題目(和文)	
Title(English)	Development of streamflow and sediment load estimation methods and the application to hydro-environmental assessment of the Tonle Sap Lake Basin
著者(和文)	ANGRAKSMEY
Author(English)	Raksmey Ang
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12441号, 授与年月日:2023年3月26日, 学位の種別:課程博士, 審査員:木内 豪,吉村 千洋,中村 恭志,中村 隆志,VARQUEZ ALVIN CHRIST
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12441号, Conferred date:2023/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Development of streamflow and sediment load estimation methods and the application to hydro-environmental assessment of the Tonle Sap Lake Basin

by

Raksmey Ang

A thesis submitted in partial fulfillment of the requirements for the
degree of Doctor of Engineering in
Global Engineering for Development, Environment and Society (GEDES)

Examination Committee:	Prof. Tsuyoshi Kinouchi (Chairperson) Prof. Yoshimura Chihiro Assoc. Prof. Takashi Nakamura (Suzukakedai) Assoc. Prof. Takashi Nakamura (Ookayama) Assoc. Prof. Alvin C.G. Varquez

Tokyo Institute of Technology
School of Environment and Society

Japan
March 2023

This page intentionally left blank.

Preface

This dissertation is an outcome of a three-year doctoral program conducted at the Department of Transdisciplinary Science and Engineering, School of Environment and Society, Tokyo Institute of Technology, Japan, from April 2020 to March 2023. The three-year program was financially supported by the Japanese Government (Monbukagakusyo: MEXT).

The content of this dissertation is fully or partially based on works either published or forthcoming publication by Raksmey Ang. Although these works have co-authors, the substantial contribution (intellectual, research, and writing) was that of the first author, and any specific contributions of co-authors have either not been presented in the dissertation or have been duly acknowledged. Some publications accomplished by Raksmey Ang during the three-years program are:

Ang, R., Kinouchi, T., Zhao, W., *Sediment load estimation using a novel regionalization sediment-response similarity method for ungauged catchments*, Journal of Hydrology, Vol. 618, 129198, 2023. *In support of Chapter 3*

Ang, R., Kinouchi, T., Zhao, W., *Evaluation of daily gridded meteorological datasets for hydrological modeling in data-sparse basins of the largest lake in Southeast Asia*, Journal of Hydrology: Regional Studies, Vol. 42, 101135, 2022. *In support of Chapter 2*

Ang, R., Kinouchi, T., *Decadal climatic trends, variability and hydrological influences in the Tonle Sap Lake Basin, Cambodia*, Water Security and Climate Change Conference, Dec. 2022. *In support of Chapter 4*

Ang, R., Kinouchi, T., *Statistical evaluation of daily gridded precipitation datasets for the Tonle Sap Lake, Cambodia*, AGU Fall Meeting, Dec. 2021. *In support of Chapter 2*

Acknowledgments

To make this thesis exist, the contribution, support and encouragement from several people are needed. First, I may take this opportunity to give full respect and gratefulness to Prof. Tsuyoshi Kinouchi, who is my academic supervisor, for guiding me during this thesis completion period. His insights, encouragement and ability to keep me on the track were instrumental throughout the duration of this research work. He has always offered me useful advice and comments on how to best improve my work. His support and inspiring suggestion have been precious for the development of this thesis content. Moreover, he is one of the excellent advisors, mentors and instructors. Therefore, it is my pleasure to acknowledge him once again, for his professional and personal involvement in my doctoral program study.

I would also like to thank my thesis committee members— Prof. Yoshimura Chihiro, Assoc. Prof. Takashi Nakamura (Suzukakedai), Assoc. Prof. Takashi Nakamura (Ookayama) and Assoc. Prof. Alvin C.G. Varquez, for all of their assistance and guidance, encouragement, indispensable advice, provisions of reference materials, meticulous guidance, checking and editing this thesis. My sincere appreciation is also extended to the Japanese Government (Monbukagakusyo: MEXT) for the financial support of my doctoral program study.

More importantly, my sincere gratefulness also flows to my family members – especially my mom and grandparents for their full love and endless support. This achievement would not have been attained without their continuous support and unconditional encouragement. Finally, my sincere thank goes to my research team members, lab mates and friends for their never-ending concerns, support, helps and encouragement during the period of my doctoral program. I also would like to thank and give my gratitude to one and all, who are directly or indirectly involved in my study.

Abstract

The state of knowledge of hydro-environmental impact due to the development and climate variation in a particular river basin is significant for sustainable environmental management of the basin. Reliable climate datasets and frameworks play vital roles in conducting hydro-environmental impact assessment through the application of regional hydrological models. However, due to the financial constraints of the Kingdom of Cambodia, the sparse and uneven distribution of hydro-meteorological gauge stations in the Tonle Sap Lake (TSL) Basin hinders the development of reliable hydrological models and accurate simulations of the hydrological impacts of anthropogenic activities and climate change. The TSL Basin in Cambodia is the largest freshwater body in Southeast Asia and one of the most productive ecosystems in the world, playing a crucial role in livelihood and sustainable development in Cambodia and the Lower Mekong region. Despite its significant value, the lake ecosystem is widely under threat from climate change together with anthropogenic activities inside and outside the TSL basin. Although these impacts could be attributed to multiple driving factors such as the influence of the Mekong mainstream and anthropogenic activities inside the lake, limited information is available regarding hydro-environmental impacts from the tributary basins of the TSL. This necessitates a comprehensive hydrological assessment of the TSL tributary basins, owing to land-use change, population growth, and climate change, for environmental conservation of the TSL Basin.

The overall objective of this study is to develop a feasible method or framework for estimating streamflow and sediment load in data-sparse or poorly gauged basins of the Tonle Sap Lake, for sustainable management and conservation of the lake ecosystems. First, a comprehensive framework was used to determine an ideal alternative meteorological dataset for hydrological modeling, considering the spatiotemporal characteristics of each climate dataset and ensuring a reliable estimate of streamflow and evapotranspiration. Next, a new regionalization method was proposed to estimate sediment load in ungauged catchments, considering spatiotemporal variability and the sediment load relation to rainfall characteristics of individual catchments. Finally, the application of the above two mentioned methods for hydro-environmental impact assessment, owing to decadal climate variation and basin development during the last few decades.

In the beginning, a comprehensive framework was introduced to assess seven gridded precipitation and air temperature products by statistically comparing these datasets with gauge-based datasets and applying the Soil and Water Assessment Tool (SWAT) model for daily streamflow and evapotranspiration (ET) simulations over the TSL Basin. The precipitation data from Asian Precipitation-Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE), European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5), Tropical Rainfall Measuring Mission (TRMM) and Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM) (IMERG) were found to have high correlations with rain-gauged data and the lowest estimation errors, and the Southeast Asia-Observation (SA-OBS) and Climate Prediction Center (CPC) were found to match the observed air temperature data well. In addition, the results of the hydrological simulation showed that the rainfall data from APHRODITE, TRMM and IMERG, combined with SA-OBS-based air temperature data, provided improved estimations of daily streamflow and mean runoff depth. The ET estimated using the TRMM and IMERG datasets showed a better temporal and spatial pattern agreement with ET from Moderate Resolution Imaging Spectroradiometer (MODIS) and Global Land Evaporation Amsterdam Model (GLEAM). This suggests that TRMM and IMERG, in conjunction with SA-OBS air temperature, are reliable for providing streamflow through the SWAT model application and other water balance components. This comprehensive evaluation framework was found to be effective in selecting reliable gridded meteorological datasets for hydrological simulation in data-sparse river basins, especially when large uncertainty existed in the spatiotemporal distribution of rainfall. These findings also showed that statistical comparisons with gauge data and hydrological evaluation of streamflow are not enough to justify the reliability of each gridded dataset.

Although satellite-based or gridded meteorological data could serve as the hydrological model inputs, model outputs (i.e., streamflow or sediment load) need to be optimized using ground-based observation. However, model calibration or validation cannot be performed in the ungauged catchment (e.g., the catchment is not monitored in terms of water level or sediment concentration). To solve this problem, a common method known as regionalization, in which model parameters from well-monitored catchments are transferred to ungauged, was used to estimate hydrological variables such as streamflow and sediment load. Therefore, a novel Sediment-Response Similarity (SRS) regionalization method has been proposed, using the SWAT model and Self-Organizing Map clustering technique to

overcome the limitation of the critical attributes of a catchment favoring sediment similarity, which usually exists in the conventional regionalization approaches. It considered the spatiotemporal variations of sediment response and its relationship with rainfall characteristics as a catchment attribute and showed the potential to ideally determine hydrological and sediment similarities between gauged and ungauged catchments. The results indicated the comprehensive performance of the SRS regionalization method for estimating sediment load in the ungauged catchments. The SRS approach obtained an estimation error reduction of up to 7%, compared with the Physical Similarity regionalization method. The SRS regionalization method proposed in this study is a global alternative method for estimating sediment, as well as other hydrological variables and rainfall-driven phenomena such as streamflow and nutrient transport, in ungauged catchments.

Eventually, the above two frameworks were used to quantify the impacts of climatic variability and land-use change on streamflow and sediment load in the TSL Basin during the last few decades. Climate analysis showed an increasing trend in the basin temperature and a downward trend in rainfall between 2001 and 2020. The results of the land-use change analysis revealed that there was a substantial decrease found in forest cover, in which cumulative loss of natural forest area was around 45% or the area decreased from 37,052 in 1995 to 20,408 km^2 in 2018. On the other hand, the cropland area increased by about 23% from around 30,400 to 37,324 km^2 in 1995 and 2018, respectively. The hydrological analysis depicts that the early rainy season flows of the TSL Basin were lower (max. 26% decrease) for 2011-2020 compared to the 2001-2010 time horizons. However, after the wet monsoon season, the streamflow was observed to receive some increases due to the rise in rainfall amount during September and November. A similar change in the seasonal pattern forced by climate variability was found between flow and sediment load. The land-use change had a positive and negative impact on streamflow sediment loading during the rainy and dry seasons, respectively. The streamflow and sediment load were seen to increase by up to 5.83 and 19.57%, respectively between May and November, while a decrease of up to 4.54 and 5.80%, respectively was found between December and April. The basin streamflow and sediment load showed a greater impact from climate variation than land-use change. Based on the results found in this study, climate variability and land-use change have noticeable impacts on streamflow and sediment load, suggesting that planning and management of their

impacts are crucial for the sustainable management of water resources and lake ecosystem conservation in the TSL Basin.

The feasible framework developed in this research would serve as a central approach in estimating streamflow and sediment load in data-sparse or poorly gauged basins. Besides that, the study also contributes a basic methodology to estimate other hydro-meteorological parameters of interest. For a similar purpose, this methodology can be applied as well in other river basins around the world.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

2-D	Two-Dimensional
3S	Sekong, Sesan and Srepok
AM	Arithmetic Mean
APHRODITE	Asian Precipitation-Highly Resolved Observational Data Integration Towards Evaluation of Water Resources
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
CFSR	Climate Forecast System Reanalysis
CHIRPS	Climate Hazards Group Infrared Precipitation with Station
CMORPH	Climate Prediction Center Morphing algorithm
CPC	Climate Prediction Center
CRRS	Catchment Runoff-Response Similarity
DEM	Digital Elevation Model
DPR	Dual-Frequency Precipitation Radar
ECMWF	European Centre for Medium-Range Weather Forecasts
ENSO	El Niño Southern Oscillation
ERA5	European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5
ET	Evapotranspiration
FAO	Food and Agricultural Organization of the United Nations
GES DISC	Goddard Earth Sciences Data and Information Services Center
GIS	Geographic Information System
GLEAM	Global Land Evaporation Amsterdam Model
GLUE	Generalized likelihood Uncertainty Estimation
GMI	Global Precipitation Measurement Microwave Imager
GPCC	Global Precipitation Climatology Center
GPM	Global Precipitation Measurement
GSMaP	Global Satellite Mapping of Precipitation
GTS	Global Telecommunication System
HRU	Hydrological Response Unit
IMERG	Integrated Multi-satellitE Retrievals for Global Precipitation Measurement

IMERG-E	Integrated Multi-satellitE Retrievals for Global Precipitation Measurement Early run
IMERG-F	Integrated Multi-satellitE Retrievals for Global Precipitation Measurement Final run
IMERG-L	Integrated Multi-satellitE Retrievals for Global Precipitation Measurement Late run
JAXA	Japan Aerospace Exploration Agency
LC	Land Cover
LMRB	Lower Mekong River Basin
Loa PDR	Lao People's Democratic Republic
LOADEST	LOAD ESTimator
MCMC	Markov Chain Monte Carlo
MODIS	Moderate Resolution Imaging Spectroradiometer
MRC	Mekong River Commission
MRC-WQMN	Mekong River Commission-Water Quality Monitoring Network
MUSLE	Modified Universal Soil Loss Equation
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NOAA	American National Oceanic and Atmospheric Administration
ParaSol	Parameter Solution
PERSIANN-CDR	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record
PET	Potential Evapotranspiration
PGF	Princeton University Global Meteorological Forcing
PS	Physical Similarity approach
PSO	Particle Swarm Optimization
RG	Regression approach
RISH	Research Institute for Sustainable Humanosphere
RNMI	Royal Netherlands Meteorological Institute
SA-OBS	Southeast Asia-Observation
SOM	Self-Organizing Map
SPX	Spatial Proximity approach
SRS	Sediment-Response Similarity
SUFI-2	Sequential Uncertainty Fitting version 2

SWAT	Soil and Water Assessment Tool
SWAT-CUP	Soil and Water Assessment Tool-Calibration and Uncertainty Programs
TRMM	Tropical Rainfall Measuring Mission
TSL	Tonle Sap Lake
TSS	Total Suspended Solids
U-matrix	Unified distance matrix
UNESCO	United Nations Educational, Scientific and Cultural Organization
USA	United States of America
USLE	Universal Soil Loss Equation
WFD	WATer and global CHange Forcing Data
WMO	World Meteorological Organization

List of Symbols

%	Percentage
°C	Degree Celsius
∞	Infinity
°E	Degree East
°N	Degree North
CA	Catchment attribute value
C_{USLE}	Universal Soil Loss Equation Land Cover and Management Factor
EVI	Enhanced Vegetation Index
K_{USLE}	Universal Soil Loss Equation Soil Erodibility Factor
LSWI	Land Surface Water Index
MAM	March-April-May
MBE	Mean Bias Error
md	Modified index of agreement
MJJ	May-June-July
N	number of data at daily time series
NSE	Nash-Sutcliffe Efficiency
$\overline{O}$	Observed mean daily precipitation or temperature
O_i	Observed daily precipitation or temperature of the ith day
P_i	Predicted daily precipitation or temperature of the ith day
$\overline{P}$	Predicted mean daily precipitation or temperature
PBIAS	Percent Bias
Q_{mean}	Mean daily streamflow
Q^{obs}	Observed streamflow of the ith day
Q^{sim}	Simulated streamflow of the ith day
R	Correlation coefficient
R^2	Coefficient of determination
RMSE	Root Mean Square Error
SON	September-October-November
Tmax	Maximum temperature
Tmin	Minimum temperature

List of Units

km	Kilometer
km^2	Square kilometer
km^3	Cubic kilometer
m	Meter
m^3	Cubic meter
m^3/s	Cubic meter per second
mm	Millimeter
mm/month	Millimeter per month
Mt	Million ton
Mt/year	Million tons per year
ton/km^2	Ton per square kilometer of land area

Chapter 1 Introduction

This chapter first explains why this research is important and the research motivations then list specific objectives to be performed in subsequent chapters, and finally outlines the dissertation.

1.1. Background and research motivations

The state of knowledge of hydro-environmental impact due to the development and climate variation in a particular river basin is significant for sustainable environmental management of the basin. Reliable climate datasets and frameworks play vital roles in conducting hydro-environmental impact assessment through the application of regional hydrological models. However, the sparse hydro-meteorological stations in the river basin located in developing countries like Cambodia hinder the development of reliable hydrological models and accurate simulations of the assessment of the hydro-environmental impact in the basin (Ang and Oeurng, 2018; Heng et al., 2020; Sao et al., 2020; Touch et al., 2020). Therefore, the hydro-environmental impact assessment of a particular river basin owing to climate change and basin development has not been fully understood. The hydro-environmental impact assessment could serve as a scientific report and be beneficial for adaptation policies for sustainable water resource management and ecosystem conservation of the river basin.

The Tonle Sap Lake (TSL) in Cambodia is the largest freshwater body in Southeast Asia and one of the most productive ecosystems in the world (Uk et al., 2018), playing a crucial role in livelihood and sustainable development in Cambodia and the Lower Mekong region (Shivakoti and Bao, 2020). The lake is an integral part of the Mekong River Basin, which flows through six riparian countries: China, Myanmar, Thailand, Lao People's Democratic Republic (PDR), Cambodia, and Vietnam (Figure 1.1). At the same time, the high productivity and rich biodiversity in the TSL have been maintained through “flood pulses”, which are seasonal inundation during the monsoon floods of the Mekong River (Kummu et al., 2014; Uk et al., 2018). This unique reverse flow through the Tonle Sap River brings in a huge mass of water, sediments, nutrients and migratory fish. The productivity of TSL and the state of the ecosystem are critical for several people who are dependent on fisheries for their livelihood. People living in floating villages are directly impacted by changes in the environmental condition of the lake due to their high reliance on the fisheries (Shivakoti and Bao, 2020). Similarly, the fertile floodplains in TSL and its basin are important for many farming communities for growing rice, vegetables, and cropland.

Figure 1.1. Location map of the Tonle Sap Lake in Cambodia. The inset indicates the location of the TSL Basin in the Lower Mekong River and the Indochina Peninsular.

Recent studies have revealed that the lake has been confronting various environmental changes. For example, Kallio and Kummu (2021) reported that the maximum inundation area of the TSL during the rainy season showed a decreasing trend of 128 km^2 per year. Wang et al. (2020) reported an increasing trend of sediment concentration in the lake based on the analysis of satellite images. Fish catches have also revealed a decline in various fish species over time (Ngor et al., 2018). Although these changes could be attributed to multiple driving factors such as the influence of the Mekong mainstream and anthropogenic activities inside the lake, limited information is available regarding hydro-environmental impacts from the tributary basins of the TSL (Ang and Oeurng, 2018; Kummu et al., 2014). This necessitates a comprehensive hydrological assessment of the TSL tributary basins, owing to land-use change, population growth, and climate change, for environmental conservation of the TSL Basin.

1.2. Research objectives, scope and limitation

This research aims to develop a feasible method or framework for estimating streamflow and sediment load in data-sparse or poorly gauged basins of the Tonle Sap Lake, for sustainable management and conservation of the lake ecosystems.

The specific objectives of the study are:

(1) To determine an ideal alternative meteorological dataset for hydrological modeling in data-sparse or poorly gauged catchments, considering the spatiotemporal characteristics of each climate dataset and ensuring a reliable estimate of streamflow and evapotranspiration.
(2) To develop a new regionalization method to estimate sediment load in ungauged catchments, considering spatiotemporal variability and the sediment load relation to rainfall characteristics of individual catchments.
(3) To apply the above two mentioned-methods for hydro-environmental impact assessment, owing to decadal climate variation and basin development during the last few decades.

The scope and limitations of this study are listed below:

- The study was conducted exclusively in Tonle Sap Lake, Cambodia for two main reasons. First is the viability of the research, meaning that some ground-based data are accessible for analysis. Second is the eager need for hydro-environmental information on the TSL's tributaries for lake ecosystem conservation.
- Only (11) gauging stations in each tributary were considered because of data availability, thus, the catchment area of each tributary was modeled based on the gauged area. Approximately 50% of the total catchment area was gauged and modeled for streamflow.
- For Objective 1, only seven potentially promising gridded precipitation and air temperature datasets were selected for analysis including their spatio-temporal representations and hydrological quantification performances.
- For Objective 2, due to data shortage, the sediment concentration data from two catchments outside the TSL Basin were used for proposing a new regionalization method for sediment load estimation in ungauged catchments.
- A number of global and regional datasets were used in this study. Topography or elevation data (90-m resolution) download from NASA Shuttle Radar Topographic Mission (SRTM). Land-use distribution was obtained from the Mekong River Commission (MRC) and SERVIR-Mekong. Information on soil types was obtained from a 1:250,000 scale map created by the MRC.

1.3. General research framework

This section describes the overall framework of the research. The study procedure can be divided into three steps (Figure 1.2). Detailed methodologies of each step are introduced in the subsequent chapters.

The first step is the data pre-processing. Raw data obtained from different sources were pre-processed before they can be used for each analysis. Gauged data of hydro-meteorology include water level/discharge, rainfall, air temperature and sediment concentration. These datasets were processed manually, except for sediment data which a sediment estimation tool was used to get daily sediment load data. Information on topography, land use/cover and soil type is based on global and regional datasets. They were processed with the aid of geographic information system (GIS) software. Land cover and management factor and soil erodibility factor were determined based on the Universal Soil Loss Equation (USLE) obtained from their databases.

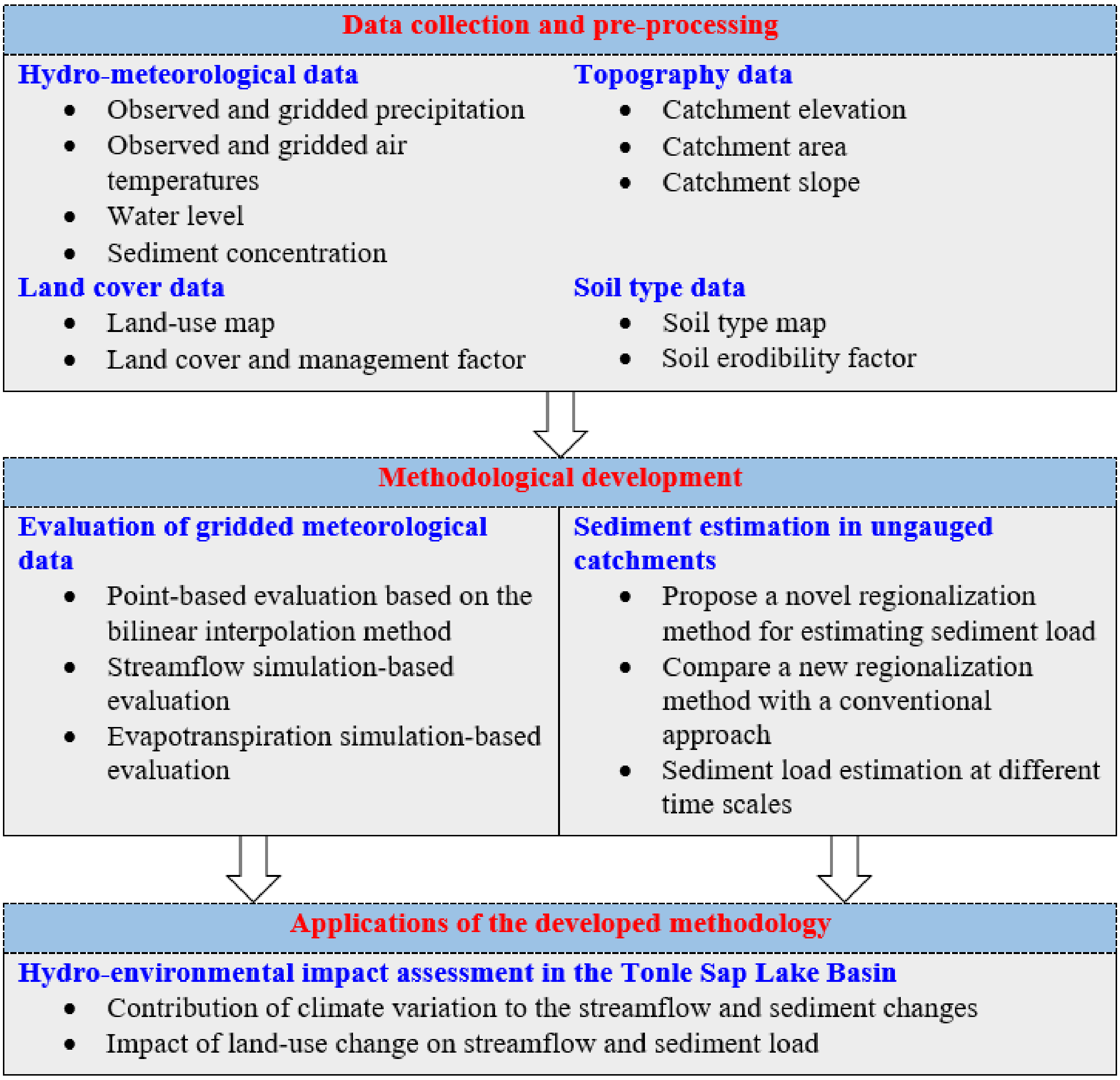


Figure 1.2. General research framework.

The second step is the methodology study (i.e., estimating streamflow and sediment load in data-sparse catchments).

- First, several gridded meteorological datasets were evaluated for hydrological modeling – especially for streamflow and evapotranspiration simulation. The linear interpolation method was used to interpolate gridded climate data to the ground-based location of observation for comparing the spatiotemporal characteristics of each climate product to the gauged data. Then, the outstanding meteorological datasets were selected as the climate input data into the hydrological model for streamflow and ET simulations.

- Next, a new regionalization method was developed to estimate sediment load in ungauged catchments. The new method was compared with the conventional approach and eventually was applied to estimate sediment load in ungauged tributaries of the TSL Basin due to its superior performance to the conventional method.

The third step is the application of the above two mentioned-methods for hydro-environmental impact assessment, owing to decadal climate variation and basin development. First, the decadal climatic variation was analyzed during the last few decades. Then, the contribution of climate variation to streamflow and sediment change was investigated by comparing streamflow and sediment in different time horizons. Finally, four different land-use change scenarios were analyzed and input into the hydrological modeling for streamflow and sediment load impact assessment.

1.4. Dissertation organization

There are, in total, five chapters present to constitute this dissertation. Each chapter is briefly described below. Chapters 2 to 4 present the main research results of this study.

- Chapter 1 begins with the background information and research motivation by describing some facts and the importance of climate datasets and a framework for conducting the hydro-environmental impact assessment in a particular river basin. After pointing out research needs and challenges in the target river basin, it indicates specific objectives to be obtained in this research work. Lastly, it outlines the dissertation chapter by chapter, such that readers can take things in at a glance.
- Chapter 2 focuses on the evaluation of alternative meteorological datasets for hydrological modeling in the data-sparse basins. Different gridded or satellite-based climate datasets were evaluated by comparing them with the available gauged data at different spatiotemporal scales. Next, the ideal gridded products were selected and input into the hydrological modeling for streamflow and evapotranspiration

simulations. The simulated streamflow of each climate dataset was compared with observed data, while the simulated ET was compared with the reanalysis and satellite-based ET for evaluation. Finally, the best alternative climate datasets were found and further used for the next work in Objectives 2 and 3 (i.e., Chapters 3 and 4).

- Chapter 3 focuses on estimating sediment load in ungauged catchments. A new regionalization method was proposed to simulate sediment load in ungauged catchments and its performance was also compared with a conventional approach. The new regionalization method considers spatiotemporal variability and the sediment load relation to rainfall characteristics of individual catchments as the catchment's attributes. Due to the outperformance of the new method over the traditional one, it was used to estimate the sediment load in ungauged tributaries of the TSL Basin.
- Chapter 4 brings altogether two methodologies or frameworks from the previous chapters for the application to hydro-environmental impact assessment, owing to climate variation and basin development during the last few decades. First, the decadal climatic variation (i.e., temperature and precipitation) was analyzed from 2001 to 2020, covering extreme events (i.e., flooding and drought) occurring over the TSL Basin. Then, the impact of climate variation on streamflow and sediment load was investigated by comparing streamflow and sediment load of the 2001-2010 to 2011-2020 periods. Finally, four different land cover (LC) scenarios (i.e., LC_1995, LC_2002, LC_2010 and LC_2018) were analyzed and input into the hydrological modeling for streamflow and sediment load impact assessment.
- Chapter 5 presents a summary of research results and contributions. It lists a set of remaining challenges for future research at the end.

Chapter 2 Evaluation of daily gridded meteorological datasets for hydrological modeling in data-sparse basins

The study in this chapter attempts to determine an ideal alternative meteorological dataset for hydrological modeling in data-sparse catchments. The final output of this section is the target of Objective 1. This model was carried on to Chapter 4 for application in hydro-environment impact assessment.

2.1. Introduction

Reliable climate datasets play vital roles in conducting hydro-environmental impact assessment through the application of regional hydrological models. However, due to the financial constraints of the Kingdom of Cambodia, the sparse and uneven distribution of rain gauge stations in the TSL Basin hinders the development of reliable hydrological models and accurate simulations of the hydrological impacts of anthropogenic activities and climate change. For example, some studies attempted to simulate the daily streamflow exclusively for the TSL Basin (Oeurng et al., 2019) and its tributary basins (Ang and Oeurng, 2018; Heng et al., 2020; Sao et al., 2020; Touch et al., 2020) using the Soil and Water Assessment Tool (SWAT) model. The limited agreement of simulated and observed streamflow values was shown in these studies, and it was mainly attributed to data sparsity and limited available information on the spatio-temporal variations of meteorological variables. Considering the current sparse monitoring network of the meteorological conditions in Cambodia, particularly over the TSL Basin, alternative reliable climate datasets are highly necessary. Furthermore, getting ground-based data at a good spatial-temporal resolution is very challenging in developing countries, and the data sometimes cannot be freely accessed by public users due to data sharing policy. Gauge observations are subject to limitations, such as reporting time delays and sparse gauge networks (Bai & Liu, 2018). The open-accessed or satellite-based datasets, on the other hand, can support near-real-time monitoring or forecasting and be merged with the in-situ observations to improve the data quality.

We can find various studies globally that utilize regional and global gridded meteorological datasets to serve as alternative inputs for hydrological model simulations for hydro-environmental assessments and water resource management (Fuka et al., 2014; Ur

Rahman et al., 2020; Zhao et al., 2022). In recent years, many different gridded rainfall and temperature datasets based on observation, satellite measurement, or reanalysis have been evaluated for their performance in climate and hydrological studies targeting different river basins (Tan et al., 2021). For instance, Xu et al. (2016) compared WATer and global CHange (WATCH) Forcing Data (WFD) and Asian Precipitation-Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) data with ground-based precipitation data for flood simulation in the Xiangjiang River Basin, China. They found that APHRODITE data had a better correlation with gauged precipitation and a relatively high flow prediction accuracy than WFD precipitation data. The satellite-based precipitation data, Tropical Rainfall Measuring Mission (TRMM), is among the best-performing datasets with lower bias from the interpolated rain-gauge observations (Duan et al., 2016). Similarly, Satgé et al. (2020) evaluated different 23 gridded precipitation datasets and revealed that TRMM precipitation data showed relatively better performance for both daily and monthly time-step. Having compared with the gauged dataset, Amjad et al. (2020) found that European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5) and Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM) (IMERG) had minor estimation errors and high correlation for daily and monthly timescales analysis. Another study similarly found that among five satellite-precipitation products, TRMM and IMERG showed the best performance when compared with ground-based monitored data (Islam et al., 2020). Zhang et al. (2020) demonstrated that the combination of air temperature and precipitation from Climate Forecast System Reanalysis (CFSR) and the Climate Hazards Group Infrared Precipitation with Station (CHIRPS), respectively, provided satisfactory performance in streamflow simulation. Moreover, Fuka et al. (2014) found that for five watersheds representing different hydro-climate regimes, the hydrological model simulation of river flow driven by the CFSR precipitation and temperature data provided results as good as those obtained using observation-based weather data. Aslam et al. (2020) compared the maximum and minimum temperatures over a data-scarce transboundary river basin of India and Pakistan, obtained from the Climate Prediction Center (CPC) and Princeton University Global Meteorological Forcing (PGF) datasets with the ground-based observation. They found that CPC had a better performance regarding station-based time series analysis. Ge et al. (2019) analyzed the performance of the newly developed meteorological dataset, Southeast Asia-Observation (SA-OBS), in climatological means and surface air temperature trends of the Indochina Peninsula on a yearly and seasonal basis, and SA-OBS product was found to have a good overall agreement with observation and reanalysis datasets.

Despite these studies, it is challenging to select one superior product conclusively owing to the peculiar characteristics and uncertainties in each dataset depending on the data generation algorithm used, the factors inherent to study regions such as climate and topography, and the quality and quantity of the observed data used in developing the datasets (Fu et al., 2016; Lauri et al., 2014; Sun et al., 2014; Vu et al., 2012; Worqlul et al., 2014). Furthermore, a limited number of studies have evaluated climate data sources covering the Southeast Asia region, especially for the Mekong River Basin. Gunathilake et al. (2020) evaluated six different gridded precipitation products and showed that IMERG had a good performance in simulating the streamflow in the Upper Nan River Basin, Northern Thailand. Four different high-resolution gridded precipitation datasets were evaluated in the Philippines and it was shown that TRMM has the least bias errors and most closely resembles the rainfall distribution observed at gauged stations (Peralta et al., 2020). Similarly, the TRMM was found to be the best-performing dataset among five different precipitation products in Malaysia (Ayoub et al., 2020). Over the Mekong River Basin, most of these studies have validated gridded datasets by only focusing on the mainstream of the Mekong River Basin (e.g., Lauri et al., 2014; Try et al, 2022; Try et al., 2020). Try et al. (2020) validated the river discharges using five different gridded precipitation datasets—APHRODITE, TRMM, Global Satellite Mapping of Precipitation (GSMaP), Global Precipitation Climatology Center (GPCC), and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR)—at four gauging stations along the mainstream of the Mekong River in Lao PDR and Cambodia. Similarly, Lauri et al. (2014) utilized two temperature datasets (CFSR and ERA-Interim) and five precipitation products (TRMM 3B42 v6, TRMM 3B42 v7, APHRODITE, CFSR, and ERA-Interim) to evaluate river flow simulation at five gauging stations along the mainstream of the Lower Mekong River. Furthermore, Try et al. (2022) employed GPCC precipitation to study on identification of the spatio-temporal and fluvial-pluvial sources of flood inundation in the Lower Mekong Basin, and the river discharges were validated at the Kratie gauging station located on the mainstream of Mekong River. No previous studies have assessed the spatio-temporal characteristics and quantitative accuracy of gridded hydro-meteorological datasets over the TSL Basin; hence, the performance of individual datasets in hydrological modeling, including other water balance components, has not been clarified yet. The majority of studies reviewed above rarely reported the evaluation of alternative air temperature datasets for hydrological simulations. Accordingly, the assessment of air temperature data or the optimal data combinations (e.g., APHRODITE-precipitation with CFSR-temperature) from different alternative climate products should also be considered in the validation schemes (Tan et al., 2021). Conventional statistical approaches using simple correlation or error analyses are often not enough to justify the variability and distribution of gridded precipitation data (Ahmed et al., 2019). For further studies, Bai and Liu (2018) suggested that hydrological evaluation of gridded products based only on streamflow simulations could not validate the reliability of gridded climate datasets, but other hydrological variables, such as evaporation and soil moisture, should be considered for more comprehensive evaluation.

Previous studies (Ang and Oeurng, 2018; Oeurng et al., 2019; Uk et al., 2018) have highlighted the demand for more reliable hydrological assessment over the TSL Basin and the importance of identifying the factors that affect environmental and ecological changes in the TSL Basin. This study aimed to extensively evaluate alternative climate datasets for their reliability in representing spatio-temporal characteristics and estimating streamflow in the TSL Basin during the past decades, using robust statistical and hydrological modelling approaches. Seven potentially promising gridded precipitation and air temperature datasets for analysis were selected; these included APHRODITE, CFSR, TRMM, IMERG, ERA5, SA-OBS, and CPC. We presented a comprehensive framework to evaluate their spatio-temporal representations and hydrological quantification performances for all the TSL tributary basins where meteorological stations are sparsely located. More specifically, we first compared these gridded products with the observed datasets based on the bilinear interpolation method and examined the suitability of each dataset in estimating the daily streamflow in the eleven TSL tributary basins. Furthermore, we evaluated each climate dataset by comparing satellite-based and model-derived actual evapotranspiration (ET) in the whole TSL Basin which further quantitatively validated the annual and seasonal rainfall amounts. This approach enabled us to avoid the selection of hydro-meteorological datasets with significant bias and uncertainty (e.g., deficient basin-wide rainfall) that could affect the accuracy of simulations of sediment erosion and nutrient transport despite parameter optimization, which could yield acceptable ranges for streamflow estimation irrespective of the dataset used. Finally, we identified the gridded datasets most reliable for hydrological and environmental impact assessments as well as the projections of future climate change impacts over the TSL Basin.

2.2. Materials and methods

2.2.1. Study area

The TSL (the largest lake in Southeast Asia) is one of the most important natural resources in Cambodia (Kummu et al., 2008) and symbolically called “the heart of Cambodia”, covering over 45% (approximately 85,850 km^2) of the national land surface area and approximately 5% of the total basin area being shared with Thailand (Figure. 2.1). The Tonle Sap Lake Basin consists of 11 major tributary basins and is connected to the Tonle Sap River, which flows to the mainstream of the Mekong River Delta. The seasonal flow direction reversal in the Tonle Sap River and the resultant inundation of the TSL floodplain during the flood season (between June and October) causes the lake's total surface area and water level to swell from approximately 2,600 km^2 and 2 m to 12,000 km^2 and 10 m (Arias et al., 2014; Kummu et al., 2014; Oeurng et al., 2019), respectively. This creates a unique aquatic environment and a diverse and productive ecosystem (Arias et al., 2014; Kummu et al., 2014; Sabo et al., 2017). Kummu et al. (2014) estimated that 34% (29.1 km^3/year) of the water volume in the lake originates from the 11 tributary basins, approximately 53.5% (41.8 km^3/year) comes from the Mekong River, and 12.5% (10.4 km^3/year) is from precipitation over the lake surface. The TSL Basin is covered by alluvial deposits composed of unconsolidated silt, sand, and gravel overlaying on the older bedrock (Tsukawaki et al., 1994). The climate in this region is characterized by two monsoon periods. High humidity conditions prevail and frequent-heavy rains are typically observed during the south-west monsoon season from mid-May to early October (rainy season), while the north-east monsoon lasts from early November to March (dry season), bringing drier and cooler air (World Bank, 2011). The tributary basins predominantly consist of gently sloping lowlands with an elevation of less than 100 m above the mean sea level. The elevation increases to over 1700 m in the Cardamom Mountains to the southwest (see Figure. 2.1). The basin area, delineated area (a basin boundary representing the contributing area for a gauging station of streamflow, see Figure. A1), and land-use area of each tributary basin are summarized in Table 2.1. The total delineated area is 49.52 % (approximately 41,158 km2) of the total basin area. Most of the tributary basins (seven out of eleven) are predominantly forested land (> 50%), while agricultural land is dominant in the Mongkol Borey, Baribo, Dauntri, and Siem Reap tributary basins (Oeurng et al., 2019).

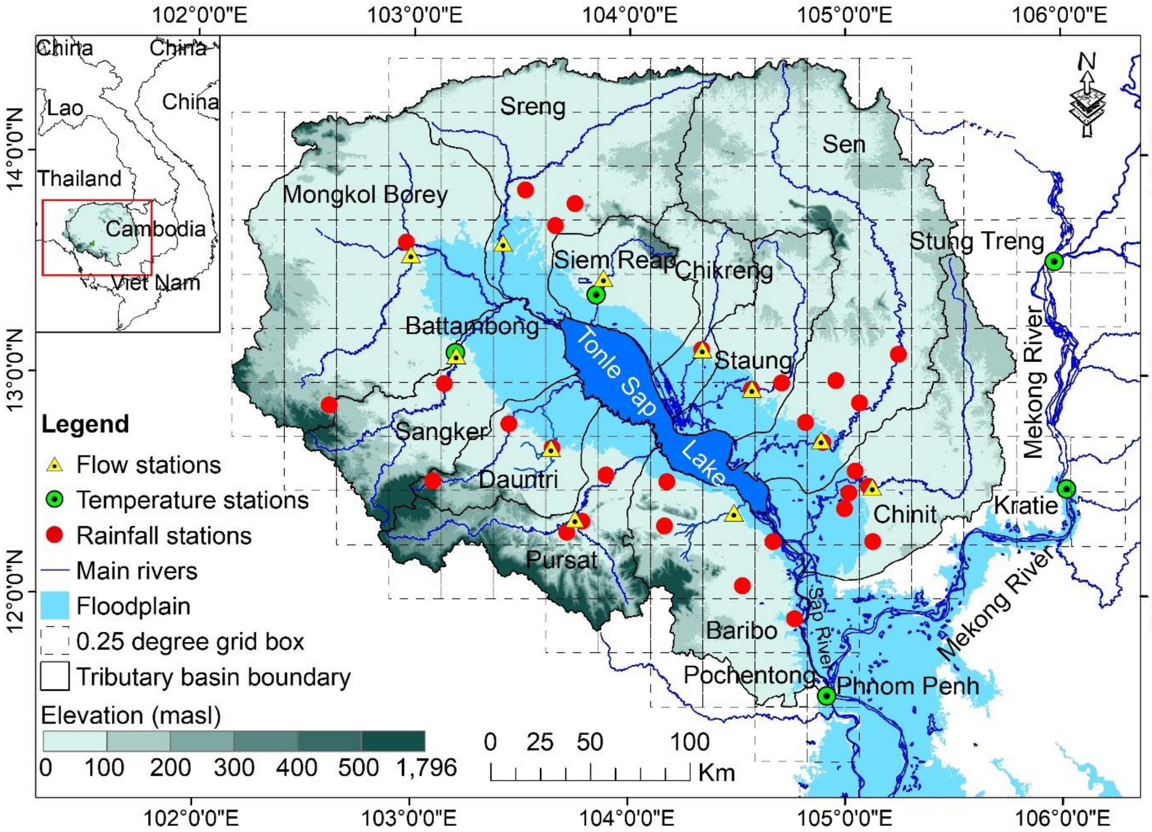


Figure. 2.1. Location map of the TSL Basin. The dash-line box shows a 0.25 × 0.25°C grid cell. The rainfall and temperature stations are the same in the Siem Reap tributary basin. A total of 156 grid cells are covering the whole TSL Basin area. The inset indicates the location of the TSL Basin in the Indochina Peninsular.

Table 2.1. The total drainage area, delineated area and land-use of each tributary basin

Nº	River Name	Area (km^2)	Delineated area (km^2 / %)	Forest land (%)	Agricultural land (%)	Urban land (%)
1	Baribo	7,153.78	875.33 / 12.24	25.01	74.61	0.39
2	Chikreng	2,713.90	1719.12 / 63.35	78.38	21.62	0.00
3	Chinit	8,236.86	4055.91 / 49.24	62.70	37.26	0.04
4	Dauntri	3,695.97	519.13 / 14.05	43.42	56.58	0.00
5	Mongkol Borey	14,966.42	3964.94 / 26.49	14.67	85.14	0.19
6	Pursat	5,964.77	4118.59 / 69.05	76.15	23.76	0.09
7	Sangker	6,052.78	2596.69 / 42.90	53.76	46.09	0.14
8	Sen	16,359.58	14129.33 / 86.37	85.37	14.61	0.02
9	Siem Reap	3,618.98	609.75 / 16.85	26.28	73.57	0.15
10	Sreng	9,986.27	6691.42 / 67.01	61.68	38.32	0.00
11	Staung	4,357.39	1878.31 / 43.11	75.03	24.97	0.00
	Total land area	83,106.70	41158.52 / 49.52	54.77	45.14	0.09
	Tonle Sap Lake	2,743.80	-	-	-	-

A limited number of rainfall and water-level gauging stations are located in each tributary basin. Most monitoring stations are situated in the downstream areas (Figure. 2.1), while gauges are sparsely located in the upstream areas, particularly in the northeastern part of the TSL Basin (Oeurng et al., 2019). Since a limited number of temperature monitoring stations are available in the TSL Basin, we used observed temperature data at five stations, including three located outside the basin to the east (Figure. 2.1). The daily observed rainfall and temperature data between 1985 and 2011, in which there is no missing value in these two datasets, were used as a benchmark to validate gridded meteorological data.

2.2.2. Gridded precipitation and temperature datasets

Among the available gridded precipitation and temperature datasets, we selected six datasets for precipitation and four datasets for air temperature (Table 2.2), namely APHRODITE, CFSR, TRMM, IMERG, ERA5, SA-OBS, and CPC, considering their fine temporal (daily) and spatial resolutions (0.1-0.5°) and potential applicability in Asian basins evaluated in previous studies (Amjad et al. (2020); Aslam et al., 2020; Fuka et al., 2014; Ge et al., 2019; Jiang et al., 2021; Lauri et al., 2014; Odusanya et al., 2018; Saha et al., 2014; Singh and Xiaosheng, 2019; Try et al., 2020; Vu et al., 2012; Xu et al., 2016b; Zhang et al., 2020a). A brief description of each dataset is given in the following sections and summarized in Table 2.2.

Table 2.2. Description of precipitation and temperature datasets used in the study

Parameter	Dataset*	Spatial resolution	Temporal resolution	Time-domain	Spatial domain	Data source	Source institution*
Precipitation	APHRODITE	0.25°	Daily	1951-2015	Asia	Observational	RISH, Kyoto University
	TRMM	0.25°	Daily	1998-Present	50°N-S	Satellite	NASA/GES DISC
	IMERG	0.1°	Daily	Jun-2000 - Present	60°N-S	Satellite	NASA/GES DISC
Precipitation & Temperature	CFSR	0.30°	Daily	1979-2014	Global land	Reanalysis	NCEP
	ERA5	0.25°	Hourly	1979-Present	Global land	Reanalysis	ECMWF
	SA-OBS	0.25°	Daily	1981-2017	South-east Asia	Observational	RNMI
Temperature	CPC	0.5°	Daily	1979-Present	Global land	Reanalysis	NCEP/CPC

*APHRODITE: Asian Precipitation - Highly-Resolved Observational Data Integration Towards Evaluation, TRMM: Tropical Rainfall Measuring Mission, IMERG: Integrated Multi-satellitE Retrievals for GPM (GPM = Global Precipitation Measurement), CFSR: Climate Forecast System Reanalysis, SA-OBS: Southeast Asia-Observational, RISH = Research Institute for Sustainable Humanosphere, NASA = National Aeronautics and Space Administration, GES DISC = Goddard Earth Sciences Data and Information Services Center, NCEP = National Centers for Environmental Prediction, ECMWF = European Centre for Medium-Range Weather Forecasts, RNMI = Royal Netherlands Meteorological Institute and CPC = Climate Prediction Center

2.2.2.1. APHRODITE

The APHRODITE rainfall dataset is primarily created with rainfall data observed via a rain-gauge-observation network over Asia (Yatagai et al., 2012). APHRODITE Version V1101 covers the period from 1951 to 2007, and its updated version (V1901) is available for the period between 1998 and 2015. We used both versions to cover the whole study period between 1985 and 2011, with daily temporal and spatial resolutions of 0.25° × 0.25°. There are a total of 13 rainfall stations in the TSL Basin used to create the APHRODITE dataset, some of which were not used in our study due to inconsistent measurement periods (Figure. A2). The performance of the APHRODITE dataset was compared to that of other products in the discharge modelling study for the Mekong River Basin (Lauri et al., 2014), which indicated that the computed discharge based on APHRODITE is similar to the observation data. Another study for the lower Mekong River Basin showed that the APHRODITE dataset was suitable for rainfall-runoff and flood inundation modeling (Try et al., 2018). Guan et al. (2020) evaluated six different precipitation datasets using four hydrological models to simulate streamflow over the Upper Yellow River Basin, China and found that the APHRODITE dataset was one of the superior products.

2.2.2.2. CFSR

CFSR was developed by the Environmental Modelling Center at the National Center for Environmental Prediction (NCEP). It is a fully coupled atmospheric model representing the interactions among Earth’s atmosphere, oceans, land, and sea-ice systems to provide the best prediction of the state of these coupled domains from 1979 to 2014 (Saha et al., 2014). The CFSR global atmospheric data have a spatial resolution of approximately 0.30° × 0.30° with daily resolution. CFSR Version 2 was used in this study. In addition, the CFSR dataset is commonly used in the SWAT modeling users’ group, as this dataset provides complete data for several climate variables (precipitation, air temperature, wind speed, solar radiation, and relative humidity) with a convenient data format that is compatible with SWAT model requirements (Zhang et al., 2020a). CFSR precipitation was found to reproduce an adequate flow in the Yangtze River Basin, China (Lu et al., 2019). The monthly coefficient of rainfall and air temperature determination was higher than 0.6 and 0.9, respectively, compared with the ground-based station data. Additionally, the air temperature of CFSR products could be a promising alternative data source to serve as input into the hydrological model for streamflow simulation in the data-scarce river basins (Zhang et al., 2020b).

2.2.2.3. TRMM

TRMM is a joint project between the North American Space Agency (NASA) and the Japan Aerospace Exploration Agency (JAXA) to study rainfall over the tropical region. The TRMM precipitation is estimated from a combination of multiple satellites and rain-gauge analyses at a fine spatial resolution of 0.25° × 0.25° (Huffman et al., 2007). The 3B42V7 product was adjusted by monthly rain gauge precipitation data from the GPCC and is superior to the TRMM 3B42RT V7. The number of rain gauges used in GPCC data is documented in earlier publications by Vu et al. (2018); however, there is only one station in TSL Basin that was used in interpolating the GPCC dataset during the study period. The TRMM-3B42V7 dataset from 1998 to 2011 with the daily resolution was used in this study. TRMM-3B42V7 precipitation data were compared with observed rainfall at 342 rain gauges over Malaysia (Tan et al., 2015). There was a tendency for TRMM-3B42V7 to yield accurate and unbiased rainfall estimations. TRMM, the Climate Prediction Center Morphing algorithm (CMORPH), and Global Satellite Mapping of Precipitation (GSMaP) Reanalysis were evaluated via a comparison with the rain gauge-based data over Central Vietnam (Trinh-Tuan et al., 2019), and TRMM was found to outperform the other datasets. Moreover, the TRMM precipitation was determined to be among the best-performing products that gave better streamflow simulation results than the other four gridded datasets over the Upper Yellow River Basin, China (Guan et al., 2020). Nhi et al. (2019) evaluated five gridded rainfall datasets in simulating streamflow and found that TRMM was one of the best datasets showing a better match to rain gauge data in simulating the streamflow in the upper Dong Nai river basin, Vietnam.

2.2.2.4. IMERG

Following the success of TRMM, NASA and JAXA jointly launched the GPM observatory satellite on 28 February 2014 to provide more accurate precipitation data with fine spatiotemporal resolution using an algorithm combining information from the GPM satellite constellation from partner nations (Huffman et al., 2020; Liu, 2016). Recently, IMERG version 6 (V06) extended its temporal coverage to the TRMM era and provides datasets from June-2000 to the present. GPM not only extended TRMM's spatial coverage but also carries advanced sensors such as the Dual-Frequency Precipitation Radar (DPR) and GPM Microwave Imager (GMI), which quantify precipitation more accurately, particularly for light and solid precipitation (Huffman et al., 2020). Like TRMM, monthly rain gauge precipitation data from the GPCC were used to adjust IMERG precipitation. Furthermore, there are three types of IMERG products. The Early run (IMERG-E) and Late run (IMERG-L) are available with a latency of 4 hours and 14 hours respectively (Huffman et al., 2020), which could serve as a potential data source for flood forecasting and real-time disaster management. The Final run (IMERG-F) is available approximately 4 months after observation, and it is mainly aimed for research purposes. The IMERG-F V06 dataset from 2001 to 2011 with the daily temporal resolution and spatial resolution of 0.1° × 0.1° was used in this study. Lu et al. (2021) found that IMERG-F precipitation has a great potential for application to hydroclimatic research and water resources assessment in the complex climatic and topographic conditions, in Western China. Generally, the IMERG-F shows high accuracy and good performance in hydrological simulations with a high correlation coefficient (0.63) and low relative bias (0.92%) when compared with the ground observations (Wang et al., 2017). Gunathilake et al. (2020) statistically and hydrologically evaluated the capacity of different six satellite-based precipitation products in the Upper Nan River Basin, Northern Thailand, showing that IMERG-precipitation demonstrates good performance for future hydrometeorological applications in this region. Huang et al. (2019) evaluated the performance of IMERG-F during six extremely heavy precipitation events caused by powerful typhoons from 2016 to 2017, and they stated that the IMERG-F shows good capability in extreme rainfall events (i.e., typhoons, rainfall storm) applications.

2.2.2.5. ERA5

ERA5 is the fifth-generation atmospheric reanalysis of the global climate by the European Centre for Medium-Range Weather Forecasts (ECMWF), and it provides hourly estimates of a large number of atmospheric, land, and oceanic climate variables from 1979 to the present with a spatial resolution of 31 km (0.25°) (Hersbach et al., 2020). Tarek et al. (2020) found that the performance of hydrological modelling based on ERA5 precipitation is equivalent to that based on the observed rainfall. Jiao et al. (2021) showed that ERA5 data is in good agreement with the annual and seasonal patterns of observed precipitation in China and have correlation coefficient values ranging from 0.79 to 0.94. ERA5 temperature data was found to be consistent with the monthly observations, and the trend of ERA5-based temperature and observation data showed a strong linear relationship with the monthly observations (Zhu et al., 2021).

