schoder, A., Habersack, H., 2015. A review of hydropower in Austria: Past, present and future development. Renewable and Sustainable Energy Reviews, 50: 304-314.

Wagner, P.D., Fiener, P., Wilken, F., Kumar, S., Schneider, K., 2012. Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions. Journal of Hydrology, 464: 388-400. DOI:10.1016/j.jhydrol.2012.07.026

Walter, M.T. et al., 2005. Process-based snowmelt modeling: does it require more input data than temperature-index modeling? Journal of Hydrology, 300(1-4): 65-75. DOI:10.1016/j.jhydrol.2004.05.002

Wang, X., Melesse, A.M., 2005. Evaluation of the swat model's snowmelt hydrology in a northwestern Minnesota watershed. Transactions of the ASAE, 48(4): 1359-1376.

Warscher, M. et al., 2013. Performance of complex snow cover descriptions in a distributed hydrological model system: A case study for the high Alpine terrain of the Berchtesgaden Alps. Water Resources Research, 49(5): 2619-2637. DOI:10.1002/wrcr.20219

White, K.L., Chaubey, I., 2005. Sensitivity analysis, calibration, and validations for a multisite and multivariable SWAT model. Journal of the American Water Resources Association, 41(5): 1077-1089. DOI:10.1111/j.1752-1688.2005.tb03786.x

Wilheit, T., 1994. Algorithms for the retrieval of rainfall from passive microwave measurements. Remote Sensing Reviews, 11(1-4): 163-194.

Woznicki, S.A. et al., 2016. Ecohydrological modeling for large-scale environmental impact assessment. Science of The Total Environment, 543, Part A: 274-286. DOI:10.1016/j.scitotenv.2015.11.044

Xie, P., Joyce, R., Wu, S., 2013. A 15-year high-resolution gauge–satellite merged analysis of precipitation. 27th Conf. on Hydrology, Austin, TX, Amer. Meteor. Soc.

Xie, P., Xiong, A.Y., 2011. A conceptual model for constructing high-resolution gauge-satellite merged precipitation analyses. Journal of Geophysical Research Atmospheres, 116(21). DOI:10.1029/2011JD016118

Xu, H.M. et al., 2010. Hydrological modeling of River Xiangxi using SWAT2005: A comparison of model parameterizations using station and gridded meteorological observations. Quaternary International, 226(1-2): 54-59. DOI:10.1016/j.quaint.2009.11.037

Yang, J., Reichert, P., Abbaspour, K.C., Xia, J., Yang, H., 2008. Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. Journal of Hydrology, 358(1-2): 1-23. DOI:10.1016/j.jhydrol.2008.05.012

Yang, X. et al., 2016. Spatiotemporal patterns and source attribution of nitrogen load in a river basin with complex pollution sources. Water Research, 94: 187-199. DOI:10.1016/j.watres.2016.02.040

Yang, Y., Onishi, T., Hiramatsu, K., 2014. Improving the performance of temperature index snowmelt model of SWAT by using MODIS land surface temperature data. The Scientific World Journal, 2014: 1-20. DOI:10.1155/2014/823424

Yang, Y., Onishi, T., Hiramatsu, K., 2015. Impacts of different spatial temperature interpolation methods on snowmelt simulations. Hydrological Research Letters, 9(2): 27-34. DOI:10.3178/hrl.9.27

Yong, B. et al., 2015. Global view of real-time TRMM multisatellite precipitation analysis: Implications for its successor global precipitation measurement mission. Bulletin of the American Meteorological society, 96(2): 283-296. DOI:10.1175/BAMS-D-14-00017.1

Yong, B. et al., 2010. Hydrologic evaluation of Multisatellite Precipitation Analysis standard precipitation products in basins beyond its inclined latitude band: A case study in Laohahe basin, China. Water Resources Research, 46(7). DOI:10.1029/2009WR008965

Zampieri, M., Scoccimarro, E., Gualdi, S., 2013. Atlantic influence on spring snowfall over the Alps in the past 150 years. Environmental Research Letters, 8(3): 034026. DOI:10.1088/1748-9326/8/3/034026

Zanotti, F., Endrizzi, S., Bertoldi, G., Rigon, R., 2004. The GEOTOP snow module. Hydrological Processes, 18(18): 3667-3679. DOI:10.1002/hyp.5794

Zappa, M., Kan, C., 2007. Extreme heat and runoff extremes in the Swiss Alps. Natural Hazards and Earth System Sciences, 7(3): 375-389. DOI:10.5194/nhess-7-375-2007

Zhang, X.S., Srinivasan, R., 2009. GIS-Based Spatial Precipitation Estimation: A Comparison of Geostatistical Approaches(1). Journal of the American Water Resources Association, 45(4): 894-906. DOI:10.1111/j.1752-1688.2009.00335.x

Zhang, X.S., Srinivasan, R., Debele, B., Hao, F.H., 2008. Runoff simulation of the headwaters of the Yellow River using the SWAT model with three snowmelt algorithms. Journal of the American Water Resources Association, 44(1): 48-61. DOI:10.1111/j.1752-1688.2007.00137.x

Zhang, Y.Y., Xia, J., Chen, J.F., Zhang, M.H., 2011. Water quantity and quality optimization modeling of dams operation based on SWAT in Wenyu River Catchment, China. Environmental Monitoring and Assessment, 173(1-4): 409-430.

Zierl, B., Bugmann, H., 2005. Global change impacts on hydrological processes in Alpine catchments. Water Resources Research, 41(2). DOI:10.1029/2004wr003447

Zolezzi, G., Bellin, A., Bruno, M.C., Maiolini, B., Siviglia, A., 2009. Assessing hydrological alterations at multiple temporal scales: Adige River, Italy. Water Resources Research, 45. DOI:10.1029/2008wr007266