2.2.2.6. SA-OBS

SA-OBS is an observational gridded dataset for precipitation and temperature in the Southeast Asia region based on the Southeast Asian Climate Assessment & Dataset (SACA&D) project (Van den Besselaar et al., 2017). The full dataset is estimated at a 0.25° × 0.25° spatial resolution from 1981 to 2017. Since SA-OBS is a newly developed daily gridded dataset for Southeast Asia, there are limited studies to evaluate the performance of this dataset. The daily precipitation and temperature data were collected from the meteorological institutes of Australia, Indonesia, Malaysia, the Philippines, Singapore, Thailand, and Vietnam. There are 1393 precipitation stations, including 365 stations with minimum and maximum temperatures and 274 stations with a daily mean temperature (Van den Besselaar et al., 2017). However, during the creation of SA-OBS, no gauged rainfall data were introduced from the TSL Basin or Cambodia; therefore, no data accuracy for this area has been clarified. Detailed information on gauged stations used in SA-OBS data generation can be found in the study by Van den Besselaar et al. (2017). A high correlation was found between station observations and the corresponding grid squares of SA-OBS (Van den Besselaar et al., 2017), while SA-OBS slightly underestimated high rainfall events (Singh and Xiaosheng, 2019).

2.2.2.7. CPC

The CPC dataset has been developed by the American National Oceanic and Atmospheric Administration (NOAA), using optimal interpolation of quality-controlled gauge records of the Global Telecommunication System (GTS) network (Fan and Van den Dool, 2008), to provide global gridded daily temperature datasets covering the period from 1979 to the present with a spatial resolution of 0.5° × 0.5°. Like SA-OBS, the gauged data was not introduced from the TSL Basin for CPC data creation, and more information on gauged stations used in interpolating CPC data is provided in the study by Vu et al. (2018). The daily hydrological evaluation results suggested that the CPC temperature combined with precipitation data is a more accurate option for estimating hydrological performance in the

Upper Mekong River Basin (Dinh et al., 2020). A similar or lower performance compared to that of other gridded datasets was observed in the downstream areas. Owing to their reliability, CPC temperature data were used as the reference data to assess the impact of climate change on the hydrology of a transboundary river between Bhutan and India (Zam et al., 2021).

2.2.3. Performance evaluation of gridded datasets of rainfall and temperature

The performance of individual gridded precipitation and temperature products was evaluated by comparing them to gauge-based datasets on the annual, monthly, and daily timescales from 1985 to 2011 for APHRODITE, ERA5, CFSR, CPC, and SA-OBS, and from 1998 or 2001 to 2011 for TRMM and IMERG, respectively. The bilinear interpolation method (Figure A3) was used to interpolate the gridded datasets to the 31 rainfall and five temperature ground-based station locations, based on the distance-weighted average of the four nearest grid values. Several previous studies have applied this grid-to-point interpolation method to compare gridded products with observation data (Bao and Zhang, 2013; Bromwich and Fogt, 2004; Ebrahimi et al., 2017; Mayor et al., 2017; Meher and Das, 2019). This simplified method with two-dimensional interpolation enables a more accurate comparison than the direct comparison between point observations and gridded data (Caroletti et al., 2019; Uddin et al., 2008).

The accuracy of gridded datasets was quantified using four statistical indicators, listed in Table A1, i.e., mean bias error (MBE), root mean square error (RMSE), correlation coefficient (*R*), and modified index of agreement (*md*). MBE is a statistical index to assess the mean difference between two data products (interpolated and observed rainfall or temperature), and RMSE represents the standard deviation of the two datasets. The lower absolute values of MBE and RMSE indicate the better performance of the gridded dataset. *R* measures the degree of linear correlation between the interpolated and observed datasets. *md* is used as a standardized measure of the degree of model prediction error, which is the modified version of the index of agreement proposed by Willmott (1981) to detect additive and proportional differences in the data (Ahmed et al., 2019; Pereira et al., 2018; Willmott, 1984). The higher values of *R* and *md* (closed to one) indicate the higher accuracy of gridded products in estimating precipitation and air temperature. The equation, range, and optimal value of each statistical index are presented in Table A1.

2.2.4. Performance evaluation based on the hydrological model application

The comparison of hydro-meteorological data at individual locations has a particular limitation in assessing the basin-wide applicability of gridded datasets for hydrological simulations, mainly if the gauging stations are sparsely distributed within an area like the TSL Basin. Therefore, we evaluated the performance of selected gridded datasets in reproducing the streamflow of each tributary basin based on the SWAT model application, and the simulated daily streamflow was compared with the observed data. Furthermore, considering the equifinality issue in the hydrological model application under different combinations of gridded datasets and model parameters, we compared the simulated and satellite-based actual ET to confirm the accuracy of each dataset from the perspective of water balance as described in section 2.2.4.3.

2.2.4.1. Hydrological SWAT model

Quantitative representation of rainfall and temperature by each gridded dataset was evaluated using the SWAT model, a widely used hydrological model to simulate streamflow, sediments, and other hydrological variables under diverse environments and scales (Arnold et al., 1998) (Figure 2.2). SWAT has gained international acceptance as a comprehensive watershed modeling tool (Gassman et al., 2007) and has been applied throughout the world, including the United States (Yuan and Forshay, 2021), Europe (Bärlund et al., 2007), Cambodia (Ang and Oeurng, 2018), Thailand (Shrestha et al., 2018), Lao PDR (Vilaysane et al., 2015), Vietnam (Vu et al., 2012), China (Hao et al., 2004), India (Kaur et al., 2004), Australia (Sun and Cornish, 2006), and Africa (Schuol and Abbaspour, 2006) among many others. No matter what type of problem is studied with SWAT, water balance is the driving force behind everything that happens in the watershed. To accurately predict the movement of pesticides, sediments, or nutrients, the hydrologic cycle as simulated by the model must conform to what is happening in the watershed. Simulation of the hydrology of a watershed can be separated into two major divisions. The first division is the land phase of the hydrologic cycle controls the amount of water, sediment, nutrient and pesticide loadings to the main channel in each subbasin. Figure 2.3 shows the general sequence of processes used by SWAT to model the land phase of the hydrological cycle. The second division is the water or routing phase (Figure 2.4) of the hydrologic cycle which can be defined as the movement of water, sediments, etc. through the channel network of the watershed to the outlet. Once SWAT determines the loadings of water, sediment, nutrients and pesticides to the main channel, the loadings are routed through the stream network of the watershed.

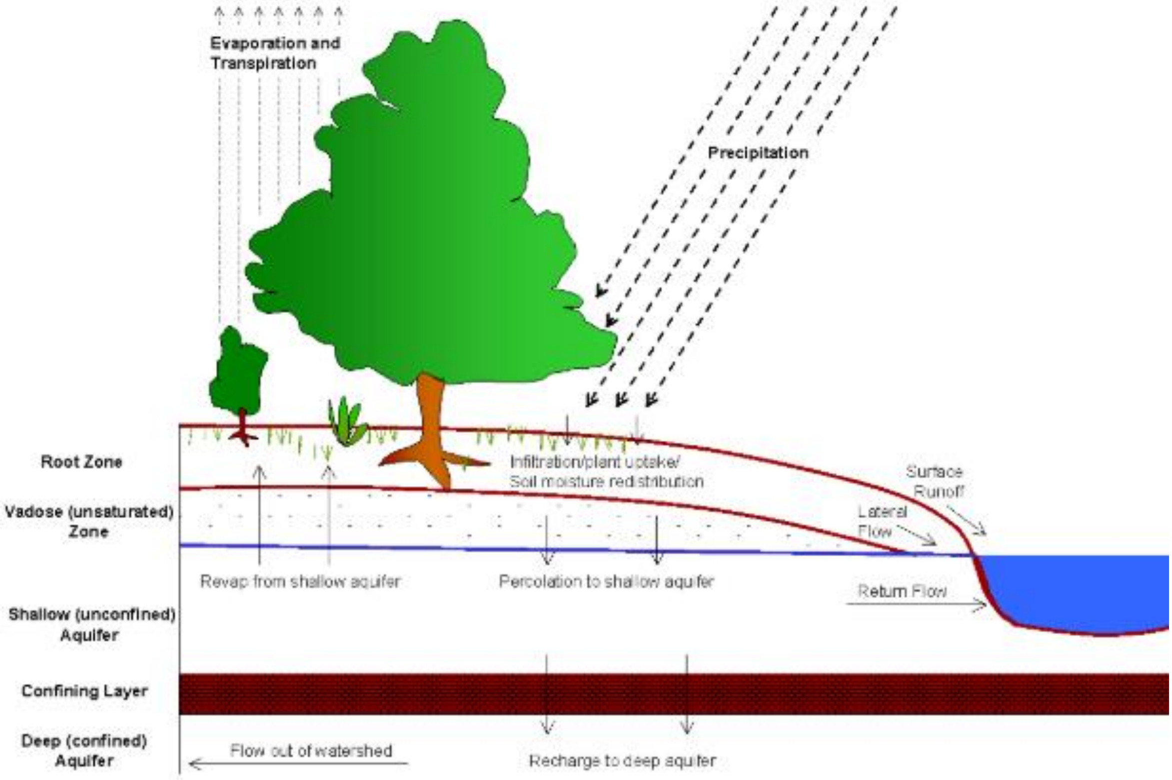


Figure 2.2. Schematic representation of the hydrologic cycle in the SWAT model (Arnold et al., 1998).

SWAT requires topography, soil, and land-use information to account for heterogeneities and divide the greater basin into a group of sub-basins, which are further subdivided into several Hydrological Response Units (HRUs). Each HRU is created from a unique combination of land cover, slope, soil properties, and management practices. The surface runoff from each HRU is simulated based on the SCS curve number method (USDA, 1972), which requires a watershed coefficient (i.e., Curve Number (CN)) with daily rainfall defined at the centroid of the sub-basin as input. Detailed information on the SCS curve number method could be found in the official theoretical documentation (Neitsch et al., 2011). The rainfall is obtained by applying the bilinear interpolation method to the specific daily gridded precipitation dataset, instead of the default SWAT model setting, which only uses rainfall data from the gauging station located closest to the centroid of each sub-basin (Arnold et al., 1998). Since the gridded precipitation data cannot be directly input into the SWAT model, the approach using interpolated areal precipitation was used, and some studies also found the performance improvement of model calibration compared to the default method of the SWAT model (Abbas and Xuan, 2020; Masih et al., 2011). More details about SWAT can be found in the official theoretical documentation (Neitsch et al., 2011) and the review paper by Gassman et al. (2007).

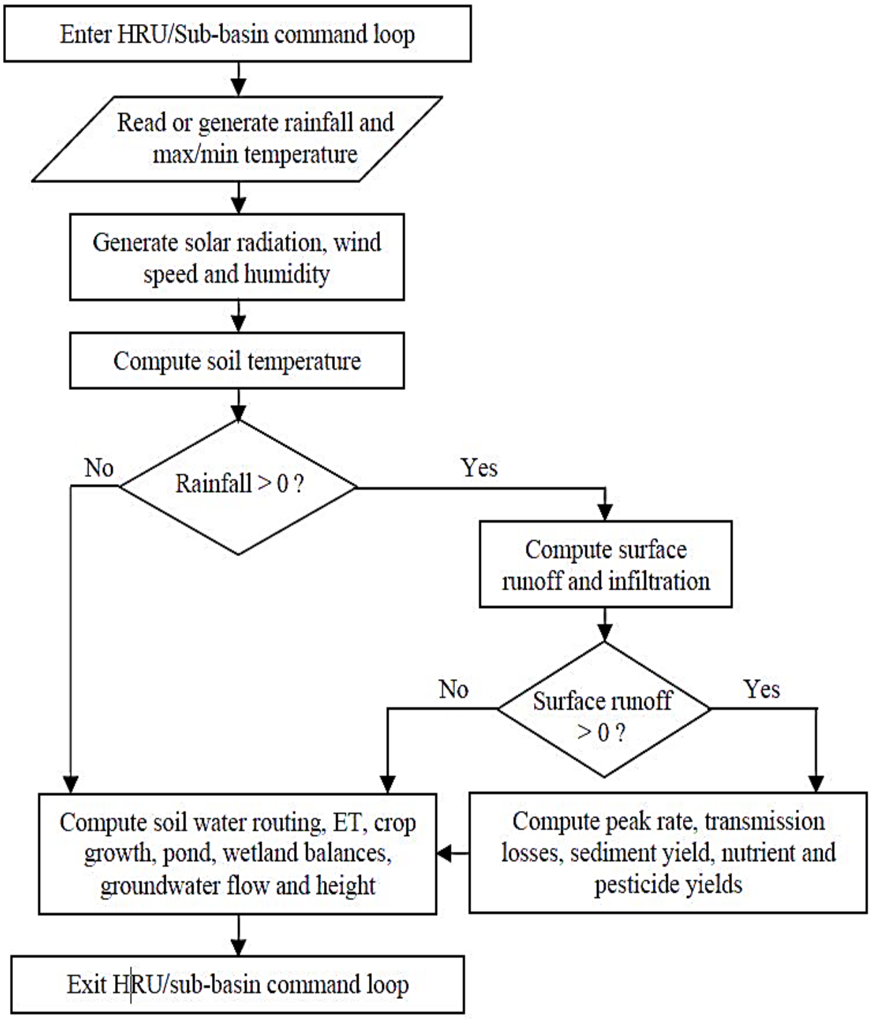


Figure 2.3. The general sequence of processes used by SWAT to model the land phase of the hydrologic cycle (Arnold et al., 1998)

Each dataset used for the SWAT model setup is summarized in Table 2.3. The topography data required to define flow accumulation, stream networks, and basin boundaries are based on a 90 m × 90 m resolution digital elevation model (DEM) from the NASA Shuttle Radar Topographic Mission (SRTM). Land-use distribution was obtained from the Mekong River Commission (MRC), which developed a land-use map based on satellite imagery collected from 1993 to 1999. Information on soil types was obtained from a 1:250,000 scale map created by the MRC that contains the required physical properties of each soil type based on the Food and Agriculture Organization/United Nations Educational, Scientific and Cultural Organization (FAO/UNESCO) 1988 classification scheme (Oeurng et al., 2019). The daily rainfall and maximum/minimum air temperature data were obtained from individual gridded datasets. The streamflow in each tributary basin was calculated using rating curves developed by Kummu et al. (2014) from the observed water levels collected by the Department of Hydrology and River Works, Ministry of Water Resources, and Meteorology of Cambodia (Oeurng et al., 2019).

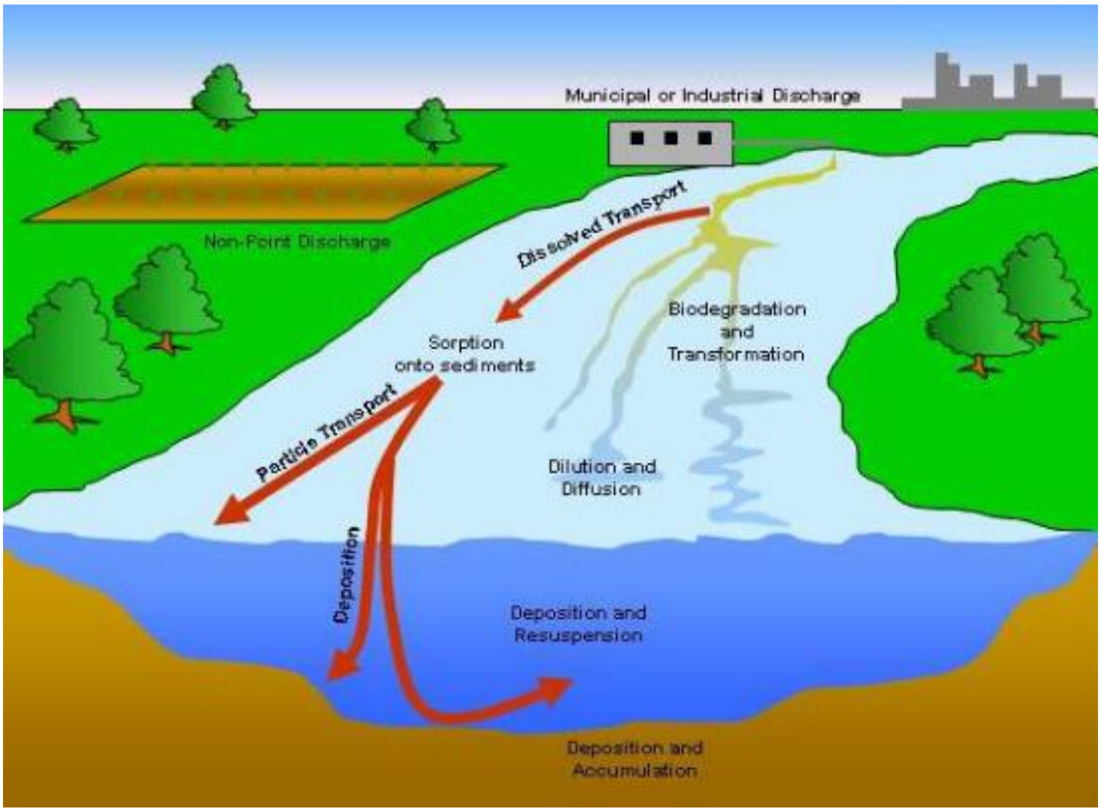


Figure 2.4. In-stream processes are modeled by the SWAT model (Arnold et al., 1998).

Table 2.3. Summary of required data used for SWAT model setup

Data type	Time period	Temporal resolution	Spatial Resolution	Sources
Topography (DEM)	-	-	90m	http://srtm.csi.cgiar.org
Land cover map	2002	-	250m	Mekong River Commission
Soil types map	2002	-	250m	Mekong River Commission
Total rainfall Maximum and minimum air temperatures	1985-2011	Daily	-	Gridded datasets
Water level	1995-2011	Daily	-	Department of Hydrology and River Works (Cambodia)

2.2.4.2. Model calibration, validation, and evaluation

Model calibration was performed with the SWAT-Calibration and Uncertainty Programs (SWAT-CUP) (Abbaspour, 2015), and the performance of each gridded dataset in the hydrological model application was evaluated for the calibration and validation periods. Model calibration involves the modification of parameter values and comparison of the predicted output of interest to measured data until a defined objective function (i.e., Nash-Sutcliffe efficiency (NSE), percent bias (PBIAS), or the coefficient of determination (R^2)) or optimum output is achieved. The purpose of model validation is to establish whether the calibrated model can predict the output compared to observed data for independent periods without making further adjustments to parameters that may adjust during the calibration process. Figure 2.5 shows that SWAT-CUP provides several methods for calibration to deal with parameter uncertainty, including the Sequential Uncertainty Fitting version 2 (SUFI-2) method (Nkonge et al., 2014; Wu and Chen, 2015), which was selected to provide more reasonable daily streamflow simulations at the outlet of each delineated drainage area (Table 2.1 and Figure. A1). Considering the availability of continuous discharge data in each tributary basin, the model parameters were calibrated with available streamflow data between 1995 and 2002, and simulation results were validated for the period between 2003 and 2011. Based on previous studies (Ang and Oeurng, 2018; Oeurng et al., 2019; Sao et al., 2020; Sok et al., 2020; Vilaysane et al., 2015; Vu et al., 2012), 23 parameters were selected and calibrated using the SWAT-CUP model (Table A2) for each gridded precipitation dataset.

For evaluating model performance in watershed simulations, we employed three quantitative metrics, i.e., Nash-Sutcliffe efficiency (NSE), percent bias (PBIAS), and coefficient of determination (R^2), as shown in Table A1 (Moriasi et al., 2007). NSE indicates how well the result of a simulation fits the observed time series (Nash and Sutcliffe, 1970). PBIAS measures the average tendency of the simulated values to be larger or smaller than their observed counterparts (Gupta et al., 1999). R^2 examines to what extent differences in one variable can be explained by the difference in a second variable. In other words, R^2 assesses how strong the linear relationship is between two variables. Furthermore, the performance ratings suggested by Moriasi et al. (2007) were applied to evaluate the SWAT model simulation's integrated performance (Table A3).

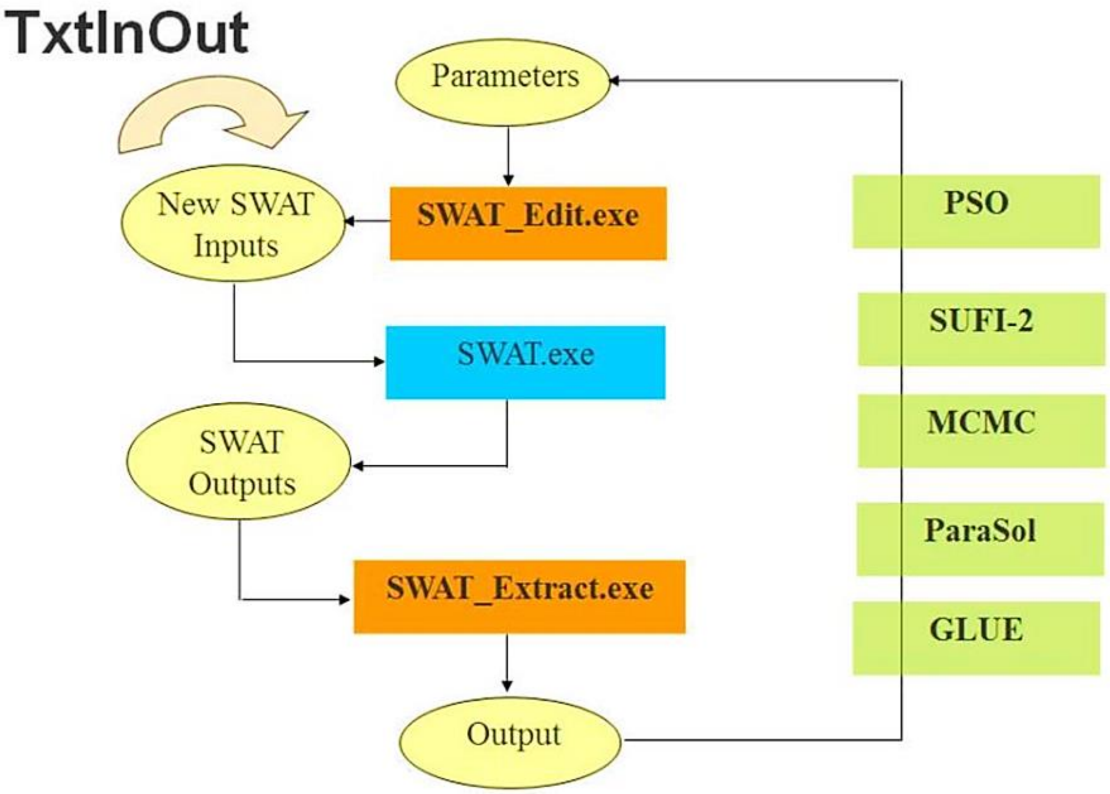


Figure 2.5. The general framework of SWAT-CUP. There are five different methods of parameter uncertainty: Sequential Uncertainty Fitting Version 2 (SUFI-2), Particle Swarm Optimization (PSO), Generalized likelihood Uncertainty Estimation (GLUE), Parameter Solution (ParaSol), and Markov Chain Monte Carlo (MCMC).

2.2.4.3. Evapotranspiration comparison

SWAT determines the actual evapotranspiration (ET) from potential evapotranspiration (PET) using one of the three built-in empirical equations (Neitsch et al., 2011)—Penman-Monteith (Allen et al., 1989; Monteith, 1965), Priestley-Taylor (Priestley and Taylor, 1972) and Hargreaves (Hargreaves and Samani, 1985). These three methods require different input variables. The Penmen-Monteith method requires solar radiation, air temperature, relative humidity, and wind speed, and the Priestley-Taylor method needs solar radiation, air temperature, and relative humidity. Considering limited data availability, we used the Hargreaves method, which requires precipitation and air temperature only (Neitsch et al., 2011). Once the PET is determined, ET must be calculated. SWAT estimates total ET from four components: canopy evaporation, transpiration, soil evaporation, and Revap-ET (Abiodun et al., 2018). SWAT models the movement of water into overlying unsaturated layers as a function of water demand for ET. This process has been termed “Revap” to avoid confusion with soil evaporation and transpiration (Neitsch et al., 2011). As SWAT calculates the Revap separately, the Revap should be added for ET calculations (Abiodun et al., 2018; Fenta Mekonnen et al., 2018). More detailed information and a comprehensive outline of ET calculations in the SWAT model can be found in the theoretical manual of Neitsch et al. (2011) and the studies by Abiodun et al. (2018) and Fenta Mekonnen et al. (2018).

The spatio-temporal variation of ET was obtained through SWAT model estimations driven by different precipitation products. Then, we compared this information to satellite-based ET data for a more comprehensive evaluation of the gridded precipitation datasets. For this purpose, the SWAT model was applied to simulate daily ET together with streamflow for the whole drainage area of each tributary basin with the parameters previously calibrated for the delineated area of each tributary basin from 2001 to 2011. The simulated ET based on different gridded rainfall products was temporally and spatially aggregated to compare with satellite-based ET, i.e., ET from MODIS (Moderate Resolution Imaging Spectroradiometer) and GLEAM (Global Land Evaporation Amsterdam Model), at monthly and annual timescales for the whole basin, excluding the flooded areas.

Satellite-based ET has been used to validate the application of hydrological models and assess water balance closure in several studies (Dash et al., 2021; Jin and Jin, 2020; López et al., 2017; Odusanya et al., 2018; Sirisena et al., 2020; Wong et al., 2021), but the satellite ET is rarely used for validating the gridded rainfall products. MODIS-ET provides evaporation from global terrestrial surfaces and transpiration from the vegetation canopy (Mu et al., 2011). The algorithm to estimate ET from MODIS data was initially developed by Cleugh et al. (2007). The MODIS-ET algorithms are based on the Penman-Monteith equation (Monteith, 1965). We used MOD16A2, which is available at an 8-day interval and a spatial resolution of 1 km over a period extending from 2000 to 2014. Mu et al. (2013) compared the MODIS-ET with observed ET across 46 sites from 2000 to 2006 and found that the average ET biases and mean absolute errors between ET observations and MODIS-ET are -0.11 mm and 0.33 mm, respectively. Similarly, Kim and Hogue (2008) validated MODIS-ET at a point-scale by comparing MODIS-ET with measured ET at four monitoring stations from 2001 to 2004, and results showed that at the daily-time step MODIS-ET show a good correlation (R = 0.89) and low bias (0.34 mm) with measured ET. The actual GLEAM-ET data are estimated as the sum of evaporation and ET from three different land surface types, i.e., bare soil, short vegetation, and vegetation with a tall canopy (Miralles et al., 2011). The Priestley and Taylor equation (Priestley and Taylor, 1972) was adopted to estimate evaporation in the GLEAM-ET algorithms. We used the daily ET of the GLEAM v3.2a dataset, spanning 41 years from 1980 to 2020 with a spatial resolution of 0.25° × 0.25°. The GLEAM-ET was validated at 91 in-situ sites in the United States and Australia between 2011–2015, and it was found that the mean correlation coefficient and root mean square difference between GLEAM- and gauge-ET are 0.78 and 0.71mm, respectively (Martens et al., 2017). The GLEAM-ET was also validated in China using in-situ measurements at three different timescales. The results showed a high average correlation coefficient with in-situ observations at all sites, on daily (R = 0.71), monthly (R = 0.86), and annual (R = 0.79) time scales (Yang et al., 2017).

2.3. Results and discussion

2.3.1. Performance of precipitation datasets

The results of all statistical indices (MBE, RMSE, R, and md) applied for each station are presented in Figures 2.6 (a)-(x). The positive mean values of MBE in CFSR, ERA5, SA-OBS, TRMM and IMERG indicate overestimation, while APHRODITE underestimated the observed data with a negative mean value of MBE (Table A4). The lower values of RMSE derived from APHRODITE (Figure 2.6 (g)), ERA5 (Fig. 2.6 (i)), TRMM (Figure 2.6 (k), and IMERG (l)) indicate that these datasets have a better performance than the other two products, except for the locations in the southwestern and western regions in the TRMM and IMERG, which are attributed to the overestimated rainfall over these areas. Figure 2.6 (m)-(x) shows a better performance regarding the correlation coefficient (R) and modified index of agreement (md) in APHRODITE, ERA5, TRMM and IMERG-based rainfall.

Figure 2.7 shows the mean monthly rainfall between 2001 and 2011 in each tributary basin, which was calculated by averaging the monthly rainfall observed or interpolated (bilinear interpolation) at all rain-gauge locations in the tributary basin. ERA5 and APHRODITE and the observed results showed a similar seasonal pattern to each other. CFSR tended to overestimate the mean monthly rainfall for most tributary basins throughout the year, and SA-OBS overestimated the mean monthly rainfall during the wet season (Figures 2.7 and A4). The TRMM and IMERG precipitation values overestimated the rain gauged values in four tributary basins (Dauntri, Mongkol Borey, Pursat, and Sangker), which is consistent with the higher spatial gradient of rainfall over the southwestern terrain (Figure. 2.9).

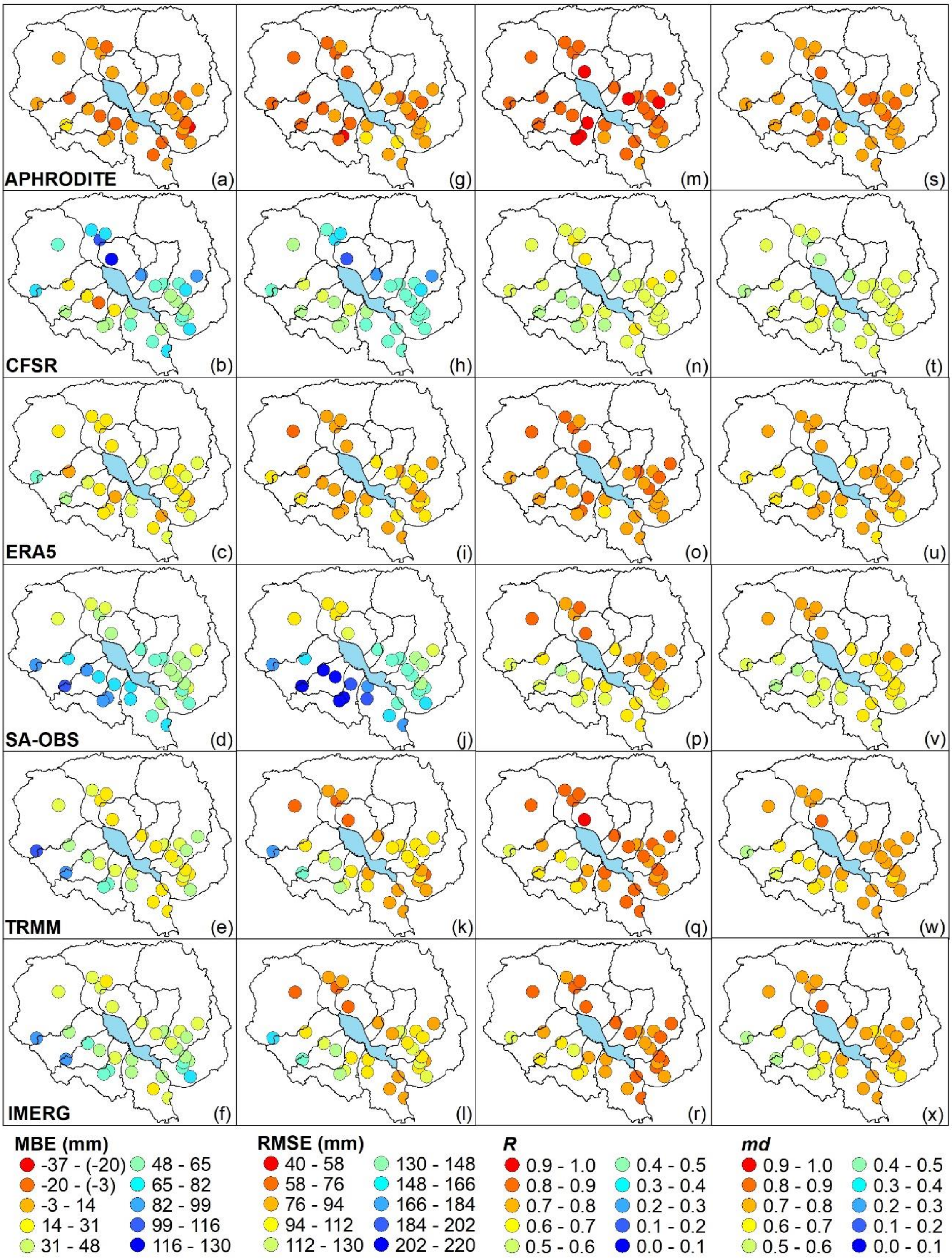


Figure. 2.6. Spatial distribution of mean bias error (a-f), root mean square error (g-l), correlation coefficient (m-r), and modified index of agreement (s-x) between gridded products and gauged observation at monthly time scale during 19985 – 2011 (1998 and 2001– 2011 for TRMM and IMERG, respectively). The circles denote rainfall gauging stations located in the TSL Basin.

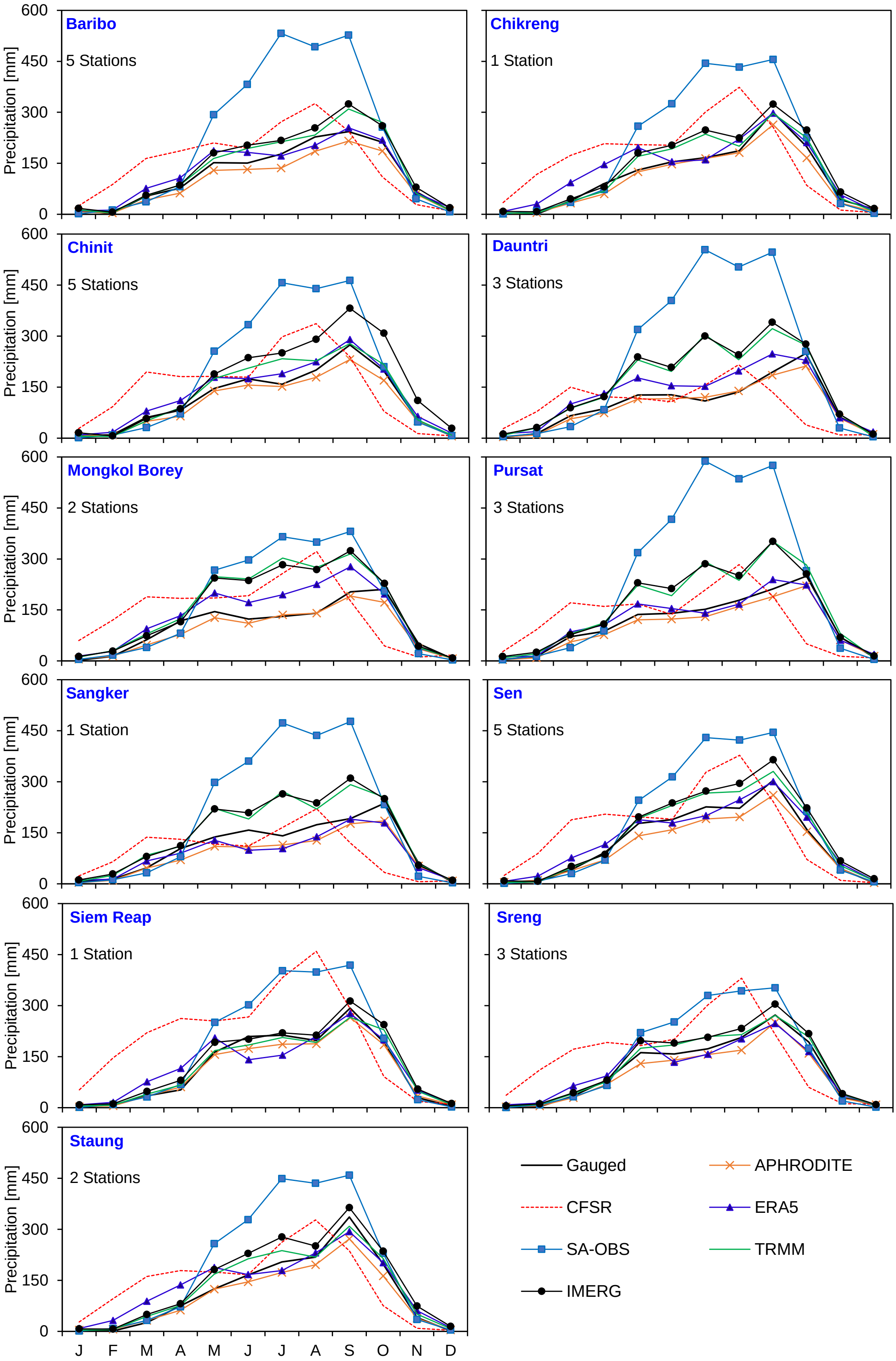


Figure 2.7. Comparison of long-term mean monthly rainfall averaged for all the gauging stations located in each tributary basin from 2001 to 2011. The number of gauging stations used for averaging monthly rainfall was included in each plot.

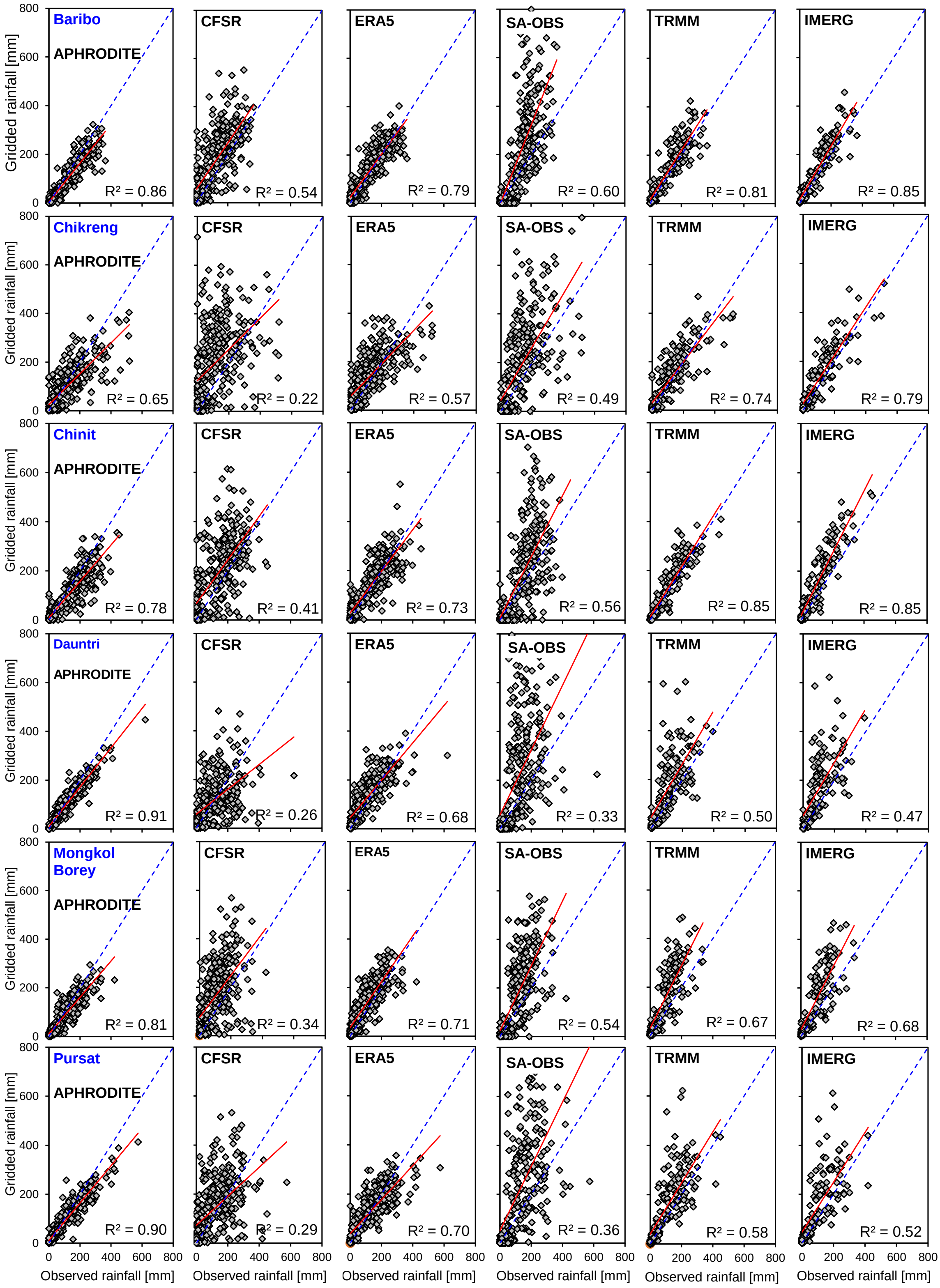


Figure 2.8. Scatter plots of observed and interpolated monthly rainfall in each tributary basin from 1985 to 2011 (1998 and 2001 to 2011 for TRMM and IMERG, respectively). For the basin with multiple gauging stations (Figs. 2.6 and 2.7), the monthly rainfall was obtained as the arithmetic mean for all gauging stations located in the basin. The blue dash lines are 1:1 line, and the red solid lines denote linear regressions.

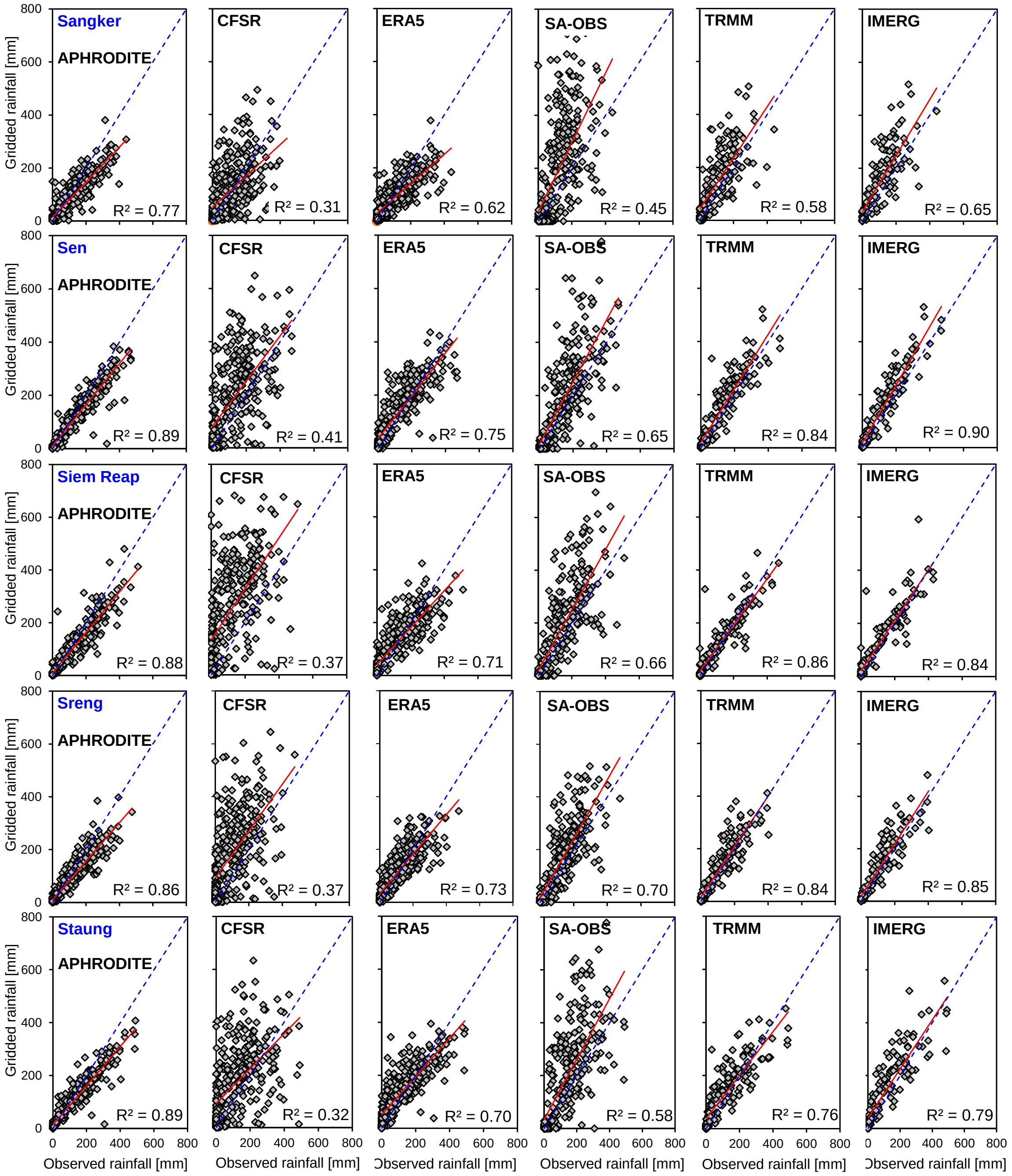


Figure 2.8. *Cont.*

Figure 2.8 shows the scatter plots between the interpolated monthly precipitation of each gridded product and the gauged observations for each tributary basin. The higher values of coefficient of determination (i.e., R^2) in APHRODITE, TRMM, IMERG and ERA5 indicate that these four datasets have a strong linear relationship with the observed monthly data, while CFSR and SA-OBS show lower correlations with the observed data. TRMM and IMERG show a good relationship with the observation data in eight tributary basins (the monthly R^2 values are higher than 0.65), while the other three basins located over the southwestern terrain of the TSL (i.e., the Dauntri, Pursat, and Sangker tributary basins) exhibit R^2 values between 0.5 and 0.58 for TRMM and between 0.47 and 0.65 for IMERG. As seen in the scatter plots, IMERG outperformed TRMM in some tributaries, while the other tributaries showed equal or smaller correlation coefficients (Fig. 2.8).

2.3.2. Spatial distribution of mean annual precipitation

The spatial distribution of mean annual precipitation from 2001 to 2011 over the TSL Basin is shown using the respective gridded dataset (Figures 2.9 (a)-(f)) and data from all 31 rain gauging stations (Figure 2.9 (g)).

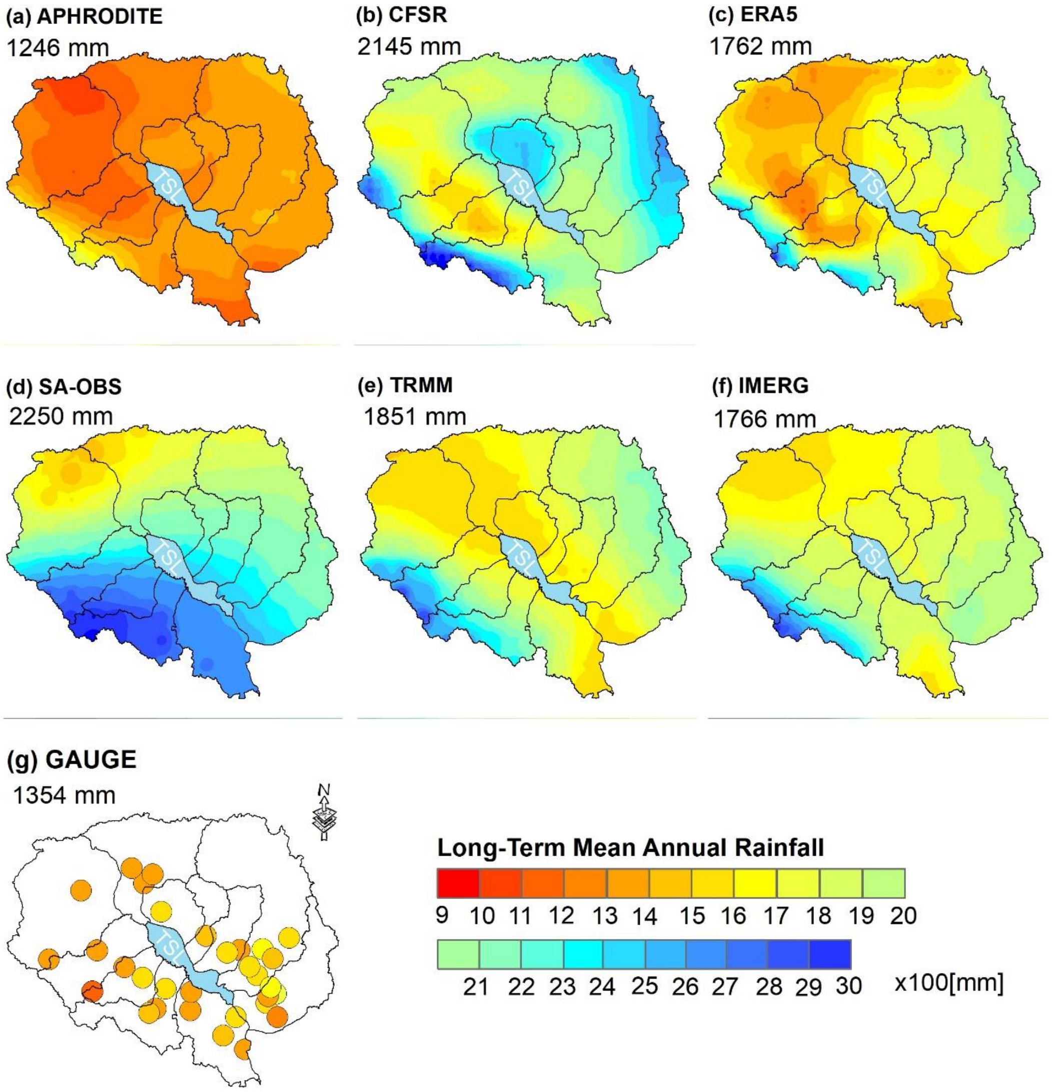


Figure 2.9. Spatial distribution of mean annual rainfall of APHRODITE (a), CFSR (b), ERA5 (c), SA-OBS (d), TRMM (e), IMERG (f) and gauged observation (g) over the TSL Basin for the period from 2001 to 2011. The areal average (for gridded datasets) and the arithmetic mean (for gauging stations) of the mean annual precipitation were included at the top left corner of each map.

The areal average of mean annual precipitation was 1246, 2145, 1762, 2250, 1851 and 1766 mm for APHRODITE, CFSR, ERA5, SA-OBS, TRMM and IMERG, respectively. There were substantial differences in the areal average and spatial distribution of the mean annual precipitation between the six gridded products. APHRODITE revealed the smallest mean annual rainfall over the TSL Basin, while ERA5, TRMM and IMERG showed comparably similar annual rainfall and spatial distribution with larger rainfall over the hilly regions in the south and northeast, which are not visible in the observations (Figure 2.9 (g)), and smaller rainfall over the lower elevation zone extending from the northwest to the southeast, including the TSL.

Although the overall spatial distribution of mean annual rainfall is similar between TRMM and IMERG, the rainfall at the edge of the basin in the south and northeast areas was noticeably smaller in IMERG than that of TRMM (Figure 2.8), which could be a source of improvement in correlation coefficient between gauged and IMERG precipitation in some tributaries (Figure 2.8). Despite both APHRODITE and SA-OBS datasets being based on the observation data, their spatial distributions are very different and reflect the different number of gauged datasets used to create them.

2.3.3. Performance of temperature datasets

Figure 2.10 illustrates the comparison values of mean monthly averaged daily maximum and minimum temperatures derived from the results of bilinear interpolation of gridded data and those derived from the data of gauging stations from 1985–2011. Overall, the SA-OBS dataset displayed better estimates of seasonal patterns and magnitudes for both minimum and maximum temperatures (Figures 2.10 and A5). CFSR products exhibited huge differences in magnitude for both maximum and minimum temperatures, especially between April and December. ERA5, CPC, SA-OBS, and the observed results showed overall similarity in their seasonal patterns for minimum temperature (Figures 2.10 and A5).

The statistical indices evaluating the performance of maximum and minimum temperatures at the monthly timescale are given in Table 2.4. The results of four different indices for individual stations and their mean values are presented separately for maximum and minimum temperatures. CFSR maximum and minimum temperatures had the poorest performance compared to other gridded products, as typically indicated by a large RMSE.

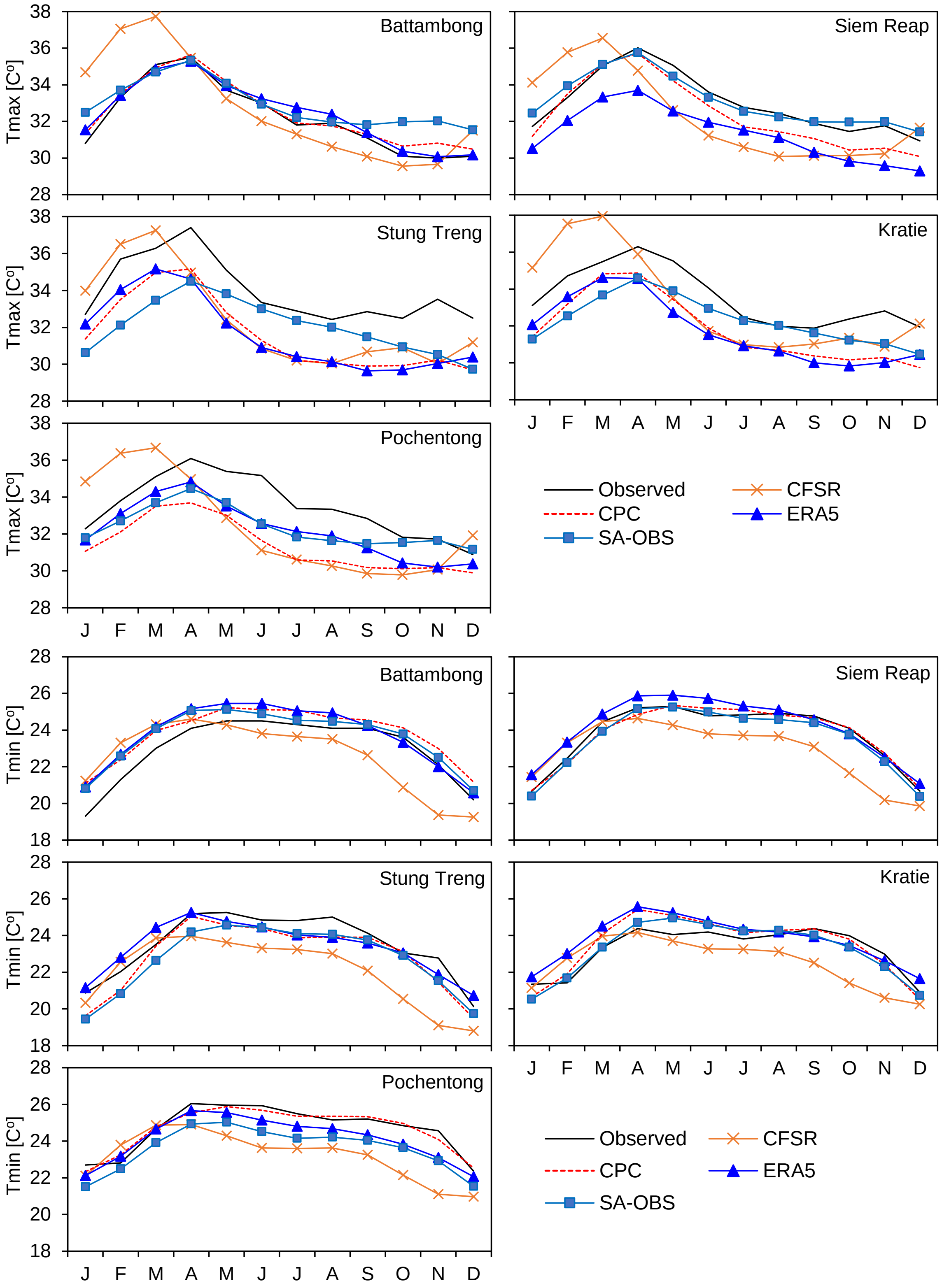


Figure 2.10. Comparison of mean monthly averaged daily maximum (top) and minimum (bottom) temperatures at individual stations from 19985 – 2011.

The CPC and SA-OBS minimum temperatures showed the highest mean correlation coefficients, which were equal to 0.84 (Table 2.4). SA-OBS maximum temperature values outperformed the CPC products, as indicated by the mean values of all indices (Table 2.4).

Therefore, the SA-OBS product is the most favorable alternative temperature data source over the TSL Basin.

From the statistical analysis, the CPC and SA-OBS-based air temperature performed comparably well with the gauged data. This result was consistent with the previous work (Van den Besselaar et al., 2017) that found a high correlation between SA-OBS and APHRODITE (gauge-based dataset) for mean temperature over Southeast Asia. Similarly, Khadka et al. (2021) indicated that CPC shows a high correlation with gauged data and well captures the variation of maximum and minimum temperature in northeast Thailand.

Table 2.4. Performance of statistical indices for maximum and minimum temperatures at the monthly time scale

Station	Dataset	Tmax				Tmin			
		MBE	RMSE	*R*	*md*	MBE	RMSE	*R*	*md*
Battambong	CFSR	0.54	2.66	0.63	0.52	-0.36	1.85	0.56	0.55
	CPC	0.26	1.17	0.83	0.72	0.81	1.09	0.91	0.72
	ERA5	0.26	1.17	0.83	0.72	0.74	1.18	0.87	0.70
	SA-OBS	0.70	1.30	0.80	0.64	0.64	0.98	0.90	0.76
Siem Reap	CFSR	-0.67	2.66	0.44	0.42	-0.89	1.74	0.65	0.50
	CPC	-0.68	1.91	0.69	0.62	-0.01	0.94	0.86	0.76
	ERA5	-1.68	2.04	0.77	0.48	0.41	1.12	0.82	0.72
	SA-OBS	0.10	1.05	0.79	0.69	-0.22	0.92	0.87	0.76
Stung Treng	CFSR	-1.52	2.96	0.59	0.45	-1.43	2.08	0.68	0.42
	CPC	-2.34	2.65	0.83	0.45	-0.65	1.13	0.88	0.70
	ERA5	-2.32	2.70	0.77	0.44	-0.13	1.00	0.82	0.69
	SA-OBS	-1.89	2.36	0.64	0.45	-0.78	1.12	0.90	0.70
Kratie	CFSR	-0.31	2.72	0.54	0.46	-0.72	1.82	0.40	0.40
	CPC	-1.74	2.39	0.68	0.50	0.21	1.27	0.71	0.61
	ERA5	-1.81	2.29	0.71	0.47	0.52	1.42	0.60	0.54
	SA-OBS	-1.26	1.75	0.70	0.53	0.00	1.20	0.71	0.63
Pochentong	CFSR	-1.04	3.07	0.35	0.37	-1.45	2.12	0.45	0.33
	CPC	-2.11	2.83	0.53	0.42	-0.05	0.81	0.83	0.71
	ERA5	-1.30	1.85	0.72	0.54	-0.55	1.04	0.79	0.60
	SA-OBS	-1.14	1.68	0.71	0.53	-1.07	1.04	0.82	0.56
Average	CFSR	-0.60	2.81	0.51	0.44	-0.97	1.92	0.55	0.44
	CPC	-1.32	2.19	0.71	0.54	0.06	1.05	0.84	0.70
	ERA5	-1.37	2.01	0.76	0.53	0.20	1.15	0.78	0.65
	SA-OBS	-0.70	1.63	0.73	0.57	-0.29	1.05	0.84	0.68

2.3.4. Evaluation of performance based on the streamflow simulation

Through the analyses described in sections, 2.3.1–2.3.3, APHRODITE, ERA5, TRMM and IMEERG were found to be the preferred alternative precipitation data sources, and SA-OBS was found to be the favored alternative source for air temperature data. Therefore, the four selected gridded precipitation products and the SA-OBS air temperature data were used in the SWAT model application to simulate streamflow in 11 tributary basins. Model performance was evaluated based on three statistical indicators: NSE, PBIAS, and R^2. Although the model parameters were optimized independently for each precipitation dataset, there were some differences in the model performance results due to inherent biases and uncertainties included in each precipitation dataset (Table A5).

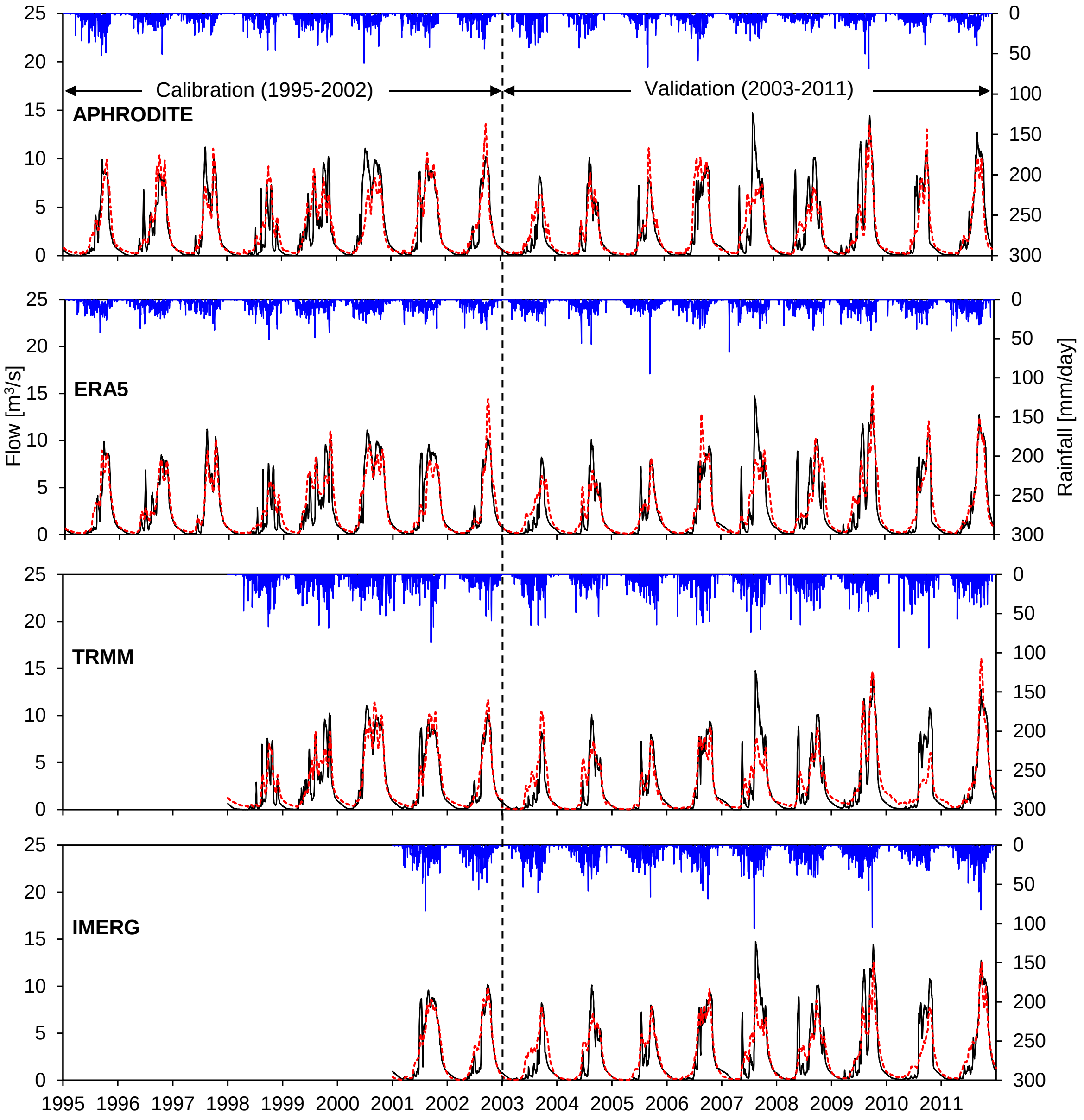


Figure 2.11. Daily observed (black solid line) and simulated (red dashed line) flow of the Sen tributary basin during calibration and validation periods from 1995 to 2002 (1998 and 2001 to 2002 for TRMM and IMERG, respectively) and 2003 to 2011, respectively. The secondary plots (blue solid line) indicate daily rainfall corresponding to each product. The black dashed line marks the end of the calibration period and the beginning of the validation period.

Figure 2.11 compares the time series of observed and simulated daily streamflow in the Sen River Basin, which has the largest drainage area, during the calibration (from 1995 or 1998 or 2001 to 2002) and validation periods (from 2003 to 2011). The hydrograph of APHRODITE, ERA5, TRMM and IMERG displayed a similar seasonal pattern to the observed data, showing very good agreement and performance during both calibration and validation periods (NSE and $R^2 \geq 0.75$, and PBIAS < ±10%) (Tables A3 and A5). The monthly observed and simulated flows of all other tributary basins during the calibration and validation periods are shown in Figures A6 and A7, respectively. Most of the results based on the four precipitation products reproduced the seasonal patterns well for both calibration and validation periods (Figs. A6 and A7), mainly due to the good representation of monthly rainfall and temporal patterns in these datasets (Figures 2.7 and 2.8).

Statistical indices summarized for the calibration and validation periods (Table A5) demonstrate that APHRODITE, ERA5, TRMM and IMERG products performed well on daily streamflow simulations, as NSE and R^2 in the calibration and validation periods were higher than 0.5 and PBIAS values were within ±15% in most tributary basins. On average, the ERA5 streamflow showed weaker performance in the 11 tributary basins than the TRMM, IMERG and APHRODITE streamflow, especially during the calibration period. The average performance of IMERG-based models showed a slightly better performance than that of the TRMM product for both calibration and validation phases (Table A5).

The comparison of observed and simulated mean annual runoff from the delineated area of each tributary basin and their total values indicated that APHRODITE, TRMM and IMERG precipitation datasets showed better overall performances in hydrological simulations during both calibration and validation periods, whereas the total runoff based on ERA5-rainfall overestimated the runoff by more than 50 mm in both calibration and validation periods (Figure 2.12). Although ERA5 precipitation showed good spatial (Figure 2.6) and temporal (Figure 2.7) performance when compared with gauged data, the bias and uncertainty in rainfall variability and distribution in the upstream areas, where gauged data are not available, could influence the performance of streamflow simulation at the gauged station. The rainfall from APHRODITE, TRMM and IMERG combined with SA-OBS air temperature provided more reliable estimates of mean runoff in each tributary basin and the total delineated basin areas during calibration and validation periods. Although APHRODITE, TRMM and IMERG performed well in reproducing the streamflow, a large difference in the annual precipitation recorded in these datasets (Figure 2.9) warrants further investigation with an additional water balance component, i.e., ET, which is discussed in Section 2.3.5.

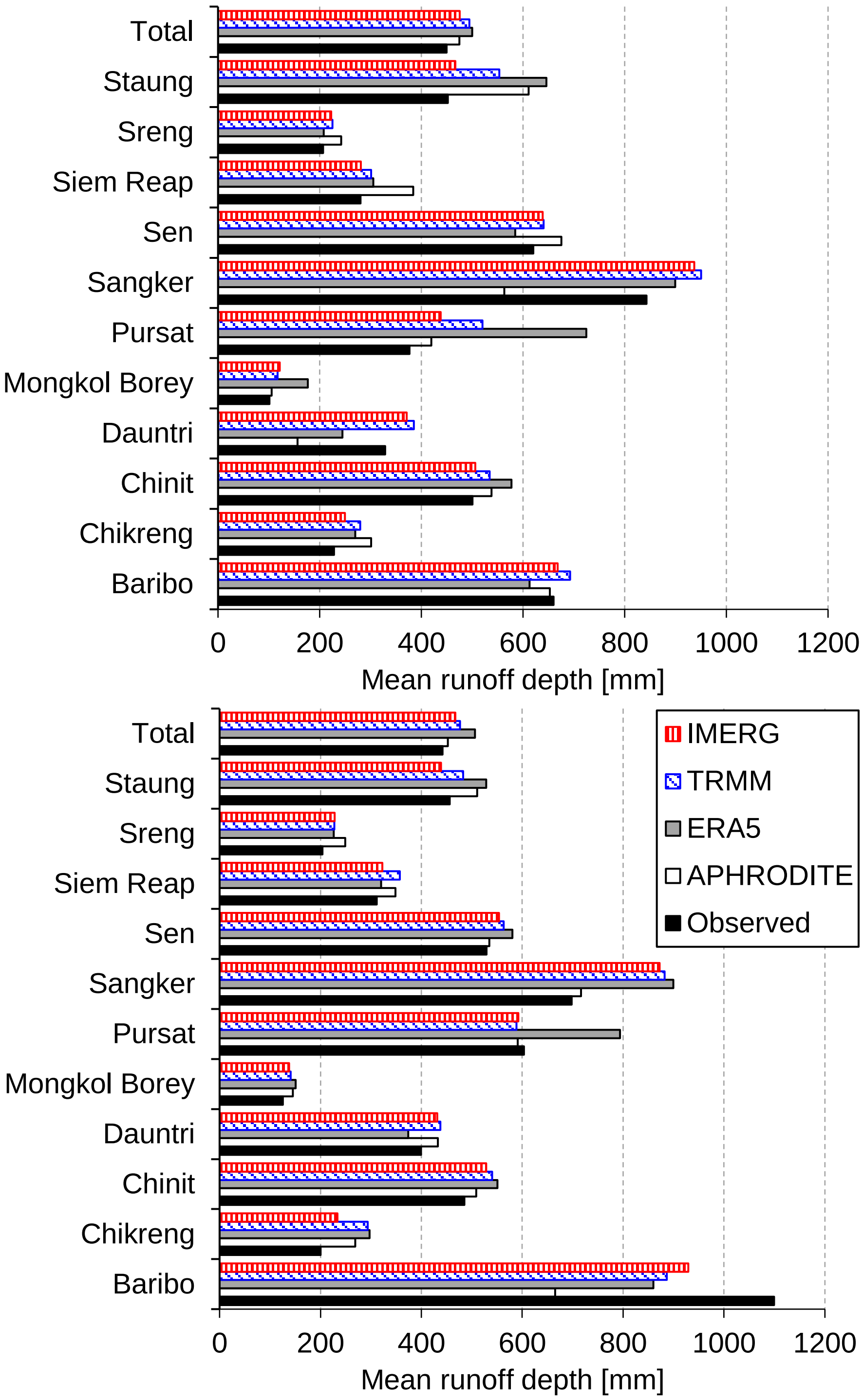


Figure 2.12. Comparison of observed and simulated mean annual runoff depth in the calibration (top) and validation (bottom) periods. The mean runoff depth for each tributary basin and the total represents the runoff from the delineated area of each tributary basin and all of the delineated basin area, respectively.

The average daily precipitation of SA-OBS overestimates the gauged data and is higher than that of APHRODITE and TRMM over Southeast Asia (Van den Besselaar et al., 2017), which agrees with the findings from our study. Errors in precipitation input can cause significant uncertainties in hydrological simulation. Consistent with the result of this study, the CFSR data overestimated the gauged rainfall which resulted in the considerable overestimation of simulated streamflow in the upper Dong Nai River, Vietnam (Nhi et al., 2019). Dinh et al. (2020); Thom et al. (2017); Vu et al. (2012) indicated that the TRMM and APHRODITE precipitation data have potential applications in driving hydrological model and water resources management in the Mekong River Basin, while our study showed that APHRODITE gives much underestimation of basin-wide precipitation. Similarly, Kawai et al. (2021) have pointed out in their study that the APHRODITE precipitation is about 33-38% less than the gauged values over Lao PDR. Like our study, Gunathilake et al. (2020) showed that IMERG had a high correlation and the lowest estimation error with gauged rainfall leading to good performance in simulating the streamflow in the Upper Nan River Basin, Northern Thailand. Therefore, Mohammed et al. (2018) revealed that the models utilizing TRMM and IMERG forcing data simulated streamflow well in the Lower Mekong River, which is also found in our study.

2.3.5. Evaluation of performance based on ET throughout the TSL Basin

The spatial distribution of mean annual ET calculated in the SWAT model application was compared with MODIS-based and GLEAM-based ET for the whole basin, excluding the flooded areas (Figures 2.13 (a)-(f)). The mean annual ET of each dataset was 759, 935, 1093, 1084, 1138 and 1036 mm for APHRODITE, ERA5, TRMM, IMERG, MODIS, and GLEAM, respectively, suggesting a better agreement between TRMM and IMERG-based and satellite-based ET, i.e., MODIS and GLEAM. The ET maps based on TRMM, IMERG and MODIS (Figures 2.13 (c), (d) and (c) showed high spatial similarities, except for the upstream areas in the Pursat tributary basin. Different algorithms used in ET calculations (discussed in 2.2.4.3) might be the source of dissimilarity between simulated ET and satellite-based ETs in some parts of the tributary basins.

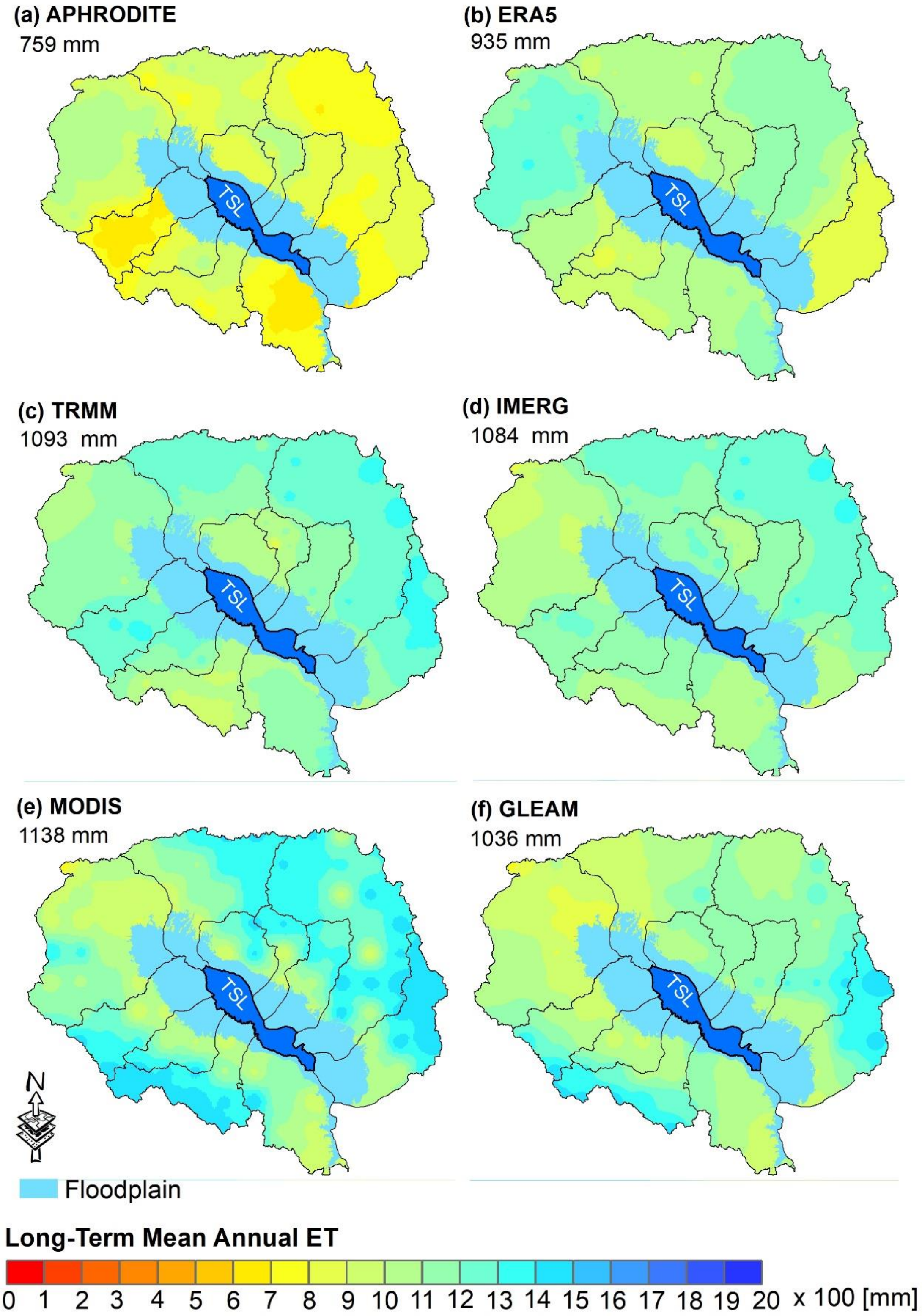


Figure 2.13. Comparison of the spatial distribution of long-term mean annual ET of APHRODITE (a), ERA5 (b), TRMM (c), IMERG (d) MODIS(e) and GLEAM (f) over the TSL Basin. The mean was calculated based on the long-term mean annual ET from 2001 to 2011. The mean annual ET of each dataset was mentioned at the top left corner of each map.

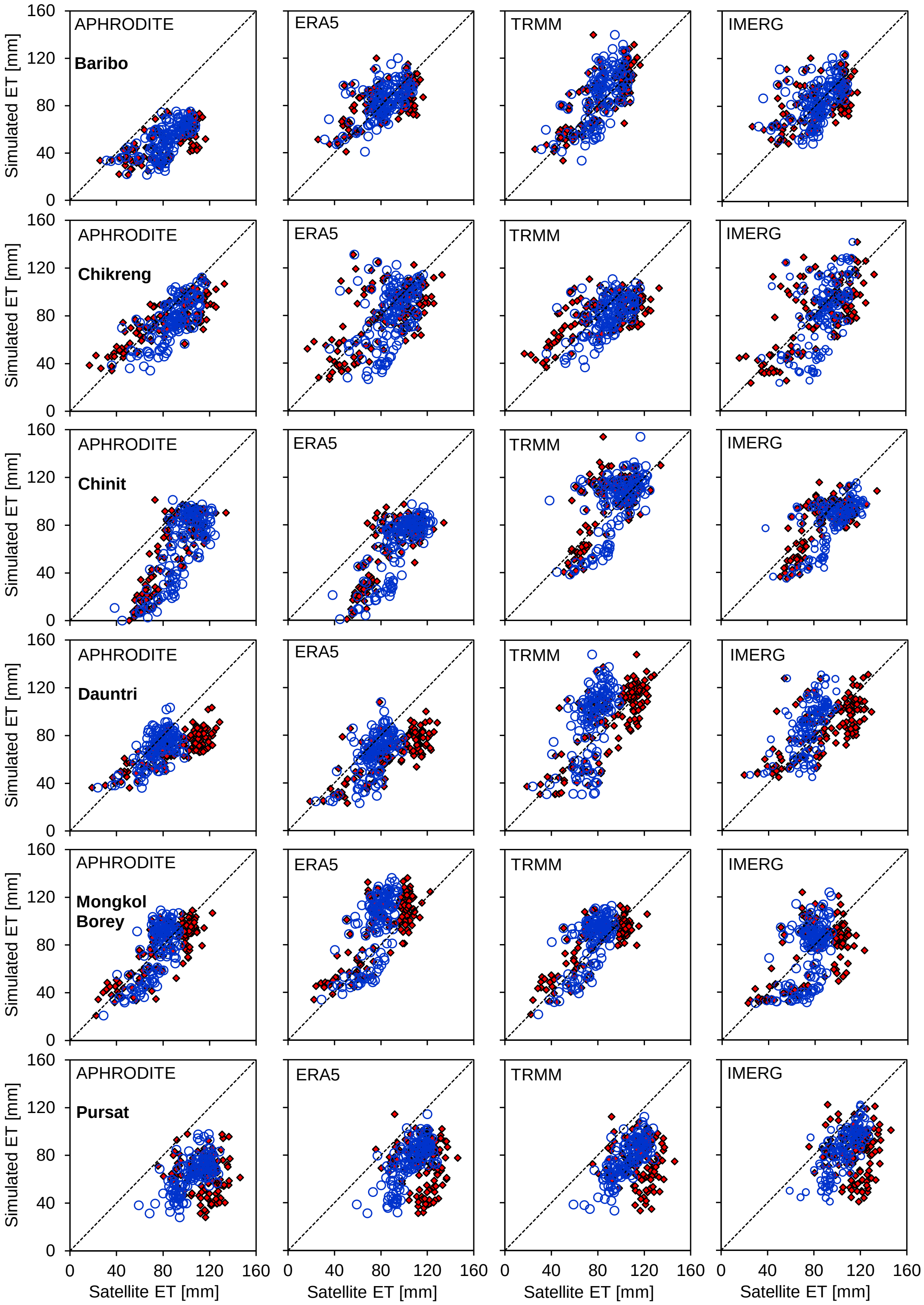


Figure 2.14. Scatter plots of tributary basin-averaged monthly simulated ET with MODIS-ET (red diamond) and GLEAM-ET (blue circle) for the individual dataset from 2001 to 2011. Each plot represents monthly ET averaged over each tributary basin excluding flooded areas.

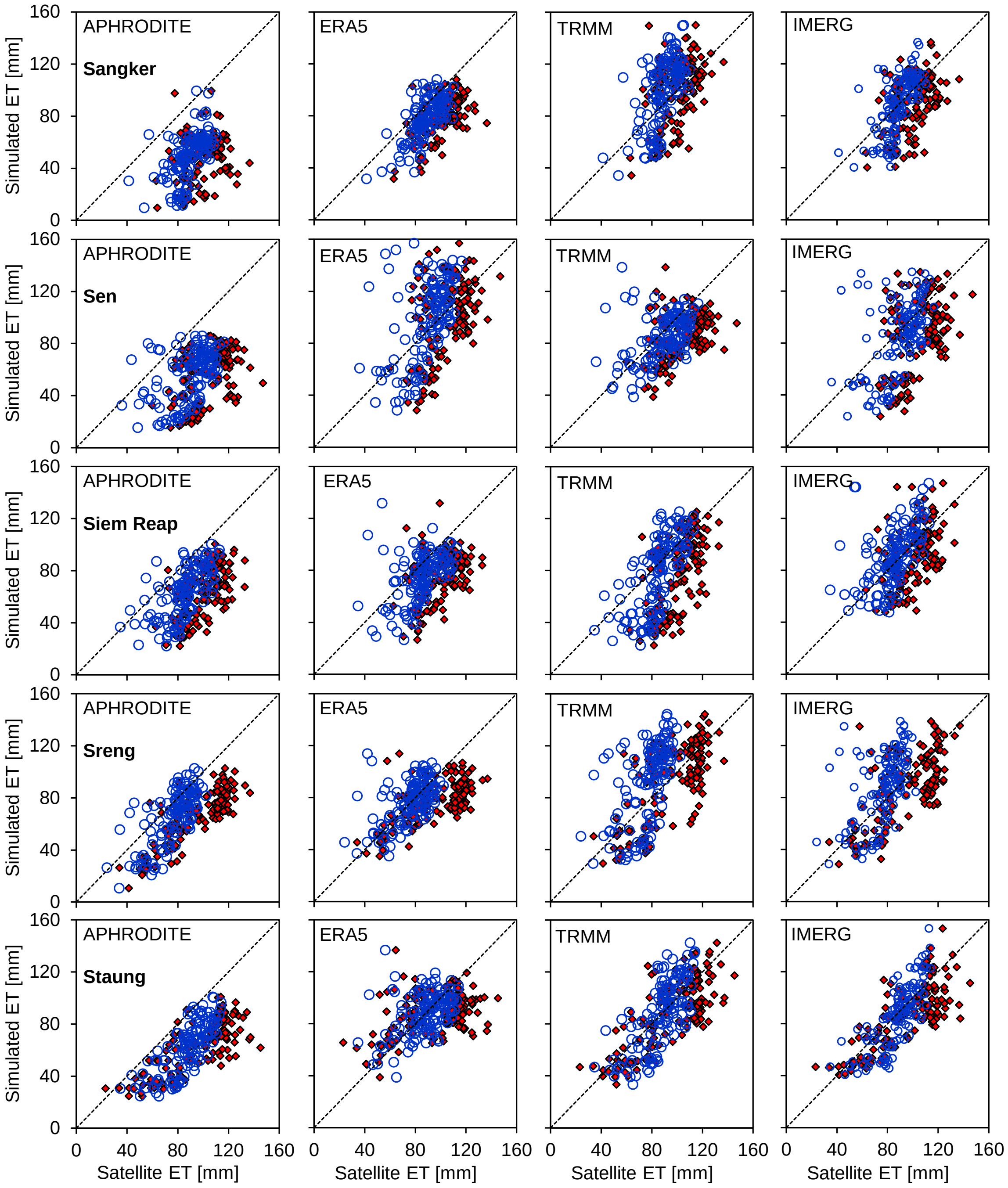


Figure 2.14. *Cont.*

The monthly ET's time series and scatter plots averaged over each tributary basin (excluding the flooded areas) as shown in Figures A8 and 2.14, respectively, the ET from the four precipitation products showed some similar seasonal patterns with MODIS- and GLEAM-ET, but significant differences were found for APHRODITE- ET. APHRODITE-ET was much lower than other models- and satellite-based ET, while RMSE values were found to be relatively smaller between those of TRMM and IMERG-ET and the satellite-based ETs (Table A6).

Additionally, the average RMSE between IMERG- and MODIS-ET is slightly smaller than that of TRMM-ET and MODIS-ET (Table A6). The limited gauge information (e.g., rainfall monitoring stations exclusively located at the downstream areas, Figure 2.1) used for gauge adjustment algorithms and resultant precipitation underestimations (Figures 2.6 and 2.7) could be the leading causes for the relatively poor performance of APHRODITE-based SWAT in ET estimations, indicating that this product could not be suitable for basin-wide hydrological applications. On average, the ET values simulated by the SWAT model using each precipitation product attributed to approximately 62% of the total precipitation (Figure A9).

Figure A10 compares mean simulated annual runoff depth values from each tributary basin and the total drainage area between 2001 and 2004, which were obtained by applying SWAT modeling and excluding the flooded regions in each tributary basin. The simulated total runoff depth from 11 tributary basins was 330, 414, 395 and 393 mm/year for APHRODITE, ERA5, TRMM and IMERG, respectively. In the study by Kummu et al. (2014), the total runoff depth from the same 11 tributaries of the TSL Basin during the same period (2001-2004) was estimated to be 394 mm/year, which is much closer to our results obtained using TRMM and IMERG precipitation.

The smaller RMSE (Table A6) and better temporal (Figure A8) and spatial (Figures 2.13 and 2.14) patterns of the IMERG- and TRMM-ET with GLEAM- and MODIS-ET indicates that TRMM and IMERG provide more reliable rainfall values over the TSL Basin, resulting in a more accurate estimation of streamflow and ET and a better prediction of other water balance components in the study area. Therefore, TRMM and IMERG precipitation, combined with SA-OBS air temperature, is a superior meteorological input for SWAT model simulations for comprehensive hydrological impact assessments over the TSL Basin.

2.4. Conclusion

This study provided a comprehensive assessment of seven gridded precipitation and air temperature products, including APHRODITE, TRMM, IMERG, CFSR, ERA5, SA-OBS, and CPC, by statistically comparing these datasets with gauge-based datasets and applying the SWAT model for daily streamflow and ET simulations over the TSL Basin. The precipitation data from APHRODITE, ERA5, TRMM and IMERG were found to have high correlations with rain-gauged data and the lowest estimation errors, whereas CFSR and SA-OBS products showed marked overestimation during most of the study period. The SA-OBS and CPC were found to match the observed air temperature data well, while the CFSR and ERA5 data showed larger estimation errors. The results of the hydrological simulation showed that the rainfall data from APHRODITE, TRMM and IMERG, combined with SA-OBS-based air temperature data, provided improved estimations of daily streamflow and mean runoff depth for most of the delineated tributary basins during the calibration and validation periods. The ET estimated using the TRMM and IMERG datasets showed a better temporal and spatial pattern agreement with GLEAM- and MODIS-ET. This suggests that TRMM and IMERG, in conjunction with SA-OBS air temperature, are reliable for providing the streamflow through the SWAT model application and other water balance components. Although TRMM and IMERG datasets provided comparable performance in streamflow and ET simulation in the study area, a slight improvement found for IMERG in statistical analysis and streamflow and ET simulations, in addition to the spatial resolution and latency time of IMERG, suggest that IMERG and SA-OBS provide the most favorable meteorological input datasets for comprehensive basin-wide hydrological impact assessments.

The comprehensive evaluation framework used in this study was found to be effective in selecting reliable gridded meteorological datasets for hydrological simulation in data-sparse river basins, especially when large uncertainty existed in the spatio-temporal distribution of rainfall. The findings also showed that applying gridded data that had only been verified through statistical comparisons with gauged data and hydrological simulation of streamflow could result in large uncertainties when quantifying other water balance components and assessing the hydrological responses, including rainfall-driven phenomena such as soil erosion and nutrient transport, due to climate change and land-use change. Our comprehensive evaluation approach results are a valuable guide for selecting alternative meteorological data for hydrological applications when the existing gauging networks are limited, not only in the TSL Basin but also in the Mekong River Basin and other river basins located within a similar climate region.

Chapter 3 Sediment load estimation using a novel regionalization sediment-response similarity method for ungauged catchments

The study of chapter 2 attempts to determine an ideal alternative meteorological dataset as the input data for hydrological modeling. Although satellite-based or gridded meteorological data could serve as the model inputs, model outputs (i.e., streamflow or sediment load) need to be optimized using ground-based observation. However, model calibration or validation cannot be performed in the ungauged catchment (e.g., the catchment is not monitored in terms of water level or sediment concentration). To solve this problem, in this chapter, a common method known as regionalization, in which model parameters from well-monitored catchments are transferred to ungauged, was used to estimate hydrological variables such as streamflow and sediment load. The final output of this section is the target of Objective 2. This model was carried on to Chapter 4 for application in hydro-environment impact assessment.

3.1. Introduction

Ground-monitored and satellite-based hydrometeorological and water quality datasets play vital roles in developing reliable hydrological models for regional environmental assessments of land-use change and anthropogenic threats such as climate change (Ang et al., 2022; Guo et al., 2020; Zhao et al., 2022b). However, the sparse and uneven distribution of hydrometeorological and water quality monitoring stations typically found in developing countries hinders the development of reliable hydrological models and accurate simulations of the hydro-environmental impacts (Ang et al., 2022). To solve this problem, a common method known as regionalization, in which hydrological information (model parameters) from well-monitored catchments (donors) are transferred to ungauged or sparsely gauged catchments (receptors), is used to predict the time series of hydrological variables such as streamflow and sediment load.

Several regionalization methods have been developed for modeling hydrological variables – especially for runoff simulation in ungauged catchments. The most common regionalization approaches include physical similarity (PS), spatial proximity (SPX), arithmetic mean (AM) and regression-based approach (RG) (Guo et al., 2020; Heng and

Suetsugi, 2014; Tegegne and Kim, 2018a; Zhao et al., 2022a). In the PS approach, one or more donor catchments with similar physical characteristics (i.e., climate, land-use, and geological conditions) to ungauged catchments are identified. Subsequently, a complete parameter set from the donor catchment is transferred to the corresponding ungauged catchment (Asurza-Véliz and Lavado-Casimiro, 2020; Guo et al., 2020). The performance of this technique depends on the selection of descriptors to explain the similarity. The SPX approach assumes that the nearby catchments should have similar characteristics and hydrological behaviors, and the entire parameter set is transferrable from the neighboring donors to the ungauged catchment, by interpolation or averaging (Swain and Patra, 2017). However, the model performance is inaccurate if strong heterogeneity is not present in the catchment characteristics as assumed (Parajka et al., 2005). In the AM approach, each calibrated model parameter from gauged catchments located in a given region is averaged and then transferred to the ungauged catchments in the same region. Consequently, the value of each corresponding parameter is identical for all ungauged catchments, which might not be realistic for heterogeneous catchments (Tegegne and Kim, 2018). The RG method aims to find the relationships between catchment attributes and model parameters calibrated in the gauged catchments. The established relationships are further used to quantify the model parameters for ungauged catchments (Heng and Suetsugi, 2014; Oudin et al., 2008). Heng and Suetsugi (2014); Zhang and Chiew (2009) mentioned that the equifinality problem of model parameterization is a major weakness of this method since the model parameters are used directly to build catchment relationships and define catchment clusters.

Numerous studies have compared the performance of different regionalization approaches for modeling streamflow and sediment in river basins and have reported the outperformance of the PS and SPX approaches over other regionalization methods (Guo et al., 2020; Heng and Suetsugi, 2014; Oudin et al., 2008; Parajka et al., 2013, 2005; Tegegne and Kim, 2018; Zhang and Chiew, 2009). Heng and Suetsugi (2014) compared three regionalization approaches (PS, RG, and SPX), to estimate sediment rating curves in ungauged catchments, for predicting instantaneous suspended sediment yield in 16 catchments of the lower Mekong River basin (LMRB). They found that the PS approach provided the best performance, while the SPX technique produced the least effective regionalization solution, owing to a sparse network of monitoring stations in the study area. Parajka et al. (2005) conducted runoff modeling across 320 Austrian catchments and revealed that the PS and SPX approaches perform similarly and were close to the calibration results. Oudin et al. (2008) compared three regionalization methods (PS, SPX, and RG) for runoff modeling in 913 catchments in France and found that wherever a dense network of gauging stations was available, the SPX provided the best regionalization solution, while the performance of PS approach is intermediary. Parajka et al. (2013) compared four regionalization techniques (as mentioned earlier) for runoff prediction based on a review of 34 past studies for 3,874 catchments; results indicated a lower performance of the RG method than other methods used in those studies. Additionally, it was found that the SPX and PS techniques performed best in humid catchments, while the PS and RG methods performed slightly better in arid catchments. However, in studies with dense streamflow gauge networks, there was a tendency for the SPX technique to outperform the RG and AM methods (Parajka et al., 2013). A common conclusion from this research is that the SPX approach has a relatively enhanced performance when dense streamflow gauge networks are available. In other conditions, the PS method is more likely to provide the best regionalization solution for the hydrological estimation of ungauged catchments.

The PS method assumes that the similarity in the input attributes, e.g., such as rainfall, land use, and geological conditions between gauged and ungauged catchments is connected to the corresponding similar responses, such as sediment or runoff. However, the relationship between the inputs and outputs can be inherently different between gauged and ungauged catchments, owing to the spatiotemporal variability of the hydrologic process in individual catchments (Tegegne and Kim, 2018). Moreover, the proximity among catchment centroids introduced in the SPX method is not necessarily translated into hydrologic similarity (Kokkonen et al., 2003). To overcome the problems associated with the SPX and PS methods, Tegegne and Kim (2018) proposed the new catchment runoff-response similarity (CRRS) regionalization method, to identify the attributes of a critical catchment and more accurately represent the hydrological similarities. The CRRS is based on the idea that catchments with similar runoff responses have similar hydrological processes, and therefore, similar model parameter values. The CRRS outperformed other regionalization methods, including the next best regionalization approach, the PS method, with 10–14% and 4–7% during the calibration and validation tests, respectively (Tegegne and Kim, 2018).

Overall, the regionalization approaches examined in previous studies assume that the catchment classification or clustering is possible, based on several static catchment descriptors (i.e., land-use and geological conditions) with long-term mean climatic descriptors (e.g., mean annual rainfall), or a combination. However, spatiotemporal variability in climate within the catchment is neglected (Samaniego et al., 2010). Wagener et al. (2007) also found that the prediction of hydrological variables (streamflow and sediment load) in ungauged catchments is a challenging task, primarily because the hydrologic processes occurring within a catchment occur over a wide range of spatiotemporal scales. Therefore, catchment attributes (descriptors), that represent the spatiotemporal variability of climate and hydrological process in gauged and ungauged catchments, are required for the regionalization of catchments. Therefore, in this study, a new regionalization method was proposed. This sediment-response similarity (SRS) method aims to find appropriate attributes that can ideally transfer the response of gauged catchments (associated with sediment erosion and transport) to that of ungauged catchments, considering spatiotemporal variability of the hydrologic processes occurring within the catchments. The method was developed by using a conventional regionalization method (an AM) and two statistical indices of catchment attributes: the coefficient of variation (CV) and correlation coefficient (R). These describe the spatiotemporal variations of sediment response and its relationship with rainfall characteristics in gauged and ungauged catchments. The values of these indices can characterize and compare the sediment-response behavior in terms of catchment attributes between gauged and ungauged catchments. Briefly, the SRS steps are as follows: 1) Apply the AM regionalization approach to transfer the calibrated parameters used in the soil and water assessment tool (SWAT) model from the gauged to ungauged catchment and simulate model output (i.e., sediment response); 2) Calculate the CV for observed rainfall (input) and the simulated sediment response obtained in the first step, and calculate the R between the rainfall and the sediment response, for each gauged and ungauged catchment; 3) Identify the catchment clusters (gauged and ungauged catchments), which are based on the sediment-response similarities, using the similar CV and R values; 4) Update the final parameter values from the same clusters, from the gauged to ungauged catchments (i.e., similar sediment responses represented by the CV and R values), and 5) Rerun the SWAT model to simulate the time series of sediment load in the ungauged catchments. To examine the validity and effectiveness of this proposed SRS approach, the method was applied to tributary basins located in the LMRB. The performance of the SRS method was compared with the conventional regionalization method, i.e., physical similarity. Finally, the sediment load time series was estimated using the SRS method for the ungauged tributaries of the Tonle Sap lake (TSL) basin (a sub-basin of the LMRB), which has a limited number of monitoring stations.

3.2. Materials and methods

3.2.1. Hydrological model setup

Figure 3.1 (top dash-box) shows the framework of the SWAT model setup (Section 3.5). First, the spatial and meteorological data were input into the SWAT model to set up hydrological models for both gauged and ungauged catchments. Next, the whole catchment was delineated into sub-catchments. The sub-catchment boundaries were delineated based on information from the digital elevation model (DEM) for the area. A combination of land cover, slope, soil properties, and management practices of the sub-catchments was classified and overlaid to form hydrologic response units (HRUs). Therefore, the gauged and ungauged sub-catchments were clustered using SRS (Section 3.2.2.1) and PS (Section 3.2.2.2) regionalization methods.

3.2.2. Regionalization methods

3.2.2.1. Sediment-response similarity

In this study, the new SRS regionalization method was based on a two-phase approach of transferring parameters (Figure 3.1). In the first phase, the conventional regionalization methods were used to tentatively transfer the calibrated model parameters from the gauged to ungauged catchments. The AM regionalization method was used in this study because the number of gauged catchments is insufficient in the two case study basins (see Figure 3.3). The number of ungauged receiver catchments will be too high compared to the number of donor catchments, if the catchment clusters were to be identified based on the PS or SPX approaches, owing to the differences between donor and receiver catchments. However, the PS or SPX approaches can be used to tentatively transfer model parameters from gauged to ungauged catchments where there are many donor catchments. The second phase of the parameter transfer approach was based on SRS. The catchments with similar rainfall characteristics and sediment responses are considered hydrologically similar. Once similar catchments were identified, the final parameter values of each gauged catchment were transferred to the ungauged catchments. The similarity of runoff response and sediment load simulated in the model implies that dominant hydrological processes within a particular ungauged catchment are well incorporated into the model.

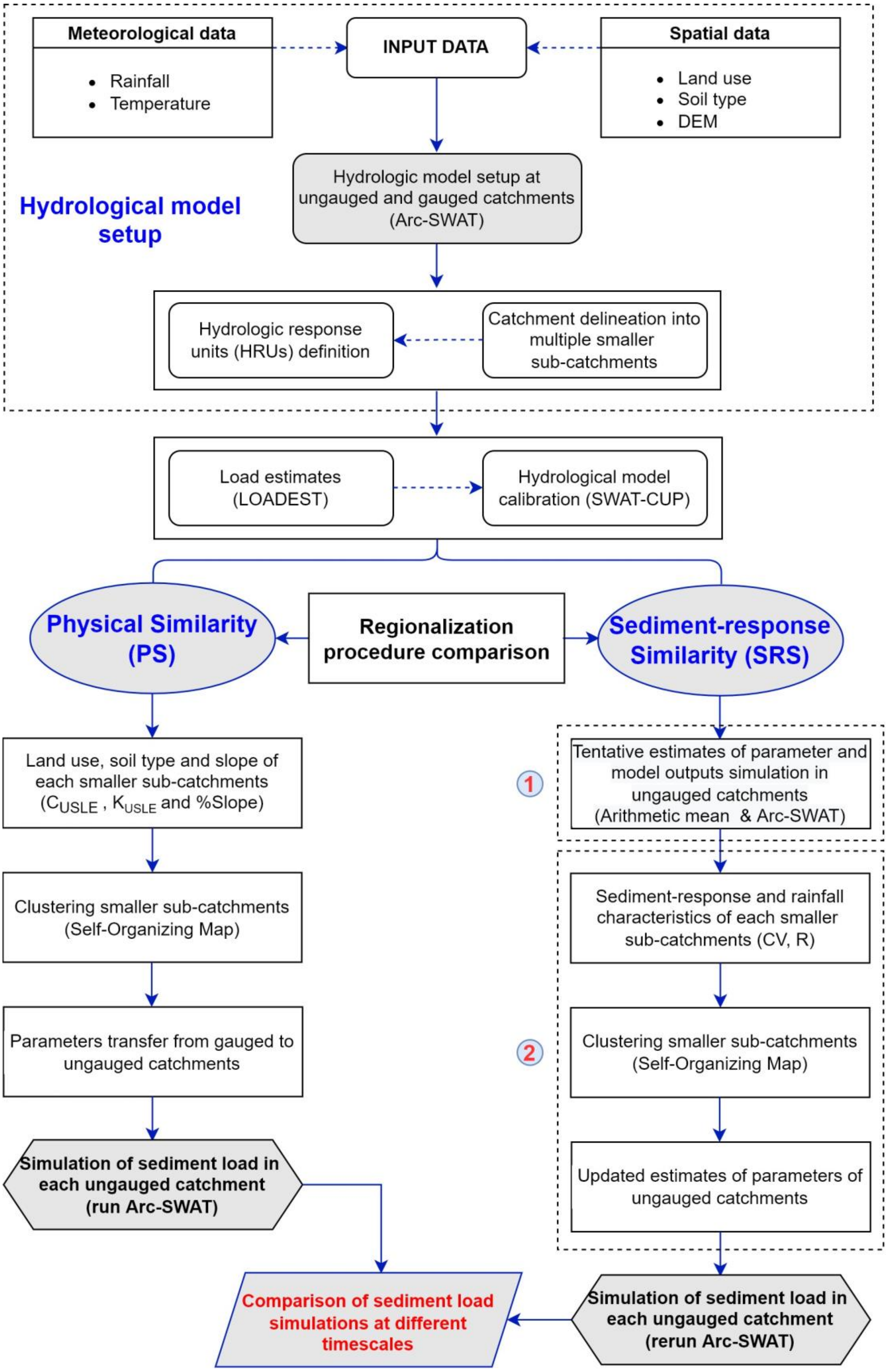


Figure 3.1. Flowchart of physical similarity (PS) and the sediment-response similarity (SRS) parameter transfer schemes. Numbers 1 and 2 denote the two-phase approach of parameter transfer of the SRS systematic procedure. The top dash-box denotes the general procedure of the SWAT model setup.

Consequently, the regionalization methods that consider the similarity in the rainfall characteristics and the sediment load response as the main attributes could reduce the hydrologic uncertainty in the predicted results of the ungauged catchments (Tegegne and Kim, 2018). The hydrological processes within a catchment occur over a wide spatiotemporal scale. Therefore, catchment attributes that represent the spatiotemporal variability of climate and hydrological processes in the catchments, are required for the regionalization of catchments. The CV index is a statistical measure of the dispersion of data points in a data series around the mean, and it is beneficial when comparing the degree of variation from one data series to another, even if the means are considerably different (Brown, 1998). The R index measures the strength of the statistical relationship, or a linear association, between two variables. The CV and R indices describe the spatiotemporal variations of sediment response and the relationship with rainfall characteristics, in gauged and ungauged catchments. This can partially characterize the unique sediment-response behavior of each catchment. Therefore, the CV and R indices could be ideal catchment attributes for identifying SRS in long-term spatiotemporal variations between gauged and ungauged catchments.

The detailed and systematic procedures of the SRS method are summarized in five steps (Fig. 1), as follows.

Step 1: During the first phase, the AM regionalization approach was applied to tentatively transfer the calibrated parameters used in the SWAT model from the gauged to ungauged catchment and simulate model output such as sediment response or load.

Step 2: In the second phase, the daily sediment load for each sub-catchment (simulated in Step 1), is normalized by the corresponding sub-catchment area. The index values of CV of daily rainfall, normalized sediment load, and R between daily rainfall and normalized sediment load, were then calculated for each sub-catchment.

Step 3: The self-organizing map (SOM) (Section 3.2.3) clustering technique was used to identify similar sub-catchments using the daily time series CV and R index values computed for each sub-catchment (see Figures 4 and 5).

Step 4: The model parameters for the ungauged catchments were updated with the final parameter values from the gauged catchments in the same cluster. This means that similar sediment responses or CV and R values were considered.

Step 5: Finally, the SWAT model was rerun to simulate the sediment load time series in the ungauged catchments.

3.2.2.2. Physical similarity

The PS method is one of the best-performing regionalization solutions reported in previous studies (Guo et al., 2020; Heng and Suetsugi, 2014; Oudin et al., 2008; Parajka et al., 2013, 2005; Tegegne and Kim, 2018; Zhang and Chiew, 2009, among others). It was used for comparison against the performance of the SRS method. The PS approach assumes that catchment physiographic characteristics predetermine the hydrological behavior. The catchment descriptors or attributes are generally used to define the similarity and are related to the topography, land cover and soil type of the catchment. In this study, the area-weighted catchment slope, the Universal Soil Loss Equation (USLE), the land cover and management factor (C_{USLE}), and the USLE soil erodibility factor (K_{USLE}), were used. In previous studies, these attributes were considered the main drivers of hydrological responses in different catchments (Merz and Blöschl, 2004; Tegegne and Kim, 2018).

The area-weighted average value of each catchment attribute of each sub-catchment was computed as follows:

$$\mathrm{CA}_i = \left(\text{area ratio} \times \mathrm{CA}\right)_1 + \left(\text{area ratio} \times \mathrm{CA}\right)_2 + \cdots + \left(\text{area ratio} \times \mathrm{CA}\right)_m \tag{3.1}$$

where, *CA* is a catchment attribute value (i.e., C_{USLE}, K_{USLE} and %Slope); *i* is the number of sub-catchments; the area ratio is the area occupied by the corresponding catchment attribute divided by the total area of the sub-catchment *i*, and *m* denotes different land use types, soil types and slope classes for C_{USLE}, K_{USLE} and %Slope, respectively.

Table A7 presents the C_{USLE} and K_{USLE} values of different land use and soil types, respectively, which were retrieved from the SWAT land-use and soil-type databases (Ang et al., 2022). The slope percentage of each sub-catchment was calculated from the DEM input data (see Table 2). The schematic framework of the PS method is illustrated using that of the SRS approach in Figure 3.1. First, the area-weighted slope, C_{USLE}, and K_{USLE} values of gauged and ungauged sub-catchments were computed and inputted into the SOM algorithm. The SOM clustering technique was then used to identify similar cluster sub-catchments, using the area-weighted slope, C_{USLE}, and K_{USLE} of each sub-catchment (see Figures 4 and 5). Finally, the model parameters for the ungauged catchments were transferred from similar gauged catchments to simulate the time series of sediment load in the ungauged catchments. This simulated time series from the PS approach was compared with that from the SRS method, to evaluate the SRS method's performance.