Supplementary material

A. Supplementary material of Chapter 3
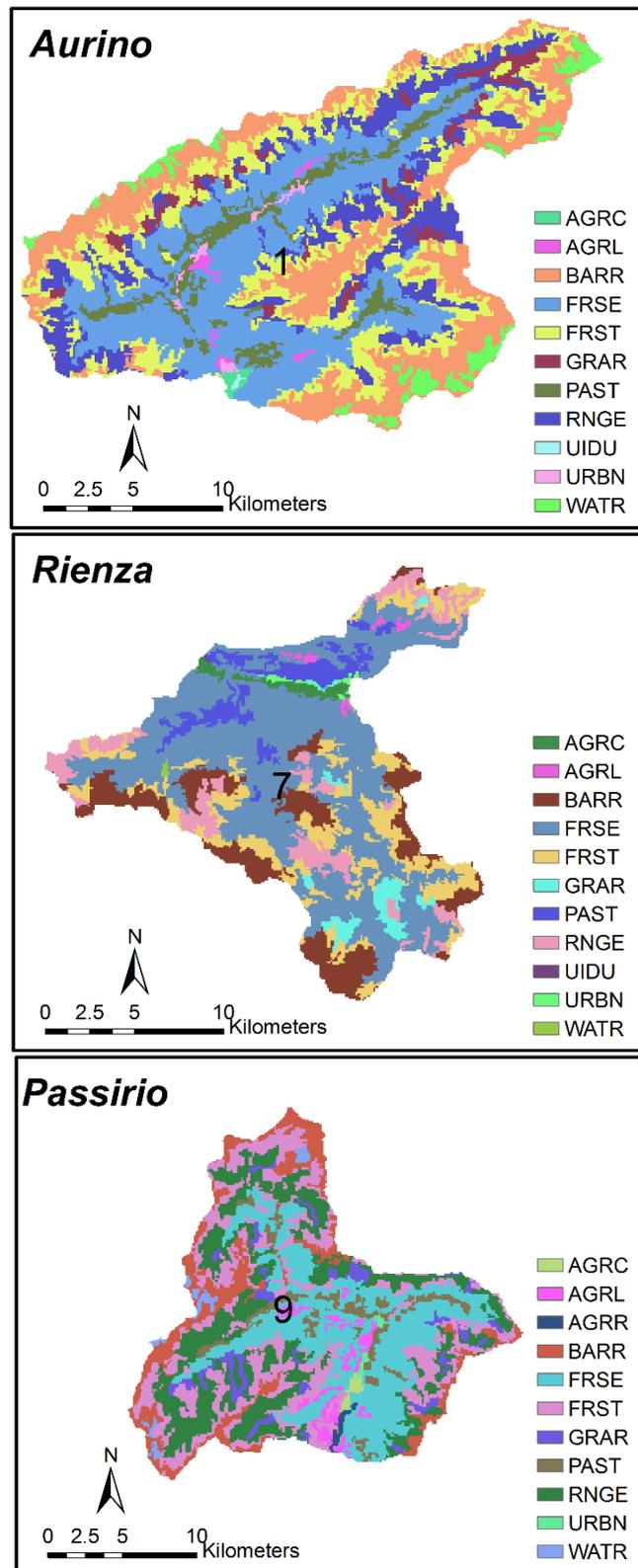
Figure S3.1 Land use map (SWAT land use classification) of the subbasin Aurino, Rienza and Passirio.
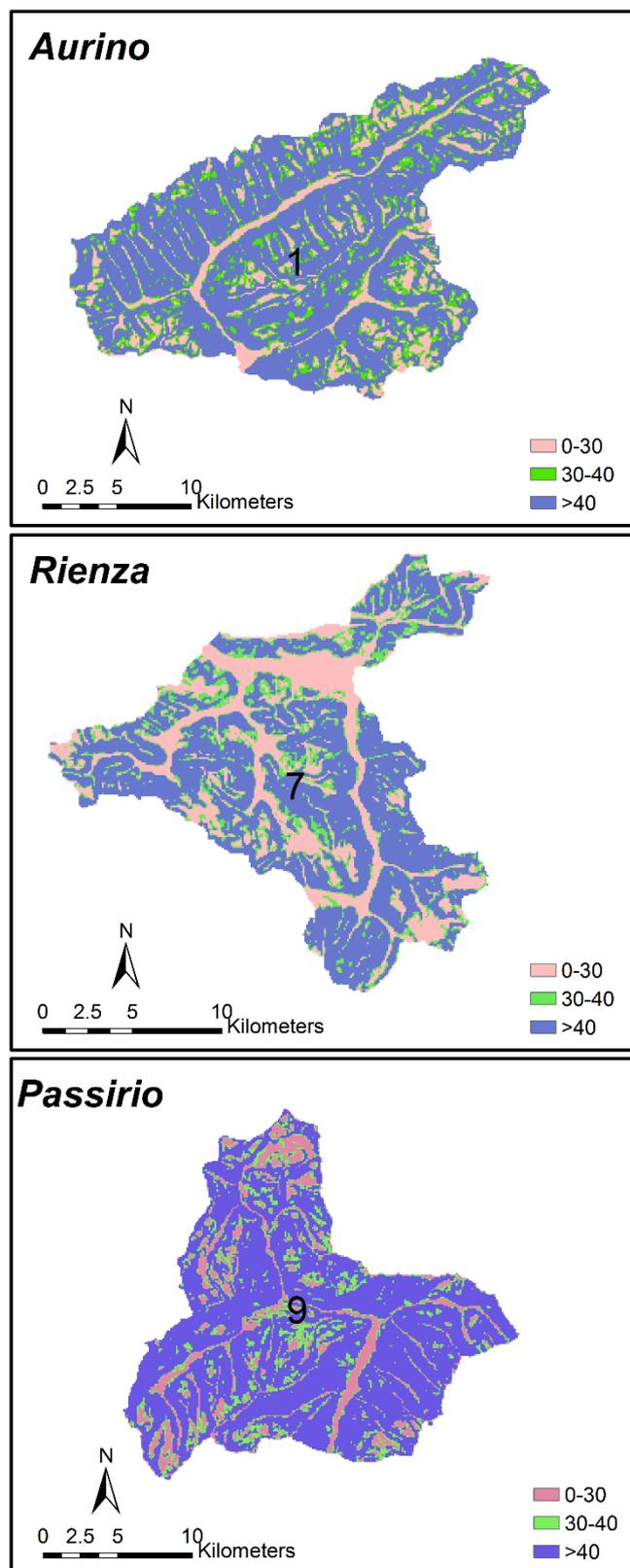
Figure S3.2 Slope distribution of the subbasin Aurino, Rienza and Passirio.
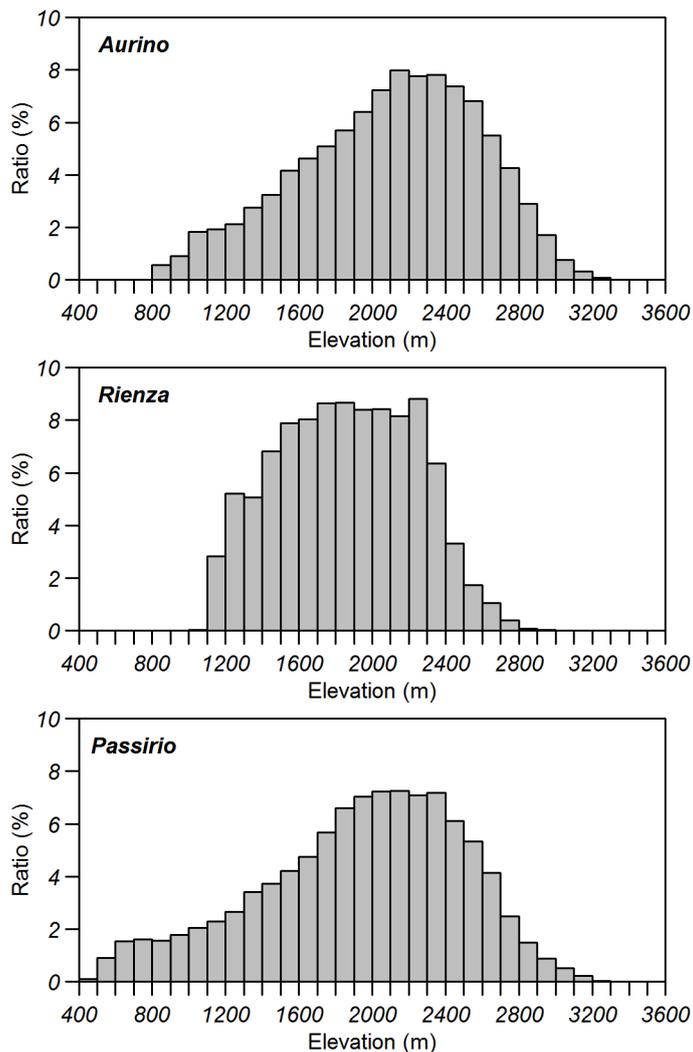
Figure S3.3 Elevation distribution of the subbasin Aurino, Rienza and Passirio.