3.2.3. Self-organizing map

An SOM is an unsupervised learning algorithm based on artificial neural networks used to cluster and visualize the structure of high-dimensional data sets (Kohonen, 1990, 1982). An SOM consists of input and output layers linked by weightings (Figure 3.2a). The input layer is formed by a set of nodes consisting of input vectors (Figure 3.2a), and the output layer is formed by nodes arranged in a two-dimensional (2-D) grid map. This 2-D grid map (also called the Unified distance matrix or U-matrix) shows the data similarities by grouping similar data into the same clusters, as indicated by the same numbers inserted in the grids (shown in Figure 3.2b). In this study, the values of C_{USLE}, K_{USLE}, and %Slope, and the values of CV and R derived from the daily rainfall and sediment response of individual sub-catchments (i.e., gauged and ungauged), were used as the input vectors to create the SOM clustering map (Figure 3.2b) for the PS and SRS methods, respectively. The input vectors were normalized to prevent unintentional weighting of vector attributes during the training process (see Box A1). The number of output neurons (i.e., the size of the output map, see Figure 3.2b) in an SOM can be selected using the empirical rule by Vesanto and Alhoniemi (2000). The optimal number of neurons, i.e., in the SOM is calculated approximately by $5\times\sqrt{L}$, where L is the number of samples in input vectors (number of sub-catchments, see Figure 3.5) (Farsadnia et al., 2014). The SOM assigns values to represent the input vector (i.e., reference vector) in each neuron, which is generated and adjusted during the process of clustering (Farsadnia et al., 2014; Hamel and Brown, 2011). During the training process, similar input vectors were close to each other, and dissimilar ones were further apart. This implies that similar input vectors between gauged and ungauged sub-catchments will be mapped to the same clusters. Then the most similar gauged to ungauged sub-catchment will be found based on the relative distance as shown in Figure 3.2b. Additionally, the maximum number of clusters identified in each catchment is beneficial, as the more clusters identified, the larger the number of donor catchments. The complete process for training the SOM can be found in Box A2. Detailed information on the SOM algorithm is documented in Kohonen (1990, 1982), and the applications of the SOM in hydrological studies can also be found in previous research (Farsadnia et al., 2014; Ley et al., 2011; Nourani et al., 2013; Tegegne and Kim, 2018; Wallner et al., 2013).

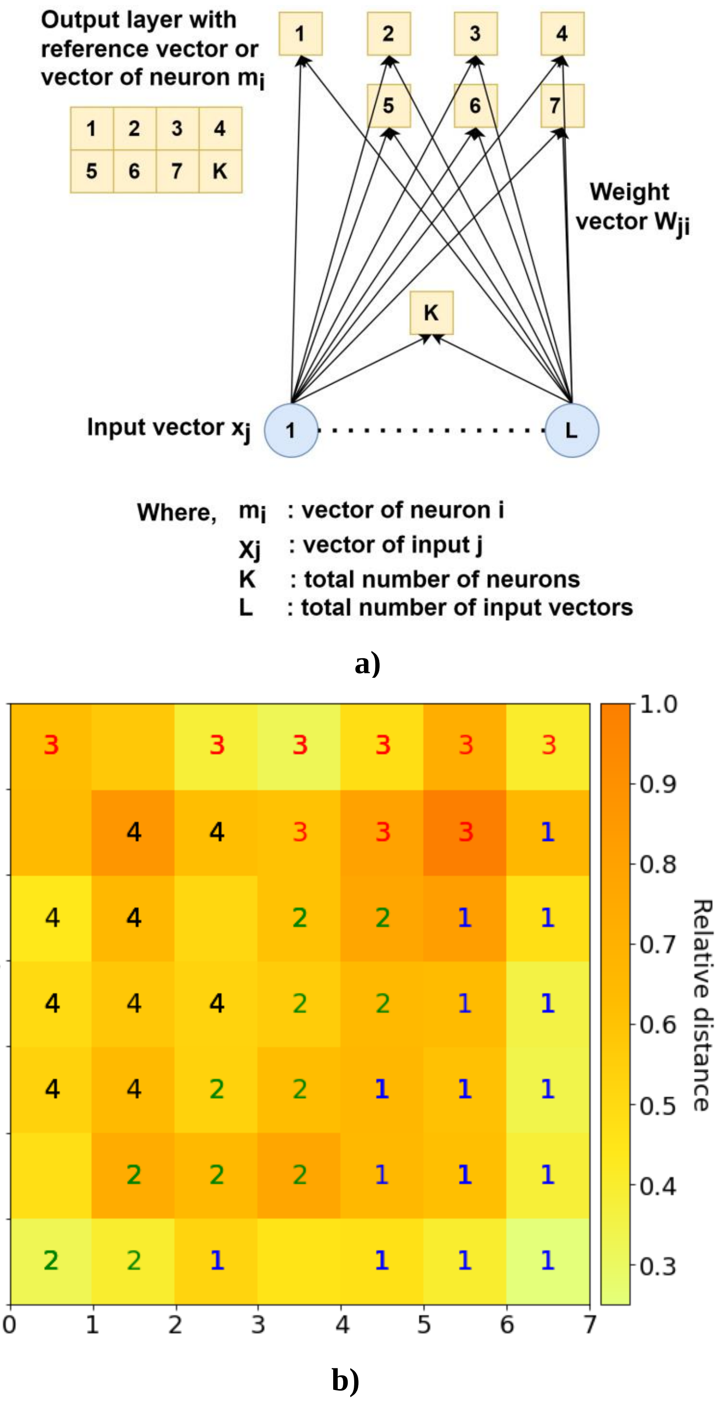


Figure 3.2. a) Schematic diagram of the self-organizing map. b) Unified distance matrix (U-matrix). In b), the number shown on the map denotes clusters. Colors denote the relative distance between the grids numbered on each axis.

3.2.4. Study area

The TSL basin, and the basins of the Sekong, Sesan and Srepok rivers, collectively known as the 3S Rivers, located in the LMRB, were selected to test the applicability of the SRS approach proposed in this study (Figure 3.3). The tropical monsoon-driven climate dominates the hydro-climatic conditions in these basins and is characterized by a rainy season (May–October) and a dry season (November–April) (B. Shrestha et al., 2017; World Bank, 2011).

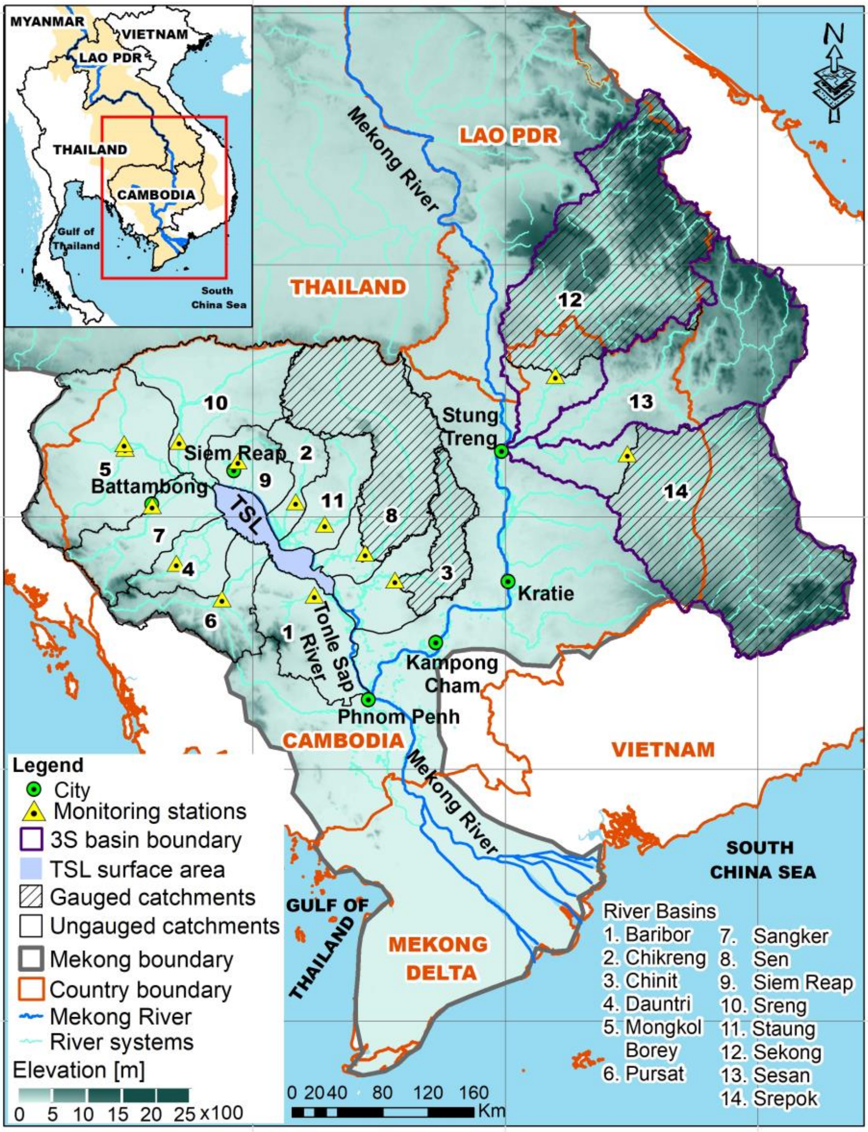


Figure 3.3. Location map of the TSL River and 3S River Basins showing the water-level and sediment monitoring stations. The gauging stations of Sen, Chinit, Sekong and Srepok catchments are for both water-level and sediment monitoring, and they are termed “gauged catchments” while others are only for water-level measurement. The inset shows the location of the TSL and 3S River Basins in the Lower Mekong River Basin and Indochina Peninsular.

The general information on the TSL Basin can be found in Chapter 2. The location map and drainage catchment characteristics of each tributary are presented in Figure 3.3 and Table 3.1, respectively.

Table 3.1. The total drainage area, gauged area, and land-use of each catchment of the TSL and 3S basins

Nº	River Name	Area (km^2)	Delineated area* (km^2 / %)	Forest land (%)	Agricultural land (%)	Urban land and others (%)
1	Baribo	7,153.78	875.33 / 12.24	25.01	74.61	0.39
2	Chikreng	2,713.90	1,719.12 / 63.35	78.38	21.62	0.00
3	Chinit	8,236.86	4,055.91 / 49.24	62.70	37.26	0.04
4	Dauntri	3,695.97	519.13 / 14.05	43.42	56.58	0.00
5	Mongkol Borey	14,966.42	3,964.94 / 26.49	14.67	85.14	0.19
6	Pursat	5,964.77	4,118.59 / 69.05	76.15	23.76	0.09
7	Sangker	6,052.78	2,596.69 / 42.90	53.76	46.09	0.14
8	Sen	16,359.58	14,129.33 / 86.37	85.37	14.61	0.02
9	Siem Reap	3,618.98	609.75 / 16.85	26.28	73.57	0.15
10	Sreng	9,986.27	6,691.42 / 67.01	61.68	38.32	0.00
11	Staung	4,357.39	1,878.31 / 43.11	75.03	24.97	0.00
	Total land area	83,106.70	4,1158.52 / 49.52	54.77	45.14	0.09
	Tonle Sap lake	2,743.80	-	-	-	-
12	Sekong	28,689.95	25,393.92 / 88.51	85.4	9.5	5.1
13	Srepok	31,179.93	27,361.41 / 87.75	83.3	9.9	6.8

The 3S River basins are transboundary rivers that flow through three countries, with a total catchment area of 78,650 km^2, of which 33% is in Cambodia, 29% is in the Lao PDR, and 38% is in Vietnam (see Figure 3.3) (López et al., 2017; B. Shrestha et al., 2017; Trang et al., 2017). The Sesan and Srepok rivers flow through Cambodia, joining the Sekong river and finally meet with the Mekong River at Stung Treng province in Cambodia (Figure 3.3). The elevation of the 3S basins ranges from less than 100 masl downstream to over 2000 masl in the upstream areas of the Sekong basin. The basin area and land use distribution of the Sekong and Srepok river basins are presented in Table 3.1, in which the areas are dominated by forest cover (> 80%) (Constable, 2015; Trang et al., 2017). Acrisols (68%) and Ferralsols (12%) with sandy clay loam and clay texture are the dominant soils in the 3S basins (M. Shrestha et al., 2017). The annual discharge from the 3S basins represents approximately 17–20% of the total annual streamflow of the Mekong River (an average of 2,886 m^3/s) (Piman et al., 2013). The 3S basins are also a primary source of sediment in the LMRB, and the annual sediment load from the 3S basins is estimated at 10–25 Mt (Million tons) (Kondolf et al., 2014).

3.2.5. Hydrological model

The SWAT model was used in this analysis. It is a semi-distributed model, a widely used hydrological model to simulate streamflow, sediments, and other hydrological variables under diverse environments and different scales (Arnold et al., 1998). The prediction of the sediment in the SWAT model is based on a modified universal soil loss equation (MUSLE) (Williams, 1975), developed initially by Wischmeier and Smith (1965), where rainfall and runoff are the main drivers of soil loss. The surface runoff from each HRU is simulated based on the SCS curve number method (USDA, 1972), with daily rainfall defined at the centroid of the sub-catchment, where the default SWAT daily rainfall (using observed rainfall at a gauging station located closest to the centroid of each sub-catchment) was not used (Arnold et al., 1998). More details about the SWAT can be found in Chapter 2 and official theoretical documentation (Neitsch et al., 2011) and the review paper by Gassman et al. (2007).

3.2.6. Data collection and pre-processing

Each dataset used for the SWAT model setup is summarized in Table 3.2. The Mekong River Commission-Water Quality Monitoring Network (MRC-WQMN) dataset has been widely used in sediment load simulation studies in the Mekong River basin (Kummu et al., 2008; Kummu and Varis, 2007; Sok et al., 2020; Wang et al., 2011). The total suspended solids (TSS) samples were taken from this dataset, collected at 0.30 m below the water surface in the center of the river cross-section, at each monitoring station (Sok et al., 2020), and were analyzed using the recommended analytical methods (MRC, 2019). There are limited water-level and sediment monitoring stations in the study basins. The monthly observed sediment data at four gauging stations (Chinit and Sen catchments in the TSL basin and Sekong and Srepok catchments in the 3S basin, see Figure 3.3) were used to calibrate and validate SWAT model parameters. Suspended sediment loads were estimated from the time series data of daily observed streamflow and monthly TSS concentrations, which were monitored at the same location, using the regression model built into the LOAD ESTimator (LOADEST) program (Runkel et al., 2004). The applicability of the LOADEST was verified through many studies in estimating daily pollutant loads for various water quality parameters, including sediment (Duan et al., 2013; Jha and Jha, 2013; Jha et al., 2007; Park and Engel, 2016, 2015, among others) and calibrating SWAT model parameters (Cakir et al., 2020;

Epelde et al., 2015; Nepal and Parajuli, 2022; Sok et al., 2020; Teshager et al., 2016). Detailed information on topography, land-use, soil type and water level data can be found in Chapter 2. The study in Chapter 2 evaluated seven gridded meteorological datasets and found that precipitation from Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM) (IMERG) and air temperature data of Southeast Asia-Observational (SA-OBS) are the optimal meteorological data for comprehensive basin-wide hydrological impact assessments in the TSL basin. Previous studies also revealed that the models using IMERG precipitation data simulated streamflow well in the LMRB and the 3S basins (Li et al., 2018; Mohammed et al., 2018; Wang et al., 2017). Therefore, the precipitation and air temperature values of IMERG and SA-OBS, respectively, were interpolated using the values at the four grids nearest to the centroid of each sub-catchment for use in the SWAT model. For the Sekong and Srepok catchments, water-level data, which were provided by the MRC (Table 3.2), from a monitoring station of each catchment (see Figure 3.3) were used as the observed data for flow model calibration (2005–2008) and validation (2009–2011). The results of the flow simulation model achieved good performance in the Sekong catchment, while good and satisfactory performance of model calibration and validation, respectively, were obtained in the Srepok catchment (Figure A11, Tables A3 and A8).

Table 3.2. Summary of required data used for SWAT model setup and calibration

Data type	Time period	Temporal resolution	Spatial Resolution	Sources
Topography (Digital Elevation Model [DEM])	-	-	90 m	http://srtm.csi.cgiar.org
Land cover map	2002	-	250 m	Mekong River Commission (MRC)
Soil types map	2002	-	250 m	MRC
Total rainfall, Maximum and minimum air temperatures	2001–2011	Daily	-	Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM) (IMERG) and Southeast Asia-Observational (SA-OBS)
Water level	1995–2011	Daily	-	MRC and Department of Hydrology and River Works (Cambodia)
Total suspended solids (TSS) concentration	2005–2008	Monthly	-	MRC

3.2.7. Model calibration, validation and evaluation

The calibration of model parameters for streamflow and sediment load was performed using the SWAT-calibration and uncertainty Programs (SWAT-CUP) (Abbaspour, 2015). Among several methods for calibration in SWAT-CUP to deal with parameter uncertainty, the Sequential Uncertainty Fitting version 2 (SUFI-2) method was selected (Nkonge et al., 2014; Wu and Chen, 2015) to provide more reasonable daily streamflow and sediment simulations at the outlet of each gauged drainage area (Table 3.1 and Figure 3.3). The streamflow in each catchment was calibrated before calibrating the sediment load (Abbaspour, 2015). Then, parameters for sediment load were calibrated while maintaining the flow-parameter ranges which were obtained through the flow calibration. Owing to the availability and consistency of continuous sediment load data in each catchment, the model parameters were calibrated with available data between 2005 and 2008. Based on previous studies (Ang et al., 2022; Ang and Oeurng, 2018; Oeurng et al., 2019; Roth et al., 2016; Sao et al., 2020; Sok et al., 2020; Vilaysane et al., 2015; Vu et al., 2012), 26 sensitive parameters were selected and calibrated for streamflow and sediment load simulation using the SWAT-CUP model (Tables A9-A12).

To validate the regionalization approach proposed in this study, the leave-one-out cross-validation approach, also known as the Jack-knife method (introduced by Quenouille (1956) and Tukey (1958)), is typically used to validate the regionalization approach (Athira et al., 2016; Gitau and Chaubey, 2010; Heng and Suetsugi, 2014; Parajka et al., 2005; Razavi and Coulibaly, 2013; Samuel et al., 2011; Tegegne and Kim, 2018). In this procedure, one of the gauged catchments is assumed as a pseudo-ungauged catchment, while the remaining catchments are regarded as the candidate donors. For example, when validation was conducted for the Chinit catchment (see Figure 3.3 for the location of the Chinit catchment), the remaining three catchments (i.e., Sen, Sekong, and Srepok) functioned as its candidate donors (see Figure 3.5). Subsequently, the PS and SRS methods were applied to select the best donor sub-catchments from the same cluster, with the smallest relative distance (see Figure 3.2b). Finally, using the calibrated model parameters transferred from the best donor sub-catchments, the time-series of sediment load in the Chinit catchment was estimated and compared with the observed data for the cross-validation phase.

For evaluating model performance in estimating the time series of sediment load, three quantitative metrics were employed. These were the Nash-Sutcliffe efficiency (NSE), percent bias (PBIAS), and coefficient of determination (R^2), as shown in Table A1 (Moriasi et al., 2007). The NSE indicates agreement between the simulated and the observed time series values (Nash and Sutcliffe, 1970). The PBIAS measures the average tendency of the simulated values to be larger or smaller than their observed counterparts (Gupta et al., 1999). The R^2 estimates the linear relationship between two variables, by assessing the extent of difference between one variable by another. Furthermore, the performance ratings suggested by Moriasi et al. (2007) were applied to evaluate the integrated performance of simulated results (Table A3).

3.3. Results and discussion

3.3.1. Interpretation of sub-catchment descriptors and clusters for each regionalization method

Based on the elevation data, the Chinit, Sen, Sekong, and Srepok catchments were delineated into 14, 29, 23 and 29 sub-catchments, respectively. The total of 95 sub-catchments is shown in Figure 3.4, which displays the spatial variation of the area-weighted values of the catchment descriptors in the four test catchments. These were subsequently used to create the SOM for generating sub-catchment clusters for each regionalization method. The area-weighted values of C_{USLE} in the Chinit, Sen, and Sekong catchments are similar, more so than that of the Srepok catchment, while similar K_{USLE} values were observed in the Chinit, Sen, and Srepok catchments (Figures 3.4a-h). The Chinit and Sen catchments are neighboring tributaries in the TSL basin; therefore, the slopes were more similar to each other (compared with Sekong and Srepok catchments) (Figures 3.4i-l). Specifically, slight slopes (10–20%) were observed in most sub-catchments of the Chinit and Sen catchments. However, steep slopes were observed in the upstream Sekong catchment (> 20%), in contrast to the low slope in a large area of the Srepok catchment (Figures 3.4i-l).

(A)

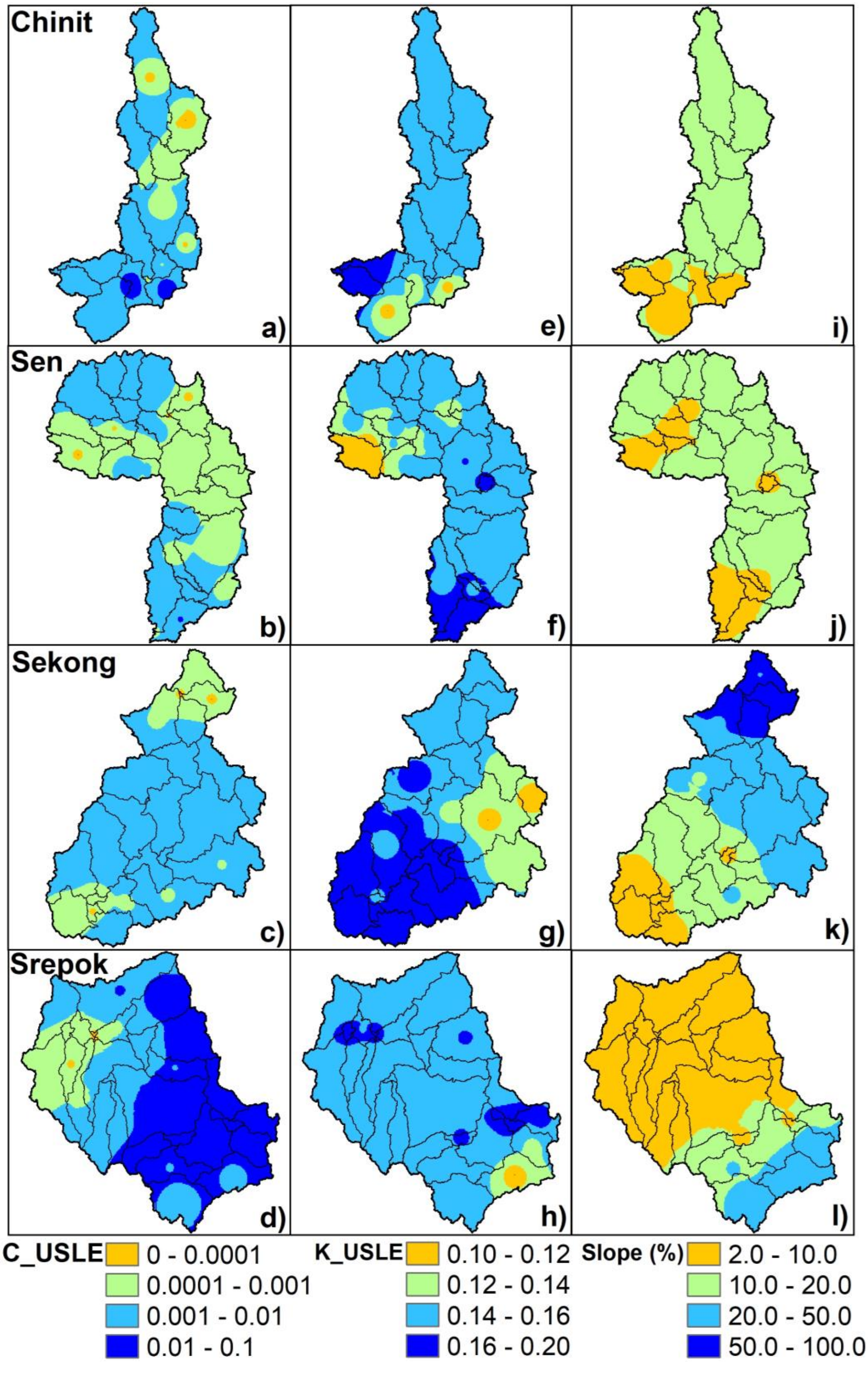


Figure 3.4. (A) Area-weighted values of C_{USLE} (a-d), K_{USLE} (e-h) and slope percentage (i-l) for the physical similarity regionalization method. (B) Coefficient of variation of rainfall (m-p), coefficient of variation of sediment load (q-t) and correlation coefficient between rainfall and sediment load (u-x) for the sediment-response similarity regionalization method.

(B)

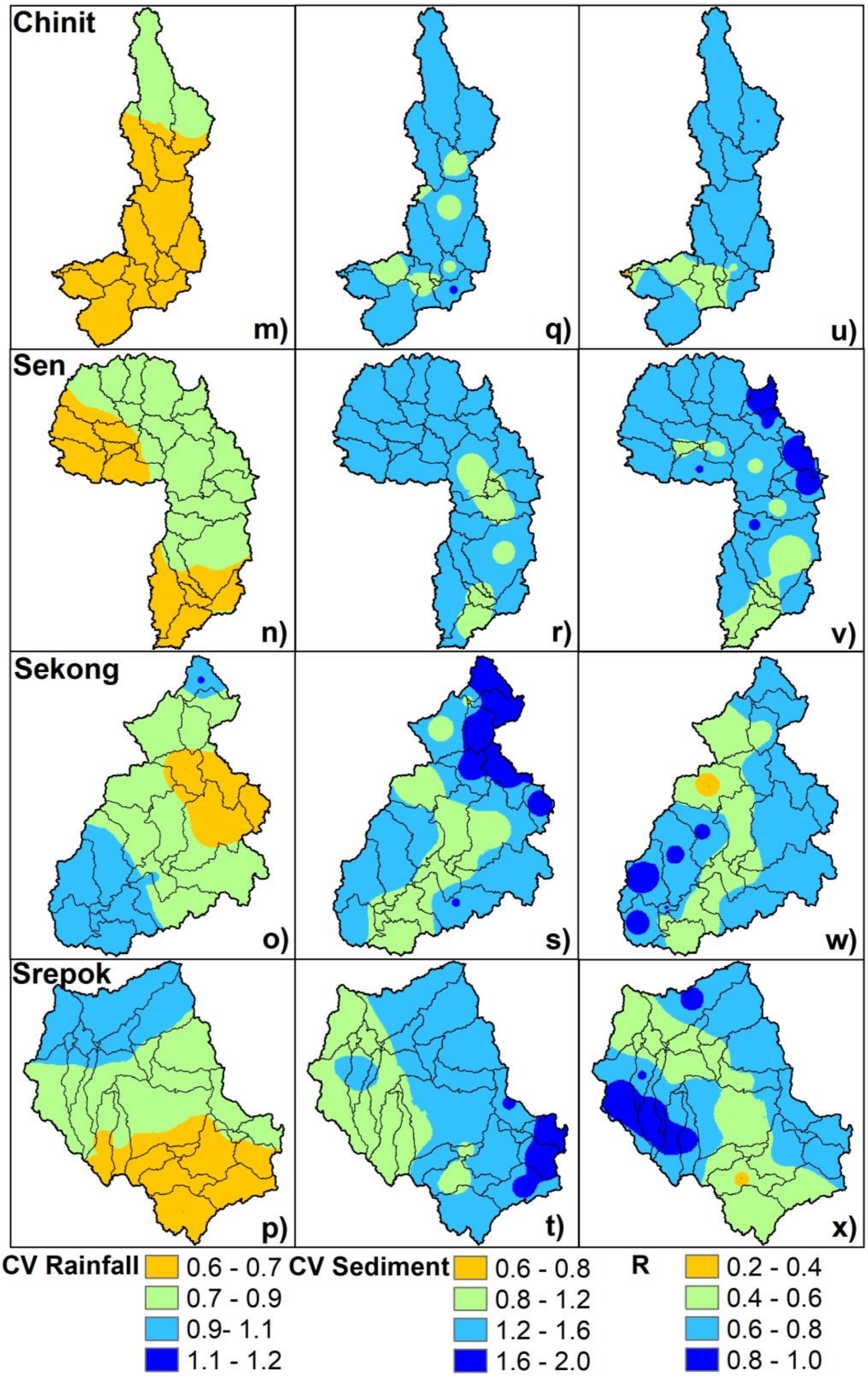


Figure 3.4. *Cont.*

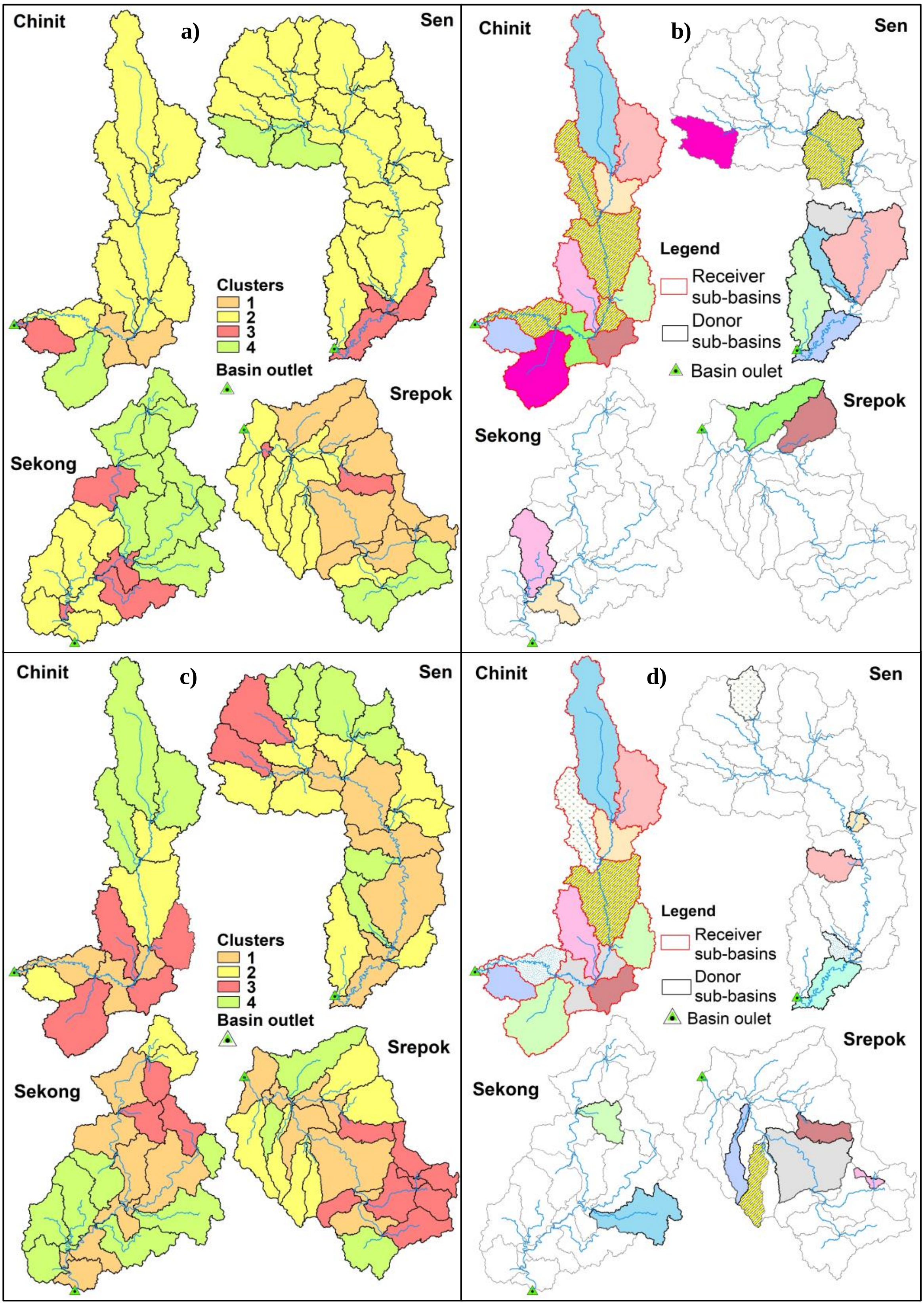


Figure 3.5. a) Physical similarity sub-catchment clusters, b) Physical similarity donor and receiver sub-catchments for Chinit catchment as a pseudo ungauged catchment, c) Sediment-response similarity sub-catchment clusters, and d) Sediment-response similarity donor and receiver sub-catchments for Chinit catchment as a pseudo ungauged catchment.

In the Chinit catchment, a relatively low variation of rainfall (CV = 0.6–0.7) was observed, which covers most of the catchment area (Figure 3.4m). The CV of rainfall is similar for the Sen, Sekong and Srepok catchments, at 0.6–0.9, except for the downstream areas of the Sekong and Srepok catchments, where the CV ≥ 0.9 (Figures 3.4n-p). The CV values of sediment were higher than that of rainfall, indicating a higher temporal variation in sediment than rainfall in each sub-catchment (Figures 3.4q-t). The relatively homogenous CV values of sediment were observed in most sub-catchments of the Chinit and Sen catchments (Figures 3.4q-r), whereas uneven CV values were observed in the Sekong and Srepok catchments, indicating the different sedimentation processes caused by slope in these two catchments (Figures 3.4s-t). A high correlation (R = 0.6–0.8) between rainfall and sediment was observed in most of all sub-catchments, demonstrating a high linear relationship between rainfall and sediment response (Figures 3.4u-x).

The sub-catchments were classified into the same four clusters for each regionalization method. Additionally, several sub-catchments within a catchment (e.g., Sen catchment) were grouped into a single cluster, through the PS approach (Figure 3.5a). However, the SRS approach provided more clusters in one catchment (Figure 3.5b). This may add to the improvement of estimating sediment load in ungauged catchments (Section 3.3.2). The donor sub-catchments, identified by the PS regionalization approach, were dominated by the sub-catchments of the Sen catchment (10 out of 14 sub-catchments) when the Chinit was considered the pseudo ungauged catchment, indicating that these two catchments have similar physical characteristics (Figures 3.5a-b). The donor and receiver sub-catchments, when the other catchments (i.e., Sen, Sekong and Srepok) were considered as the pseudo-ungauged catchment, are shown in Figures A12-14.

3.3.2. Evaluation of the performance of regionalization methods

The performance of the SRS and PS approaches were compared based on three statistical indicators (NSE, PBIAS, and R^2) for sediment load simulation, at each monitoring station in the Chinit, Sen, Sekong, and Srepok catchments with the leave-one-out cross-validation approach. The daily observed and simulated sediment load of all catchments for the model calibration and cross-validation phases from 2005 to 2008 are shown in Figure 3.6. Most results for the three model simulation schemes reproduced the seasonal patterns well for both seasons. However, there were some underestimations, specifically for extreme peak flow periods (Figure 3.6). Statistical indices summarized for the calibration model on site (Table 3.3) demonstrate a very good performance of daily sediment load simulations, with NSE and R^2 values of > 0.75, and PBIAS values within ±10% in the three catchments (i.e., Chinit, Sen, and Sekong). Satisfactory performance was found for the Srepok catchment (NSE and R^2 of 0.61), suggesting that the model parameters in each catchment were calibrated well. In the cross-validation phase, on average, the performance of the SRS approach (NSE and R^2 of 0.75 and 0.76, respectively) was slightly better than the PS approach (NSE and R^2 of 0.72 and 0.74, respectively) and close to the results of the at-site calibrated models (NSE and R^2 of 0.77) (Table 3.3).

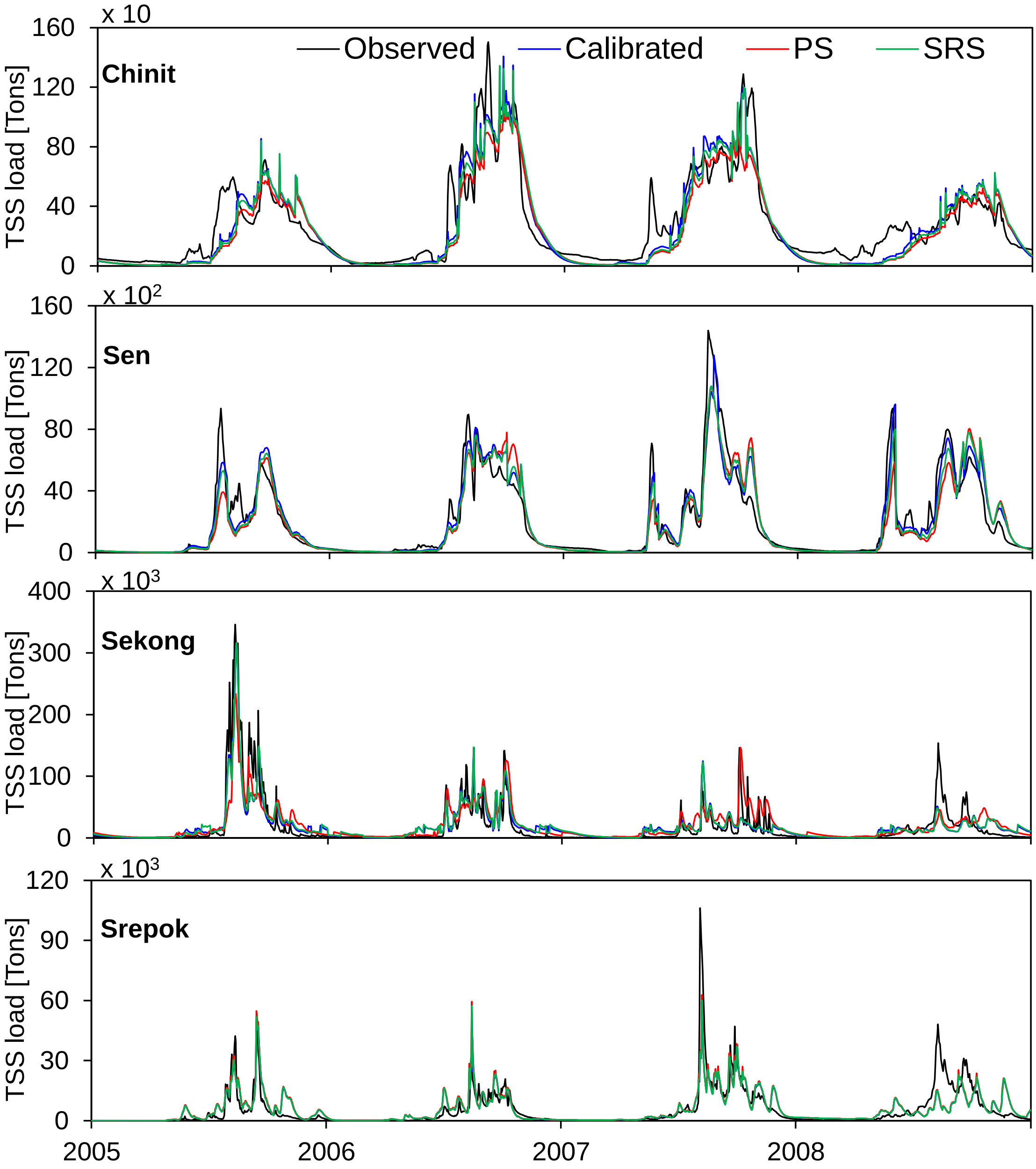


Figure 3.6. Comparison of daily sediment loads at the gauging station of each catchment showing observed (i.e., LOADEST-based), calibrated, PS-derived, and SRS-derived sediment loads.

Table 3.3. Results of statistical indices during the model calibration and cross-validation for the sediment load at the gauging station of each catchment.

At-site Calibration

River	R^2	NSE	PBIAS (%)
Chinit	0.80	0.78	8.30
Sen	0.86	0.86	-1.15
Sekong	0.82	0.82	-7.81
Srepok	0.61	0.61	-3.90
Average	*0.77*	*0.77*	*-1.14*

Cross-validation [Physical Similarity (PS)]

River	R^2	NSE	PBIAS (%)
Chinit	0.78	0.76	16.70
Sen	0.81	0.81	10.53
Sekong	0.74	0.74	-13.56
Srepok	0.61	0.59	-8.70
Average	*0.74*	*0.72*	*1.24*

Cross-validation [Sediment-Response Similarity (SRS)]

River	R^2	NSE	PBIAS (%)
Chinit	0.79	0.77	10.00
Sen	0.84	0.84	7.58
Sekong	0.82	0.79	-9.76
Srepok	0.61	0.60	-5.00
Average	*0.76*	*0.75*	*0.70*

Figure 3.7 compares the monthly sediment load simulated by each method (i.e., calibration on site, PS and SRS) and the observations for each catchment. The high values of the coefficient of determination ($R^2 > 0.70$) indicate that simulated sediments of the calibration and cross-validation approaches have a strong linear relationship with the observed monthly data. As seen in the scatter plots, the SRS method outperformed the PS method in three catchments (Chinit, Sen, and Sekong).

Additionally, the same correlation coefficients were obtained for the Srepok catchment, demonstrating more reliable sediment load estimation by the SRS regionalization method. Figure 3.8 shows the observed and simulated mean monthly sediment load for each catchment (2005–2008). The simulated sediment load obtained from the three methods showed a seasonal pattern similar to the observed results. It was found that the PS-based sediment loads in the Sekong catchment showed marked overestimation during October and December, while results from the SRS and calibrated models were close to the observed results (Figure 3.8).

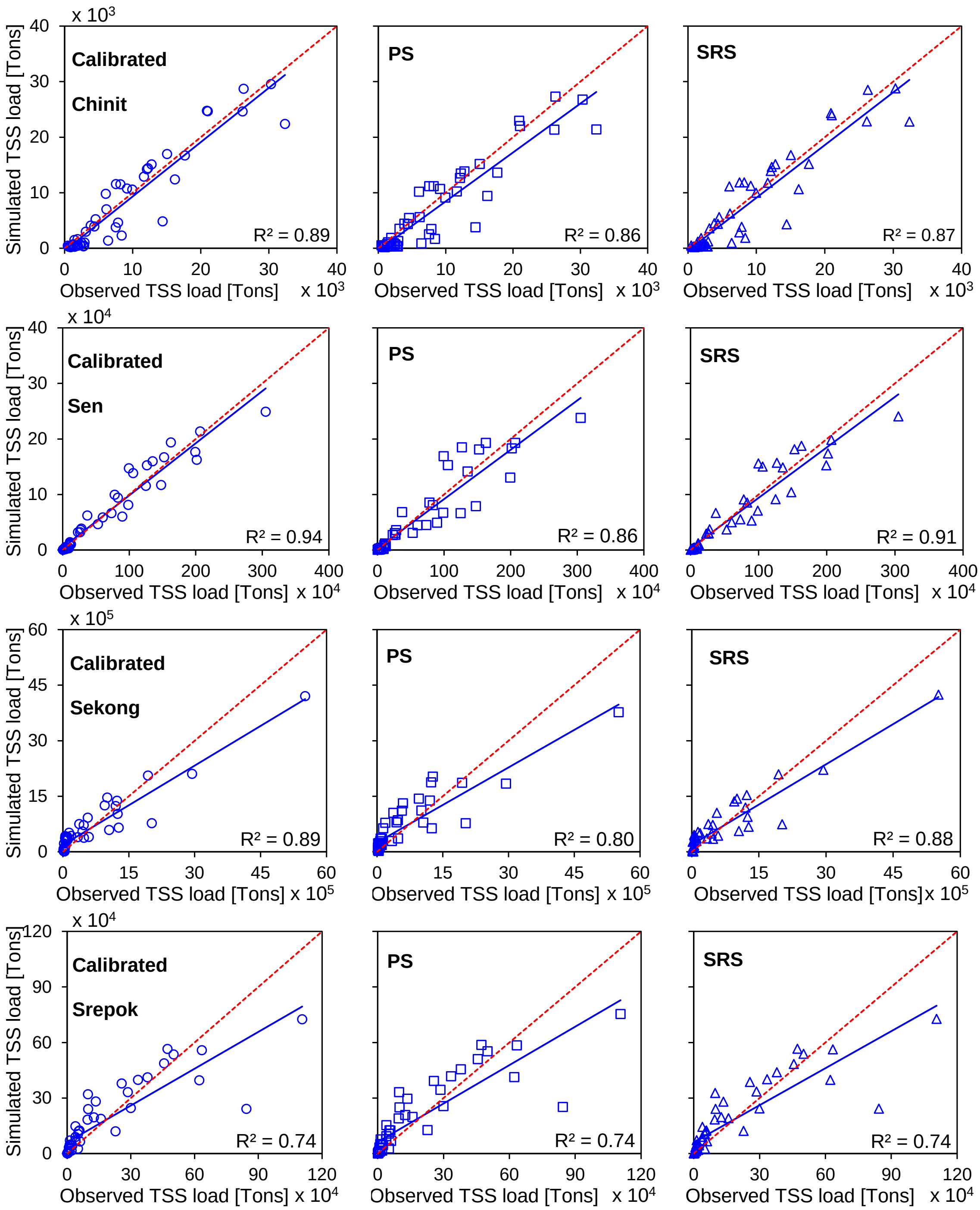


Figure 3.7. Scatter plots of monthly observed sediment load with calibrated (circles), PS-derived (squares) and SRS-derived (triangles) sediment loads at the gauging station of each catchment. The red dash lines are 1:1 line, and the blue solid lines denote linear regressions.

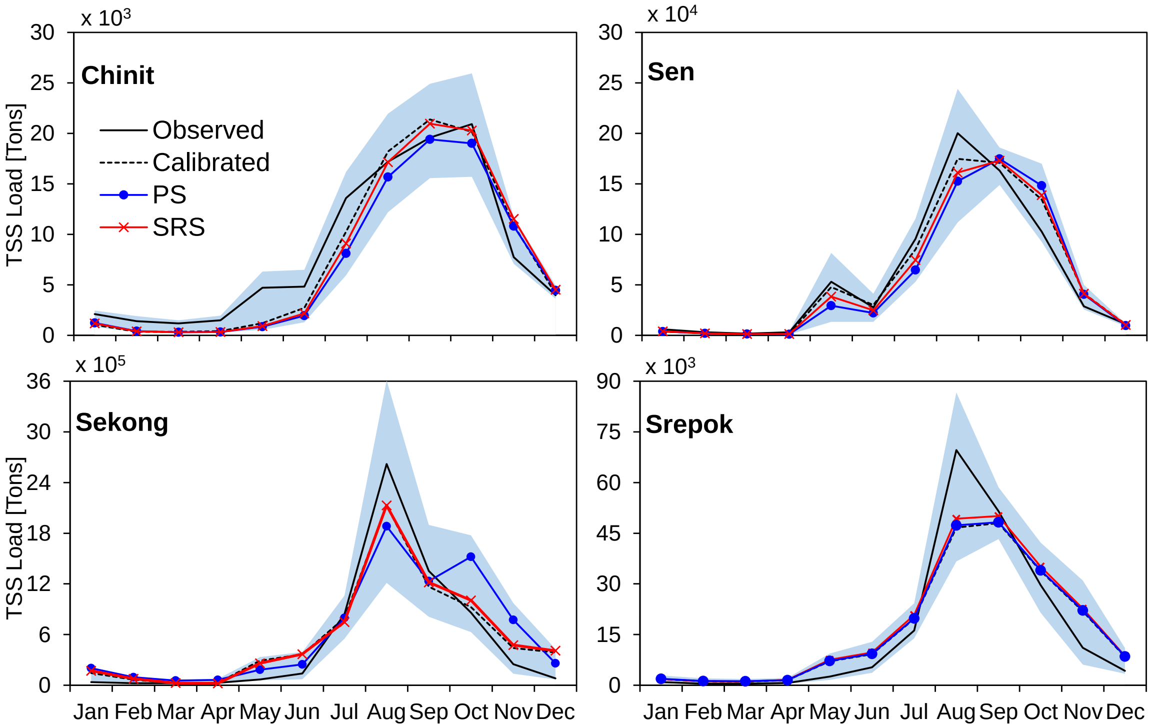


Figure 3.8. Observed and simulated the mean monthly sediment load at the gauging station of each catchment. Simulated results are based on the calibration, PS and SRS methods. The shaded area denotes the maximum and minimum standard errors of upper and lower bounds from the above-mentioned four simulation schemes.

The comparison of observed and simulated annual sediment load from each catchment indicated that the SRS cross-validation and at-site calibration showed similar and better overall performances in sediment simulations, whereas the estimated errors of total sediment load based on the PS method cross-validation were up to ± 16% (Figures 3.9 and A15). The SRS approach proved to be superior for sediment prediction in the ungauged catchments, with an improvement of up to 7%, compared with the PS method (Figure A15). The reliability of the seasonal variation of sediment simulated by the PS models was relatively lower (as shown in Figures 3.6 and 3.8), and could result in an overestimation in the total annual sediment load simulation.

More importantly, the student t-tests were carried out further to check the statistically significant improvement of model performance. Two groups of paired samples were tested by hypothesis student t-test. There are pairs of NSE and R^2 at four gauging sites from the PS and SRS regionalization methods (Table 3.4). The null hypothesis assumes that the SRS method does not significantly improve the model performance, while the alternative hypothesis means that the model performance is significantly improved. It can be inferred that the improvement of NSE achieved through sediment-response similarity is relatively more distinguished, while the enhancement of R^2 is relatively less significant. The p-values were 0.03 (< 0.05) for NSE and 0.09 (< 0.10) for R^2, demonstrating that the null hypothesis should be rejected at 95% and 90% confidence intervals, respectively (Table 3.4). According to the results of student t-tests, a statistically significant improvement in model performance was found when considering sediment-response similarity (i.e., sediment response and rainfall characteristics) as the catchment attribute for the regionalization method, consequently enhancing sediment simulation at ungauged river basins.

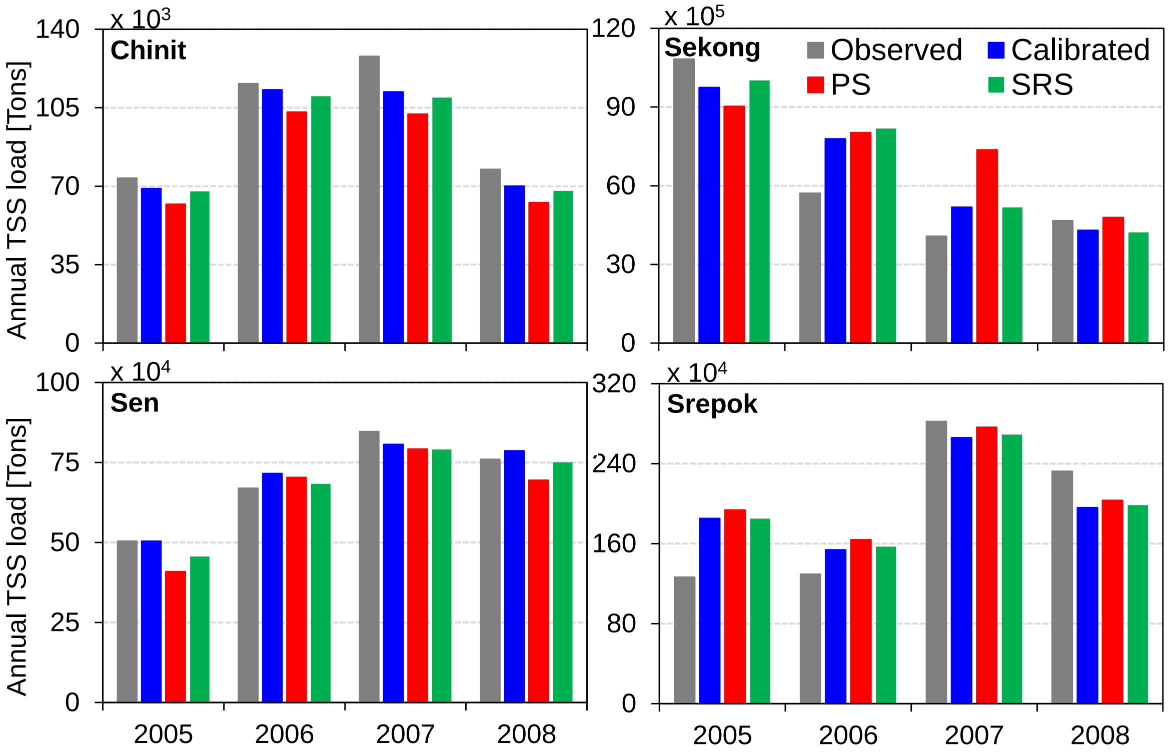


Figure 3.9. Comparison of annual calibrated, PS- and SRS-derived sediment loads at the gauging station of each catchment. The annual sediment load for each catchment represents the sediment load from the gauged area of each corresponding catchment.