Figure S3.4 Calibrated parameter distributions for the four precipitation inputs in Aurino catchment within the initial parameter range (y-axis domain): green bars show the final parameter ranges; red points represent the values of "best parameter". "Δ" means an absolute increase, without "Δ" means absolute value.
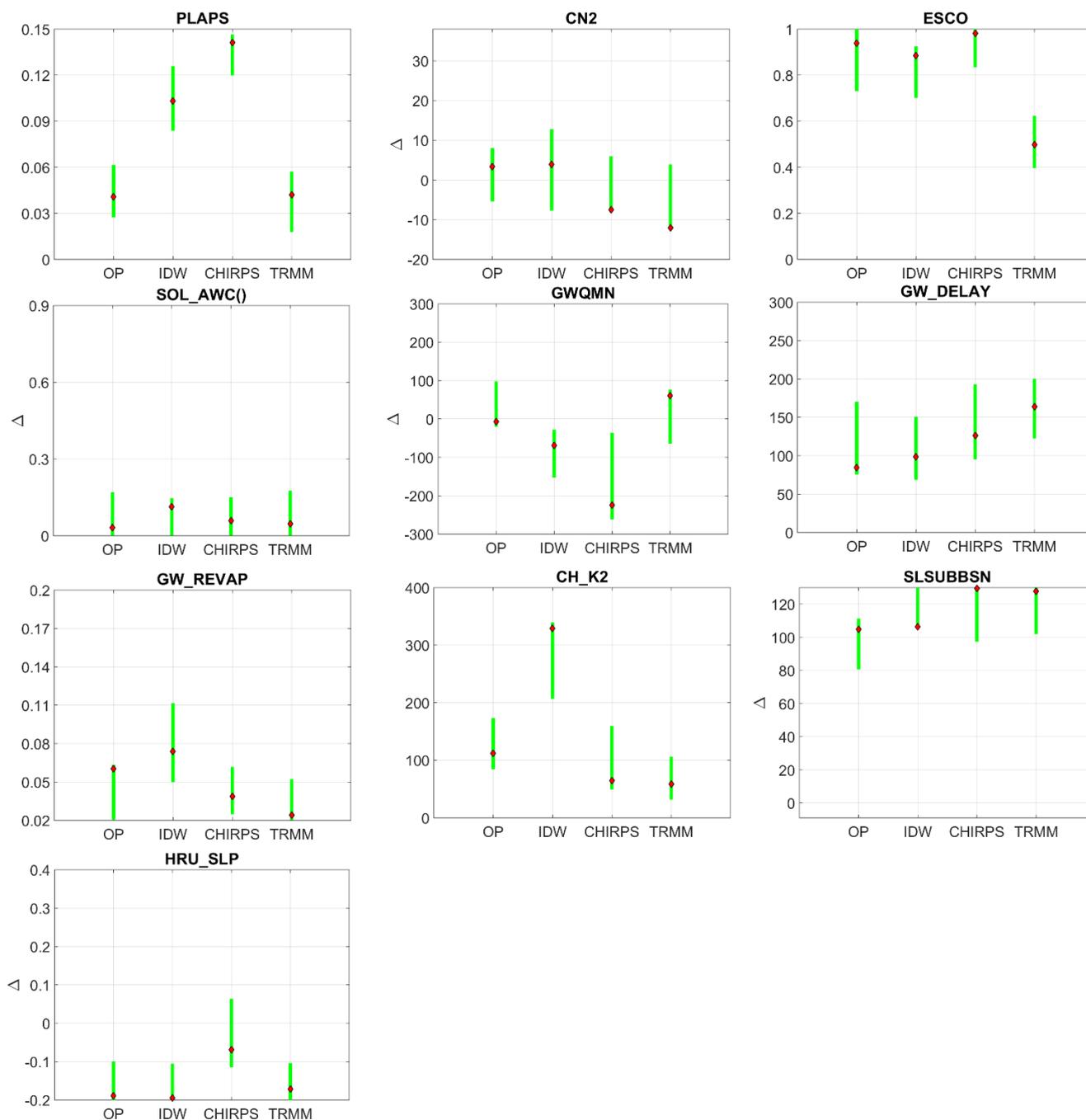
Figure S3.5 Calibrated parameter distributions for the four precipitation inputs in Rienza catchment within the initial parameter range (y-axis domain): green bars show the final parameter ranges; red points represent the values of "best parameter". "Δ" means an absolute increase, without "Δ" means absolute value.
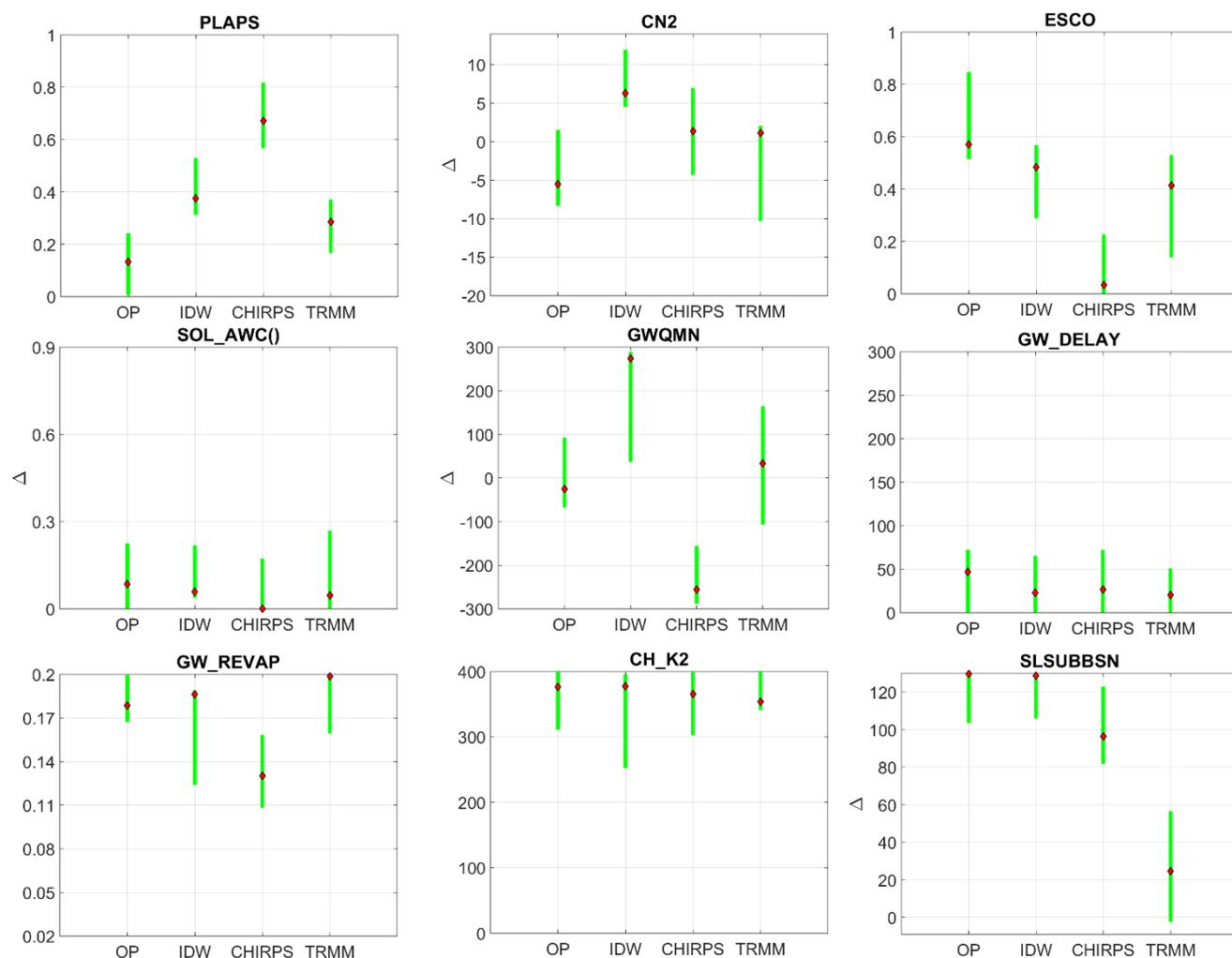
Figure S3.6 Calibrated parameter distributions for the four precipitation inputs in Passirio catchment within the initial parameter range (y-axis domain): green bars show the final parameter ranges; red points represent the values of "best parameter". "Δ" means an absolute increase, without "Δ" means absolute value.

Table S3.1 Fixed hydrological parameters (mean value of best parameters of the best simulations using different datasets at each subbasin) to get the results in Table 3.6.

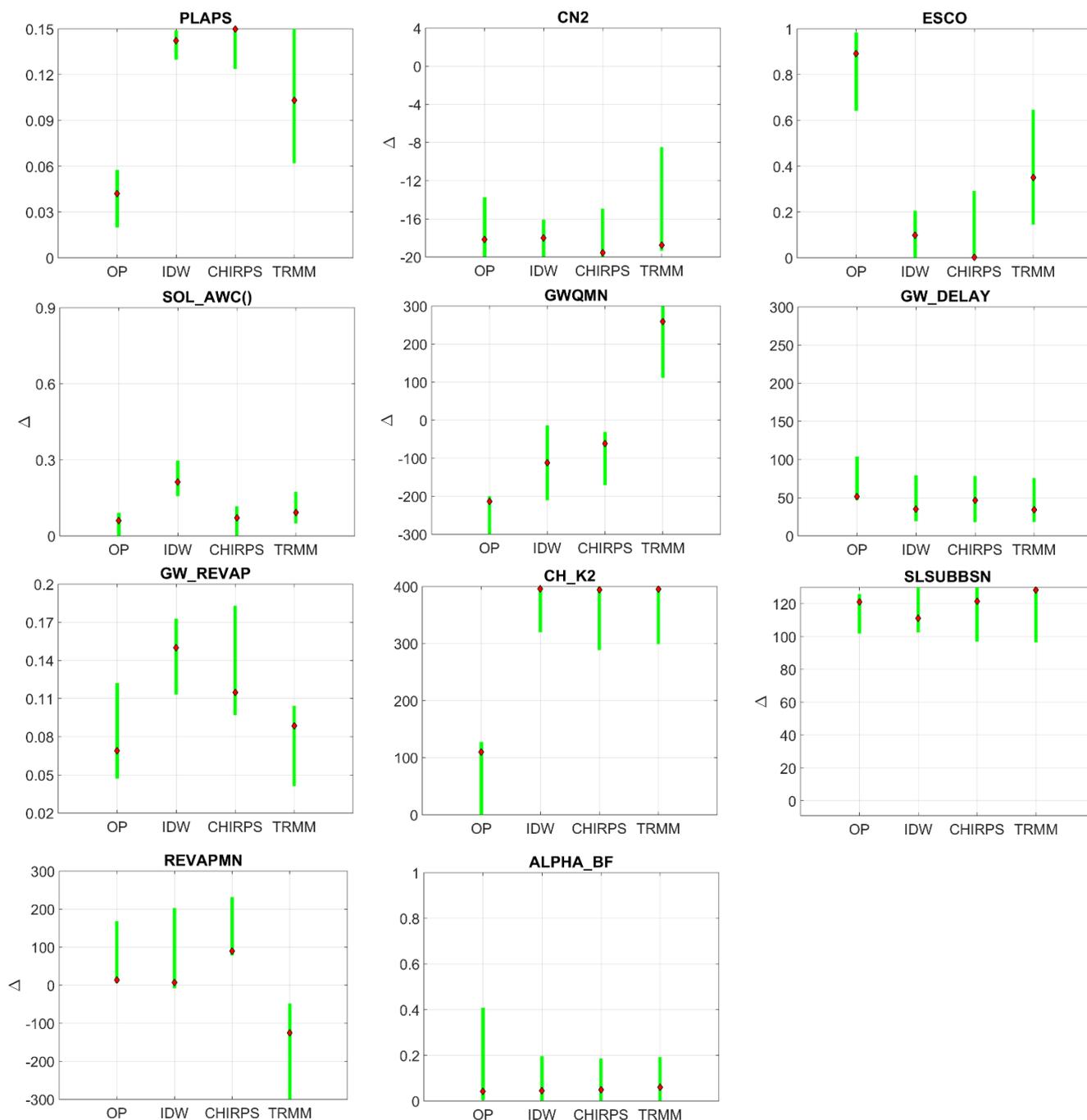| Parameter	value
Aurino	
a__SOL_AWC().sol	0.109
a__CN2.mgt	-18.615
v__ESCO.hru	0.336
a__GWQMN.gw	-31.898
a__REVAPMN.gw	-3.940
v__GW_REVAP.gw	0.106
v__GW_DELAY.gw	41.769
v__ALPHA_BF.gw	0.049
v__CH_K2.rte	323.733
a__SLSUBBSN.hru	120.483
Rienza	
a__SOL_AWC().sol	0.063
a__CN2.mgt	-3.069
v__ESCO.hru	0.826
a__SLSUBBSN.hru	117.181
a__HRU_SLP.hru	-0.156
a__GWQMN.gw	-60.072
v__GW_REVAP.gw	0.049
v__GW_DELAY.gw	118.302
v__CH_K2.rte	141.183
Passirio	
a__SOL_AWC().sol	0.048
a__CN2.mgt	0.824
v__ESCO.hru	0.376
a__SLSUBBSN.hru	94.885
a__GWQMN.gw	6.799
v__GW_REVAP.gw	0.173
v__GW_DELAY.gw	29.165
v__CH_K2.rte	368.354