Table 3.4. Significance test (student-t) for improvement in NSE and R^2

Indices	n	T	p	α
NSE	4	2.66	0.03	0.05
R^2	4	1.76	0.09	0.10

Note: "n" is the sample size, "T" is the t-statistic score, "p" is the probability, and "α" is the significant degree

Subsequent to updating the model parameters in the second phase of the SRS procedure, the sub-catchments were clustered again (as shown in Figure 3.5c) using the simulated sediment load. Therefore, the sub-catchment clusters were derived from the AM transferred parameters in the first phase, and then derived from the updated parameters in the second phase, which could be compared and checked (Fig. A16). There were a few sub-catchments that were assigned with different cluster numbers between Figures 3.3a and 3.3b. This indicated that the AM method was reliable in tentatively transferring parameters.

The proposed SRS regionalization method effectively addressed the main challenge of selecting ideal catchment attributes in conventional approaches. The conventional regionalization methods found it difficult to select the catchment's key attributes that ideally identified hydrological and sediment similarities. Thus, relatively higher uncertainty was found in the hydrological prediction at the ungauged catchments for the PS method. However, the SRS method considered the spatiotemporal variations of rainfall and sediment response through the CV and R indices, as the catchment attributes that could relatively better determine the hydrological and sediment similarities between gauged and ungauged catchments. The SRS approach generated more sub-catchment clusters in a single catchment than the conventional method (as shown in Figures 3.5 and A12-14), providing a diversity of regionalized model parameters and a better representation of hydrological and environmental conditions in the ungauged catchment. Eventually, the new regionalization SRS approach could reduce the uncertainty of sediment estimation in ungauged catchments, compared with the conventional method (as shown in Figures 3.6-3.9, A12, and Tables 3.3-3.4). Additionally, it showed the potential for catchment-wide sediment load simulations in the ungauged catchments (Section 3.3.3).

3.3.3. The application of the proposed regionalization method in ungauged catchments of the TSL basin

The newly proposed regionalization method, i.e., sediment-response similarity, needed to be validated before being applied in the actual ungauged catchments. Through the comprehensive analyses shown in Sections 3.3.1–3.3.2, the SRS method was successfully validated and outperformed the conventional method. Therefore, the SRS approach was selected to simulate the sediment load at the stations of the nine ungauged catchments of the TSL basin (see Figure 3.3) during the past decades (i.e., 2001–2011). In Chapter 2, we successfully calibrated and validated the SWAT model parameters in flow simulations at the hydrological monitoring stations of each catchment. Therefore, those flow-calibrated parameters (see Tables A9) and the regionalized-sediment parameters from the best donor catchments (i.e., Chinit, Sen, Sekong and Srepok catchments, see Figure A17 and Tables A10-A13) were inputted into the SWAT model. Subsequently, the sediment load in each ungauged catchment was simulated from 2001 to 2011. Eleven catchments drain into the lake (Figure 3.3). The total area of the TSL basin, disregarding the floodplain, is 83,107 km^2, and the flows of 41,159 km^2 (~50%) are gauged (Table 3.1). For computational purposes, the tributary basins can be divided into gauged, ungauged, and flooded areas (Kummu et al., 2008) (see Fig. A18 and Table A14).

Figure 3.10 illustrates the mean simulated annual sediment load values from the total drainage area (2001–2011), from the SRS approach and SWAT modeling, excluding the flooded regions in each catchment. The simulated mean annual total sediment load from all catchments was approximately 2.30 Mt/year from 2001–2011 (dashed line in Figure 3.10). Kummu et al. (2008) used hydrodynamic modeling to estimate the mean annual total sediment load for the same 11 catchments of the TSL basin during the period 2001–2003, and found that the annual total load was approximately 1.77 Mt/year. The current study result for the corresponding period (2001–2003) was 1.81 Mt/year, which is similar to that of the previous study. The annual variation of sediment load flux into the lake was significant. For example, it varied from 1.45 Mt in 2003 to > 3.00 Mt in 2006, as presented in Figure 3.10.

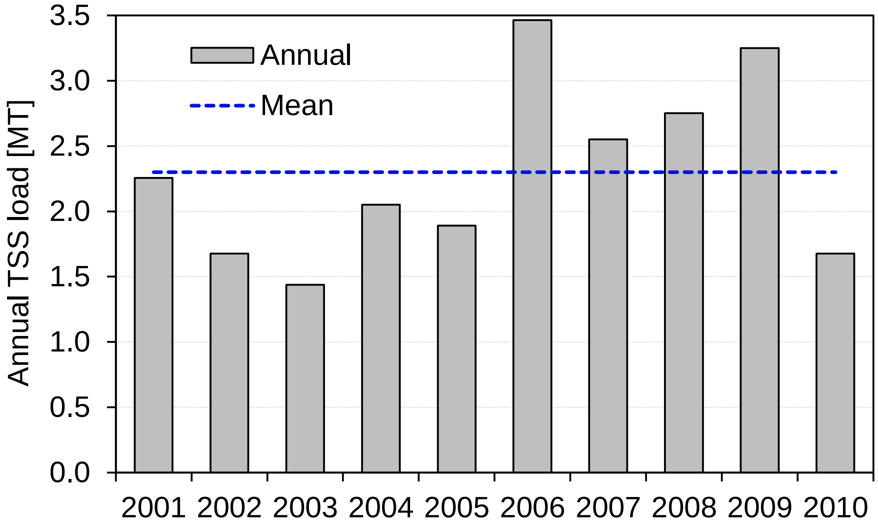


Figure 3.10. Annual total sediment load into Tonle Sap lake. The total load is the summation of load from 11 catchments. The years shown are hydrological years (from May 1, of the year indicated to April 30, of the following year, for the period 2001–2011).

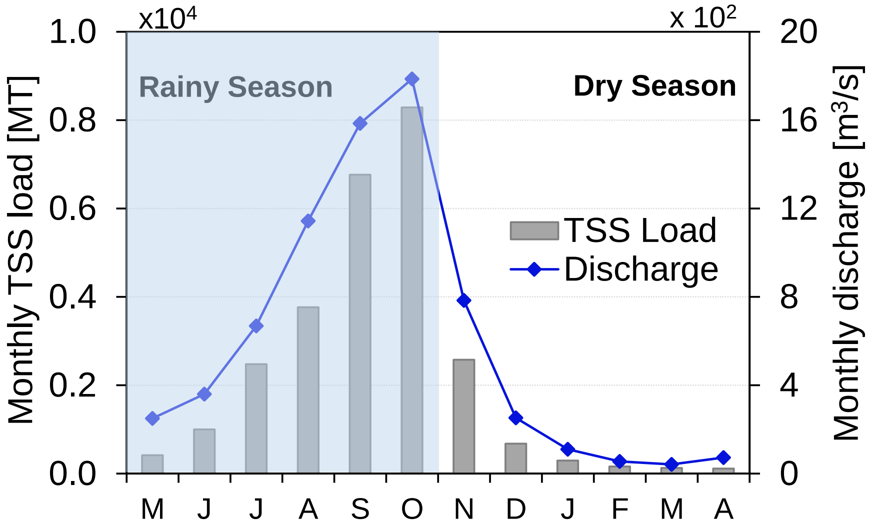


Figure 3.11. Seasonal variation of total sediment load into Tonle Sap lake between 2001–2011. The total load was calculated as per the description in Figure 3.10.

The long-term mean monthly variations in total sediment load from all catchments flowing into the lake from 2001 to 2011 are presented in Figure 3.11. Results showed that the sediment transport was dominant in the high-rainfall period (September–October), and the highest sediment load (and discharge) occurred in October, exceeding 0.80 Mt, and dropped to 0.01–0.25 Mt during the dry season, or receding flood period (November–April). The total annual sediment loads from Sen, Sangker, and Baribo catchments contributed approximately 65% (~ 1.55 Mt) of the total sediment into the lake from all catchments, while the smallest contribution (~ 0.20 Mt) was from the Dauntri and Sreng catchments (Figure A19). Additionally, the average total sediment yield from all catchments was approximately 29 t/km^2/year, between 2001 and 2011 (Figure A20).

The total annual sediment yields from the Sen, Sangker, and Baribo catchments were higher than other catchments (> 50 t/km^2/year). The annual sediment yield of the Sekong catchment was the highest (> 200 t/km^2/year). The Srepok catchment produced a sediment yield of approximately 70 t/km^2/year, which was comparable with that of some catchments of the TSL basin, including the Baribo, Sangker, and Sen catchments, which had a sediment yield of 55–98 t/km^2/year between 2005 and 2011 (Figure A21).

3.4. Conclusion

In this study, a novel SRS regionalization method has been proposed, using the SWAT model and SOM clustering technique to overcome the limitation of the critical attributes of a catchment favoring sediment similarity, which usually exists in the conventional regionalization approaches. Additionally, the SRS minimizes uncertainty when simulating sediment load in ungauged catchments. The performance of the SRS model was evaluated by comparing it with the conventional regionalization methods, and the results of model calibration in four catchments of the TSL and 3S basins. The results indicated the comprehensive performance of the SRS regionalization method for estimating sediment load in the ungauged catchments. It considered the spatiotemporal variations of sediment response and its relationship with rainfall characteristics as a catchment attribute and showed the potential to ideally determine hydrological and sediment similarities between gauged and ungauged catchments. The SRS approach obtained an estimation error reduction of up to 7%, compared with the PS regionalization method. Compared to the conventional approach, there was an improvement in statistical metrics (NSE and R^2 = 0.75 and 0.76, respectively) for the SRS method, which was close to the calibration model results (NSE and R^2 = 0.77) thereby outperforming the conventional approach. The SRS method was also selected to estimate sediment load in the ungauged tributaries of the TSL basin. Results showed that the annual total sediment flux into the lake from its tributaries is approximately 2.3 Mt/year, with the highest sediment load occurring in October, exceeding 0.8 Mt.

The proposed SRS regionalization method effectively addressed the main challenge in conventional approaches, which is the selection of ideal catchment attributes that favor hydrologic similarity between gauged and ungauged catchments. Thus, the SRS regionalization method proposed in this study is a global alternative method for estimating sediment, as well as other hydrological variables and rainfall-driven phenomena such as streamflow and nutrient transport, in ungauged catchments. However, in this study, the applicability of the SRS regionalization method is tested on a limited number of catchments and gauged datasets. Therefore, it is recommended to incorporate more gauged catchments, especially those with a larger scale and more sufficient in-situ data, to widely extend the method's applicability and minimize the associated uncertainty.

Chapter 4 Decadal climatic variability, land-use change and hydro-environmental impact assessment

This chapter brings altogether the two methodologies or frameworks developed in Chapters 2 and 3 for the application to hydro-environmental impact assessment, owing to climate variation and land-use change during the last few decades. The final output of this section is the target of Objective 3.

4.1. Introduction

Human population growth, anthropogenic activities, and land use and land cover changes, exacerbated by climate variation, have led to the freshwater lakes and river basins being identified as some of the most severely endangered ecosystems (Rosenzweig et al., 2007). As the result, freshwater resources in many regions of the world have deteriorated in terms of quality and quantity due to the above-mentioned environmental-related stressors (Bastia and Equeenuddin, 2016; Phi Hoang et al., 2016). Among them, climate and land-use change have been identified as two important drivers affecting water resources at both regional and global scales (Khoi et al., 2022; Shrestha et al., 2018). Climate variability can significantly impact the spatiotemporal characteristics of precipitation and temperatures and as a result, cause changes in hydrological processes and river flow regimes (Oeurng et al., 2019; Zhao et al., 2022). The changes in hydrological processes then lead to changes in the transport of sediment and nutrient yields (Shrestha et al., 2018). Moreover, changes in land-use types can affect the hydrological components, including evapotranspiration, infiltration, base flow, and surface runoff, as well as soil erosion and degradation (Khoi et al., 2022; Li et al., 2009; Shrestha et al., 2018; Wan et al., 2014).

Numerous studies have examined the impacts of climate variability and land-use change on hydrology and water quality in many different river basins. For instance, Shrestha et al. (2018) found a maximum of 24 and 15.25% decrease in streamflow and nitrate nitrogen, respectively under climate change in Songkhram River, Thailand. Whereas land-use change projection is found to be responsible for a maximum 6.35% increase in streamflow and an 11.60% decrease in nitrate nitrogen loading in the same river basin. Climate variability was found to cause an increase of 1.00% in streamflow, and 2.91% in sediment loading, while the effect of land-use change increased by 0.01%, and 3.70% in streamflow and sediment load, respectively in Sekong, Sesan and Srepok (3S) rivers (Khoi et al., 2022). Fu et al. (2019) reported that the streamflow changed by −39.1% due to climate change and 2.2% due to land-use change in the Woken River, China. Furthermore, climate change caused an increase in the annual streamflow of the Johor River in Malaysia by 2.9%, while land-use change only contributed to an increase of 0.1% (Tan et al., 2015).

The TSL Basin in Cambodia is the largest freshwater body in Southeast Asia and one of the most productive ecosystems in the world (Uk et al., 2018), playing a crucial role in livelihood and sustainable development in Cambodia and the Lower Mekong region (Shivakoti and Bao, 2020). Despite its significant value, the lake ecosystem is widely under threat from climate change together with anthropogenic activities inside and outside the TSL basin (Oeurng et al., 2019; Shivakoti and Bao, 2020; Siev et al., 2018; Uk et al., 2018). According to the impact assessments, TSL and its floodplain are vulnerable to hydrological changes, attributed mainly to the development of water infrastructure (Arias et al., 2014; Johnstone et al., 2013; Masumoto et al., 2008; Morovati et al., 2023), climate change (Frappart et al., 2018; Johnstone et al., 2013; Morovati et al., 2023; Oeurng et al., 2019; Phi Hoang et al., 2016), water quality degradation (Chea et al., 2016; Soum et al., 2021) and land use/ land cover change (Chen et al., 2022; Niu et al., 2022; Senevirathne et al., 2010).

Previous research works paid more attention to the Mekong mainstream and anthropogenic activities along the Mekong River, whereas limited information is available regarding hydro-environmental impacts from the tributary basins of the TSL (Ang and Oeurng, 2018; Kummu et al., 2014). A better and more appropriate assessment of the climate variability and land-use change impacts within the TSL Basin will facilitate better management and environmental conservation of the TSL Basin. The overall objective of this study is to quantify the impacts of climate variability and land-use change on the streamflow and sediment load of the TSL Basin during the past decades, using hydrological modeling approaches. First, the spatiotemporal variation of decadal climatic was analyzed from 2001 to 2020. Furthermore, the contribution of climate variation to streamflow and sediment change was investigated by using the SWAT model to simulate streamflow and sediment during this period. Finally, four different land-use change scenarios were analyzed and input into the hydrological modeling for streamflow and sediment load impact assessment. The findings of this study can support policymakers and planners in proposing suitable management practices to cope with the impact of climate and land-use changes on water quantity and quality in the TSL Basin.

4.2. Materials and methods

4.2.1. Study area

The study was conducted in the TSL Basin. The general information on the TSL Basin can be found in Chapter 2. The location map showing eleven tributaries of the TSL Basin is presented in Figure 4.1.

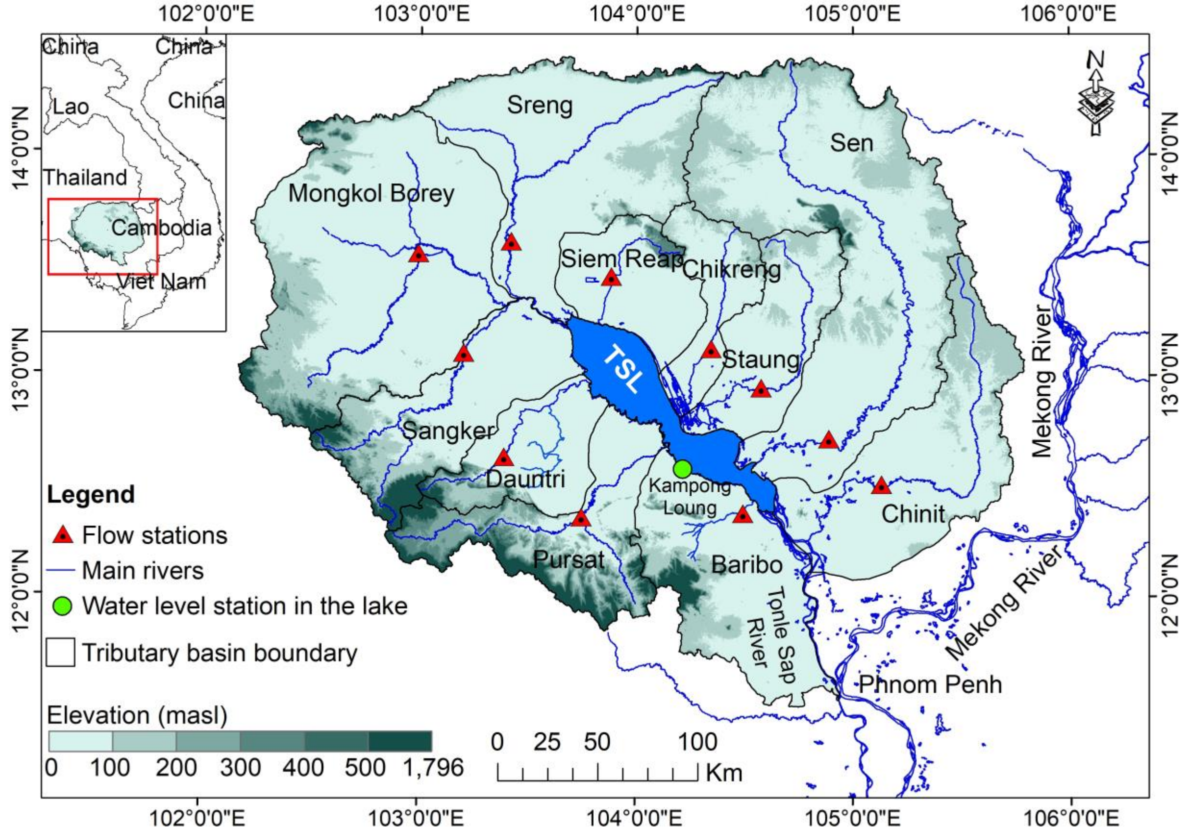


Figure 4.1. Location map of the TSL Basin. The green circle denotes the station monitoring the water level in the lake area. The inset indicates the location of the TSL Basin in the Indochina Peninsular.

4.2.2. Hydrological model

The SWAT model was used in this analysis. It is a semi-distributed model, a widely used hydrological model to simulate streamflow, sediments, and other hydrological variables under diverse environments and different scales (Arnold et al., 1998). More details about the SWAT can be found in Chapter 2 and official theoretical documentation (Neitsch et al., 2011) and the review paper by Gassman et al. (2007).

To study the impact of climate and land-use changes on sediment load, the regionalization method (i.e., sediment-response similarity) proposed in Chapter 3 was used to transfer optimized parameters (i.e., regionalized parameters of model optimization using the 2002 land-use map) from gauged to ungauged tributary basins for sediment load simulation. Climate variability impacts were analyzed by keeping the land-use constant (i.e., 2002 land-use condition), while the impacts of land-use change were estimated by driving the regionalized SWAT model with 1995, 2010 and 2018 land-use datasets. As shown in Chapter 3, the SRS approach provided a diversity of regionalized model parameters and a better representation of hydrological and environmental conditions in the ungauged catchment. As the result, the accuracy in parameter regionalization of the SRS method was better than that of the conventional method; therefore, a more reliable assessment of climate and land-use change impact on sediment load in ungauged tributary basins could be estimated by using the SRS regionalization approach.

4.2.3. Data collection and pre-processing

Each dataset used for the SWAT model setup is summarized in Table 4.1. Four different land-use periods: 1995, 2002, 2010, and 2018 land cover maps, together with the meteorological data from 2001 to 2020 were used to analyze spatiotemporal changes in climate and land use and input into the SWAT model to simulate their impacts on streamflow and sediment loading in the TSL Basin. The study in Chapter 2 found that precipitation from Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM) (IMERG) and air temperature data of Southeast Asia-Observational (SA-OBS) and Climate Prediction Center (CPC) are the optimal meteorological data for comprehensive basin-wide hydrological impact assessments in the TSL basin. Although the SA-OBS temperature is the most outstanding dataset, for three years (2018-2020), the CPC-based temperature was used since the data availability of SA-OBS is only up to 2017. Since the land-use map used in this study is different from that used in Chapter 2, model optimization needs to be performed again using land-use conditions in 2002 before impact assessment. Water-level data, which were provided by the MRC (Table 4.1), from a monitoring station of each catchment (see Figure 3.3) were used as the observed data for flow model calibration (2001–207) and validation (2008–2011). The results of the flow calibration and validation show that on average, the model achieved satisfactory performance (Tables A3 and A15).

After calibration and validation of the SWAT model using the 2002 land cover map, the impacts of the three land-use change scenarios on the streamflow and sediment load were simulated by driving the calibrated SWAT model with 1995, 2010 and 2018 land-use datasets. The SWAT model was not optimized or calibrated for different land-use conditions for the land-use change impact assessment. If the model were optimized (i.e., the model is calibrated to minimize the error and get similar values to the observation) for different land-use data, the results from model optimization would not reflect the impact of land-use change since the simulation values were optimized to be similar with the observed values.

Table 4.1. Summary of required data used for SWAT model setup and calibration

Data type	Time period	Temporal resolution	Spatial Resolution	Sources
Topography (Digital Elevation Model [DEM])	-	-	90 m	http://srtm.csi.cgiar.org
Land cover map	1995, 2002, 2010 and 2018	-	30 m	SERVIR-Mekong
Soil types map	2002	-	250 m	Mekong River Commission
Total rainfall, Maximum and minimum air temperatures	2001–2020	Daily	-	Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM) (IMERG), Southeast Asia-Observational (SA-OBS) and Climate Prediction Center (CPC)
Water level	1995–2011	Daily	-	MRC and Department of Hydrology and River Works (Cambodia)

On the other hand, the impact assessment in this study was conducted based on the assumption that how much the streamflow and sediment load in the catchment would change after the certain area of land cover in the corresponding catchment is changed and input into the SWAT model. Therefore, the SWAT model was run for each scenario using the climate data for the period 2001–2020, and the results under each land-use scenario were compared to the corresponding streamflow and sediment load values for the 1995 land-use condition for the land-use change impact assessment. The same approach of studying the land-use change impact using the SWAT model was also found in the previous studies (e.g., He et al., 2021; Hu et al., 2021; Larbi et al., 2020; Samal and Gedam, 2021; Shrestha et al., 2018, among others).

4.2.4. Lake surface area and volume

The approach developed by (Sakamoto et al., 2007) and simplified by Normandin et al. (2018) to monitor flood extent in the Mekong Basin and the TSL Basin by (Frappart et al., 2018) was adopted in this study (processing steps are presented in Figure A22). It is based on the thresholding of the Enhanced Vegetation Index (EVI), the Land Surface Water Index (LSWI), and the Difference Value between EVI and LSWI (DVEL) to determine the status (non-flooded or flooded) of any pixel in an 8-day composite MODIS image of surface reflectance. The two indexes are defined as follows (Huete et al., 1997):

$$EVI = 2.5 \times \frac{\rho_{NIR} - \rho_R}{\rho_{NIR} + 6 \times \rho_R - 7.5 \times \rho_B + 1} \tag{4.1}$$

$$LSWI = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{WSIR}} \tag{4.2}$$

where ρ_{NIR} is the surface reflectance value in the near-infrared (841–875 nm, band 2), ρ_R is the surface reflectance value in the red (621–670 nm, band 1), ρ_B the surface reflectance value in the blue (459–479 nm, band 3), and ρ_{SWIR} the surface reflectance in the short-wave infrared (1628–1652 nm, band 6). Spatio-temporal variations of the flood were determined in the TSL Basin from January 2001 to December 2020.

The lake volume is calculated based on the water level at Kompong Luong station (H_{KL}). We followed the same approach reported by Kummu et al. (2014) to calculate the temporal lake volume (V) from 2001 to 2020. Kummu et al. (2014) used the digital bathymetric model (DBM) of the lake together with H_{KL} to develop the following relationships on a daily time scale which give a correlation coefficient greater than 0.99.

$$V(km^3) = 0.7307 \times H^2_{KL} - 0.354 \times H_{KL} + 0.9127 \qquad r^2 > 0.99 \tag{4.3}$$

4.3. Results and discussion

4.3.1. Spatiotemporal climatic variability

The temporal variation and trends in basin-scale mean annual temperatures (i.e., Tmax and Tmin) from 1990 to 2020 in the TSL basin are shown in Figure 4.2. The annual variation of annual Tmax was significant. For example, it varied from approximately 33.5 °C in 1998 to < 32.0 °C in 1999, as presented in Figure 4.2. The difference in Tmin between dry and rainy seasons was more obvious than that in Tmax. The increase of Tmin during the rainy season is higher than that in the dry season, indicating relatively warmer nights in the rainy season.

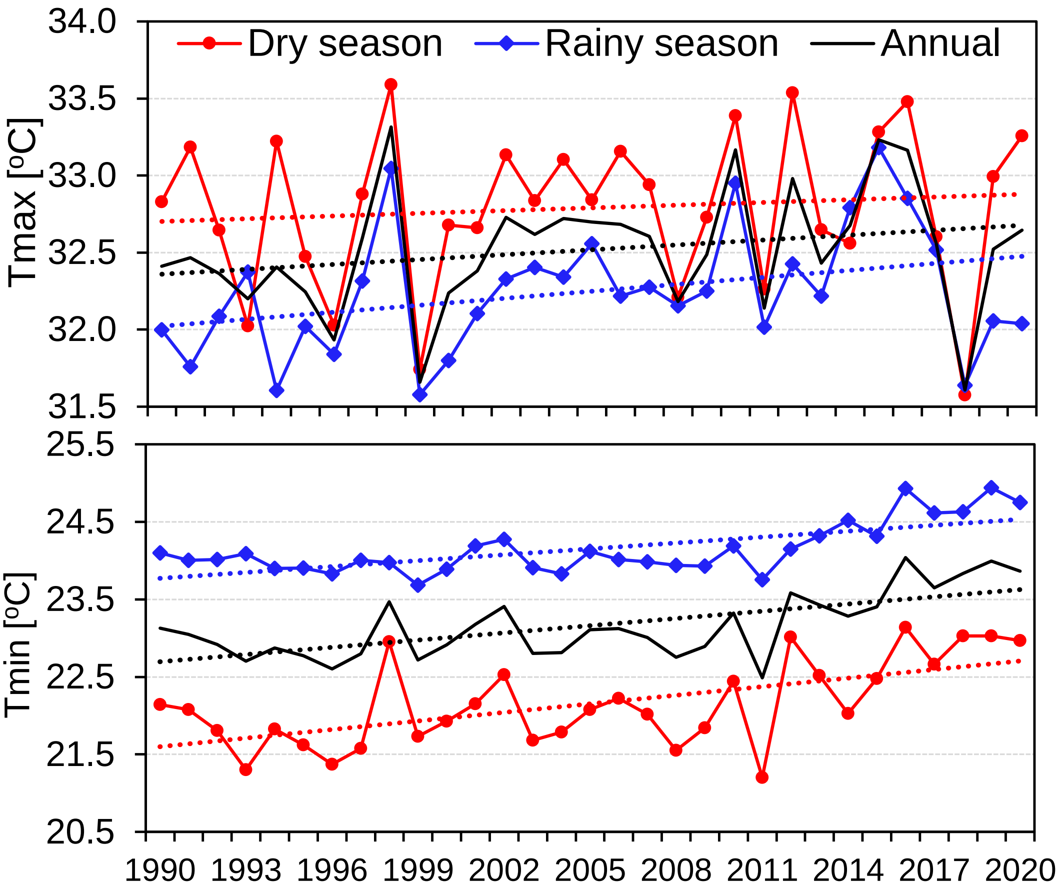


Figure 4.2. Temporal trends and variability of basin-scale seasonal and annual maximum (top) and minimum (bottom) temperatures from 1990 to 2020. The rainy and dry seasons are from May to October and November to April, respectively.

On average, maximum and minimum temperature increases have been observed, with an approximate increase of 0.15°C per decade since the 1990s. For maximum temperature, it was observed that the increases have risen relatively higher between the 2000s and 2010s, increasing by around 0.35°C (Figure 4.2).

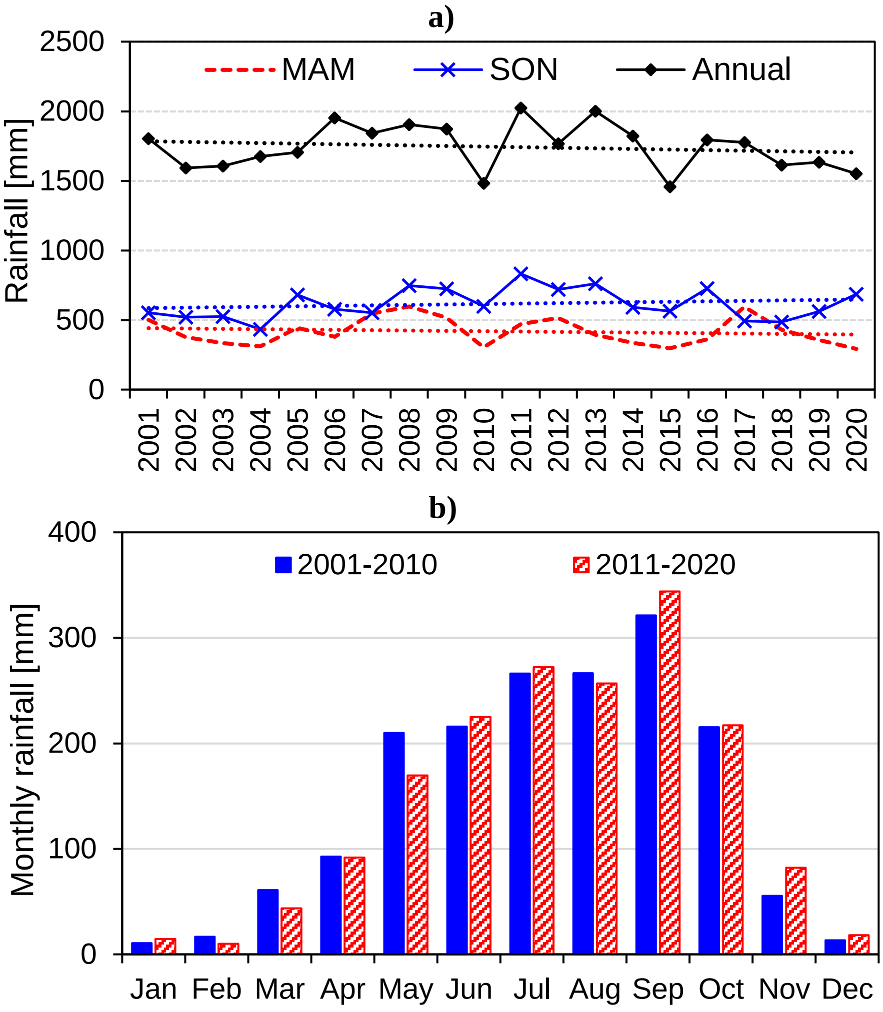


Figure 4.3. a) Temporal trends and variability of basin-scale MAM (March-April-May) rainfall, SON (September-October-November) rainfall, and annual rainfall from 2001 to 2017. b) Comparison of long-term mean monthly rainfall between the 2001–2009 and 2010–2017 periods.

The basin-scale annual rainfall was observed to slightly decrease from 2001-2020 (Figure 4.3). Figure 4.3a indicates that the annual rainfall over the TSL Basin varied from 1500 to 2000 mm between 2001 to 2020. Additionally, there was a relationship between precipitation variability and the El Niño Southern Oscillation (ENSO) phenomenon, with years recording strong El Niño correlated with years of low annual precipitation–especially during the last decade (Kabeya et al., 2021). The most recent El Niño that significantly altered Cambodia’s weather patterns was from 2014–2016, which began in late–2014 and lasted until mid–2016 causing higher temperatures (Figure 4.2), declined rainfall (Figure 4.3a), and eventually drought (Sutton et al., 2019). The rainfall was likely to decrease and increase before (March-May) and after (September-November) the wet monsoon, respectively, which would increase river discharge and flooding during the wet season, while extended droughts are likely to occur during the dry season. The March-May and September-November periods are the growing and harvesting seasons, respectively, of farming–especially for the paddy fields in Cambodia (Bunthan et al., 2018). Therefore, changes in monsoon rainfall amount and timing could impact people's livelihoods and the national food security of the country.

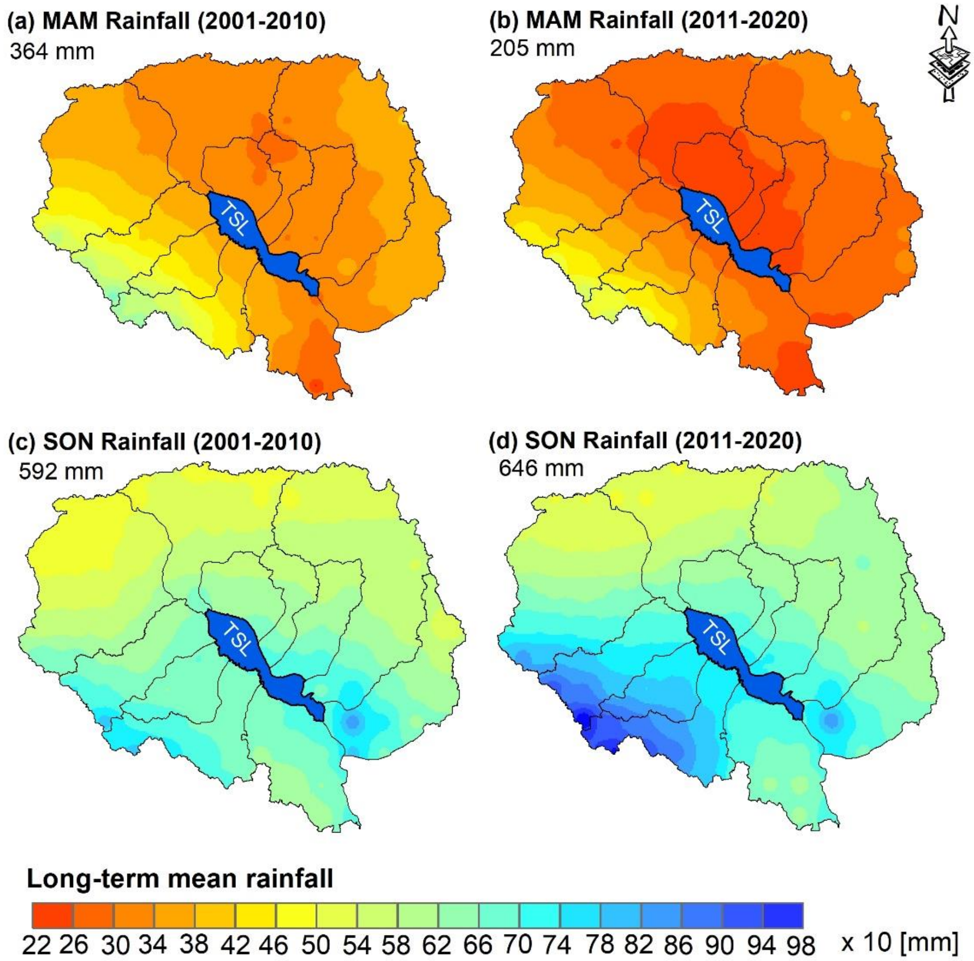


Figure 4.4. Spatial distribution of long-term mean MAM and SON rainfall between the 2001-2010 and 2011-2020 periods. The full name of MAM and SON are as per the description in Figure 4.3. The areal average of the mean rainfall is included at the top left corner of each map.

The long-term mean monthly rainfall between the 2001–2010 and 2011–2020 periods was compared and shown in Figure 4.3b. It was observed that rainfall decreased in some months between February and August but increased between October and January in 2011–2020 compared to that in 2001–2010, which would alter the seasonal pattern of streamflow and the amplitude of flood peaks in the TSL basin (as discussed in Section 4.3.3).

The spatial distribution of mean MAM and SON rainfall shows larger rainfall over the hilly regions in the southwest and northeast and smaller rainfall over the lower elevation zone extending from the northwest to the southeast, including the TSL (Figure 4.4). It was observed that from March to May, rainfall decreased across the TSL basin area between 2001–2010 and 2011–2020, decreasing by approximately 160 mm (Figure 4a-b). From September to November, on the other hand; rainfall increased by around 50 mm (Figure 4c-d). Since the decreased amount of rainfall before the wet monsoon was higher than the increased rainfall during September and November, there was a decreasing trend in annual rainfall between 2001 and 2020 (see Figure 4.3a).

4.3.2. Land-use/Land cover change analysis

The land cover maps of 1995, 2002, 2010 and 2018 were utilized to compute the area of all land cover classes to examine the variations which took place over time from 1995 to 2018 (Table 4.2 and Figure 4.5). From the 1995 land cover map, there are two major land cover classes in the TSL Basin: 56% of forest area (i.e., flooded forest, forest, plantation forest, evergreen broadleaf, and mixed forest) and 35% of cropland area (Table 4.2). The results presented in Figure 4.5 and Table 2.4 revealed that there is a substantial decrease found in forest-related areas while some growth in cropland areas was found. For example, the flooded forest was much deforested between 1995 and 2018, and the occupied area decreased from approximately 6,992 to 1700 km^2, or 76% of the flooded forest area was cut down (Table 4.2). On the other hand, the cropland area increased by about 23% from around 30,400 to 37,324 km^2 in 1995 and 2018, respectively. Table 4.2 indicates that there was a dramatic loss of natural forest area (i.e., flooded forest, forest, evergreen broadleaf, and mixed forest) between 1995 and 2002, in which around 40% cumulative loss was found. Additionally, there is another 5% of the cumulative loss from 2002 to 2018; thus, totally the cumulative loss of natural forest area is around 45% or the area decreased from 37,052 in

1995 to 20,408 km^2 in 2018 (Table 4.2). The change in natural forest cover found in this study is comparable with studies by Lohani et al. (2020; Lovgren (2020). The urban and grassland areas increased by around 12% and 229%, respectively between 1995 and 2018. The spatial changes in land cover would alter the seasonal pattern of streamflow and sediment load in the TSL basin (as discussed in Section 4.3.4).

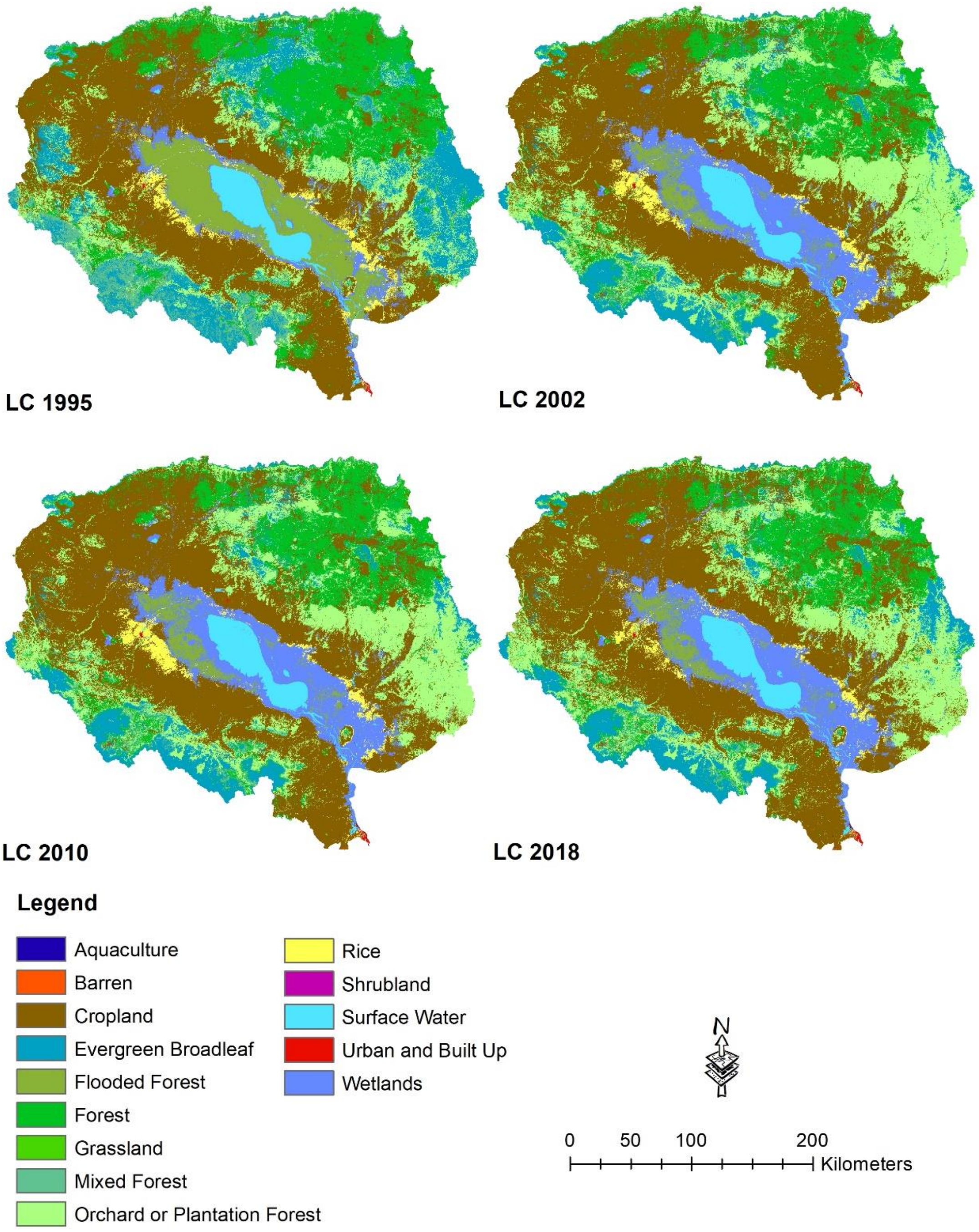


Figure 4.5. Land-use map of 1995, 2002, 2010 and 2018 of the TSL Basin.

Table 4.2. Summaries of land cover classes in the TSL Basin from 1995 to 2018

Land Cover Typology	1995		2002			2010			2018		
	Area [km^2]	Area [%]	Area [km^2]	Area [%]	*Change [%]	Area [km^2]	Area [%]	Change [%]	Area [km^2]	Area [%]	Change [%]
Surface Water	2,694.56	3.10	2,670.78	3.07	-0.88	2,642.80	3.04	-1.92	2,604.95	3.00	-3.33
Flooded Forest	6,991.94	8.05	1,814.72	2.09	-74.05	1,795.62	2.07	-74.32	1,699.52	1.96	-75.69
Forest	15,004.59	17.27	12,617.45	14.52	-15.91	12,115.66	13.95	-19.25	11,107.22	12.79	-25.97
Plantation Forest	11,714.43	13.49	18,044.46	20.77	54.04	17,364.01	19.99	48.23	17,119.27	19.71	46.14
Evergreen Broadleaf	11,155.79	12.84	5,893.67	6.78	-47.17	5,904.11	6.80	-47.08	6,488.81	7.47	-41.83
Mixed Forest	3,900.24	4.49	1,385.30	1.59	-64.48	1,301.41	1.50	-66.63	1,112.54	1.28	-71.48
Urban and Built Up	67.67	0.08	75.54	0.09	11.63	78.91	0.09	16.60	75.64	0.09	11.77
Cropland	30,399.46	34.99	34,773.45	40.03	14.39	36,041.51	41.49	18.56	37,324.14	42.97	22.78
Rice	1,868.38	2.15	1,561.16	1.80	-16.44	1,614.08	1.86	-13.61	1,252.61	1.44	-32.96
Barren	24.71	0.03	30.87	0.04	24.91	40.74	0.05	64.84	57.77	0.07	133.78
Wetlands	3,005.58	3.46	7,911.77	9.11	163.24	7,866.97	9.06	161.75	7,844.76	9.03	161.01
Grassland	18.28	0.02	45.53	0.05	149.08	38.61	0.04	111.24	60.06	0.07	228.58
Shrubland	13.82	0.02	33.63	0.04	143.28	53.68	0.06	288.33	112.88	0.13	716.59
Aquaculture	10.22	0.01	11.33	0.01	10.85	11.55	0.01	13.03	9.49	0.01	-7.16
Total	86869.66	100.00	86869.66	100.00	-	86869.66	100.00	-	86869.66	100.00	-

**change is calculated by comparing to the 1995 land cover area.*

4.3.3. Impact of climate variability

4.3.3.1. Contribution of climate variability to streamflow changes

Climate variability impacts on streamflow or sediment load were estimated by keeping the land-use constant with the baseline (i.e., land cover 2002). The main effect of climate variability on water resources is to alter streamflow quantity and timing, which is caused by changes in components of the water balance, mainly rainfall and air temperature. In this study, the spatiotemporal changes in temperature and rainfall characteristics impact the hydrological cycle in the TSL basin, which was evident in the changing streamflow pattern. There was a downtrend of annual and SON-flow volume, which was caused by the decrease in the annual and MAM rainfall amount, respectively (Figures 4.3-4.4, and 4.6a).

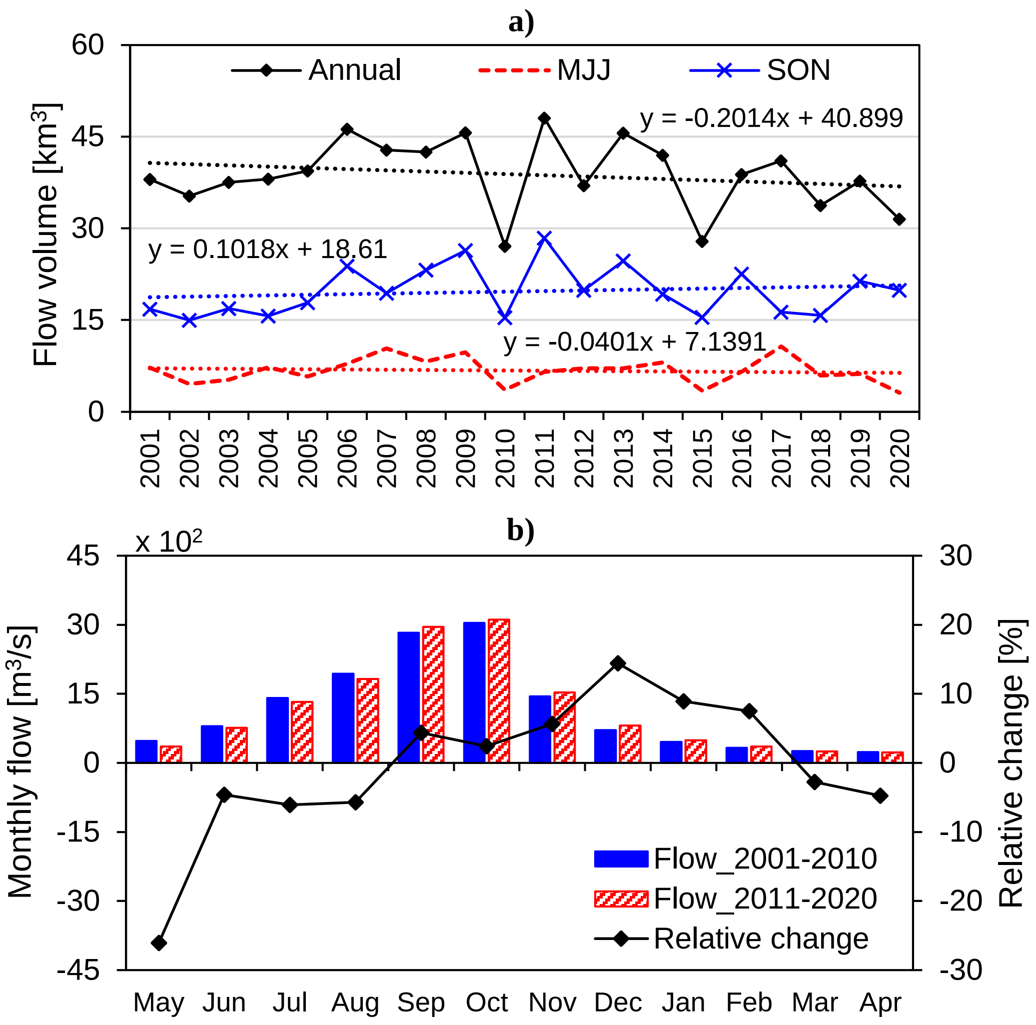


Figure 4.6. a) Temporal trends and variability of annual, MJJ (May-June-July) and SON (September-October-November) flow volume into the lake. b) Comparison of long-term mean monthly streamflow between the 2001–2010 and 2011–2020 periods. The flow was simulated from the total area of the TSL basin excluding the flooded area.

The decrease in streamflow can be attributed to an increase in evapotranspiration as a result of the increase in temperature and decrease in rainfall. The annual total and MJJ-flow volume into the lake experienced decreasing trend, with a decrease of 3.83 (9.40%) and 0.76 km^3 (10.73%), respectively since the 2000s (Figure 4.6a). On the other hand, an increasing trend was found in the SON-flow volume, with a rising of 1.93 km^3 (10.34%) between 2001 and 2020. During the last decade, i.e., 2012–2020, the impact of El Niño causes a drop in rainfall (Figure 4.3a) consequently leading to a decrease in streamflow and flow volume in the TSL basin (Figures 4.6a-b). The long-term mean monthly streamflow between the 2001–2010 and 2011–2020 periods was compared and shown in Figure 4.6b. The streamflow was likely to decline (max. 26% decrease in May) between May and August in 2011–2020 compared to that in 2001–2010, mainly because of the decreased rainfall between March and May (Figures 4.3-4.4 and 4.6). In addition, the peak flow and the streamflow increased (max. 14% increase in December) between September and December, which was caused by changes in rainfall after the wet monsoon (September to November) as shown in Figures 4.3 and 4.4. Therefore, the shifting of monsoon rainfall and the increasing amplitude of flood peaks would impact agricultural activity and fishery productivity in the TSL Basin. The basin streamflow showed a greater impact from climate variation than land-use change (discussed in Section 4.3.4).

4.3.3.2. Changes in lake surface area and volume and water level

Figure 4.7a presents temporal variations of surface water extent in the TSL drainage area between January 2001 and December 2020 using MODIS images. They exhibit a well-marked seasonal cycle, with minima occurring in April and May and maxima between September and November. There is a strong interannual variability of the maximum flood extent ranging from 5,035 km^2 in October 2015 up to 15,102 km^2 in October 2011. The minimum lake surface varies between 2142 km^2 in 2016 and 2647 km^2 in 2001 during the low stage. Water stage changes accordingly, with an annual amplitude (i.e. difference between the maximum and the minimum over the year) varying from 4.57 m in 2015 to 8.90 m in 2011 (Figure 4.7a). Temporal trends and variability of annual maximum water surface area and water level are presented in Figure 4.7b, in which decreasing trends of water extent and water level were found between 2001 and 2020, indicating a shrinking rate of lake volume over the years (see Figure 4.7c).

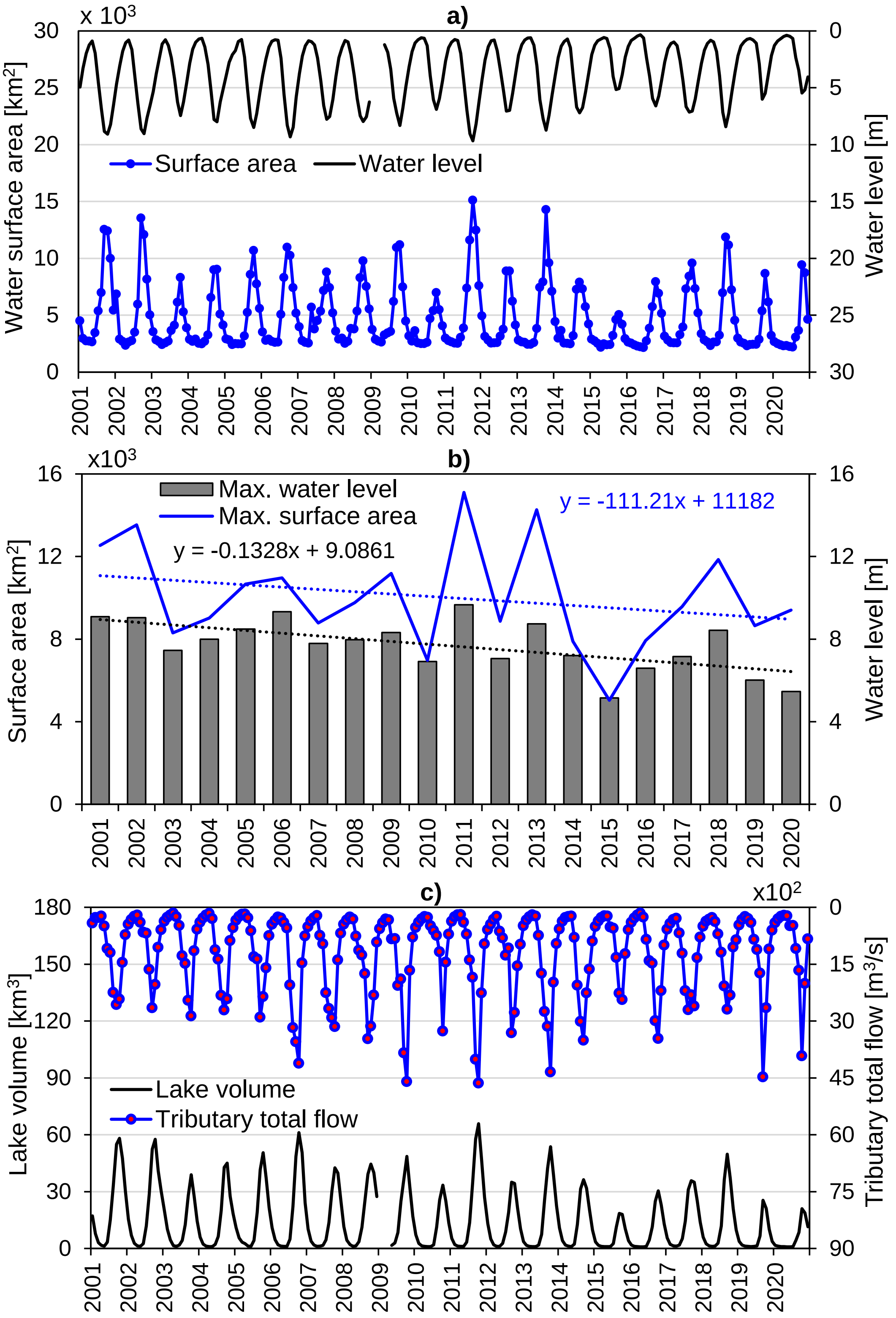


Figure 4.7. a) Monthly timer series and b) Annual maximum of surface water extent in the Tonle Sap basin based on the processing of MODIS images, and measured water level in the lake at Kampong Luong station over 2001–2020. c) Monthly time series of total streamflow from 11 tributary basins and lake volume. Refer to Figure 4.1 for the location of the Kampong Luong station. In a) and c) the discontinued lines denote missing data on water level and lake volume, respectively during this period.