B. Supplementary material of Chapter 4
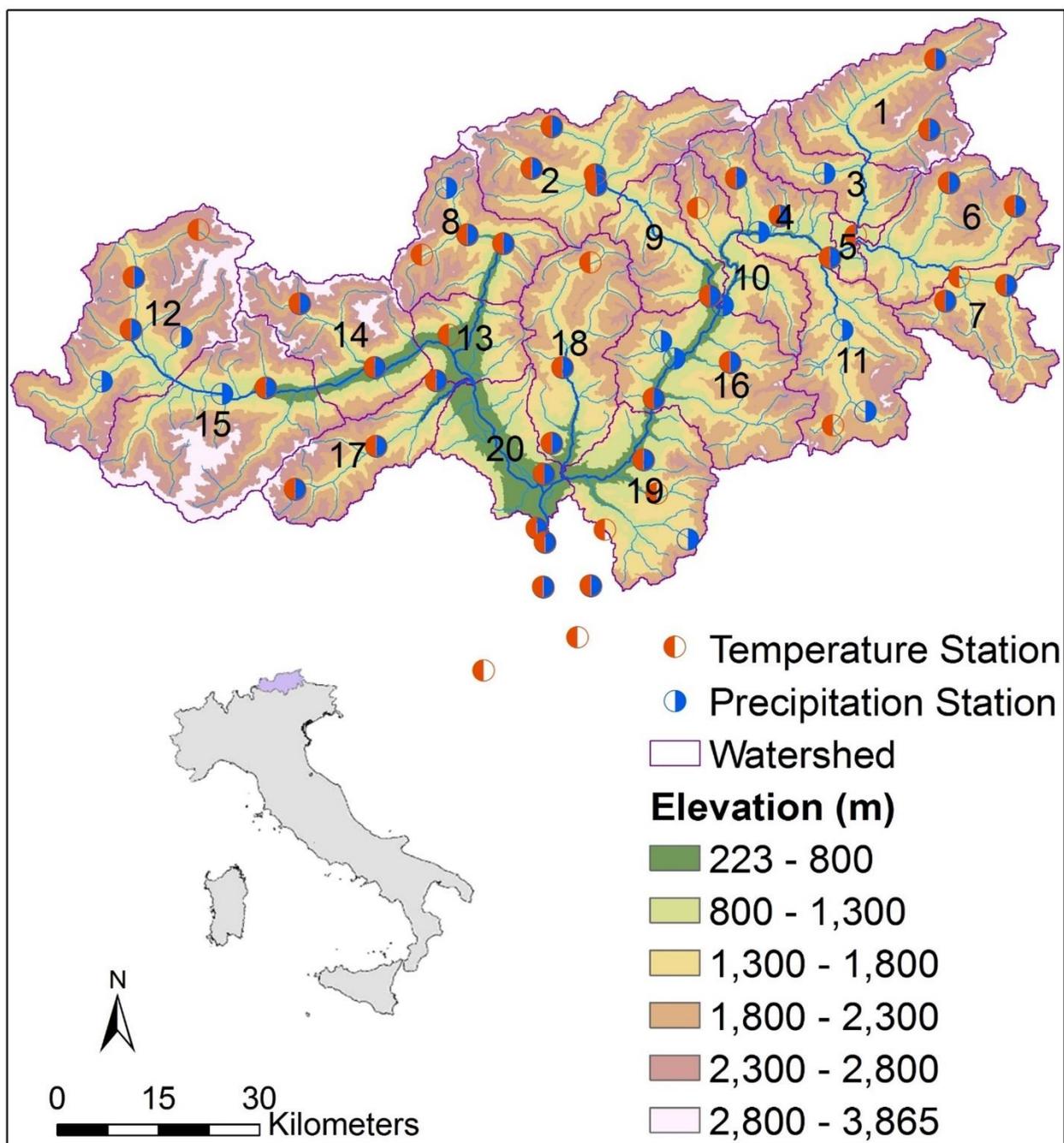
Figure S4.1 Location of temperature and precipitation stations in the Upper Adige River basin.
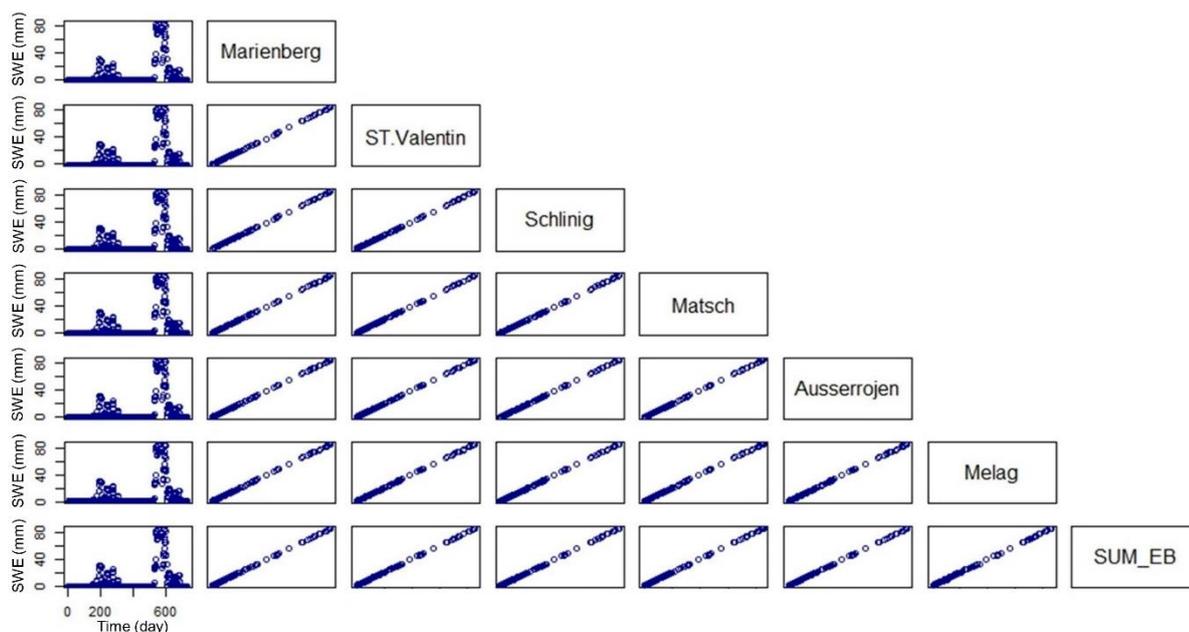
Figure S4.2 Left column: Simulated SWE of each HRU where the snow station is located (name of the station in the box on the right) and the weighting sum of the SWE value of the elevation bands (SUM_EB). The perfect correlation between all pairs of stations and SUM_EB is shown in the other plots.

Figure S4.3 Comparison between the values of sub-basin-scale snow parameters (histogram) and those of the basin-scale snow parameters (blue lines).

Table S4.1 Information about the snow stations of each sub-basin.

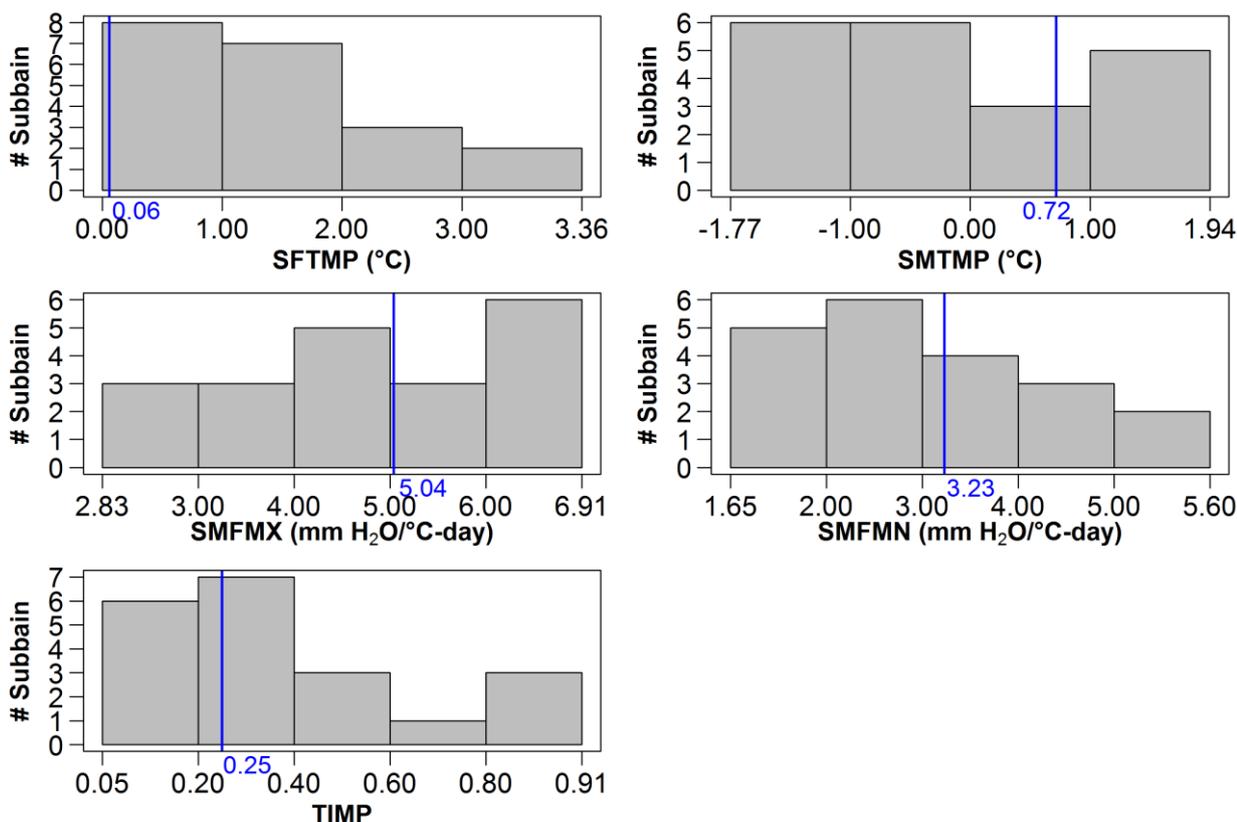| Sub-basin	Station	LAT	LONG	ELEVATION
1	PRETTAU	47.03658867	12.09912722	1450
2	WEHR_IN_PFITSCH	46.92499219	11.5246721	1365
	RIDNAUN	46.90908292	11.30677415	1350
	INNERRATSCHINGS	46.8660261	11.31719654	1260
3	MUEHLEN	46.90025669	11.94743948	851
4	TERENTEN	46.83445756	11.78557637	1349
	PFUNDERS	46.88718005	11.7029428	1159
5	BRUNECK	46.8042795	11.93148052	821
6	ST.MAGDALENA_IN_GSIES	46.83526049	12.24271916	1398
	STAUSEE_OLANG	46.75973668	12.0536567	1057
	MITTERTAL	46.86072197	12.10089661	1236
7	TOBLACH	46.72997843	12.21915906	1219
8	PLATT	46.82249293	11.17808117	1147
	ST.MARTIN_IN_PASSEIER	46.78268262	11.227909	588
9	FRANZENSFESTE	46.77830419	11.63224641	725
10	LUESEN	46.74503125	11.76389069	981
11	STERN	46.5836304	11.90707588	1390
	CAMPILL	46.63934499	11.8608778	1396
12	AUSSERROJEN	46.8095884	10.48411317	1833
	MELAG	46.83743215	10.65658211	1915
	ST.VALENTIN	46.77587237	10.52890216	1499
	MARIENBERG	46.70568911	10.5212676	1310
	SCHLINIG	46.70413104	10.47392043	1690
	MATSCH	46.69431625	10.61752601	1570
13	SCHENNA	46.68959619	11.19083111	680
	MERAN	46.68805798	11.13656021	330
	OBERMAIS	46.6578197	11.18323354	334
14	VORDERKASER	46.73510025	10.92669711	1705
15	ALLITZ	46.65020601	10.69607321	1754
	STAUSEE_ZUFRITT	46.50906285	10.72507183	1851
16	FLITZ	46.62357402	11.66281009	1350
	LATZFONS	46.67240083	11.54613904	1190
	ST.ULRICH	46.57404443	11.67296415	1180
17	WEISSBRUNN	46.48684723	10.83181103	1900
	STAUSEE_ZOGGL	46.54167092	10.98978931	1142
	STAUSEE_ST.PANKRAZ	46.56489465	11.05568468	810
	PAWIGL	46.62776162	11.10933866	1400
18	REINSWALD	46.70104927	11.41450083	1365
19	VOELS	46.51340578	11.50595194	840
20	MOELTEN	46.58577237	11.25168816	1133
Nearby	ALTREI	46.277346	11.367751	1209
	SEXTEN	46.703054	12.349964	1310
	DEUTSCHNOFEN	46.419133	11.412535	1415
	RADEIN	46.346545	11.396822	1562

Description of the elevation band method

As shown in Eqs. (S4.1) - (S4.4), the method modifies the precipitation and temperature inputs for each sub-basin by weighting the elevation difference between the rain gauge/climate station and the bands.