The annual maximum lake surface area and water level experienced decreasing trend, with a decrease of 2113 km^2 (19%) and 2.52 m (28%), respectively (Figure 4.7b). Figure 4.7c shows that there is a relationship between lake volume and total streamflow from its tributary basins except for a few years in the beginning and at the end of the 2001-2020 period. The maximum lake volume and tributary flow are in 2011, while the minimum values were found in 2015. Although there is the influence of flow from the Mekong mainstream to the lake volume change, there was good agreement between lake volume and tributary flow shown in Figure 7c, suggesting the influence of decreasing trend of streamflow from tributary basins to the declining rate of lake volume–especially during the last decade.

4.3.3.3. Climate variability and sediment load changes

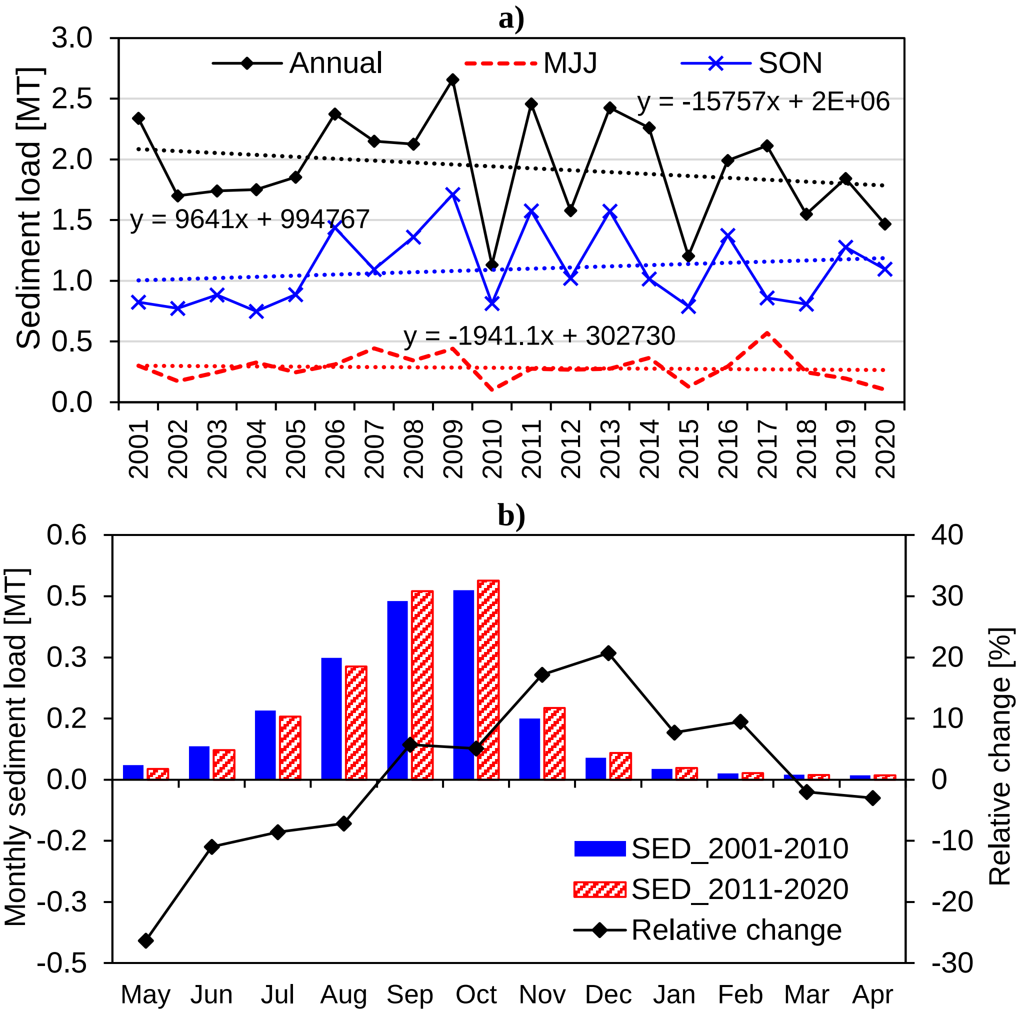


Figure 4.8. a) Temporal trends and variability of annual, MJJ- and SON-total sediment load. b) Comparison of long-term mean monthly total sediment load between the 2001–2010 and 2011–2020 periods. The full name of MJJ and SON are as per the description in Figure 4.6. The total sediment load was simulated from the total area of the TSL basin excluding the flooded area.

The sediment load in the TSL Basin was changing over 2001-2020 due to the streamflow alteration caused by spatiotemporal changes in rainfall and temperature over the basin area during this period (Figure 4.8). There was a decrease of about 15% (~0.3 Mt) and 12% (~0.037 Mt) in annual and MJJ-sediment load, respectively from 2001 to 2020 (Figure 4.8a). Similar to rainfall and streamflow; on the other hand, an increasing trend was found in the SON-sediment load, with a rising of around 0.18 Mt or 18% between 2001 and 2020.

Figure 4.8b illustrates the comparison of long-term mean monthly sediment load between the 2001–2010 and 2011–2020 periods. Due to the decline of streamflow between May and August, the sediment load during these months decreased up to 26% in 2011–2020 compared to that in 2001–2010. Conversely, the sediment load was found to increase (max. 21% increase in December) between September and December, which was triggered by hydrological cycle alternation (Figures 4.6b and 4.8b).

4.3.4. Impact of land-use change on streamflow and sediment load

Land-use maps for the years 2002, 2010, and 2018 were used to estimate the potential impacts of land-use change on the total streamflow and sediment load of the TSL Basin compared to the 1995 land cover map. Each land-use map (i.e., 1995, 2002, 2010 and 2018 land-use maps) was input together with the same climate data (i.e., rainfall and temperature) from 2001 to 2020 into the SWAT model for streamflow and sediment load impact simulations due to land-use change.

The changes in annual streamflow due to land-use changes were found to maximum increase by 3.66, 4.41, and 5.53% under 2002, 2010, and 2018 land cover scenarios, respectively between 2001 and 2020 (Figure 4.9a). On a seasonal scale, there is an increase in streamflow of up to 3.55, 4.48 and 5.83% between May and November, while a decrease of up to 3.45, 3.86 and 4.54% between December and April for 2002, 2010, and 2018 land cover scenarios, respectively compared to the 1995 land cover (Figure 4.9b). The decrease in streamflow during December and April is caused by changes in the hydrological cycle, in which groundwater recharge is increasing in the dry season due to less forest or vegetation cover. Although basin streamflow is more sensitive to climate change than land-use change, climate and land-use change are concurrent phenomena and, therefore, the land-use change impact should not be overlooked.

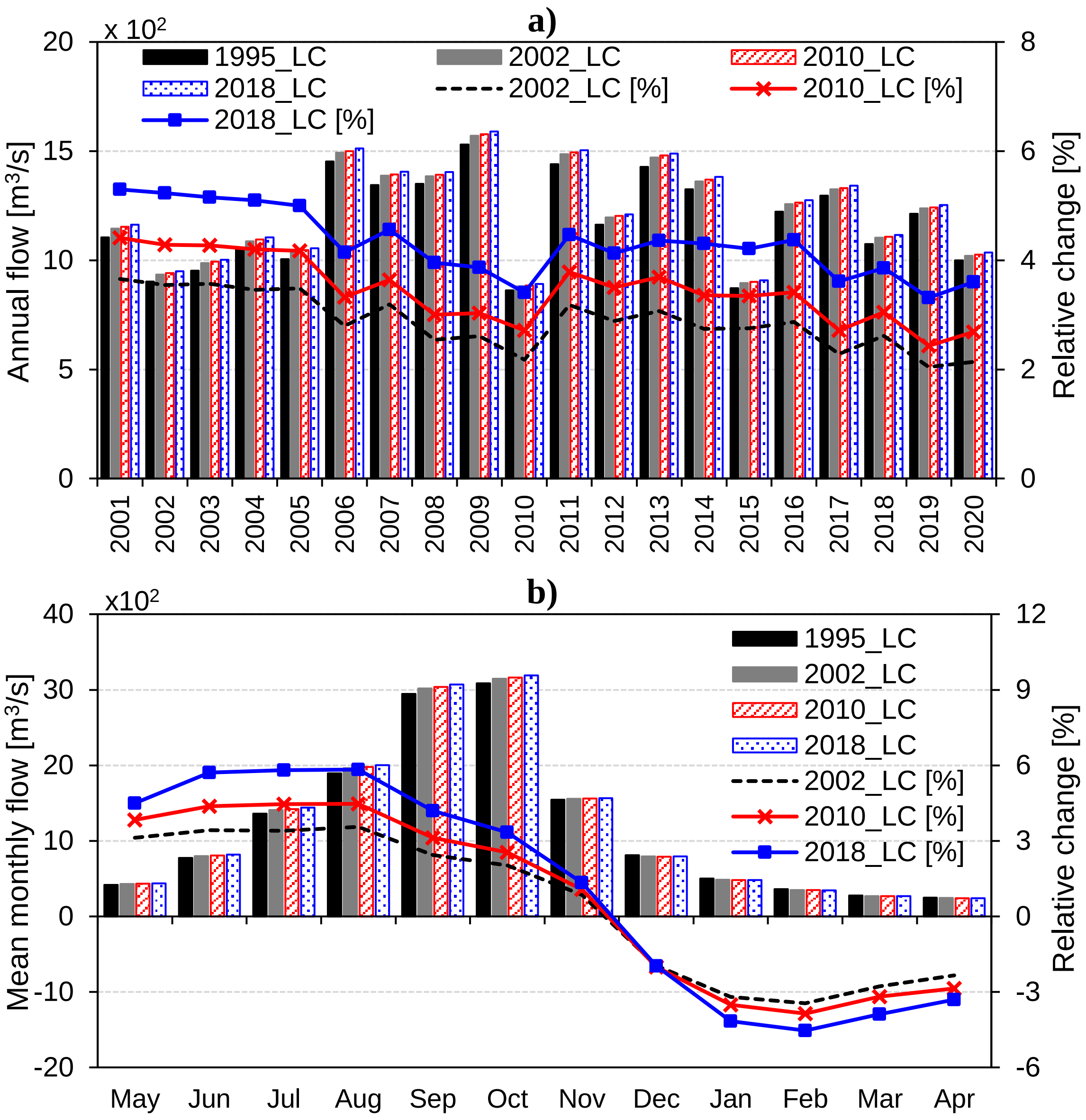


Figure 4.9. Impacts of land-use change on a) annual and b) long-term monthly streamflow under 2002, 2010, and 2018 land cover scenarios.

Sediment load is more sensitive than streamflow to land-use change in the TSL Basin (Figure 4.10a-b). High seasonal variability in loading can be seen, with the rainy season (May to October) accounting for around 74% of the annual sediment load whereas only 24% of the annual load was observed during the dry season (November to April) (Figure 4.10b). Loading is the function of streamflow, rainfall, topography, land-use type, etc. with flow playing a vital role in generating loading; therefore, the same seasonal pattern was found between flow and sediment loading (Figures 4.9b and 4.10b). The TSL annual total sediment loading can be seen to increase by 8.16, 9.97, and 12.91% under 2002, 2010, and 2018 land-use scenarios, respectively compared to the baseline land-use period (i.e., 1995 land cover map). Likewise, the land-use change forced an increase in sediment loading during May and November by a maximum of 10.11, 13.51, and 19.57% of the monthly sediment load, while the maximum relative decrease in loading is 4.56, 5.09, and 5.80% between December and April under 2002, 2010, and 2018 land-use scenarios, respectively.

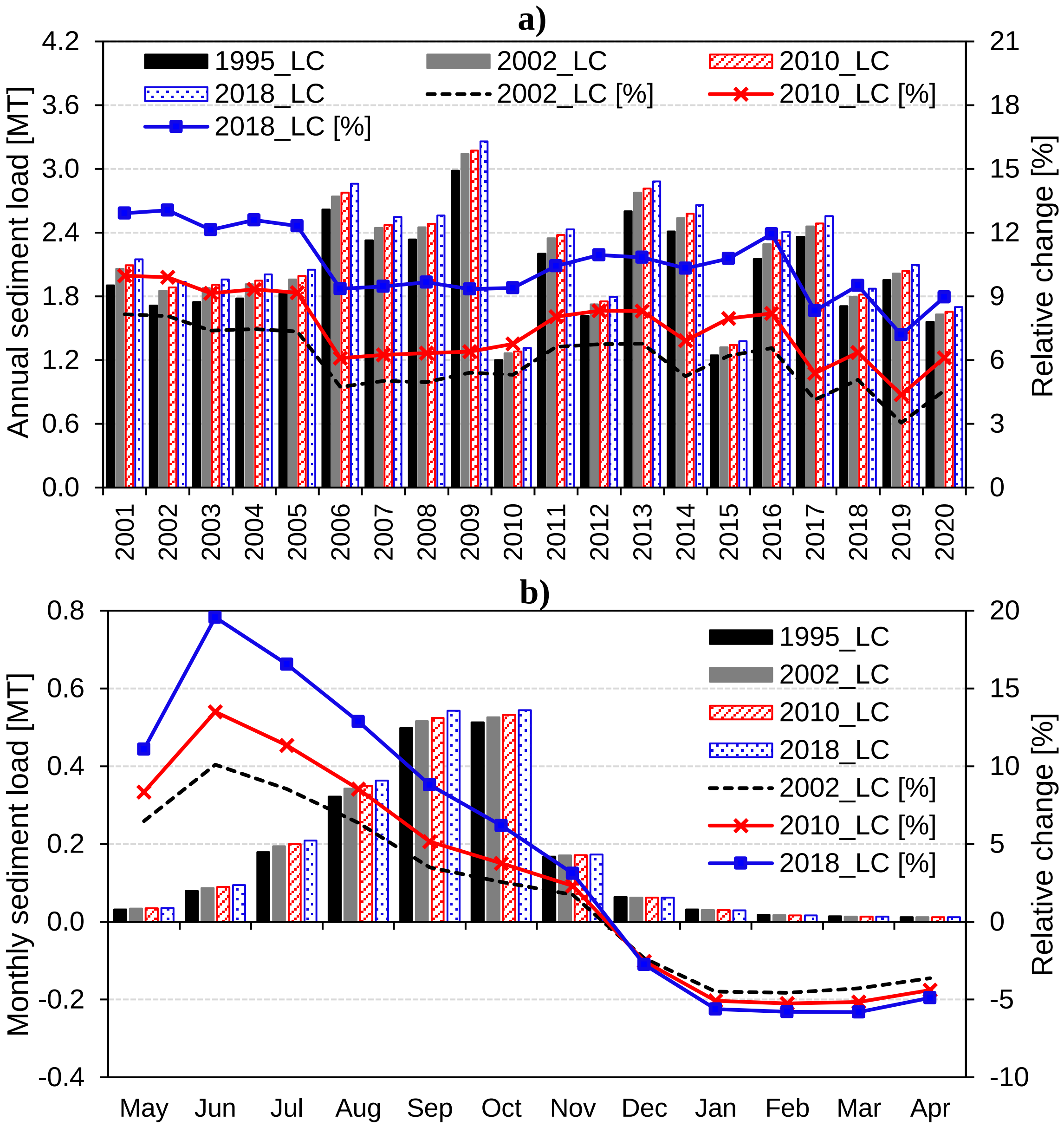


Figure 4.10. Impacts of land-use change on a) annual and b) long-term monthly sediment load under 2002, 2010, and 2018 land cover scenarios.

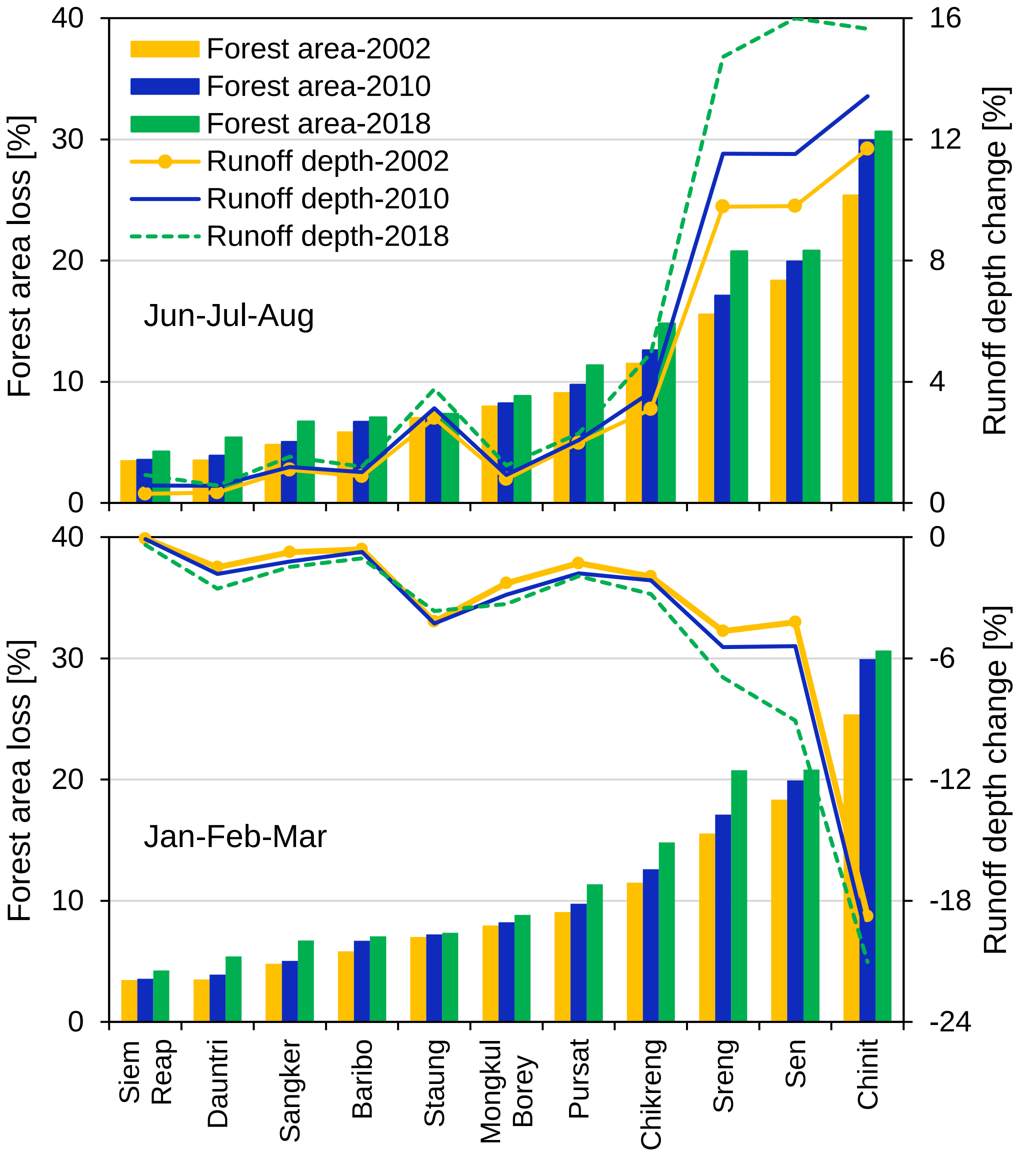


Figure 4.11. Comparison of runoff depth change during June-July-August (top) and January-February-March (bottom) response to forest area loss under 2002, 2010 and 2018 land cover scenarios. Forest area change refers to a relative loss of forest cover to the total area of the corresponding tributary basin.

Figures 4.11 and 4.12 show that runoff depth and sediment yield, respectively in the Chinit tributary basin were found to be the most sensitive to land-use change impact due to a large relative decrease in the forest area to the catchment area (approximate 30%). A 2518 km^2 decrease in forest area was found between 1995 and 2018, which gave runoff depth an increase of about 11.69, 13.43 and 15.85% during June and August, and a decrease of around 18.77, 20.04 and 21.12% during January and March for 2002, 2010, and 2018 land cover scenarios, respectively (Figure 4.11 and Table A16). Likewise, a decrease in forest cover in the Chinit tributary basin enforced sediment yield with an increase of approximately 12.81, 15.60 and 16.75% during June and August and a decline of around 18.76, 22.18 and 23.16%

during January and March for 2002, 2010, and 2018 land cover scenarios, respectively (Figure 4.12). Furthermore, relatively high changes in runoff depth and sediment yield were found in the other two tributary basins (i.e., Sreng and Sen), which are consistent with the high percentage loss of forest area (up to 20%) shared with the catchment area. The runoff depth and sediment yield were seen to increase by up to 11.52 and 17.35%, respectively during June and August, while a decrease of up to 9.09 and 13.27%, respectively was found during January and March (Figures 4.11 and 4.12). The relatively lower forest area loss indicated a minimum land-use impact in the other eight tributary basins. There were increases in runoff depth and sediment yield from 0.30 to 4.95% and 0.54 to 3.70%, respectively during June and August, while the decreases from 0.07 to 2.81 and 0.24 to 3.49 %, respectively were found during January and March (Figures 4.11 and 4.12).

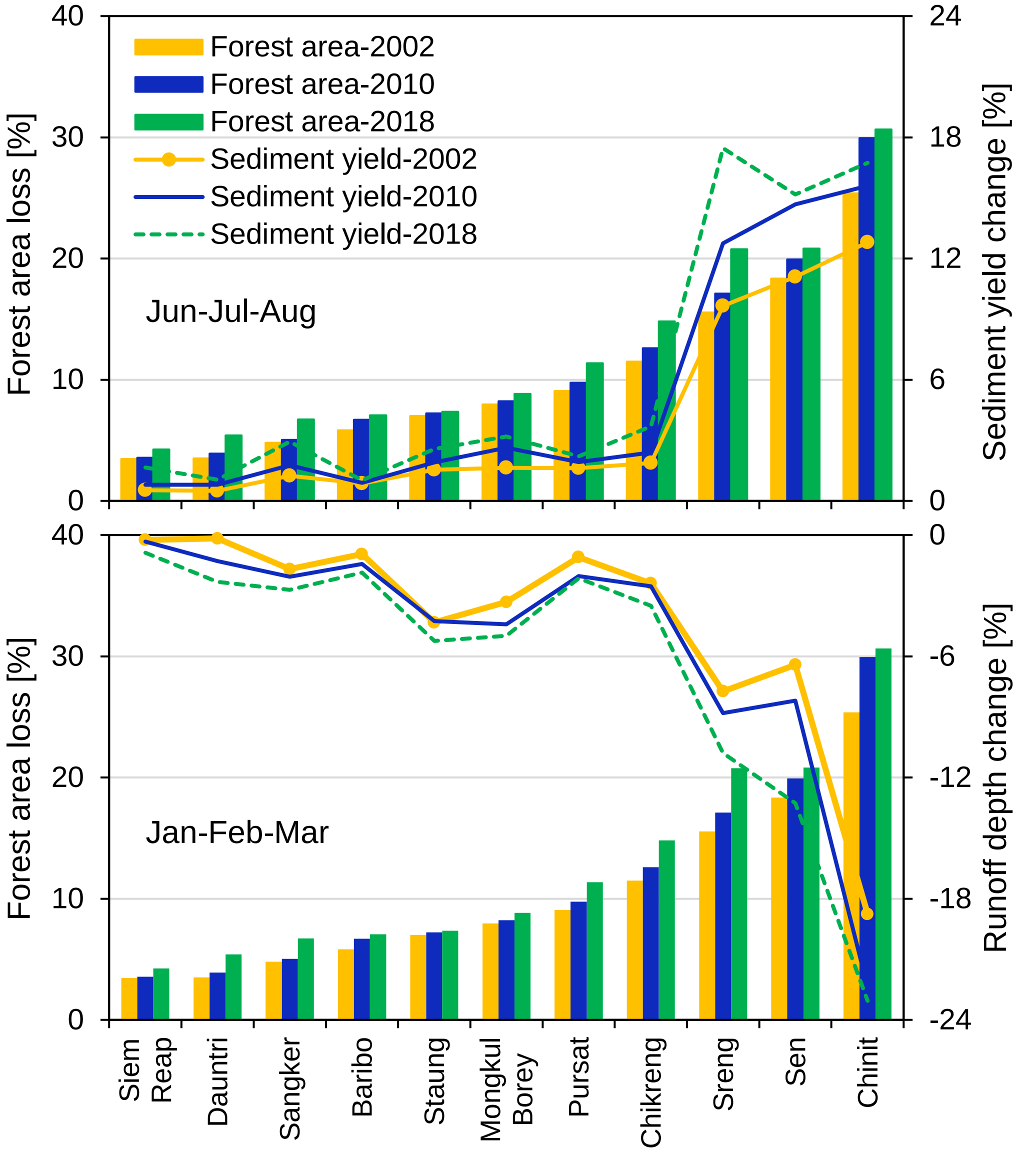


Figure 4.12. Same as Figure 4.11 but for sediment yield change.

4.4. Conclusion

This study investigates the decadal climatic variability and land-use change and quantifies their impacts on streamflow and sediment load in the TSL Basin during the last few decades. Four different land-use periods: 1995, 2002, 2010, and 2018 land cover maps, together with the meteorological data from 2001 to 2020 were used to analyze spatiotemporal changes in climate and land use and input into the SWAT model to simulate their impacts on streamflow and sediment loading in the TSL Basin.

Climate analysis showed an increasing trend in the basin temperature and a downward trend in rainfall between 2001 and 2020. The hydrological analysis depicts that the early rainy season flows of the TSL Basin were lower (max. 26% decrease in May) for 2011-2020 compared to the 2001-2010 time horizons. However, after the wet monsoon season, the streamflow was observed to receive some increases due to the rise in rainfall amount during September and November. Since flow plays a vital role in generating loading, a similar change in the seasonal pattern forced by climate variability was found between flow and sediment load. The basin streamflow and sediment load showed a greater impact from climate variation than land-use change. From the MODIS images analysis and measured water level in the lake, the annual maximum lake surface area and water level showed decreasing trends between 2001 and 2020, indicating a shrinking rate of lake volume over the years. The results of the land-use change analysis revealed that there was a substantial decrease found in forest-related areas while some growth in cropland areas was found. The flooded forest area decreased from approximately 6,992 to 1700 km^2 or 76%, while the cropland area increased by about 23% from around 30,400 to 37,324 km^2 in 1995 and 2018, respectively. The land-use change had a positive and negative impact on streamflow sediment loading during the rainy and dry seasons, respectively. The streamflow and sediment load were seen to increase by up to 5.83 and 19.57%, respectively between May and November, while a decrease of up to 4.54 and 5.80%, respectively was found between December and April. Sediment load is more sensitive than streamflow to land-use change in the TSL Basin. Furthermore, the highest impact of land-use changes on streamflow and sediment load was found in the Chinit tributary basin due to the largest decrease in the forest area shared with the total area of the catchment.

Based on the results found in this study, climate variability and land-use change have noticeable impacts on streamflow and sediment load in the TSL Basin, suggesting that planning and management of their impacts are crucial for the sustainable management of water resources and lake ecosystem conservation in the TSL Basin.

Chapter 5 Summary of the research

This chapter summarizes the research results and contributions, together with some remaining challenges for future research.

5.1. Summary of research results

This research successfully developed a feasible method or framework for estimating streamflow and sediment load in data-sparse or poorly gauged basins of the Tonle Sap Lake, serving for sustainable management and conservation of the lake ecosystems. First, a comprehensive framework was used to determine an ideal alternative meteorological dataset for hydrological modeling, considering the spatiotemporal characteristics of each climate dataset and ensuring a reliable estimate of streamflow and evapotranspiration. Next, a new regionalization method was proposed to estimate sediment load in ungauged catchments, considering spatiotemporal variability and the sediment load relation to rainfall characteristics of individual catchments. Finally, the application of the above two mentioned methods for hydro-environmental impact assessment, owing to decadal climate variation and basin development during the last few decades.

The study begins by introducing a comprehensive assessment of seven gridded precipitation and air temperature products by statistically comparing these datasets with gauge-based datasets and applying the SWAT model for daily streamflow and ET simulations over the TSL Basin. The precipitation data from APHRODITE, ERA5, TRMM and IMERG were found to have high correlations with rain-gauged data and the lowest estimation errors, and the SA-OBS and CPC were found to match the observed air temperature data well. The results of the hydrological simulation showed that the rainfall data from APHRODITE, TRMM and IMERG, combined with SA-OBS-based air temperature data, provided improved estimations of daily streamflow and mean runoff depth. The ET estimated using the TRMM and IMERG datasets showed a better temporal and spatial pattern agreement with GLEAM- and MODIS-ET. This suggests that TRMM and IMERG, in conjunction with SA-OBS air temperature, are reliable for providing streamflow through the SWAT model application and other water balance components. The comprehensive evaluation framework used in this study was found to be effective in selecting reliable gridded meteorological datasets for hydrological simulation in data-sparse river basins, especially when large uncertainty existed in the spatiotemporal distribution of rainfall. These findings also showed that statistical comparisons with gauge data and hydrological evaluation of streamflow are not enough to justify the reliability of each gridded dataset.

Next, a novel SRS regionalization method has been proposed, using the SWAT model and SOM clustering technique to overcome the limitation of the critical attributes of a catchment favoring sediment similarity, which usually exists in the conventional regionalization approaches. It considered the spatiotemporal variations of sediment response and its relationship with rainfall characteristics as a catchment attribute and showed the potential to ideally determine hydrological and sediment similarities between gauged and ungauged catchments. The results indicated the comprehensive performance of the SRS regionalization method for estimating sediment load in the ungauged catchments. The SRS approach obtained an estimation error reduction of up to 7%, compared with the PS regionalization method. The SRS regionalization method proposed in this study is a global alternative method for estimating sediment, as well as other hydrological variables and rainfall-driven phenomena such as streamflow and nutrient transport, in ungauged catchments.

Eventually, the above two frameworks were used to quantify the impacts of climatic variability and land-use change on streamflow and sediment load in the TSL Basin during the last few decades. The decadal climatic variability between 2001 and 2020 and land-use change patterns in different periods were investigated as well. Climate analysis showed an increasing trend in the basin temperature and a downward trend in rainfall between 2001 and 2020. The results of the land-use change analysis revealed that there was a substantial decrease found in natural forest areas, in which cumulative loss of natural forest area was around 45% or the area decreased from 37,052 in 1995 to 20,408 km^2 in 2018. On the other hand, the cropland area increased by about 23% from around 30,400 to 37,324 km^2 in 1995 and 2018, respectively. The hydrological analysis depicts that the early rainy season flows of the TSL Basin were lower (max. 26% decrease) for 2011-2020 compared to the 2001-2010 time horizons. However, after the wet monsoon season, the streamflow was observed to receive some increases due to the rise in rainfall amount during September and November. A similar change in the seasonal pattern forced by climate variability was found between flow and sediment load. The land-use change had a positive and negative impact on streamflow sediment loading during the rainy and dry seasons, respectively. The streamflow and sediment load were seen to increase by up to 5.83 and 19.57%, respectively between May and November, while a decrease of up to 4.54 and 5.80%, respectively was found between December and April. The basin streamflow and sediment load showed a greater impact from climate variation than land-use change. Based on the results found in this study, climate variability and land-use change have noticeable impacts on streamflow and sediment load, suggesting that planning and management of their impacts are crucial for the sustainable management of water resources and lake ecosystem conservation in the TSL Basin.

The feasible framework developed in this research would serve as a central approach in estimating streamflow and sediment load in data-sparse or poorly gauged basins. Besides that, the study also contributes a basic methodology to estimate other hydro-meteorological parameters of interest. For a similar purpose, this methodology can be applied as well in other river basins around the world.

5.2. Contributions of research results

The main contributions of this research are summarized as follows:

- A novel regionalization method for estimating sediment load in ungauged catchments.
- A comprehensive framework for evaluating satellite-based or gridded meteorological datasets as alternative climate data for hydrological modeling in data-sparse river basins.
- Two ideal alternative meteorological data sources (i.e., IMERG and SA-OBS) for basin-wide hydro-environmental assessment in the TSL Basin and beyond.
- A coupled approach (the streamflow and sediment load estimating frameworks) for assessing the potential impact of climatic variability and land-use change on streamflow and sediment loading in the TSL Basin.

5.3. Future challenges

Although the current works could make some contributions, future works need to be paid attention to the following:

- A framework for evaluating alternative climate datasets should pay further attention to sub-daily timescales (i.e., hourly), which would be beneficial for flood assessment and near-real-time forecasting. Additionally, this framework could be extended to evaluate alternative products covering the whole Mekong River basin.
- Using the current framework, more and new climate products could be validated for basin-wide hydrological modeling and hydrological impact assessment.
- The applicability of the SRS regionalization method is tested on a limited number of catchments. Therefore, it is recommended to incorporate more gauged catchments, especially those with a smaller and larger scale, to widely extend the method's applicability.
- In the SRS method, if observed data is available, the method could be applied every 5 years with different land-use conditions (e.g., 2002, 2010, and 2018 land-use maps), then the regionalized parameters would better represent the hydrological and environmental conditions of ungauged catchments for the whole period from 2001 to 2020. Eventually, the temporal changes in sediment load in the ungauged catchments between 2001 and 2020 would be well simulated by the SRS method.

References

Abbas, S.A., Xuan, Y., 2020. Impact of precipitation pre-processing methods on hydrological model performance using high-resolution gridded dataset. Water (Switzerland) 12. https://doi.org/10.3390/w12030840

Abbaspour, K.C., 2015. SWAT-CUP 2012: SWAT Calibration and Uncertainty Programs: A User Manual.

Abiodun, O.O., Guan, H., Post, V.E.A., Batelaan, O., 2018. Comparison of MODIS and SWAT evapotranspiration over a complex terrain at different spatial scales. Hydrol. Earth Syst. Sci. 22, 2775–2794. https://doi.org/10.5194/hess-22-2775-2018

Ahmed, K., Shahid, S., Wang, X., Nawaz, N., Najeebullah, K., 2019. Evaluation of gridded precipitation datasets over arid regions of Pakistan. Water (Switzerland) 11. https://doi.org/10.3390/w11020210

Allen, R.G., Jensen, M.E., Wright, J.L., Burman, R.D., 1989. Operational Estimates of Reference Evapotranspiration. Agron. J. 81, 650–662. https://doi.org/10.2134/agronj1989.00021962008100040019x

Amjad, M., Yilmaz, M.T., Yucel, I., Yilmaz, K.K., 2020. Performance evaluation of satellite- and model-based precipitation products over varying climate and complex topography. J. Hydrol. 584, 124707. https://doi.org/10.1016/j.jhydrol.2020.124707

Ang, R., Kinouchi, T., Zhao, W., 2022. Evaluation of daily gridded meteorological datasets for hydrological modeling in data-sparse basins of the largest lake in Southeast Asia. J. Hydrol. Reg. Stud. 42, 101135. https://doi.org/10.1016/j.ejrh.2022.101135

Ang, R., Oeurng, C., 2018. Simulating streamflow in an ungauged catchment of Tonlesap Lake Basin in Cambodia using Soil and Water Assessment Tool (SWAT) model. Water Sci. 32, 89–101. https://doi.org/10.1016/j.wsj.2017.12.002

Arias, M.E., Cochrane, T.A., Kummu, M., Lauri, H., Holtgrieve, G.W., Koponen, J., Piman, T., 2014. Impacts of hydropower and climate change on drivers of ecological productivity of Southeast Asia’s most important wetland. Ecol. Modell. 272, 252–263. https://doi.org/10.1016/j.ecolmodel.2013.10.015

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large Area Hydrologic Modeling and Assessment Part I: Model Development 34, 73–89.

Aslam, R.A., Shrestha, S., Pal, I., Ninsawat, S., Shanmugam, M.S., Anwar, S., 2020.

Projections of climatic extremes in a data poor transboundary river basin of India and Pakistan. Int. J. Climatol. 40, 4992–5010. https://doi.org/10.1002/joc.6501

Asurza-Véliz, F.A., Lavado-Casimiro, W.S., 2020. Regional parameter estimation of the SWAT model: Methodology and application to river basins in the Peruvian Pacific drainage. Water (Switzerland) 12, 1–25. https://doi.org/10.3390/w12113198

Athira, P., Sudheer, K.P., Cibin, R., Chaubey, I., 2016. Predictions in ungauged basins: an approach for regionalization of hydrological models considering the probability distribution of model parameters. Stoch. Environ. Res. Risk Assess. 30, 1131–1149. https://doi.org/10.1007/s00477-015-1190-6

Ayoub, A.B., Tangang, F., Juneng, L., Tan, M.L., Chung, J.X., 2020. Evaluation of gridded precipitation datasets in Malaysia. Remote Sens. 12. https://doi.org/10.3390/rs12040613

Bai, P., Liu, X., 2018. Evaluation of five satellite-based precipitation products in two gauge-scarce basins on the Tibetan Plateau. Remote Sens. 10. https://doi.org/10.3390/RS10081316

Bank, W., 2011. Climate Rick and Adapation Country Profiles: Vulnerability, Risk Reduction, and Adoptation to Climate Change. World Dev. 1–12.

Bao, X., Zhang, F., 2013. Evaluation of NCEP-CFSR, NCEP-NCAR, ERA-Interim, and ERA-40 reanalysis datasets against independent sounding observations over the Tibetan Plateau. J. Clim. 26, 206–214. https://doi.org/10.1175/JCLI-D-12-00056.1

Bärlund, I., Kirkkala, T., Malve, O., Kämäri, J., 2007. Assessing SWAT model performance in the evaluation of management actions for the implementation of the Water Framework Directive in a Finnish catchment. Environ. Model. Softw. 22, 719–724. https://doi.org/10.1016/j.envsoft.2005.12.030

Bastia, F., Equeenuddin, S.M., 2016. Spatio-temporal variation of water flow and sediment discharge in the Mahanadi River, India. Glob. Planet. Change 144, 51–66. https://doi.org/10.1016/j.gloplacha.2016.07.004

Bromwich, D.H., Fogt, R.L., 2004. Strong trends in the skill of the ERA-40 and NCEP-NCAR reanalyses in the high and midlatitudes of the southern hemisphere, 1958-2001. J. Clim. 17, 4603–4620. https://doi.org/10.1175/3241.1

Brown, C.E., 1998. Coefficient of Variation. In: Applied Multivariate Statistics in Geohydrology and Related Sciences. Springer, Berlin, Heidelberg, pp. 155–157. https://doi.org/10.1007/978-3-642-41714-6_33027

Bunthan, S., Takahashi, Y., Izumida, Y., 2018. A Study on Cambodian Rice Farming: Comparative Analysis on Aromatic and Non-Aromatic Rice Farming in Voatkor Commune, Battambang Province. Int. J. Environ. Rural Dev. 9, 71–76.

Cakir, R., Raimonet, M., Sauvage, S., Paredes-Arquiola, J., Grusson, Y., Roset, L., Meaurio, M., Navarro, E., Sevilla-Callejo, M., Lechuga-Crespo, J.L., Pascual, J.J.G., Bodoque, J.M., Sánchez-Pérez, J.M., 2020. Hydrological alteration index as an indicator of the calibration complexity of water quantity and quality modeling in the context of global change. Water (Switzerland) 12. https://doi.org/10.3390/w12010115

Caroletti, G.N., Coscarelli, R., Caloiero, T., 2019. Validation of satellite, reanalysis and RCM data of monthly rainfall in Calabria (Southern Italy). Remote Sens. 11. https://doi.org/10.3390/rs11131625

Chea, R., Grenouillet, G., Lek, S., 2016. Evidence of water quality degradation in lower mekong basin revealed by self-organizing map. PLoS One 11, 1–19. https://doi.org/10.1371/journal.pone.0145527

Chen, Aifang, Chen, Anping, Varis, O., Chen, D., 2022. Large net forest loss in Cambodia’s Tonle Sap Lake protected areas during 1992–2019. Ambio 51, 1889–1903. https://doi.org/10.1007/s13280-022-01704-4

Cleugh, H.A., Leuning, R., Mu, Q., Running, S.W., 2007. Regional evaporation estimates from flux tower and MODIS satellite data. Remote Sens. Environ. 106, 285–304. https://doi.org/10.1016/j.rse.2006.07.007

Constable, D., 2015. Atlas of the 3S Basins. Thailand: IUCN, Bangkok.

Dash, S.S., Sahoo, B., Raghuwanshi, N.S., 2021. How reliable are the evapotranspiration estimates by Soil and Water Assessment Tool (SWAT) and Variable Infiltration Capacity (VIC) models for catchment-scale drought assessment and irrigation planning? J. Hydrol. 592, 125838. https://doi.org/10.1016/j.jhydrol.2020.125838

Dinh, K.D., Anh, T.N., Nguyen, N.Y., Bui, D.D., Srinivasan, R., 2020. Evaluation of grid-based rainfall products and water balances over the Mekong river Basin. Remote Sens. 12. https://doi.org/10.3390/rs12111858

Duan, W., Takara, K., He, B., Luo, P., Nover, D., Yamashiki, Y., 2013. Spatial and temporal trends in estimates of nutrient and suspended sediment loads in the Ishikari River, Japan, 1985 to 2010. Sci. Total Environ. 461–462, 499–508. https://doi.org/10.1016/j.scitotenv.2013.05.022

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. Sci. Total Environ. 573, 1536–1553. https://doi.org/10.1016/j.scitotenv.2016.08.213

Ebrahimi, S., Chen, C., Chen, Q., Zhang, Y., Ma, N., Zaman, Q., 2017. Effects of temporal scales and space mismatches on the TRMM 3B42 v7 precipitation product in a remote mountainous area. Hydrol. Process. 31, 4315–4327. https://doi.org/10.1002/hyp.11357

Epelde, A.M., Cerro, I., Sánchez-Pérez, J.M., Sauvage, S., Srinivasan, R., Antigüedad, I., 2015. Application du modèle SWAT à l'évaluation de l'impact des modificationss des pratiques agricoles sur la qualité de l'eau. Hydrol. Sci. J. 60, 825–843. https://doi.org/10.1080/02626667.2014.967692

Fan, Y., van den Dool, H., 2008. A global monthly land surface air temperature analysis for 1948-present. J. Geophys. Res. Atmos. 113, 1–18. https://doi.org/10.1029/2007JD008470

Farsadnia, F., Rostami Kamrood, M., Moghaddam Nia, A., Modarres, R., Bray, M.T., Han, D., Sadatinejad, J., 2014. Identification of homogeneous regions for regionalization of watersheds by two-level self-organizing feature maps. J. Hydrol. 509, 387–397. https://doi.org/10.1016/j.jhydrol.2013.11.050

Fenta Mekonnen, D., Duan, Z., Rientjes, T., Disse, M., 2018. Analysis of combined and isolated effects of land-use and land-cover changes and climate change on the upper Blue Nile River basin's streamflow. Hydrol. Earth Syst. Sci. 22, 6187–6207. https://doi.org/10.5194/hess-22-6187-2018

Frappart, F., Biancamaria, S., Normandin, C., Blarel, F., Bourrel, L., Aumont, M., Azemar, P., Vu, P.L., Le Toan, T., Lubac, B., Darrozes, J., 2018. Influence of recent climatic events on the surface water storage of the Tonle Sap Lake. Sci. Total Environ. 636, 1520–1533. https://doi.org/10.1016/j.scitotenv.2018.04.326

Fu, Q., Shi, R., Li, T., Sun, Y., Liu, D., Cui, S., Hou, R., 2019. Effects of land-use change and climate variability on streamflow in the Woken River basin in Northeast China. River Res. Appl. 35, 121–132. https://doi.org/10.1002/rra.3397

Fu, Y., Xia, J., Yuan, W., Xu, B., Wu, X., Chen, Y., Zhang, H., 2016. Assessment of multiple precipitation products over major river basins of China. Theor. Appl. Climatol. 123, 11–22. https://doi.org/10.1007/s00704-014-1339-0

Fuka, D.R., Walter, M.T., Macalister, C., Degaetano, A.T., Steenhuis, T.S., Easton, Z.M.,

2014. Using the Climate Forecast System Reanalysis as weather input data for watershed models. Hydrol. Process. 28, 5613–5623. https://doi.org/10.1002/hyp.10073

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. Trans.

Ge, F., Peng, T., Fraedrich, K., Sielmann, F., Zhu, X., Zhi, X., Liu, X., Tang, W., Zhao, P., 2019. Assessment of trends and variability in surface air temperature on multiple high-resolution datasets over the Indochina Peninsula. Theor. Appl. Climatol. 135, 1609–1627. https://doi.org/10.1007/s00704-018-2457-x

Gitau, M.W., Chaubey, I., 2010. Regionalization of SWAT model parameters for use in ungauged watersheds. Water (Switzerland) 2, 849–871. https://doi.org/10.3390/w2040849

Guan, X., Zhang, J., Yang, Q., Tang, X., Liu, C., Jin, J., Liu, Y., Bao, Z., Wang, G., 2020. Evaluation of precipitation products by using multiple hydrological models over the upper yellow river basin, china. Remote Sens. 12, 1–27. https://doi.org/10.3390/rs12244023

Gunathilake, M.B., Amaratunga, Y. V., Perera, A., Karunanayake, C., Gunathilake, A.S., Rathnayake, U., 2020. Statistical evaluation and hydrologic simulation capacity of different satellite-based precipitation products (SbPPs) in the Upper Nan River Basin, Northern Thailand. J. Hydrol. Reg. Stud. 32, 100743. https://doi.org/10.1016/j.ejrh.2020.100743

Guo, Y., Zhang, Y., Zhang, L., Wang, Z., 2020. Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A comprehensive review. Wiley Interdiscip. Rev. Water 8, 1–32. https://doi.org/10.1002/wat2.1487

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1999. Status of Automatic Calibration for Hydrologic Models: Comparison With Multilevel Expert Calibration 135–143.

Hamel, L., Brown, C., 2011. Improved Interpretability of the Unified Distance Matrix with Connected Components. Proc. Adv. Meas. Sci. 338–343.

Hao, F., Zhang, X., Yang, Z., 2004. A distributed non-point source pollution model: calibration and validation in the Yellow River Basin. J. Environ. Sci. (China) 16, 646–50.

Hargreaves, G.H., Samani, Z.A., 1985. Reference Crop Evapotranspiration From Ambient

Air Temperature. Pap. - Am. Soc. Agric. Eng. 96–99.

He, J., Wan, Y.R., Chen, H.T., Wang, W.C., 2021. Study on the impact of land-use change on runoff variation trend in Luojiang river basin, China. Water (Switzerland) 13. https://doi.org/10.3390/w13223282

Heng, B., Oeurng, C., Try, S., Yuzir, A., 2020. Flow regime alteration analysis under climate change in Tonle Sap Subbasin. IOP Conf. Ser. Earth Environ. Sci. 479. https://doi.org/10.1088/1755-1315/479/1/012007

Heng, S., Suetsugi, T., 2013. Using Artificial Neural Network to Estimate Sediment Load in Ungauged Catchments of the Tonle Sap River Basin, Cambodia. J. Water Resour. Prot. 05, 111–123. https://doi.org/10.4236/jwarp.2013.52013

Heng, S., Suetsugi, T., 2014. Comparison of regionalization approaches in parameterizing sediment rating curve in ungauged catchments for subsequent instantaneous sediment yield prediction. J. Hydrol. 512, 240–253. https://doi.org/10.1016/j.jhydrol.2014.03.003

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

Hu, J., Wu, Y., Wang, L., Sun, P., Zhao, F., Jin, Z., Wang, Y., Qiu, L., Lian, Y., 2021. Impacts of land-use conversions on the water cycle in a typical watershed in the southern Chinese Loess Plateau. J. Hydrol. 593, 125741. https://doi.org/10.1016/j.jhydrol.2020.125741

Huang, C., Hu, J., Chen, S., Zhang, A., Liang, Z., Tong, X., Xiao, L., Min, C., Zhang, Z., 2019. How well can IMERG products capture typhoon extreme precipitation events over southern China? Remote Sens. 11. https://doi.org/10.3390/rs11010070

Huete, A.R., Liu, H.Q., Batchily, K., Leeuwen, W. Van, 1997. A comparison of vegetation indices over a global set of TM images for EOS-MODIS 59, 440–451. https://doi.org/https://doi.org/10.1016/S0034-4257(96)00112-5

Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E., Sorooshian, S., Tan, J., Xie, P., 2020. NASA GPM Integrated Multi-satellitE Retrievals for GPM (IMERG) Algorithm Theoretical Basis Document (ATBD) Version 06. Nasa/Gsfc 29.

Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F., Wolff, D.B., 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeorol. 8, 38–55. https://doi.org/10.1175/JHM560.1

Islam, M.A., Yu, B., Cartwright, N., 2020. Assessment and comparison of five satellite precipitation products in Australia. J. Hydrol. 590, 125474. https://doi.org/10.1016/j.jhydrol.2020.125474

Jha, B., Jha, M.K., 2013. Rating Curve Estimation of Surface Water Quality Data Using LOADEST. J. Environ. Prot. (Irvine. Calif). 04, 849–856. https://doi.org/10.4236/jep.2013.48099

Jha, M.K., Gassman, P.W., Arnold, J.G., 2007. Water quality modeling for the Raccoon River watershed using SWAT 50, 479–494. https://doi.org/10.13031/2013.22660

Jiang, Q., Li, W., Fan, Z., He, X., Sun, W., Chen, S., Wen, J., Gao, J., Wang, J., 2021. Evaluation of the ERA5 reanalysis precipitation dataset over Chinese Mainland. J. Hydrol. 595, 125660. https://doi.org/10.1016/j.jhydrol.2020.125660

Jiao, D., Xu, N., Yang, F., Xu, K., 2021. Evaluation of spatial-temporal variation performance of ERA5 precipitation data in China. Sci. Rep. 11, 1–13. https://doi.org/10.1038/s41598-021-97432-y

Jin, X., Jin, Y., 2020. Calibration of a distributed hydrological model in a data-scarce basin based on GLEAM datasets. Water (Switzerland) 12. https://doi.org/10.3390/w12030897

Johnstone, G., Puskur, R., Fabrice, D., Kosal, M., Oeur, I.I., Sithirith, M., Bunna, P., Sophat, S., Chan, S., Sochanny, H., Samnang, L., Suon, S., Kimhor, P., Sameth, R., 2013. Tonle Sap scoping report. CGIAR Research Program on Aquatic Agricultural Systems. Penang, Malaysia. Project Report: AAS-2013-28.

Kabeya, N., Shimizu, A., Shimizu, T., Iida, S., Tamai, K., Miyamoto, A., Chann, S., ARaki, M., Ohnuki, Y., 2021. Long-term Hydrological Observations in a Lowland Dry Evergreen Forest Catchment Area of the Lower Mekong River, Cambodia. Japan Agric. Res. Q. 55, 177–190. https://doi.org/10.6090/jarq.55.177

Kallio, M., Kummu, M., 2021. Comment on ‘Changes of inundation area and water turbidity of Tonle Sap Lake: Responses to climate changes or upstream dam construction?' Environ. Res. Lett. 16. https://doi.org/10.1088/1748-9326/abf3da

Kaur, R., Srivastava, R., Betne, R., Mishra, K., Dutta, D., 2004. Integration of linear programming and a watershed-scale hydrologic model for proposing an optimized land-use plan and assessing its impact on soil conservation - A case study of the Nagwan watershed in the Hazaribagh district of Jharkhand, India. Int. J. Geogr. Inf. Sci. 18, 73–98. https://doi.org/10.1080/13658810310001620915

Kaur, R., Srivastava, R., Betne, R., Mishra, K., Dutta, D., 2004. Integration of linear programming and a watershed-scale hydrologic model for proposing an optimized land-use plan and assessing its impact on soil conservation - A case study of the Nagwan watershed in the Hazaribagh district of Jharkhand, India. Int. J. Geogr. Inf. Sci. 18, 73–98. https://doi.org/10.1080/13658810310001620915

Kawai, H., Zaki, M.K., Noda, K., 2021. Geographical and seasonal characteristics of APHRODITE and GSMaP in Lao P.D.R. IOP Conf. Ser. Earth Environ. Sci. 724. https://doi.org/10.1088/1755-1315/724/1/012045

Khadka, D., Babel, M.S., Shrestha, S., Virdis, S.G.P., Collins, M., 2021. Multivariate and multi-temporal analysis of meteorological drought in the northeast of Thailand. Weather Clim. Extrem. 34, 100399. https://doi.org/10.1016/j.wace.2021.100399

Khoi, D.N., Loi, P.T., Trang, N.T.T., Vuong, N.D., Fang, S., Nhi, P.T.T., 2022. The effects of climate variability and land-use change on streamflow and nutrient loadings in the Sesan, Sekong, and Srepok (3S) River Basin of the Lower Mekong Basin. Environ. Sci. Pollut. Res. 29, 7117–7126. https://doi.org/10.1007/s11356-021-16235-w

Kim, J., Hogue, T.S., 2008. Evaluation of a MODIS-based potential evapotranspiration product at the point scale. J. Hydrometeorol. 9, 444–460. https://doi.org/10.1175/2007JHM902.1

Kohonen, T., 1982. Analysis of a simple self-organizing process. Biol. Cybern. 44, 135–140. https://doi.org/10.1007/BF00317973

Kohonen, T., 1990. The Self-organizing Map. Proceeding IEEE 78, 1464–1480. https://doi.org/10.1109/5.58325

Kokkonen, T.S., Jakeman, A.J., Young, P.C., Koivusalo, H.J., 2003. Predicting daily flows in ungauged catchments: Model regionalization from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina. Hydrol. Process. 17, 2219–2238.

https://doi.org/10.1002/hyp.1329

Kondolf, M.G., Rubin, Z.K., Minear, J.T., 2014. Dams on the Mekong: Cumulative sediment starvation. Water Resour. Res. Res. 50, 5158–5169. https://doi.org/10.1002/2013WR014651

Kummu, M., Penny, D., Sarkkula, J., Koponen, J., 2008. Sediment: Curse or blessing for Tonle Sap Lake? Ambio 37, 158–163. https://doi.org/10.1579/0044-7447(2008)37[158:SCOBFT]2.0.CO;2

Kummu, M., Tes, S., Yin, S., Adamson, P., Józsa, J., Koponen, J., Richey, J., Sarkkula, J., 2014. Water balance analysis for the Tonle Sap Lake-floodplain system. Hydrol. Process. 28, 1722–1733. https://doi.org/10.1002/hyp.9718

Kummu, M., Varis, O., 2007. Sediment-related impacts due to upstream reservoir trapping, the Lower Mekong River. Geomorphology 85, 275–293. https://doi.org/10.1016/j.geomorph.2006.03.024

Larbi, I., Obuobie, E., Verhoef, A., Julich, S., Feger, K.H., Bossa, A.Y., Macdonald, D., 2020. Water balance components estimation under scenarios of land cover change in the Vea catchment, West Africa. Hydrol. Sci. J. 65, 2196–2209. https://doi.org/10.1080/02626667.2020.1802467

Lauri, H., Räsänen, T.A., Kummu, M., 2014. Using Reanalysis and Remotely Sensed Temperature and Precipitation Data for Hydrological Modeling in Monsoon Climate: Mekong River Case Study. J. Hydrometeorol. 15, 1532–1545. https://doi.org/10.1175/jhm-d-13-084.1

Ley, R., Casper, M.C., Hellebrand, H., Merz, R., 2011. Catchment classification by runoff behaviour with self-organizing maps (SOM). Hydrol. Earth Syst. Sci. 15, 2947–2962. https://doi.org/10.5194/hess-15-2947-2011

Li, Y., Wang, W., Lu, H., 2018. Evaluation of the latest satellite-based precipitation products through pixel-point comparison and hydrological application over the Mekong River basin. Int. Geosci. Remote Sens. Symp. 2018-July, 3043–3046. https://doi.org/10.1109/IGARSS.2018.8519448

Li, Z., Liu, W. zhao, Zhang, X. chang, Zheng, F. li, 2009. Impacts of land use change and climate variability on hydrology in an agricultural catchment on the Loess Plateau of China. J. Hydrol. 377, 35–42. https://doi.org/10.1016/j.jhydrol.2009.08.007

Liu, Z., 2016. Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) monthly precipitation products:

Initial results. J. Hydrometeorol. 17, 777–790. https://doi.org/10.1175/JHM-D-15-0068.1

Lohani, S., Dilts, T.E., Weisberg, P.J., Null, S.E., Hogan, Z.S., 2020. Rapidly accelerating deforestation in Cambodia’s Mekong River basin: A comparative analysis of spatial patterns and drivers. Water (Switzerland) 12. https://doi.org/10.3390/W12082191

López, P.L., Sutanudjaja, E.H., Schellekens, J., Sterk, G., Bierkens, M.F.P., 2017. Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products. Hydrol. Earth Syst. Sci. 21, 3125–3144. https://doi.org/10.5194/hess-21-3125-2017

Lovgren, S., 2020. Cambodia’s biggest lake is running dry, taking forests and fish with it, National Geographic.

Lu, C., Ye, J., Fang, G., Huang, X., Yan, M., 2021. Assessment of GPM IMERG satellite precipitation estimation under complex climatic and topographic conditions. Atmosphere (Basel). 12. https://doi.org/10.3390/atmos12060780

Lu, J.Z., Zhang, L., Cui, X.L., Zhang, P., Chen, X.L., Sauvage, S., Sanchez-Perez, J.M., 2019. Assessing the climate forecast system reanalysis weather data driven hydrological model for the Yangtze river basin in China. Appl. Ecol. Environ. Res. 17, 3615–3632. https://doi.org/10.15666/aeer/1702_36153632

Martens, B., Miralles, D.G., Lievens, H., Schalie, R. Van Der, Jeu, R.A.M. De, 2017. GLEAM v3 : satellite-based land evaporation and root-zone soil moisture 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

Masih, I., Maskey, S., Uhlenbrook, S., Smakhtin, V., 2011. Assessing the Impact of Areal Precipitation Input on Streamflow Simulations Using the SWAT Model. J. Am. Water Resour. Assoc. 47, 179–195. https://doi.org/10.1111/j.1752-1688.2010.00502.x

Masumoto, T., Hai, P.T., Shimizu, K., 2008. Impact of paddy irrigation levels on floods and water use in the Mekong River basin. Hydrol. Process 22, 1321–1328. https://doi.org/10.1002/hyp.6941

Mayor, Y.G., Tereshchenko, I., Fonseca-Hernández, M., Pantoja, D.A., Montes, J.M., 2017. Evaluation of error in IMERG precipitation estimates under different topographic conditions and temporal scales over Mexico. Remote Sens. 9, 1–18. https://doi.org/10.3390/rs9050503

Meher, J.K., Das, L., 2019. Gridded data as a source of missing data replacement in station records. J. Earth Syst. Sci. 128, 1–14. https://doi.org/10.1007/s12040-019-1079-8

Merz, R., Blöschl, G., 2004. Regionalization of catchment model parameters. J. Hydrol. 287, 95–123. https://doi.org/10.1016/j.jhydrol.2003.09.028

Miralles, D.G., Holmes, T.R.H., De Jeu, R.A.M., Gash, J.H., Meesters, A.G.C.A., Dolman, A.J., 2011. Global land-surface evaporation estimated from satellite-based observations. Hydrol. Earth Syst. Sci. 15, 453–469. https://doi.org/10.5194/hess-15-453-2011

Mohammed, I.N., Bolten, J.D., Srinivasan, R., Lakshmi, V., 2018. Improved hydrological decision support system for the Lower Mekong River Basin using satellite-based earth observations. Remote Sens. 10. https://doi.org/10.3390/rs10060885

Monteith, J., 1965. Evaporation and environment. Symp. Soc. Exp. Biol. 205–234.

Moriasi, D.N., Arnold, J.G., Liew, M.W. Van, Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Trans. ASABE 50, 885–900. https://doi.org/10.13031/2013.23153

Morovati, K., Tian, F., Kummu, M., Shi, L., Tudaji, M., Nakhaei, P., Alberto, M., 2023. Contributions from climate variation and human activities to flow regime change of Tonle Sap Lake from 2001 to 2020. J. Hydrol. 616, 128800. https://doi.org/10.1016/j.jhydrol.2022.128800

MRC, 2019. 2017 Lower Mekong Regional Water Quality Monitoring Report 38 pp.

Mu, Q., Zhao, M., Running, S.W., 2011. Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote Sens. Environ. 115, 1781–1800. https://doi.org/10.1016/j.rse.2011.02.019

Mu, Q., Zhao, M., Running, S.W., 2013. MODIS Global Terrestrial Evapotranspiration (ET) Product (NASA MOD16A2/A3) Collection 5. Numer. Terradynamic Simul. Group, Univ. Mont. Missoula MT, USA.

Nash, j. R., Sutcliffe, J. V, 1970. River flow forecasting through conceptual models part i-a discussion of principles. J. Hydrol. 10, 282–290. https://doi.org/https://www.sciencedirect.com/science/article/abs/pii/0022169470902556

Neitsch, S.., Arnold, J.., Kiniry, J.., Williams, J.., 2011. Soil & Water Assessment Tool Theoretical Documentation Version 2009. Texas Water Resour. Inst. 1–647. https://doi.org/10.1016/j.scitotenv.2015.11.063

Nepal, D., Parajuli, P.B., 2022. Assessment of Best Management Practices on Hydrology and Sediment Yield at Watershed Scale in Mississippi Using SWAT. Agric. 12. https://doi.org/10.3390/agriculture12040518

Ngor, P.B., McCann, K.S., Grenouillet, G., So, N., McMeans, B.C., Fraser, E., Lek, S., 2018. Evidence of indiscriminate fishing effects in one of the world’s largest inland fisheries. Sci. Rep. 8, 1–12. https://doi.org/10.1038/s41598-018-27340-1

Nhi, P.T.T., Khoi, D.N., Hoan, N.X., 2019. Evaluation of five gridded rainfall datasets in simulating streamflow in the upper Dong Nai river basin, Vietnam. Int. J. Digit. Earth 12, 311–327. https://doi.org/10.1080/17538947.2018.1426647

Niu, X., Hu, Y., Lei, Z., Wang, H., Zhang, Y., Yan, H., 2022. Spatial and Temporal Evolution Characteristics of Land Use/Cover and Its Driving Factor in Cambodia during 2000–2020. Land 11. https://doi.org/10.3390/land11091556

Nkonge, L.K., Sang, J.K., Gathenya, J.M., Home, P.G., 2014. Comparison of two Calibration-uncertainty Methods for Soil and Water Assessment Tool in Stream Flow Modeling. J. Sustain. Res. Eng. 1, 40–44.

Normandin, C., Frappart, F., Lubac, B., Bélanger, S., Marieu, V., Blarel, F., Robinet, A., Guiastrennec-Faugas, L., 2018. Quantification of surface water volume changes in the Mackenzie Delta using satellite multi-mission data. Hydrol. Earth Syst. Sci. 22, 1543–1561. https://doi.org/10.5194/hess-22-1543-2018

Nourani, V., Baghanam, A.H., Adamowski, J., Gebremichael, M., 2013. Using self-organizing maps and wavelet transforms for space-time pre-processing of satellite precipitation and runoff data in neural network based rainfall-runoff modeling. J. Hydrol. 476, 228–243. https://doi.org/10.1016/j.jhydrol.2012.10.054

Odusanya, A., Mehdi, B., Schürz, C., Oke, A., Awokola, O., Awomeso, J., Adejuwon, J., Schulz, K., 2018. Multi-site calibration and validation of SWAT with satellite-based evapotranspiration in a data sparse catchment in southwestern Nigeria. Hydrol. Earth Syst. Sci. Discuss. 1–37. https://doi.org/10.5194/hess-2018-170

Oeurng, C., Cochrane, T.A., Arias, M.E., Shrestha, B., Piman, T., 2016. Assessment of changes in riverine nitrate in the Sesan, Srepok and Sekong tributaries of the Lower Mekong River Basin. J. Hydrol. Reg. Stud. 8, 95–111. https://doi.org/10.1016/j.ejrh.2016.07.004

Oeurng, C., Cochrane, T.A., Chung, S., Kondolf, M.G., Piman, T., Arias, M.E., 2019. Assessing climate change impacts on river flows in the Tonle Sap Lake Basin, Cambodia. Water (Switzerland) 11. https://doi.org/10.3390/w11030618

Oudin, L., Andréassian, V., Perrin, C., Michel, C., Le Moine, N., 2008. Spatial proximity, physical similarity, regression and ungauged catchments: A comparison of regionalization approaches based on 913 French catchments. Water Resour. Res. 44, 1–15. https://doi.org/10.1029/2007WR006240

Parajka, J., Merz, R., Blöschl, G., 2005. A comparison of regionalisation methods for catchment model parameters. Hydrol. Earth Syst. Sci. 9, 157–171. https://doi.org/10.5194/hess-9-157-2005

Parajka, J., Viglione, A., Rogger, M., Salinas, J.L., Sivapalan, M., Blöschl, G., 2013. Comparative assessment of predictions in ungauged basins-Part 1: Runoff-hydrograph studies. Hydrol. Earth Syst. Sci. 17, 1783–1795. https://doi.org/10.5194/hess-17-1783-2013

Park, Y.S., Engel, B.A., 2015. Analysis for Regression Model Behavior by Sampling Strategy for Annual Pollutant Load Estimation. J. Environ. Qual. 44, 1843–1851. https://doi.org/10.2134/jeq2015.03.0137

Park, Y.S., Engel, B.A., 2016. Identifying the correlation between water quality data and LOADEST model behavior in annual sediment load Estimations. Water (Switzerland) 8. https://doi.org/10.3390/w8090368

Peralta, J.C.A.C., Narisma, G.T.T., Cruz, F.A.T., 2020. Validation of high-resolution gridded rainfall datasets for climate applications in the Philippines. J. Hydrometeorol. 21, 1571–1587. https://doi.org/10.1175/JHM-D-19-0276.1

Pereira, H.R., Meschiatti, M.C., Pires, R.C. de M., Blain, G.C., 2018. On the performance of three indices of agreement: An easy-to-use r-code for calculating the willmott indices. Bragantia 77, 394–403. https://doi.org/10.1590/1678-4499.2017054

Phi Hoang, L., Lauri, H., Kummu, M., Koponen, J., Vliet, M.T.H.V., Supit, I., Leemans, R., Kabat, P., Ludwig, F., 2016. Mekong River flow and hydrological extremes under climate change. Hydrol. Earth Syst. Sci. 20, 3027–3041. https://doi.org/10.5194/hess-20-3027-2016

Piman, T., Cochrance, T.A., Arias, M.E., Green, A., Dat, N.D., 2013. Assessment of Flow Changes from Hydropower Development and Operations in Sekong, Sesan, and Srepok Rivers of the Mekong Basin. J. Water Resour. Plan. Manag. 139, 723–732. https://doi.org/10.1061/(ASCE)WR.1943-5452

Priestley, C.H.B., Taylor, R.J., 1972. On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. Mon. Weather Rev. 100, 81–92.

https://doi.org/10.1175/1520-0493(1972)100<0081:otaosh>2.3.co;2

Quenouille, M.H., 1956. Notes on Bias in Estimation. Biometrika 43, 353. https://doi.org/10.2307/2332914

Razavi, T., Coulibaly, P., 2013. Streamflow Prediction in Ungauged Basins: Review of Regionalization Methods. J. Hydrol. Eng. 18, 958–975. https://doi.org/10.1061/(asce)he.1943-5584.0000690

Rosenzweig, C., Casassa, G., Karoly, D.J., Imeson, a., Liu, C., Menzel, a., Rawlins, S., Root, T.L., Seguin, B., Tryjanowski, P., 2007. Assessment of observed changes and responses in natural and managed systems. Clim. Chang. 2007 Impacts, Adapt. Vulnerability. Contrib. Work. Gr. II to Fourth Assess. Rep. Intergov. Panel Clim. Chang. 79–131. https://doi.org/Milandre climat

Roth, V., Nigussie, T.K., Lemann, T., 2016. Model parameter transfer for streamflow and sediment loss prediction with SWAT in a tropical watershed. Environ. Earth Sci. 75, 1–13. https://doi.org/10.1007/s12665-016-6129-9

Runkel, R.L., Crawford, C.G., Cohn, T. a, 2004. Load Estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers. Tech. Methods. U.S. Geol. Surv. U.S. Dep. Inter. 4, 69.

Sabo, J.L., Ruhi, A., Holtgrieve, G.W., Elliott, V., Arias, M.E., Ngor, P.B., Räsänen, T.A., Nam, S., 2017. Designing river flows to improve food security futures in the Lower Mekong Basin. Science (80-.). 358. https://doi.org/10.1126/science.aao1053

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.T., Chuang, H.Y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M.P., Van Den Dool, H., Zhang, Q., Wang, W., Chen, M., Becker, E., 2014. The NCEP climate forecast system version 2. J. Clim. 27, 2185–2208. https://doi.org/10.1175/JCLI-D-12-00823.1

Sakamoto, T., Van Nguyen, N., Kotera, A., Ohno, H., Ishitsuka, N., Yokozawa, M., 2007. Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. Remote Sens. Environ. 109, 295–313. https://doi.org/10.1016/j.rse.2007.01.011

Samal, D.R., Gedam, S., 2021. Assessing the impacts of land use and land cover change on water resources in the Upper Bhima river basin, India. Environ. Challenges 5, 100251. https://doi.org/10.1016/j.envc.2021.100251

Samaniego, L., Bárdossy, A., Kumar, R., 2010. Streamflow prediction in ungauged catchments using copula-based dissimilarity measures. Water Resour. Res. https://doi.org/10.1029/2008WR007695

Samuel, J., Coulibaly, P., Metcalfe, R.A., 2011. Estimation of Continuous Streamflow in Ontario Ungauged Basins: Comparison of Regionalization Methods. J. Hydrol. Eng. 16, 447–459. https://doi.org/10.1061/(asce)he.1943-5584.0000338

Sao, D., Kato, T., Tu, L.H., Thouk, P., Fitriyah, A., Oeurng, C., 2020. Evaluation of different objective functions used in the Sufi-2 calibration process of swat-cup on water balance analysis: A case study of the Pursat River basin, Cambodia. Water (Switzerland) 12, 1–22. https://doi.org/10.3390/w12102901

Satgé, F., Defrance, D., Sultan, B., Bonnet, M.P., Seyler, F., Rouché, N., Pierron, F., Paturel, J.E., 2020. Evaluation of 23 gridded precipitation datasets across West Africa. J. Hydrol. 581, 124412. https://doi.org/10.1016/j.jhydrol.2019.124412

Schuol, J., Abbaspour, K.C., 2006. Calibration and uncertainty issues of a hydrological model (SWAT) applied to West Africa. Adv. Geosci. 2, 137–143.

Senevirathne, N., Mony, K., Hazarika, M., Samarakoon, L., 2010. Land Cover/Land Use Change Detection in Tonle Sap Watershed, Cambodia. Conf. 31st Asian Conf. Remote Sens. 2–7.

Shivakoti, B.R., Bao, P.N., 2020. Environmental Changes in Tonle Sap Lake and its Floodplain: Status and Policy Recommendations. Institute for Global Environmental Strategies (IGES), Tokyo Institute of Technology (Tokyo Tech) and Institute of Technology of Cambodia (ITC).

Shrestha, B., Cochrane, T.A., Caruso, B.S., Arias, M.E., 2017. Land use change uncertainty impacts on streamflow and sediment projections in areas undergoing rapid development: A case study in the Mekong Basin. L. Degrad. Dev. 29, 835–848. https://doi.org/10.1002/ldr.2831

Shrestha, M., Acharya, S.C., Shrestha, P.K., 2017. Bias correction of climate models for hydrological modelling – are simple methods still useful? Meteorol. Appl. 24, 531–539. https://doi.org/10.1002/met.1655

Shrestha, S., Bhatta, B., Shrestha, M., Shrestha, P.K., 2018. Integrated assessment of the climate and landuse change impact on hydrology and water quality in the Songkhram River Basin, Thailand. Sci. Total Environ. 643, 1610–1622. https://doi.org/10.1016/j.scitotenv.2018.06.306

Siev, S., Yang, H., Sok, T., Uk, S., Song, L., Kodikara, D., Oeurng, C., Hul, S., Yoshimura, C., 2018. Sediment dynamics in a large shallow lake characterized by seasonal flood pulse in Southeast Asia. Sci. Total Environ. 631–632, 597–607. https://doi.org/10.1016/j.scitotenv.2018.03.066

Singh, V., Xiaosheng, Q., 2019. Data assimilation for constructing long-term gridded daily rainfall time series over Southeast Asia. Clim. Dyn. 53, 3289–3313. https://doi.org/10.1007/s00382-019-04703-6

Sirisena, T.A.J.G., Maskey, S., Ranasinghe, R., 2020. Hydrological model calibration with streamflow and remote sensing based evapotranspiration data in a data poor basin. Remote Sens. 12, 1–24. https://doi.org/10.3390/rs12223768

Sok, T., Oeurng, C., Ich, I., Sauvage, S., Sanchez-Perez, J.M., 2020. Assessment of Hydrology and Sediment Yield in the Mekong River Basin Using SWAT Model. Water (Switzerland) 12, 1–25. https://doi.org/https://doi.org/10.3390/w12123503

Soum, S., Ngor, P.B., Dilts, T.E., Lohani, S., Kelson, S., Null, S.E., Tromboni, F., Hogan, Z.S., Chan, B., Chandra, S., 2021. Spatial and long-term temporal changes in water quality dynamics of the tonle sap ecosystem. Water (Switzerland) 13. https://doi.org/10.3390/w13152059

Sun, H., Cornish, P.S., 2006. A catchment-based approach to recharge estimation in the Liverpool Plains, NSW, Australia. Aust. J. Agric. Res. 57, 309–320. https://doi.org/https://doi.org/10.1071/AR04015

Sun, Q., Miao, C., Duan, Q., Kong, D., Ye, A., Di, Z., Gong, W., 2014. Would the “real” observed dataset stand up? A critical examination of eight observed gridded climate datasets for China. Environ. Res. Lett. 9. https://doi.org/10.1088/1748-9326/9/1/015001

Sutton, W.R., Srivastava, J.P., Koo, J., Vasileiou, I., Pradesh, A., 2019. Striking a Balance: Managing El Niño and La Niña in Cambodia’s Agriculture. Washington, DC.

Swain, J.B., Patra, K.C., 2017. Streamflow estimation in ungauged catchments using regionalization techniques. J. Hydrol. 554, 420–433. https://doi.org/10.1016/j.jhydrol.2017.08.054

Tan, M.L., Gassman, P.W., Liang, J., Haywood, J.M., 2021. A review of alternative climate products for SWAT modelling: Sources, assessment and future directions. Sci. Total Environ. 795, 148915. https://doi.org/10.1016/j.scitotenv.2021.148915

Tan, M.L., Ibrahim, A.L., Duan, Z., Cracknell, A.P., Chaplot, V., 2015. Evaluation of six high-resolution satellite and ground-based precipitation products over Malaysia.

Remote Sens. 7, 1504–1528. https://doi.org/10.3390/rs70201504

Tan, M.L., Ibrahim, A.L., Yusop, Z., Duan, Z., Ling, L., 2015. Impacts of land-use and climate variability on hydrological components in the Johor River basin, Malaysia. Hydrol. Sci. J. 60, 873–889. https://doi.org/10.1080/02626667.2014.967246

Tarek, M., Brissette, F.P., Arsenault, R., 2020. Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. Hydrol. Earth Syst. Sci. 24, 2527–2544. https://doi.org/10.5194/hess-24-2527-2020

Tegegne, G., Kim, Y.O., 2018a. Modelling ungauged catchments using the catchment runoff response similarity. J. Hydrol. 564, 452–466. https://doi.org/10.1016/j.jhydrol.2018.07.042

Teshager, A.D., Gassman, P.W., Secchi, S., Schoof, J.T., Misgna, G., 2016. Modeling Agricultural Watersheds with the Soil and Water Assessment Tool (SWAT): Calibration and Validation with a Novel Procedure for Spatially Explicit HRUs. Environ. Manage. 57, 894–911. https://doi.org/10.1007/s00267-015-0636-4

Touch, T., Oeurng, C., Jiang, Y., Mokhtar, A., 2020. Integrated modeling of water supply and demand under climate change impacts and management options in tributary Basin of Tonle Sap Lake, Cambodia. Water (Switzerland) 12. https://doi.org/10.3390/w12092462

Trang, N.T.T., Shrestha, S., Shrestha, M., Datta, A., Kawasaki, A., 2017. Evaluating the impacts of climate and land-use change on the hydrology and nutrient yield in a transboundary river basin: A case study in the 3S River Basin (Sekong, Sesan, and Srepok). Sci. Total Environ. 576, 586–598. https://doi.org/10.1016/j.scitotenv.2016.10.138

Trinh-Tuan, L., Matsumoto, J., Ngo-Duc, T., Nodzu, M.I., Inoue, T., 2019. Evaluation of satellite precipitation products over Central Vietnam. Prog. Earth Planet. Sci. 6. https://doi.org/10.1186/s40645-019-0297-7

Try, S., Lee, G., Yu, W., Oeurng, C., Jang, C., 2018. Large-Scale Flood-Inundation Modeling in the Mekong River Basin. J. Hydrol. Eng. 23, 05018011. https://doi.org/10.1061/(asce)he.1943-5584.0001664

Try, S., Sayama, T., Oeurng, C., Sok, T., Ly, S., Uk, S., 2022. Identification of the spatio-temporal and fluvial-pluvial sources of flood inundation in the Lower Mekong Basin. Geosci. Lett. 9. https://doi.org/10.1186/s40562-022-00215-0

Try, S., Tanaka, S., Tanaka, K., Sayama, T., Oeurng, C., Uk, S., Takara, K., Hu, M., Han, D., 2020. Comparison of gridded precipitation datasets for rainfall-runoff and inundation modeling in the Mekong River Basin. PLoS One 15, 1–13. https://doi.org/10.1371/journal.pone.0226814

Tsukawaki, S., Lao, K.-L., 1994. Lake Tonle Sap, Cambodia. J. Geogr. (Chigaku Zasshi) 103, Plate3–Plate4. https://doi.org/10.5026/jgeography.103.6_plate3

Tukey, J.W., 1958. Bias and Confidence in Not-Quite Large Sample. Abstract. Ann. Math. Stat. 9, 614.

Ud din, S., Al-Dousari, A., Ramdan, A., Al Ghadban, A., 2008. Site-specific precipitation estimate from TRMM data using bilinear weighted interpolation technique: An example from Kuwait. J. Arid Environ. 72, 1320–1328. https://doi.org/10.1016/j.jaridenv.2007.12.013

Uk, S., Yoshimura, C., Siev, S., Try, S., Yang, H., Oeurng, C., Li, S., Hul, S., 2018. Tonle Sap Lake: Current status and important research directions for environmental management. Lakes Reserv. Res. Manag. 23, 177–189. https://doi.org/10.1111/lre.12222

Ur Rahman, K., Shang, S., Shahid, M., Wen, Y., 2020. Hydrological evaluation of merged satellite precipitation datasets for streamflow simulation using SWAT: A case study of Potohar Plateau, Pakistan. J. Hydrol. 587, 125040. https://doi.org/10.1016/j.jhydrol.2020.125040

USDA, 1972. National engineering handbook, section 4: Hydrology. Washington, DC.

van den Besselaar, E.J.M., van der Schrier, G., Cornes, R.C., Suwondo, A., Iqbal, Tank, A.M.G.K., 2017. SA-OBS: A daily gridded surface temperature and precipitation dataset for Southeast Asia. J. Clim. 30, 5151–5165. https://doi.org/10.1175/JCLI-D-16-0575.1

Vesanto, J., Alhoniemi, E., 2000. Clustering of the self-organizing map. IEEE Trans. Neural Networks 11, 586–600. https://doi.org/10.1109/72.846731

Vilaysane, B., Takara, K., Luo, P., Akkharath, I., Duan, W., 2015. Hydrological Stream Flow Modelling for Calibration and Uncertainty Analysis Using SWAT Model in the Xedone River Basin, Lao PDR. Procedia Environ. Sci. 28, 380–390. https://doi.org/10.1016/j.proenv.2015.07.047

Vu, M.T., Raghavan, S. V., Liong, S.Y., 2012. SWAT use of gridded observations for simulating runoff - A Vietnam river basin study. Hydrol. Earth Syst. Sci. 16, 2801–2811. https://doi.org/10.5194/hess-16-2801-2012

Vu, T.M., Raghavan, S. V., Liong, S.Y., Mishra, A.K., 2018. Uncertainties of gridded precipitation observations in characterizing spatio-temporal drought and wetness over Vietnam. Int. J. Climatol. 38, 2067–2081. https://doi.org/10.1002/joc.5317

Wagener, T., Sivapalan, M., Troch, P., Woods, R., 2007. Catchment Classification and Hydrologic Similarity. Geogr. Compass 1, 901–931. https://doi.org/10.1111/j.1749-8198.2007.00039.x

Wallner, M., Haberlandt, U., Dietrich, J., 2013. A one-step similarity approach for the regionalization of hydrological model parameters based on self-organizing maps. J. Hydrol. 494, 59–71. https://doi.org/10.1016/j.jhydrol.2013.04.022

Wan, R., Cai, S., Li, H., Yang, G., Li, Z., Nie, X., 2014. Inferring land use and land cover impact on stream water quality using a Bayesian hierarchical modeling approach in the Xitiaoxi River Watershed, China. J. Environ. Manage. 133, 1–11. https://doi.org/10.1016/j.jenvman.2013.11.035

Wang, J.-J., Lu, X.X., Kummu, M., 2011. Sediment load estimation and variations in the Lower Mekong River. River Res. Appl. 30, 132–133. https://doi.org/10.1002/rra.1337

Wang, W., Lu, H., Zhao, T., Jiang, L., Shi, J., 2017. Evaluation and comparison of daily rainfall from the latest GPM and TRMM products over the Mekong River Basin. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 2540–2549. https://doi.org/10.1109/JSTARS.2017.2672786

Wang, Y., Feng, L., Liu, J., Hou, X., Chen, D., 2020. Changes of inundation area and water turbidity of Tonle Sap Lake: Responses to climate changes or upstream dam construction? Environ. Res. Lett. 15. https://doi.org/10.1088/1748-9326/abac79

Wang, Z., Zhong, R., Lai, C., Chen, J., 2017. Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility. Atmos. Res. 196, 151–163. https://doi.org/10.1016/j.atmosres.2017.06.020

Williams, J.R., 1975. Sediment-yield prediction with universal equation using runoff energy factor, in: In Present and Prospective Technology for Predicting Sediment Yield and Sources: Proceedings of the Sediment- Yield Workshop. USDA Sedimentation Lab, Oxford, pp. 244–252.

Willmott, C.J., 1981. On the validation of models. Phys. Geogr. 2, 184–194. https://doi.org/10.1080/02723646.1981.10642213

Willmott, C.J., 1984. On the Evaluation of Model Performance in Physical Geography. Spat. Stat. Model. 443–460. https://doi.org/10.1007/978-94-017-3048-8_23

Wischmeier, W.H., Smith, D.D., 1965. Predicting Rainfall-Erosion Losses From Cropland east of the Rocky Mountains: Guide for Selection of Practices for Soil and Water Conservation. Agric. Handb. no. 282 47.

Wong, J.S., Zhang, X., Gharari, S., Shrestha, R.R., Wheater, H.S., Famiglietti, J.S., 2021. Assessing Water Balance Closure Using Multiple Data Assimilation and Remote Sensing-Based Datasets for Canada. J. Hydrometeorol. 1569–1589. https://doi.org/10.1175/jhm-d-20-0131.1

World Bank, 2011. Climate Risk and Adaptation Country Profiles: Vulnerability, Risk Reduction, and Adaptation to Climate Change. World Dev. 1–12.

Worqlul, A.W., Maathuis, B., Adem, A.A., Demissie, S.S., Langan, S., Steenhuis, T.S., 2014. Comparison of rainfall estimations by TRMM 3B42, MPEG and CFSR with ground-observed data for the Lake Tana basin in Ethiopia. Hydrol. Earth Syst. Sci. 18, 4871–4881. https://doi.org/10.5194/hess-18-4871-2014

Wu, H., Chen, B., 2015. Evaluating uncertainty estimates in distributed hydrological modeling for the Wenjing River watershed in China by GLUE, SUFI-2, and ParaSol methods. Ecol. Eng. 76, 110–121. https://doi.org/10.1016/j.ecoleng.2014.05.014

Xu, H., Xu, C.Y., Chen, S., Chen, H., 2016a. Similarity and difference of global reanalysis datasets (WFD and APHRODITE) in driving lumped and distributed hydrological models in a humid region of China. J. Hydrol. 542, 343–356. https://doi.org/10.1016/j.jhydrol.2016.09.011

Yang, X., Yong, B., Ren, L., Zhang, Y., Long, D., 2017. Multi-scale validation of GLEAM evapotranspiration products over China via ChinaFLUX ET measurements. Int. J. Remote Sens. 38, 5688–5709. https://doi.org/10.1080/01431161.2017.1346400

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., Kitoh, A., 2012. Aphrodite constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. Bull. Am. Meteorol. Soc. 93, 1401–1415. https://doi.org/10.1175/BAMS-D-11-00122.1

Yuan, L., Forshay, K.J., 2021. Enhanced streamflow prediction with SWAT using support vector regression for spatial calibration: A case study in the Illinois River watershed, U.S. PLoS One 16. https://doi.org/10.1371/journal.pone.0248489

Zam, P., Shrestha, S., Budhathoki, A., 2021. Assessment of climate change impact on hydrology of a transboundary river of Bhutan and India. J. Water Clim. Chang. 12, 3224–3239. https://doi.org/10.2166/wcc.2021.338

Zhang, G., Su, X., Ayantobo, O.O., Feng, K., Guo, J., 2020a. Remote-sensing precipitation and temperature evaluation using soil and water assessment tool with multiobjective calibration in the Shiyang River Basin, Northwest China. J. Hydrol. 590, 125416. https://doi.org/10.1016/j.jhydrol.2020.125416

Zhang, G., Su, X., Ayantobo, O.O., Feng, K., Guo, J., 2020b. Evaluation of open access precipitation and temperature products using SWAT in Shiyang river basin, Northwest China. https://doi.org/10.20944/preprints202003.0294.v1

Zhang, Y., Chiew, F.H.S., 2009. Relative merits of different methods for runoff predictions in ungauged catchments. Water Resour. Res. 45. https://doi.org/10.1029/2008WR007504

Zhao, W., Abhishek, Kinouchi, T., 2022a. Uncertainty quantification in intensity-duration-frequency curves under climate change: Implications for flood-prone tropical cities. Atmos. Res. 270, 106070. https://doi.org/10.1016/j.atmosres.2022.106070

Zhao, W., Abhishek, Kinouchi, T., Ang, R., Zhuang, Q., 2022. A framework for quantifying climate-informed heavy rainfall change: Implications for adaptation strategies. Sci. Total Environ. 835, 155553. https://doi.org/10.1016/j.scitotenv.2022.155553

Zhu, J., Xie, A., Qin, X., Wang, Yetang, Xu, B., Wang, Yicheng, 2021. An assessment of ERA5 reanalysis for antarctic near-surface air temperature. Atmosphere (Basel). 12. https://doi.org/10.3390/atmos12020217

Appendices

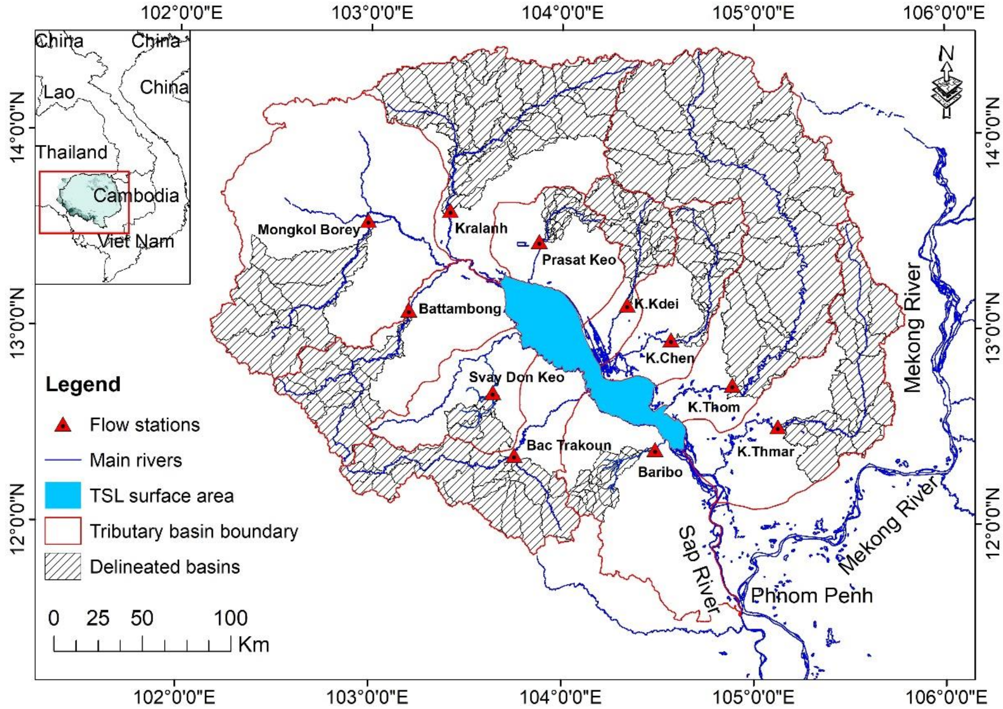


Figure A1. Map of the delineated basin area of each tributary basin. The hatched area denotes the gauging area of each tributary basin in the TSL basin.

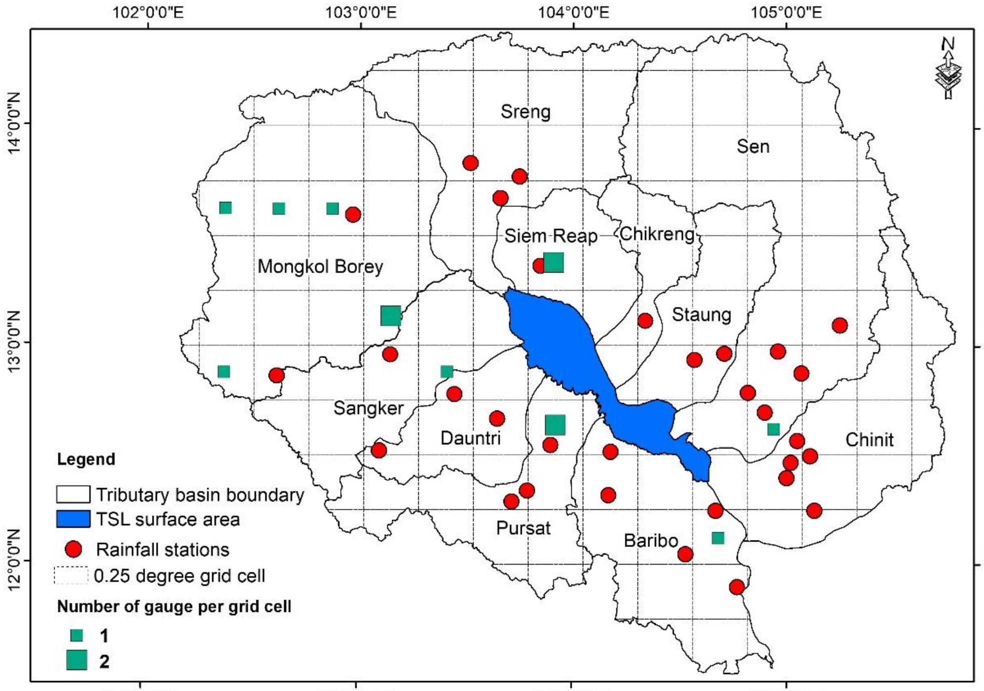
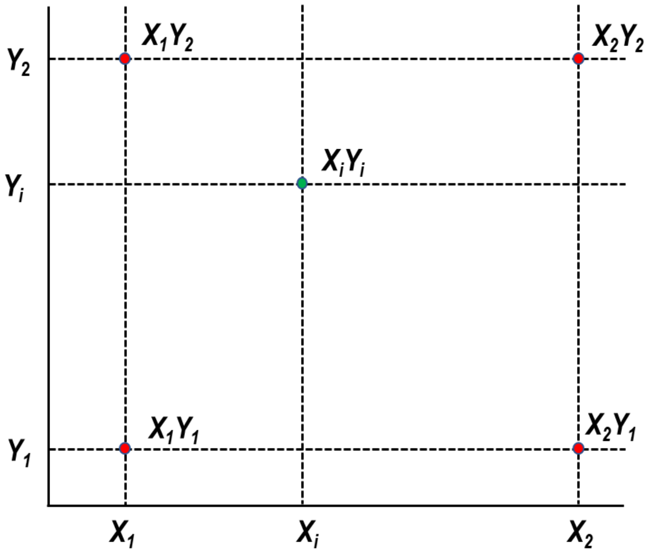


Figure A2. The average number of rain gauges per grid cell used for creating the APHRODITE dataset. The number of gauges used in APHRODITE (green square) is not necessarily the same as the number of rainfall stations (red circle, same as Figure 2.1) in each grid cell because of the different time coverage of rainfall data and data availability at each station. Each green square is plotted at the center of the grid cell.



$$f\left(x_i, y_i\right) = \frac{y_i - y_2}{y_2 - y_1}\left(\frac{x_i - x_2}{x_1 - x_2} f\left(x_1, y_1\right) + \frac{x_i - x_1}{x_2 - x_1} f\left(x_2, y_1\right)\right) + \frac{y_i - y_1}{y_2 - y_1}\left(\frac{x_i - x_2}{x_1 - x_2} f\left(x_1, y_2\right) + \frac{x_i - x_1}{x_2 - x_1} f\left(x_2, y_2\right)\right)$$

Figure A3. Schematic and equation of bilinear interpolation method.

Table A1. Equations and optimal values of statistical indices

N°	Statistical indices	Equations*	Range values	Optimal values
1	Mean Bias Error	$MBE=\frac{1}{N}\sum_{i=1}^{N}\left(P_i-O_i\right)$	-∞ to +∞	0
2	Root Mean Square Error	$RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(P_i-O_i\right)^2}$	0 to +∞	0
3	Correlation Coefficient	$R=\frac{\sum\left(P_i-\overline{P}\right)\left(O_i-\overline{O}\right)}{\sqrt{\sum\left(P_i-\overline{P}\right)^2\sum\left(O_i-\overline{O}\right)^2}}$	-1 to 1	1
4	Modified Index of Agreement	$md=1-\frac{\sum_{i=1}^{N}\left(O_i-P_i\right)}{\sum_{i=1}^{N}\left(\left\lvert P_i-\overline{O}\right\rvert+\left\lvert O_i-\overline{O}\right\rvert\right)}$	0 to 1	1
5	Nash-Sutcliffe Efficiency	$NSE=1-\left[\frac{\sum_{i=1}^{N}\left(Q_i^{obs}-Q_i^{sim}\right)^2}{\sum_{i=1}^{N}\left(Q_i^{obs}-Q_{mean}^{obs}\right)^2}\right]$	-∞ to 1	1
6	Percent Bias	$PBIAS=\frac{\sum_{i=1}^{N}\left(Q_i^{obs}-Q_i^{sim}\right)}{\sum_{i=1}^{N}Q_i^{obs}}\times 100$	-∞ to +∞	0
7	Coefficient of Determination	$R^2=\left[\frac{\sum_{i=1}^{N}\left(Q_i^{obs}-Q_{mean}^{obs}\right)\left(Q_i^{sim}-Q_{mean}^{sim}\right)}{\sqrt{\sum_{i=1}^{N}\left(Q_i^{obs}-Q_{mean}^{obs}\right)^2}\sqrt{\sum_{i=1}^{N}\left(Q_i^{sim}-Q_{mean}^{sim}\right)^2}}\right]^2$	0 to 1	1

* P_i and O_i denote predicted and observed, respectively, daily precipitation or temperature of the ith day, $\overline{P}$ and $\overline{O}$ denote predicted and observed, respectively, mean daily precipitation or temperature, and N is the number of data at daily time series. Q^{sim} and Q^{obs} denote simulated and observed, respectively, streamflow of the ith day. Q_{mean} denotes mean daily streamflow. The first four indices are used for the evaluation of precipitation and temperature, while the last three indices are used to evaluate the streamflow simulation performance.

Table A2. The selected parameters and their initial range setting for calibration using SWAT-CUP

No.	Parameter name	Description	Range		Method
			Min	Max	
1	r__CN2.mgt	SCS runoff curve number (f)	-25%	25%	Relative (1)
2	v__ALPHA_BF.gw	Baseflow alpha factor (days)	0	1	Replace (2)
3	v__GW_DELAY.gw	Groundwater delay (days)	0	500	-
4	v__GWQMN.gw	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)	0	5000	-
5	v__GW_REVAP.gw	Groundwater "revap" coefficient	0.02	0.2	-
6	v__REVAPMN.gw	Threshold depth of water in the shallow aquifer required for "revap" to occur (mm)	0	500	-
7	v__RCHRG_DP.gw	Deep aquifer percolation fraction	0	1	-
8	v__LAT_TTIME.hru	Lateral flow travel time	0	180	-
9	v__SLSOIL.hru	Slope length for lateral subsurface flow	0	150	-
10	v__CANMX.hru	Maximum canopy storage	0	100	-
11	v__ESCO.hru	Soil evaporation compensation factor	0	1	-
12	v__EPCO.hru	Plant uptake compensation factor	0	1	-
13	v__OV_N.hru	Manning's "n" value for overland flow	0.01	30	-
14	v__CH_N2.rte	Manning's "n" value for the main channel	-0.01	0.3	-
15	v__CH_K2.rte	Effective hydraulic conductivity in the main channel alluvium	-0.01	500	-
16	v__ALPHA_BNK.rte	Baseflow alpha factor for bank storage	0	1	-
17	v__CH_K1.sub	Effective hydraulic conductivity in tributary channel alluvium	0	300	-
18	v__CH_N1.sub	Manning's "n" value in the tributary channel	0.01	30	-
19	r__SOL_AWC().sol	Available water capacity of the soil layer	-25%	25%	-
20	r__SOL_BD().sol	Moist bulk density	-25%	25%	Relative (1)
21	r__SOL_K().sol	Saturated hydraulic conductivity	-25%	25%	-
22	r__SOL_ALB().sol	Moist soil albedo	-25%	25%	-
23	v__SURLAG.bsn	Surface runoff lag time	0	24	Replace (2)

Note: (1): Multiplying initial parameter by value in percentage; (2): replacing initial parameter by value

Table A3. Performance ratings of recommended statistics for streamflow simulation

Performance rating	NSE	R^2	PBIAS (%)
Unsatisfactory	NSE ≤ 0.5	$R^2 \leq 0.5$	PBIAS ≥ ±25
Satisfactory	0.5 - 0.65	0.5 - 0.65	±15 - ±25
Good	0.65 - 0.75	0.65 - 0.75	±10 - ±15
Very good	0.75 - 1	0.75 - 1	PBIAS < ±10

Table A4. The mean value of each statistical index was calculated for thirty-one gauging stations.

Indices	Datasets					
	APHRODITE	CFSR	ERA5	SA-OBS	TRMM	IMERG
MBE (mm/month)	-17.87	53.22	12.16	55.45	26.59	28.95
RMSE (mm/month)	60.46	135.32	73.44	147.22	81.87	83.05
R	0.86	0.56	0.78	0.68	0.79	0.79
md	0.77	0.56	0.71	0.63	0.72	0.71

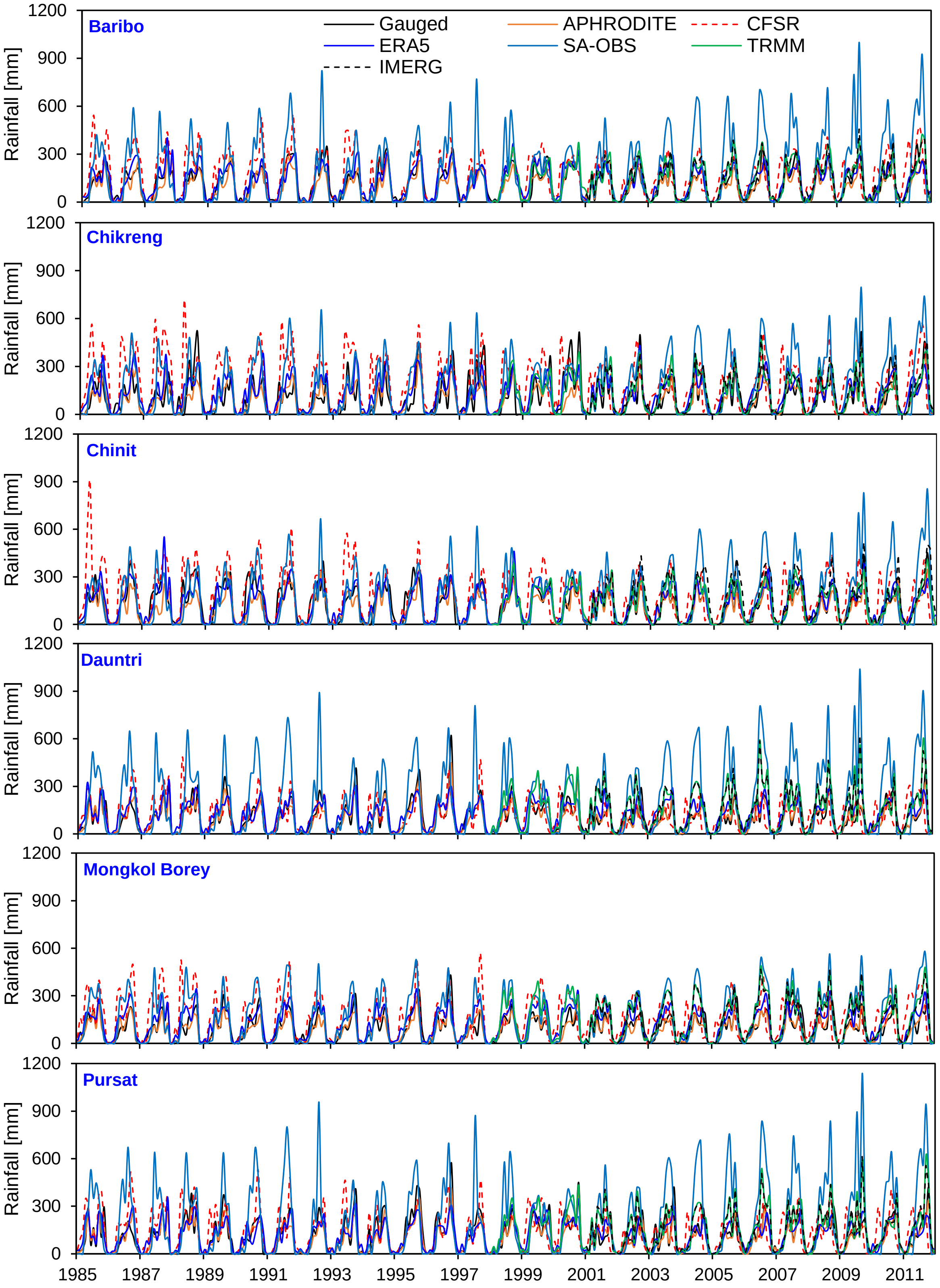


Figure A4. Comparison of monthly rainfall averaged for all the gauging stations included in each tributary basin from 1998 to 2011 for TRMM, 2001 to 2011 for IMERG and 1985 to 2011 for other products and observations.