$$R_{band} = R_{day} + \left(EL_{band} - EL_{gauge}\right) \cdot \frac{plaps}{days_{pcp,yr} \cdot 1000}, R_{day} > 0.01 \qquad (S4.1)$$

$$T_{band} = T + \left(EL_{band} - EL_{gauge}\right) \cdot \frac{tlaps}{1000} \qquad (S4.2)$$

$$R_{or} = \sum_{bnd=1}^{b} R_{band} \cdot fr_{bnd} \qquad (S4.3)$$

$$T_{or} = \sum_{bnd=1}^{b} T_{band} \cdot fr_{bnd} \qquad (S4.4)$$

where R_{band} is the precipitation in the elevation band (mm), R_{day} is the precipitation recorded at the rain gauge (mm), EL_{band} is the mean elevation at the elevation band (m), EL_{gauge} is the elevation at the recording gauge (m), plaps is the precipitation lapse rate (mm/km) and $days_{pcp,yr}$ is the average number of days of precipitation in the sub-basin in a year, fr_{bnd} is the fraction of the sub-basin area within the elevation band, and b is the total number of elevation bands in the sub-basin. T_{band} is the temperature (maximum, minimum and average temperatures) in the elevation band (°C), T is the temperature (maximum, minimum and average temperatures) recorded at the climate station (°C), tlpas is the temperature lapse rate (°C/km). R_{or} and T_{or} are the daily precipitation (mm) and temperature (°C), respectively, adjusted for orographic effects, and these temperature and precipitation values are assigned to all the HRUs in one sub-basin.

Test of SWE outputs of different HRUs in the same subbasin (output.hru)

To investigate the mechanism of calculating the HRU scale SWE outputs in SWAT, a specific model which defined a single and unique HRU for the location of each snow station was constructed. Subbasin 12 is taken as an example for this exercise, since six snow stations are located at different elevations (Table S5.1) in this subbasin. However, as shown in Figure S5.1, the SWE outputs for the HRUs related to the six stations do not display any significant difference. The value assigned to each HRU was equal to the weighted sum of the SWE of all elevation bands SUM_EB:

$$SUM\_EB = \sum_{bnd=1}^{b} SWE_{band} \cdot fr_{bnd} = Sub\text{-}SWE \qquad\qquad (S5.1)$$

where SWE_{band} is the SWE value of each band (mm), fr_{bnd} is the fraction of the subbasin area within the elevation band, and b is the total number of elevation bands in the subbasin.

Therefore, we can conclude that the SWAT model assigned the same SWE output for each HRU in one subbasin and this value represents the mean SWE of the subbasin ($Sub\text{-}SWE$). Besides, the measurement of a single monitoring station cannot be used to calibrate the model output obtained at the HRU scale.
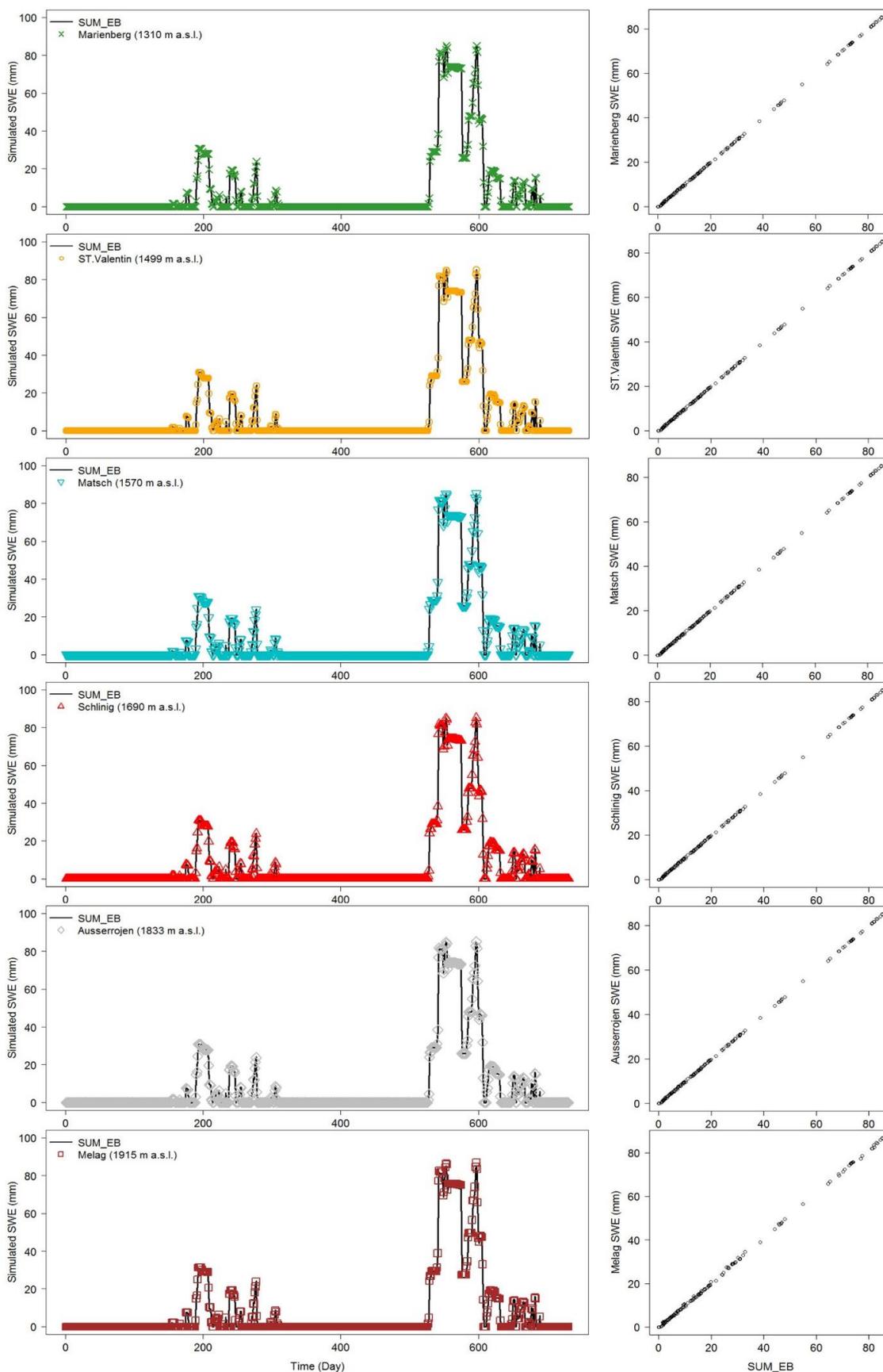
C. Supplementary material of Chapter 5

Figure S5.1 Comparison of the Simulated SWE of each HRU where the snow station is located and the weighting sum of the SWE value of the elevation bands. The SUM_EB is used as a common reference for comparing the SWE outputs.

Table S5.1 Information about the snow stations of each subbasin.

Subbasin	Station	LAT	LONG	ELEVATION
1	PRETTAU	47.03658867	12.09912722	1450
2	WEHR_IN_PFITSCH	46.92499219	11.5246721	1365
	RIDNAUN	46.90908292	11.30677415	1350
	INNERRATSCHINGS	46.8660261	11.31719654	1260
3	MUEHLEN	46.90025669	11.94743948	851
4	TERENTEN	46.83445756	11.78557637	1349
	PFUNDERS	46.88718005	11.7029428	1159
5	BRUNECK	46.8042795	11.93148052	821
6	ST.MAGDALENA_IN_GSIES	46.83526049	12.24271916	1398
	STAUSEE_OLANG	46.75973668	12.0536567	1057
	MITTERTAL	46.86072197	12.10089661	1236
7	TOBLACH	46.72997843	12.21915906	1219
8	PLATT	46.82249293	11.17808117	1147
	ST.MARTIN_IN_PASSEIER	46.78268262	11.227909	588
9	FRANZENSFESTE	46.77830419	11.63224641	725
10	LUESEN	46.74503125	11.76389069	981
11	STERN	46.5836304	11.90707588	1390
	CAMPILL	46.63934499	11.8608778	1396
12	AUSSERROJEN	46.8095884	10.48411317	1833
	MELAG	46.83743215	10.65658211	1915
	ST.VALENTIN	46.77587237	10.52890216	1499
	MARIENBERG	46.70568911	10.5212676	1310
	SCHLINIG	46.70413104	10.47392043	1690
	MATSCH	46.69431625	10.61752601	1570
13	SCHENNA	46.68959619	11.19083111	680
	MERAN	46.68805798	11.13656021	330
	OBERMAIS	46.6578197	11.18323354	334
14	VORDERKASER	46.73510025	10.92669711	1705
15	ALLITZ	46.65020601	10.69607321	1754
	STAUSEE_ZUFRITT	46.50906285	10.72507183	1851
16	FLITZ	46.62357402	11.66281009	1350
	LATZFONS	46.67240083	11.54613904	1190
	ST.ULRICH	46.57404443	11.67296415	1180
17	WEISSBRUNN	46.48684723	10.83181103	1900
	STAUSEE_ZOGGL	46.54167092	10.98978931	1142
	STAUSEE_ST.PANKRAZ	46.56489465	11.05568468	810
	PAWIGL	46.62776162	11.10933866	1400
18	REINSWALD	46.70104927	11.41450083	1365
19	VOELS	46.51340578	11.50595194	840
20	MOELTEN	46.58577237	11.25168816	1133
Nearby	ALTREI	46.277346	11.367751	1209
	SEXTEN	46.703054	12.349964	1310
	DEUTSCHNOFEN	46.419133	11.412535	1415
	RADEIN	46.346545	11.396822	1562

Table S5.2 Absolute percentage error of snow cover duration (SCD).

Subbasin	Calibration			Validation		
	I	II	III	I	II	III
1	48.8	39.9	11.0	36.4	29.4	10.1
2	45.1	64.6	12.2	37.0	54.3	12.6
3	60.4	25.6	6.7	32.8	5.3	8.4
4	19.5	9.5	18.9	7.0	6.2	10.4
5	7.3	15.5	5.8	6.4	10.9	1.1
6	26.8	30.5	0.0	23.8	28.6	0.3
7	67.1	89.0	16.5	55.7	71.1	21.8
8	24.1	23.5	1.2	35.3	32.5	10.1
9	1.8	83.5	5.5	3.1	70.9	0.0
10	94.5	58.2	7.3	67.5	45.7	5.0
11	8.8	12.8	9.8	8.1	8.4	9.0
12	62.8	54.0	7.0	51.0	52.1	9.8
13	96.6	56.4	2.7	74.2	39.8	2.8
14	64.9	47.0	0.3	40.9	39.5	0.8
15	42.1	86.3	19.5	21.3	58.8	27.5
16	17.7	10.4	11.9	26.1	5.9	6.2
17	39.0	11.6	1.2	14.8	23.2	17.6
18	39.9	9.8	7.3	27.2	10.6	1.7
19	29.6	11.9	16.8	28.6	10.4	18.5
20	83.2	13.7	7.0	33.1	4.8	2.5

Table S5.3 Absolute percentage error of snow peak value (SPV).

Subbasin	Calibration			Validation		
	I	II	III	I	II	III
1	82.9	93.7	20.0	15.5	20.7	3.1
2	56.0	83.6	13.7	19.6	34.5	16.5
3	41.7	22.4	5.8	5.4	9.4	9.2
4	29.3	23.4	0.0	15.4	5.0	4.1
5	8.8	15.7	21.5	19.9	23.6	5.3
6	24.0	18.3	3.9	1.8	2.3	10.7
7	13.7	11.2	4.1	15.0	9.2	7.6
8	60.2	65.8	2.2	26.0	31.6	36.4
9	25.5	80.5	25.4	9.7	35.4	9.5
10	94.4	16.0	21.1	79.9	11.3	0.4
11	1.1	6.9	6.7	8.7	14.0	4.2
12	20.8	5.7	16.6	4.8	15.7	8.1
13	183.4	104.8	36.4	148.4	129.8	15.4
14	84.8	78.4	29.7	76.6	58.6	2.7
15	77.7	82.1	22.3	64.9	85.1	19.1
16	4.7	28.1	31.1	25.4	9.6	13.0
17	65.8	31.6	42.2	15.4	30.7	13.5
18	48.2	27.4	8.0	26.7	48.0	38.5
19	45.1	20.8	33.6	14.0	23.2	0.9
20	61.7	49.1	41.6	24.2	25.7	21.6

Table S5.4 Absolute percentage error of snow peak day (SPD).

Subbasin	Calibration			Validation		
	I	II	III	I	II	III
1	18.0	18.0	12.5	2.5	2.5	2.5
2	18.0	13.3	4.7	1.2	1.6	1.2
3	16.7	12.1	10.5	1.5	16.1	1.1
4	10.5	12.2	10.9	16.1	16.1	1.5
5	8.9	0.1	1.5	0.8	0.5	0.0
6	13.3	12.2	5.1	1.2	2.1	2.1
7	12.7	18.6	6.1	2.7	3.6	2.7
8	18.5	18.5	3.0	2.4	1.2	0.1
9	0.7	22.9	0.7	1.6	2.5	1.6
10	0.8	5.8	1.6	14.0	0.9	0.0
11	6.6	6.6	6.6	0.5	0.5	2.5
12	18.5	17.9	65.8	0.9	0.0	0.0
13	21.8	14.4	0.2	0.6	0.7	0.8
14	3.1	3.1	4.0	12.0	13.5	0.3
15	3.7	3.7	5.9	2.6	3.6	0.3
16	8.9	0.7	0.7	0.0	0.0	0.0
17	3.4	17.2	12.3	1.3	2.5	1.3
18	0.6	13.9	5.0	0.8	0.7	1.2
19	0.7	1.1	0.7	0.3	0.4	0.3
20	0.6	0.6	0.6	0.1	0.1	0.3