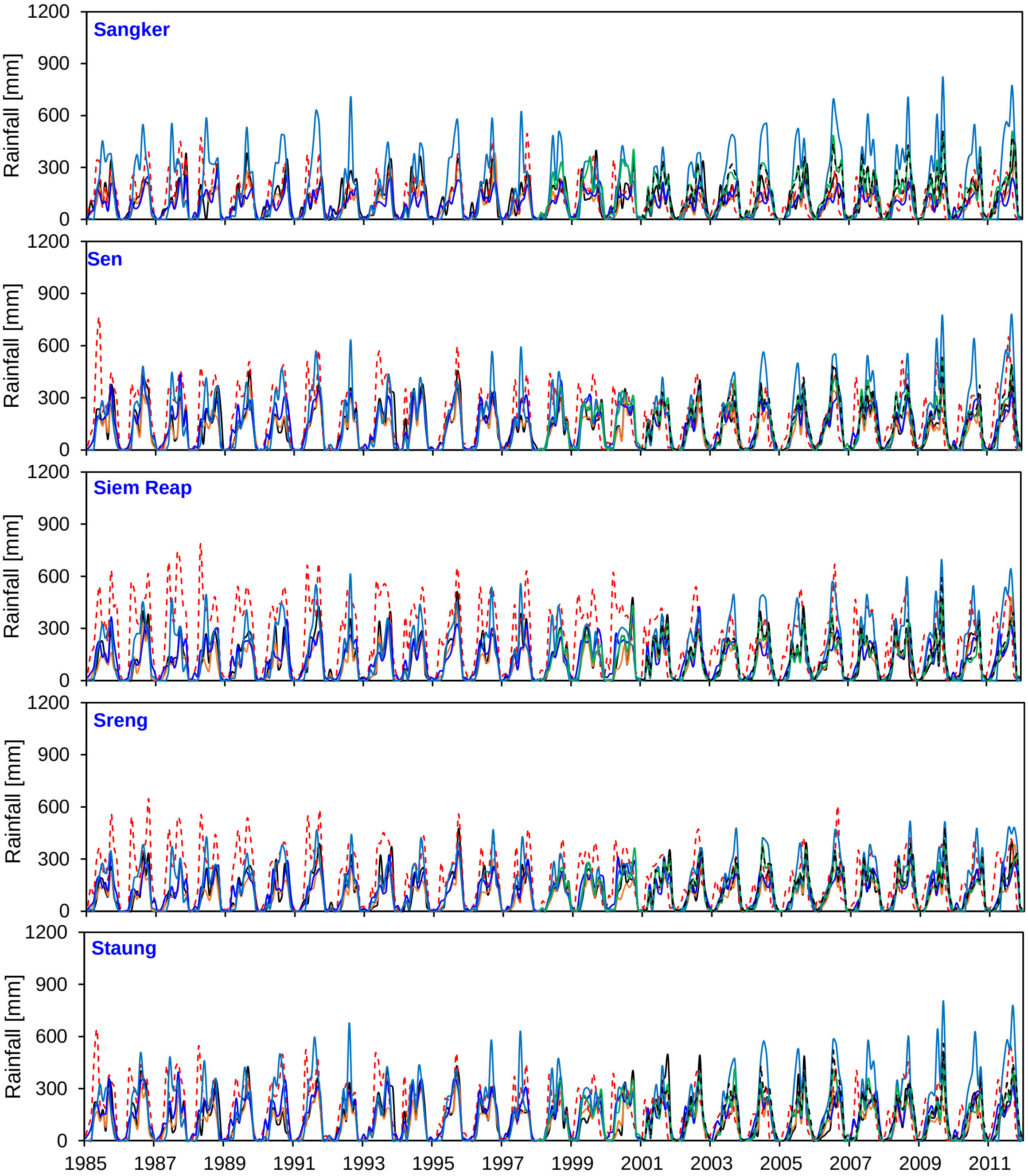


Figure A4. *Cont.*

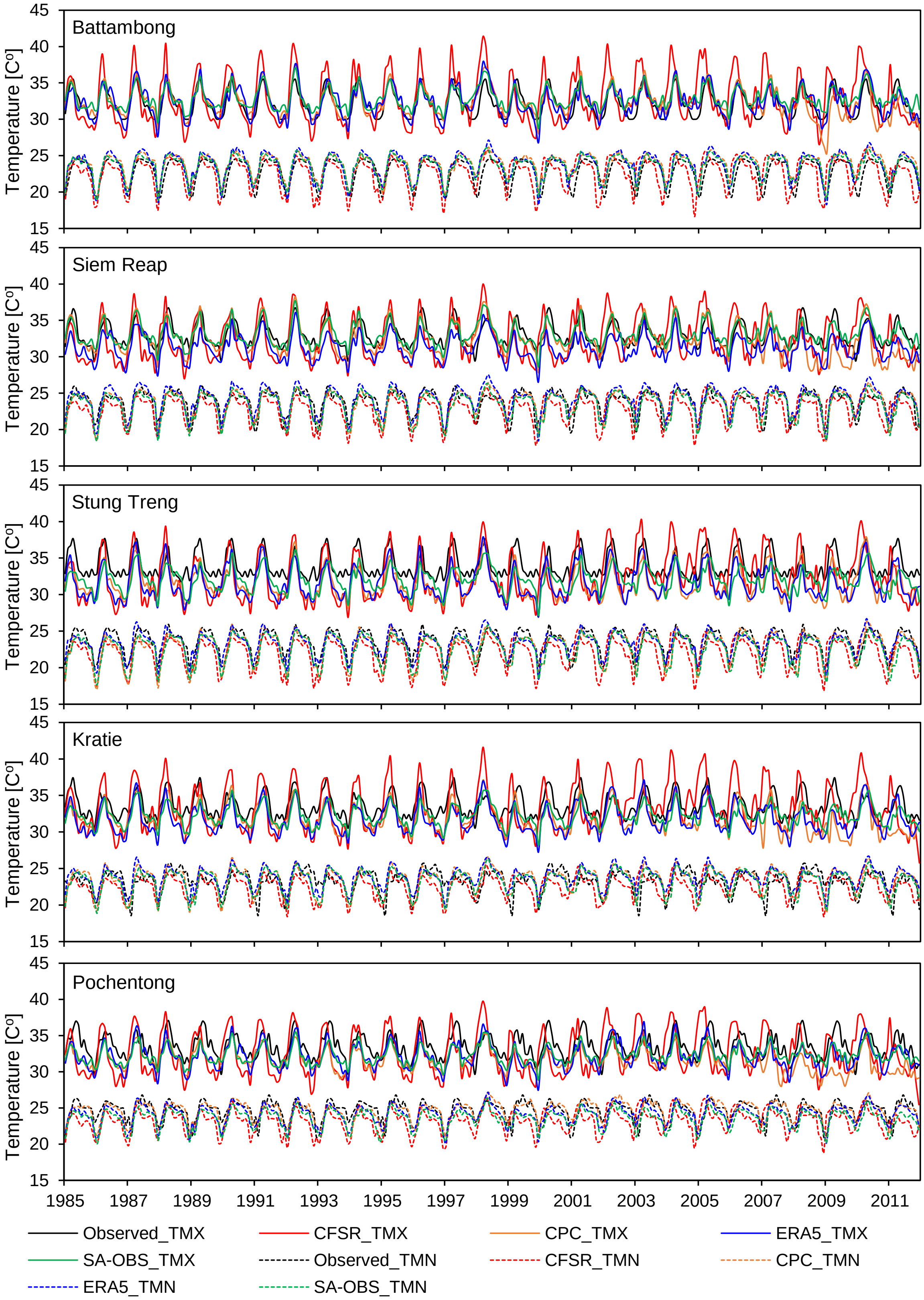


Figure A5. Comparison of monthly mean daily maximum (TMX) and daily minimum (TMN) temperatures of individual stations from 19985 – 2011. The different colors of solid- and dash-lines denote the maximum and minimum temperature, respectively, of different datasets.

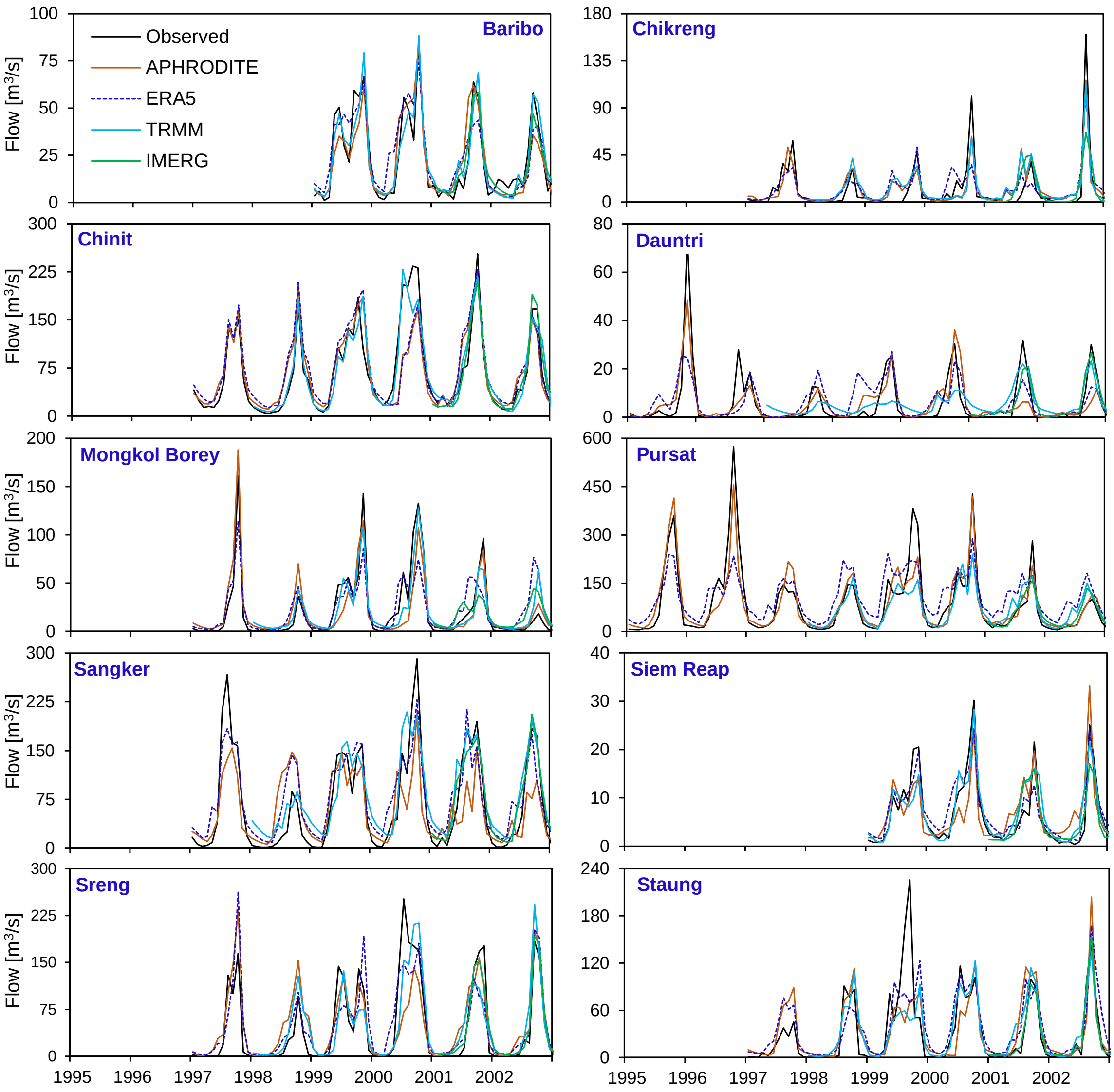


Figure A6. Monthly observed and simulated flow of each tributary basin in the calibration period. The beginning of the flow simulation depended on the flow data availability of each tributary basin.

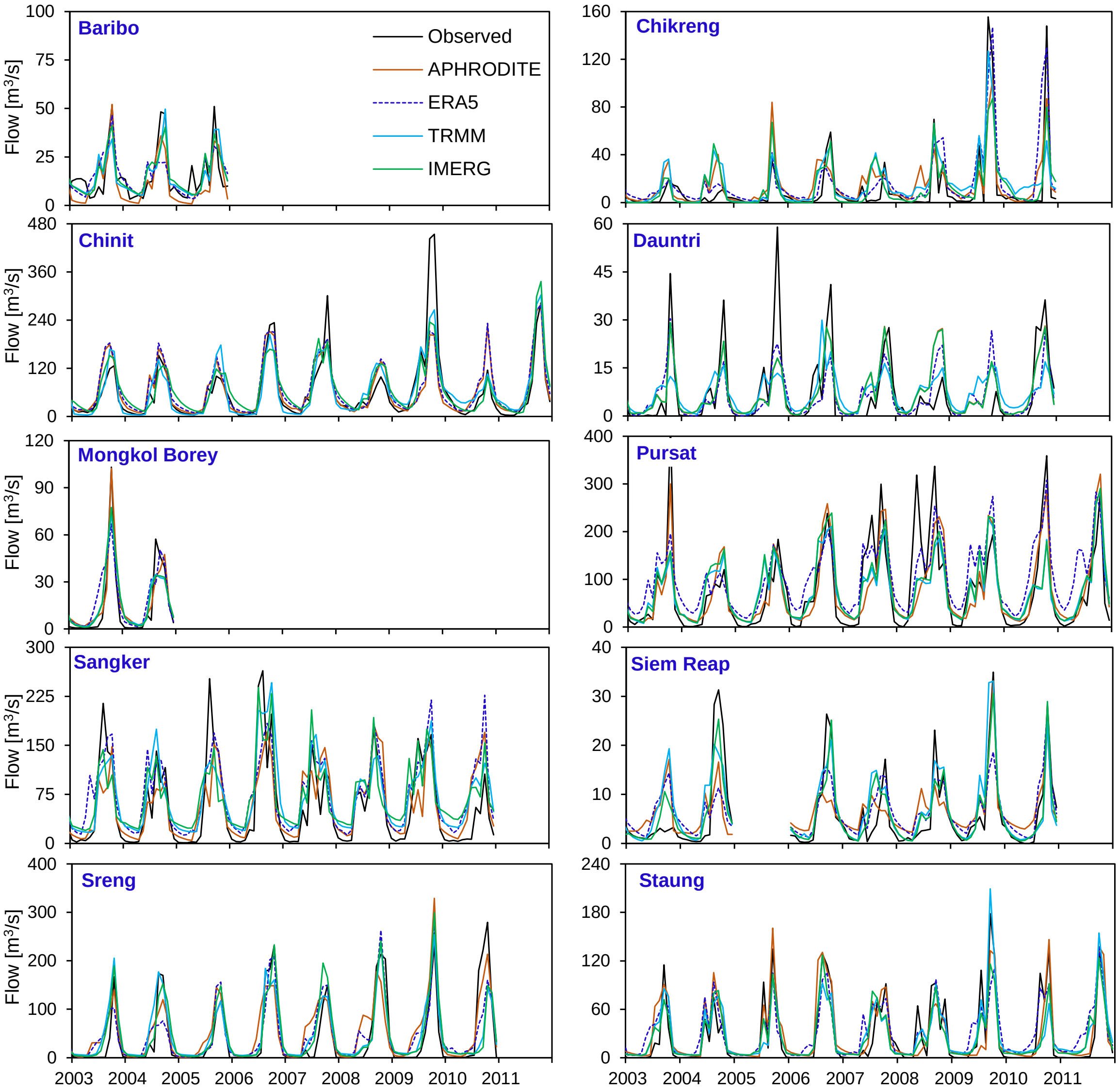


Figure A7. Same as Figure A6 but for the validation period. Due to discontinued monitoring stations in Mongkol Borey and Baribo tributary basins, the flow simulations were validated from 2003 to 2004 and 2005, respectively. The flow data of the Siem Reap tributary basin was missing in the year 2005, thus; the model was validated from 2003 to 2004 and 2006 to 2010.

Table A5. The performance of statistical indices in the calibration and validation periods.

Dataset	River name	Calibration			Validation		
		NSE	PBIAS (%)	R^2	NSE	PBIAS (%)	R^2
APHRODITE	Baribo	0.62	3.40	0.62	0.43	27.20	0.53
	Chikreng	0.54	-10.60	0.56	0.53	-34.20	0.54
	Chinit	0.72	-0.80	0.72	0.66	-4.90	0.66
	Dauntri	0.52	5.30	0.53	0.49	-8.30	0.50
	Mongkol Borey	0.79	4.60	0.79	0.73	-15.90	0.75
	Pursat	0.68	1.90	0.68	0.52	2.00	0.52
	Sangker	0.43	13.50	0.44	0.30	-2.70	0.30
	Sen	0.80	-5.40	0.80	0.76	-1.10	0.76
	Siem Reap	0.57	-1.30	0.57	0.60	-11.80	0.61
	Sreng	0.65	-6.20	0.65	0.67	-22.20	0.69
	Staung	0.45	-5.70	0.46	0.62	12.00	0.64
	Average	**0.62**	**-0.12**	**0.62**	**0.57**	**-5.45**	**0.59**
Dataset	River name	Calibration			Validation		
		NSE	PBIAS (%)	R^2	NSE	PBIAS (%)	R^2
ERA5	Baribo	0.54	-8.20	0.55	0.37	9.76	0.41
	Chikreng	0.34	-15.20	0.34	0.50	-48.40	0.52
	Chinit	0.70	-9.10	0.71	0.65	-13.60	0.66
	Dauntri	0.40	-8.70	0.40	0.42	6.40	0.43
	Mongkol Borey	0.59	-5.10	0.60	0.65	-20.60	0.67
	Pursat	0.48	-23.60	0.53	0.44	-31.60	0.49
	Sangker	0.58	-14.80	0.60	0.39	-28.80	0.42
	Sen	0.84	-6.60	0.84	0.79	-9.80	0.80
	Siem Reap	0.59	-9.10	0.59	0.52	-12.20	0.57
	Sreng	0.70	-7.10	0.70	0.67	-11.10	0.67
	Staung	0.44	-21.00	0.44	0.58	-15.90	0.60
	Average	**0.56**	**-11.68**	**0.57**	**0.54**	**-15.99**	**0.57**
Dataset	River name	Calibration			Validation		
		NSE	PBIAS (%)	R^2	NSE	PBIAS (%)	R^2
TRMM	Baribo	0.66	-2.10	0.66	0.45	6.50	0.46
	Chikreng	0.53	-24.80	0.56	0.47	-36.50	0.49
	Chinit	0.81	-0.60	0.81	0.74	-3.20	0.74
	Dauntri	0.43	-13.70	0.46	0.33	-9.60	0.34
	Mongkol Borey	0.82	0.14	0.83	0.72	-12.80	0.75
	Pursat	0.46	6.60	0.48	0.40	2.40	0.40
	Sangker	0.54	-21.50	0.57	0.38	-26.40	0.39
	Sen	0.86	-8.20	0.87	0.79	-5.20	0.79
	Siem Reap	0.61	-5.00	0.61	0.67	-9.70	0.68
	Sreng	0.70	-2.40	0.70	0.69	-11.50	0.70
	Staung	0.46	2.50	0.46	0.61	-5.80	0.62
	Average	**0.63**	**-6.28**	**0.64**	**0.57**	**-10.16**	**0.58**
Dataset	River name	Calibration			Validation		
		NSE	PBIAS (%)	R^2	NSE	PBIAS (%)	R^2
IMERG	Baribo	0.62	-1.20	0.62	0.47	-1.80	0.47
	Chikreng	0.35	-9.40	0.35	0.47	-16.50	0.48
	Chinit	0.88	-1.10	0.88	0.72	-9.00	0.73
	Dauntri	0.58	-12.80	0.59	0.49	-8.10	0.50
	Mongkol Borey	0.72	1.60	0.72	0.74	-12.20	0.76
	Pursat	0.48	-16.20	0.49	0.40	-2.00	0.40
	Sangker	0.64	-11.10	0.65	0.42	-24.90	0.44
	Sen	0.88	-3.00	0.89	0.80	-4.70	0.80
	Siem Reap	0.55	-0.10	0.56	0.68	-2.50	0.70
	Sreng	0.84	-7.60	0.85	0.68	-12.00	0.69
	Staung	0.66	-3.30	0.67	0.63	3.80	0.65
	Average	**0.65**	**-5.84**	**0.66**	**0.59**	**-8.17**	**0.60**

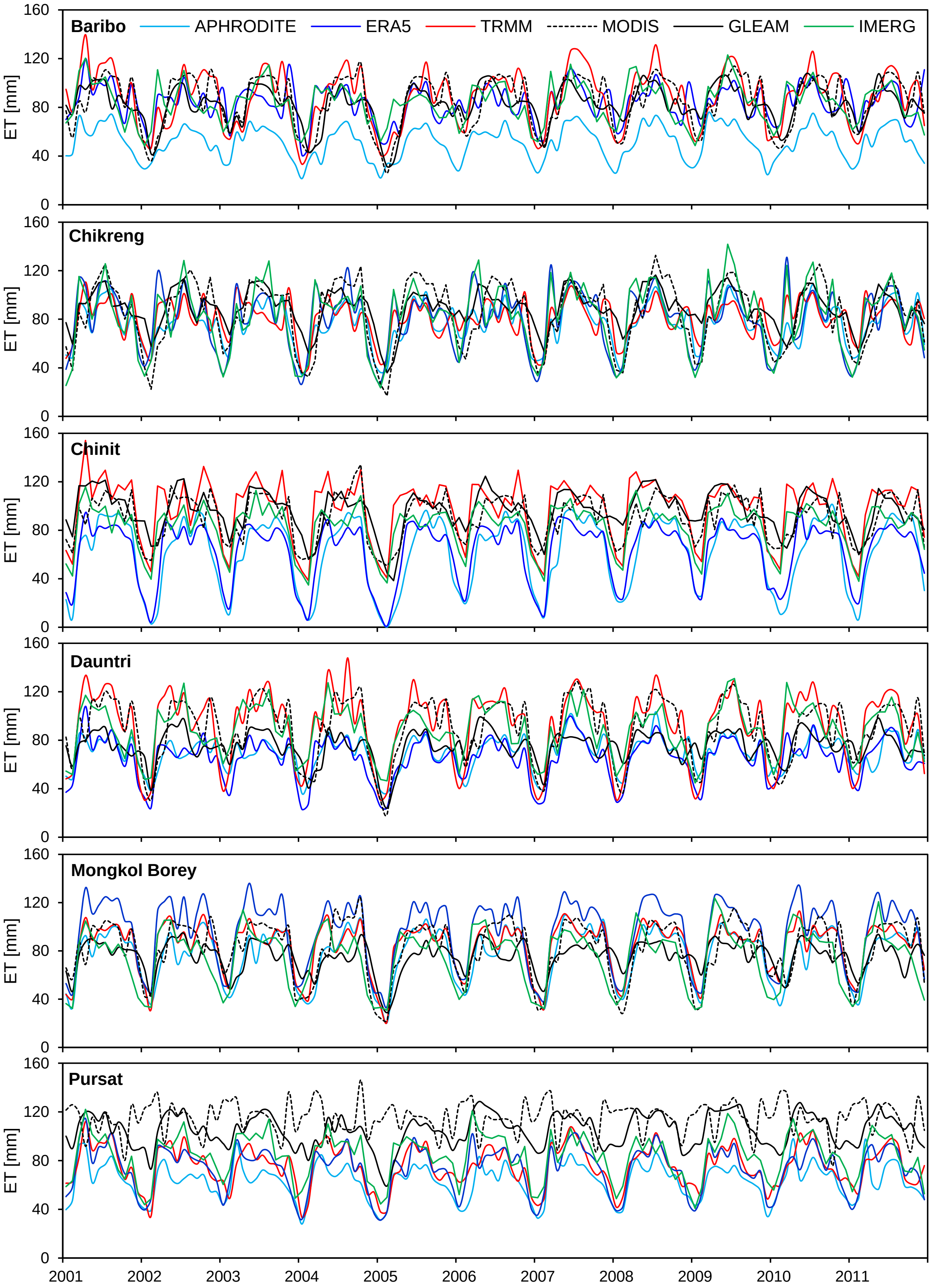


Figure A8. Comparison of tributary basin-averaged monthly evapotranspiration (ET) of the individual dataset from 2001 to 2011. The tributary basin-averaged monthly ET is the average monthly ET from all delineated sub-basins in each tributary basin excluding flooded areas.

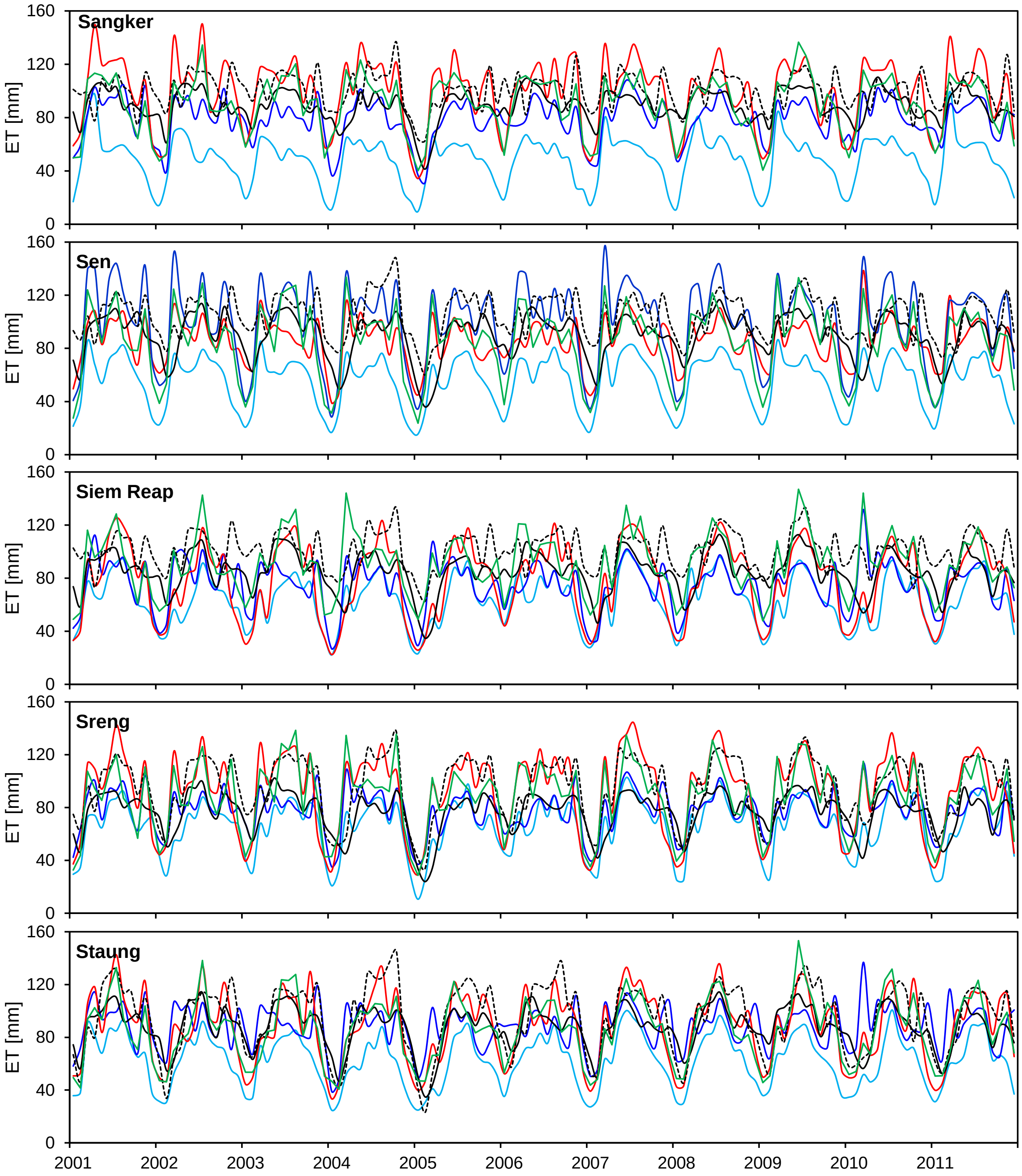


Figure A8. *Cont.*

Table A6. The RMSE (mm/month) between SWAT-based ET and satellite-based ET over each tributary basin excluding flooded areas.

River	APHRODITE		ERA5		TRMM		IMERG	
	GLEAM	MODIS	GLEAM	MODIS	GLEAM	MODIS	GLEAM	MODIS
Baribo	33.42	37.66	13.66	17.54	18.27	15.20	16.69	19.40
Chikreng	18.41	18.81	24.32	23.22	16.79	21.69	24.12	22.77
Chinit	40.08	34.01	36.44	31.37	19.94	20.46	19.88	17.26
Dauntri	13.09	26.51	17.17	30.46	29.71	19.71	21.29	19.68
Mongkol Borey	15.61	14.68	29.43	22.24	17.05	14.55	20.85	22.01
Pursat	43.19	55.89	33.74	47.54	31.50	45.14	24.66	38.63
Sangker	43.92	56.42	15.44	26.05	22.90	22.73	16.17	21.71
Sen	38.83	52.21	27.34	27.25	18.51	25.80	24.34	31.64
Siem Reap	27.09	41.32	20.17	31.42	22.24	32.09	20.33	22.26
Sreng	18.13	31.88	15.19	24.90	29.83	21.15	24.64	21.52
Staung	29.04	36.75	17.18	25.02	18.69	18.78	15.01	18.71
Average	**29.16**	**36.92**	**22.73**	**27.91**	**22.31**	**23.39**	**20.73**	**23.24**

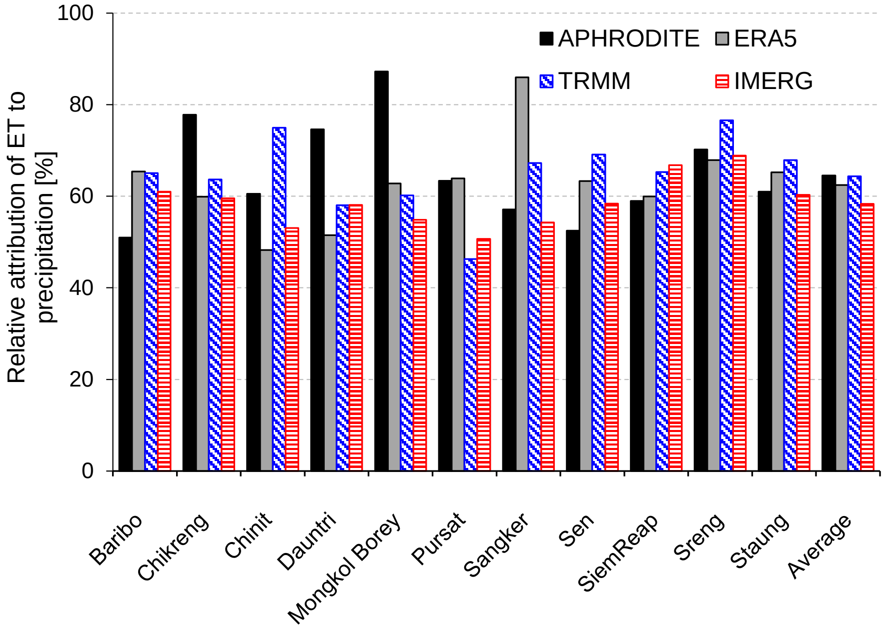


Figure A9. The relative attribution of evapotranspiration to the total precipitation for the individual product between 2001 and 2011 over the TSL Basin excluding flooded areas.

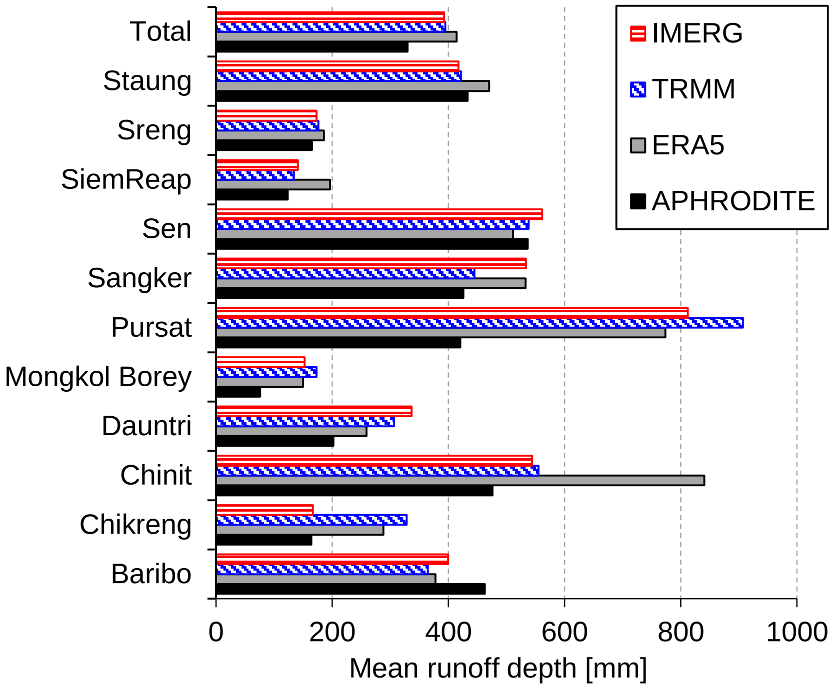


Figure A10. Comparison of simulated mean annual runoff depth between 2001 and 2004. The mean runoff depth represents the runoff from the whole drainage area of each tributary basin and the whole TSL Basin excluding the flooded areas.

Table A7. CUSLE and KUSLE values of different land use and soil types, respectively

Land use name (SWAT Database)	**SWAT code**	**CUSLE**
Evergreen, high cover density	EHCD	0.001
Evergreen, medium-low cover density	EMLD	0.001
Evergreen mosaic	EVMS	0.24
Mixed, high cover density	MEDH	0.008
Mixed, medium-low cover density	MEDM	0.008
Mixed mosaic	MXMS	0.24
Crop mosaic, cropping area <30	CMCS	0.02
Crop mosaic, cropping area >30	CMCL	0.15
Agriculture land-intensive	AGRI	0.24
Deciduous	DECD	0.048
Deciduous mosaic	DCMS	0.048
Regrowth	REGR	0.02
Regrowth, inundated	REGI	0.02
Wood- and shrubland, evergreen	WSEV	0.02
Grassland	GRAS	0.02
Wood- and shrubland, dry	WSDR	0.04
Urban or built-over area	URBN	0.15
Water	WATR	0
Lower Mekong Basin (LMB) paddy field	PDDY	0.03
LMB forest land	FRSL	0.001
LMB distributed forest land	DTFR	0.001
Soil name (SWAT Database)	**SWAT code**	**KUSLE**
Ferric acrisol	ACf/ACp	0.15
Gleyic acrisol	ACg	0.13
Gleyic acrisol / Dystric planosol	ACg/PLd	0.13
Areni-gley	ACga	0.13
Haplic acrisol	ACh	0.16
Haplic acrisol / Dystric planosol	ACh/LPd	0.16
Haplic acrisol-skeletic	ACh-C	0.16
Gleyic-plinthic	ACpg	0.27
Gleyic-plinthic-gley	ACpga	0.27
Gleyic	ARg	0.02
Dystric	CMd	0.15
Dystric cambisol / Dystric leptosol	CMd/LPd	0.15
Ferric cambisol / Ferric acrisol	CMo/ACf	0.15
Rhodic ferrasol	FRr	0.09
Eutric gleysol	GLe	0.18
Dystric leptosol	LPd	0.18
Eutric leptosol	LPe	0.18
Dystric planosol / Gleyic acrisol	PLd/ACg	0.15
Dystric plintosol	PTd	0.2
Rock out crop	R	0.15
Slope complex	SC	0.16

Box A1. The self-organizing map

A sequential regression process carries out the training process of the SOM (Wallner et al., 2013):

1. Set j = 1 and select the input vector x_j.
2. Identify the node with the reference vector m_i, most similar to the current input vector x_j. This neuron is called the winner neuron or best matching unit (BMU):

$$c = \arg\min_{i} \left\{ \left| x_i - m_i(t) \right| \right\} \qquad i = 1 \ldots K, \tag{A3}$$

where, the index i ranges over all reference vectors on the map. The quantity $m_i(t)$ refers to the reference vector at position i on the map at time step t.

3. Update the reference vector of the winner neuron m_c and the neighboring neurons m_i based on the current input vector x_j:

$$m_i(t+1) = m_i(t) + \alpha(t) \cdot h_{ci}(t) \cdot \left[x_j - m_i(t) \right] \qquad i = 1 \ldots K, \tag{A4}$$

where h_{ci} is the neighborhood function, which decreases with increasing distance to the winner neuron, and α is the learning rate, which decreases with increasing time.

4. Set $j = j + 1$ and go back to step 2 until all input vectors are considered.
5. Set $t = t + 1$ and go back to step 1 until a fixed number of iterations is reached.

The above computation is usually repeated over the available input vectors many times during the training phase of the map. Each iteration is called a training epoch.

One of the advantages of the SOM is that they have an appealing visual representation as the 2-D unified distance matrix, as seen in Figure 3.2b. As before, the colors on the map represent the relative distances between reference vectors: light colors indicate short distances, and dark colors indicate long distances. Contiguous areas of light colors represent strong clusters (Hamel and Brown, 2011).

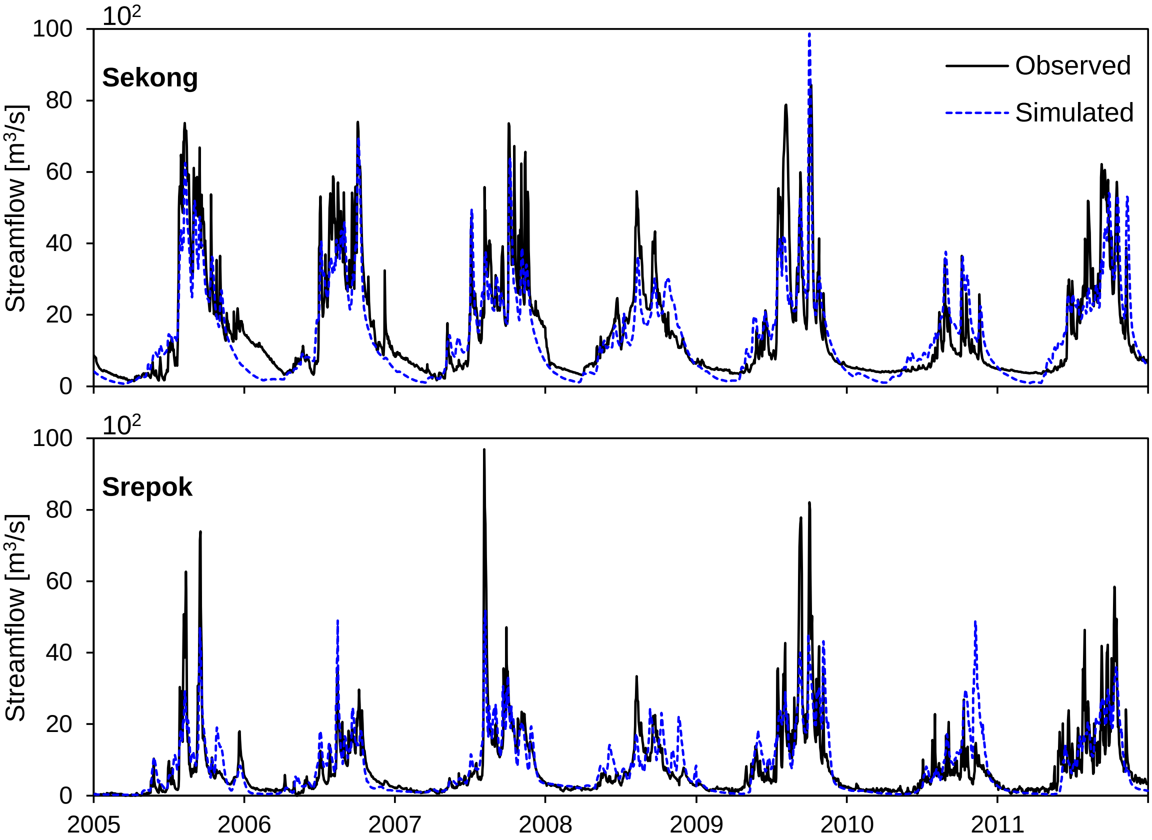


Figure A11. Comparison of daily observed (solid black) and simulated (dash blue) flows in Sekong and Srepok catchments. The calibration and validation periods were from 2005 to 2008 and 2009 to 2011, respectively.

Table A8. Results of statistical indices obtained during the model calibration and validation for the flow at the gauging stations of Sekong and Srepok catchments.

River	Calibration			Validation		
	R^2	NSE	PBIAS	R^2	NSE	PBIAS
Sekong	0.79	0.77	13.50	0.72	0.72	-5.20
Srepok	0.71	0.70	-3.00	0.58	0.57	-3.20

Table A9. The values of flow-calibrated parameters of each tributary basin. These fitted values were obtained from model optimization in Chapter 2.

Parameter Name	Calibrated Value					
	Baribo	Chikreng	Chinit	Dauntri	Mongkol Borey	Pursat
r__CN2.mgt	0.036	0.197	0.200	-0.298	-0.462	0.099
v__ALPHA_BF.gw	0.721	0.616	0.219	0.044	0.232	0.372
v__GW_DELAY.gw	20.343	148.356	164.770	19.592	11.932	0.999
v__GWQMN.gw	179.155	4588.491	2611.880	4646.876	1789.026	1026.557
v__GW_REVAP.gw	0.178	0.101	0.198	0.187	0.193	0.171
v__REVAPMN.gw	20.218	185.689	136.673	40.991	115.336	206.535
v__RCHRG_DP.gw	0.292	0.216	0.013	0.206	0.082	0.117
v__LAT_TTIME.hru		20.137		158.997		40.287
v__SLSOIL.hru		127.245		105.617		0.914
v__CANMX.hru	14.500	15.715	46.362	14.297	92.099	13.718
v__ESCO.hru	0.489	0.748	0.325	0.133	0.325	0.986
v__EPCO.hru	0.554	0.856	0.937	0.823	0.542	0.286
v__OV_N.hru		26.998		4.984		27.784
v__CH_N2.rte	0.034	0.042	0.217	0.197	0.073	0.195
v__CH_K2.rte	365.609	83.757	188.763	131.570	418.183	202.901
v__ALPHA_BNK.rte		0.518		0.303		0.916
v__CH_K1.sub		4.864		61.288		24.970
v__CH_N1.sub		9.196		25.991		10.577
r__SOL_AWC().sol	0.178	0.326	0.410	-0.277	0.257	0.159
r__SOL_BD().sol		-0.056		-0.066		0.183
r__SOL_K().sol	-0.183	-0.178	-0.150	0.331	-0.085	-0.185
r__SOL_ALB().sol	0.187	-0.294	-0.310	0.029	0.050	-0.188
v__SURLAG.bsn	15.220	5.079	4.158	10.453	17.066	13.036

Table A9. *Cont.*

Parameter Name	Calibrated Value				
	Sangker	Sen	Siem Reap	Sreng	Staung
r__CN2.mgt	-0.366	0.092	-0.271	0.035	0.189
v__ALPHA_BF.gw	0.741	0.507	0.066	0.223	0.410
v__GW_DELAY.gw	30.005	11.926	194.516	360.340	122.564
v__GWQMN.gw	4959.035	1434.173	4470.288	4392.917	3999.254
v__GW_REVAP.gw	0.199	0.189	0.054	0.151	0.156
v__REVAPMN.gw	92.278	386.447	474.308	465.885	109.103
v__RCHRG_DP.gw	0.294	0.266	0.029	0.004	0.003
v__LAT_TTIME.hru	87.238	109.855	14.474		
v__SLSOIL.hru	116.617	62.217	22.592		
v__CANMX.hru	71.645	96.416	70.088	82.399	79.324
v__ESCO.hru	0.217	0.087	0.254	0.492	0.474
v__EPCO.hru	0.723		0.803	0.728	0.309
v__OV_N.hru	1.194		28.250		
v__CH_N2.rte	0.025	0.169	0.149	0.297	0.192
v__CH_K2.rte	474.281	149.830	178.124	267.889	490.836
v__ALPHA_BNK.rte	0.589		0.959		
v__CH_K1.sub	0.104	0.020	51.020		
v__CH_N1.sub	17.455	25.484	8.185		
r__SOL_AWC().sol	0.094	0.025	0.411	-0.028	0.412
r__SOL_BD().sol	0.000		-0.027		
r__SOL_K().sol	-0.329		-0.066	0.085	0.229
r__SOL_ALB().sol	0.513		0.389	0.037	-0.001
v__SURLAG.bsn	12.285		19.563	11.729	9.404

Table A10. The selected flow and sediment parameters and their calibrated values of the Chinit catchments. These fitted values were obtained from model optimization from 2005 to 2008.

Type	Nº	Parameter name	Description	Fitted value	Method*
Flow and sediment	1	r__CN2.mgt	SCS runoff curve number (f)	0.127	Relative (1)
	2	v__ALPHA_BF.gw	Baseflow alpha factor (days)	0.114	Replace (2)
	3	v__GW_DELAY.gw	Groundwater delay (days)	340.113	-
	4	v__GWQMN.gw	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)	4628.748	-
	5	v__GW_REVAP.gw	Groundwater "revap" coefficient	0.188	-
	6	v__REVAPMN.gw	Threshold depth of water in the shallow aquifer required for "revap" to occur (mm)	338.286	-
	7	v__RCHRG_DP.gw	Deep aquifer percolation fraction	0.049	-
	8	v__CANMX.hru	Maximum canopy storage	-0.378	-
	9	v__ESCO.hru	Soil evaporation compensation factor	0.852	-
	10	v__EPCO.hru	Plant uptake compensation factor	0.623	-
	11	v__CH_N2.rte	Manning's "n" value for the main channel	0.249	-
	12	v__CH_K2.rte	Effective hydraulic conductivity in the main channel alluvium	213.555	-
	13	r__SOL_AWC().sol	Available water capacity of the soil layer	-0.231	Relative (1)
	14	r__SOL_K().sol	Saturated hydraulic conductivity	-0.230	-
	15	r__SOL_ALB().sol	Moist soil albedo	-0.022	-
	16	v__SURLAG.bsn	Surface runoff lag time	15.279	Replace (2)
Sediment	17	v__LAT_SED.hru	Sediment concentration in lateral flow	2157.10	Replace (2)
	18	v__CH_COV2.rte	Channel cover factor	1.00	-
	19	r__USLE_K().sol	Soil erodibility factor	-0.18	Relative (1)

(1): multiplying the initial parameter by value in percentage; (2): replacing the initial parameter by value. "-" refers to the method of the corresponding parameter being the same as the above parameter.

Table A11. Same as Table A10 but for the Sen catchment.

Type	Nº	Parameter name	Description	Fitted value	Method*
Flow and sediment	1	r__CN2.mgt	SCS runoff curve number (f)	0.217	Relative (1)
	2	v__ALPHA_BF.gw	Baseflow alpha factor (days)	0.372	Replace (2)
	3	v__GWQMN.gw	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)	2871.432	
	4	v__GW_DELAY.gw	Groundwater delay (days)	0.264	-
	5	v__GW_REVAP.gw	Groundwater "revap" coefficient	0.189	-
	6	v__REVAPMN.gw	Threshold depth of water in the shallow aquifer required for "revap" to occur (mm)	288.358	-
	7	v__RCHRG_DP.gw	Deep aquifer percolation fraction	0.072	-
	8	v__LAT_TTIME.hru	Lateral flow travel time	138.220	-
	9	v__SLSOIL.hru	Slope length for lateral subsurface flow	78.399	-
	10	v__CANMX.hru	Maximum canopy storage	68.347	-
	11	v__ESCO.hru	Soil evaporation compensation factor	0.476	-
	12	v__CH_N2.rte	Manning's "n" value for the main channel	0.129	-
	13	v__CH_K2.rte	Effective hydraulic conductivity in the main channel alluvium	439.889	-
	14	v__CH_K1.sub	Effective hydraulic conductivity in tributary channel alluvium	3.159	-
	15	v__CH_N1.sub	Manning's "n" value in tributary channel	25.906	-
	16	r__SOL_AWC().sol	Available water capacity of the soil layer	-0.072	Relative (1)
Sediment	17	v__LAT_SED.hru	Sediment concentration in lateral flow	3178.58	Replace (2)
	18	v__CH_COV2.rte	Channel cover factor	0.99	-
	19	r__USLE_K().sol	Soil erodibility factor	-0.23	Relative (1)

Table A12. Same as Table A10 but for the Sekong catchment.

Type	Nº	Parameter name	Description	Fitted value	Method*
Flow and sediment	1	r__CN2.mgt	SCS runoff curve number (f)	0.174	Relative (1)
	2	v__ALPHA_BF.gw	Baseflow alpha factor (days)	0.835	Replace (2)
	3	v__GW_DELAY.gw	Groundwater delay (days)	33.865	-
	4	v__GWQMN.gw	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)	127.561	-
	5	v__GW_REVAP.gw	Groundwater "revap" coefficient	0.027	-
	6	v__REVAPMN.gw	Threshold depth of water in the shallow aquifer required for "revap" to occur (mm)	74.678	-
	7	v__CANMX.hru	Maximum canopy storage	5.834	-
	8	v__ESCO.hru	Soil evaporation compensation factor	0.998	-
	9	v__EPCO.hru	Plant uptake compensation factor	0.285	-
	10	v__OV_N.hru	Manning's "n" value for overland flow	12.497	-
	11	v__CH_N2.rte	Manning's "n" value for the main channel	0.167	-
	12	v__CH_K2.rte	Effective hydraulic conductivity in the main channel alluvium	484.442	-
	13	r__SOL_AWC().sol	Available water capacity of the soil layer	-0.187	Relative (1)
	14	r__SOL_K().sol	Saturated hydraulic conductivity	-0.199	-
	15	r__SOL_ALB().sol	Moist soil albedo	0.365	-
Sediment	16	v__LAT_SED.hru	Sediment concentration in lateral flow	3692.74	Replace (2)
	17	v__CH_COV2.rte	Channel cover factor	0.99	-
	18	r__USLE_K().sol	Soil erodibility factor	0.246	Relative (1)

Table A13. Same as Table A10 but for the Srepok catchment.

Type	N°	Parameter name	Description	Fitted value	Method*
Flow and sediment	1	r__CN2.mgt	SCS runoff curve number (f)	0.168	Relative (1)
	2	v__ALPHA_BF.gw	Baseflow alpha factor (days)	0.501	Replace (2)
	3	v__GW_DELAY.gw	Groundwater delay (days)	218.852	-
	4	v__GWQMN.gw	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)	4057.173	-
	5	v__GW_REVAP.gw	Groundwater "revap" coefficient	0.117	-
	6	v__REVAPMN.gw	Threshold depth of water in the shallow aquifer required for "revap" to occur (mm)	54.452	-
	7	v__CANMX.hru	Maximum canopy storage	63.618	-
	8	v__ESCO.hru	Soil evaporation compensation factor	0.028	-
	9	v__EPCO.hru	Plant uptake compensation factor	0.707	-
	10	v__OV_N.hru	Manning's "n" value for overland flow	28.725	-
	11	v__CH_N2.rte	Manning's "n" value for the main channel	0.035	-
	12	v__CH_K2.rte	Effective hydraulic conductivity in the main channel alluvium	496.553	-
	13	r__SOL_AWC().sol	Available water capacity of the soil layer	-0.058	Relative (1)
	14	r__SOL_K().sol	Saturated hydraulic conductivity	-0.177	-
	15	r__SOL_ALB().sol	Moist soil albedo	-0.046	-
Sediment	16	v__LAT_SED.hru	Sediment concentration in lateral flow	2327.12	Replace (2)
	17	v__CH_COV2.rte	Channel cover factor	0.80	-
	18	r__USLE_K().sol	Soil erodibility factor	0.03	Relative (1)

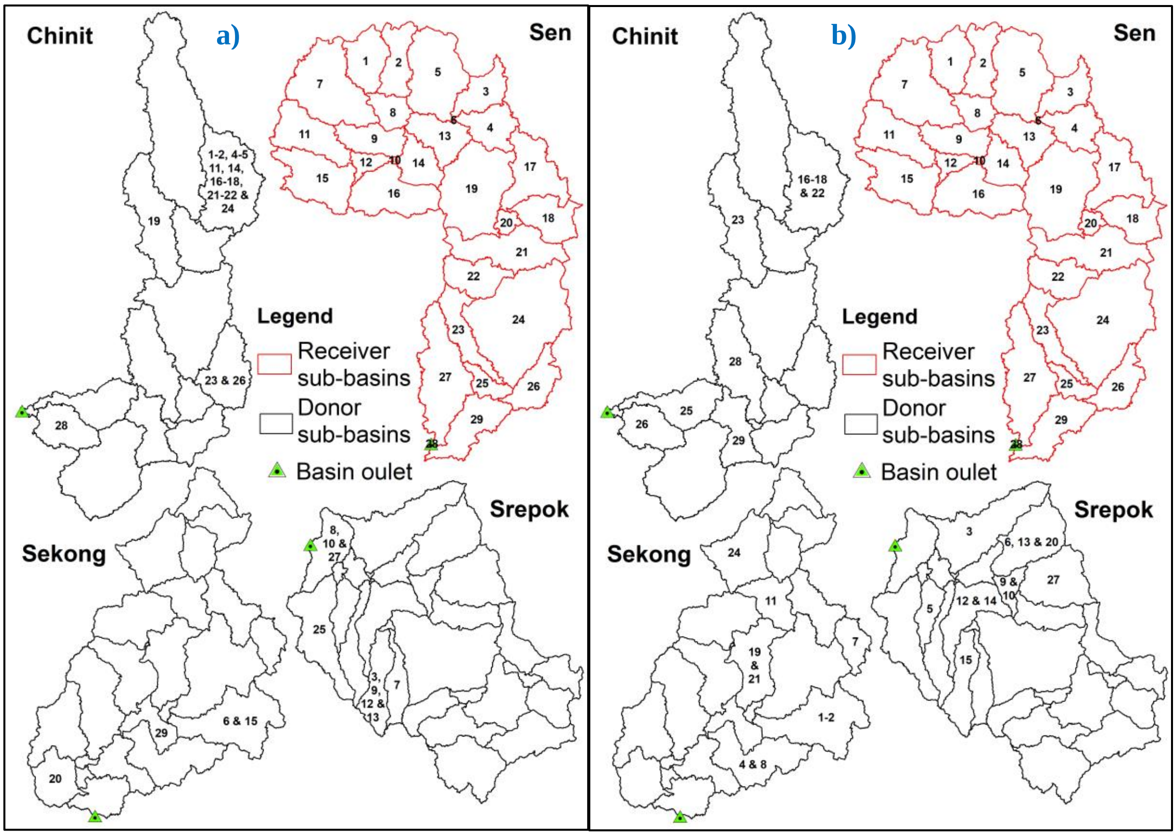


Figure A12. a) Physical similarity donor and receiver sub-catchments for Sen catchment as a pseudo ungauged catchment and b) Sediment-response similarity donor and receiver sub-catchments for Sen catchment as a pseudo ungauged catchment. The number shown on the map denotes the corresponding donor and receiver sub-catchments.

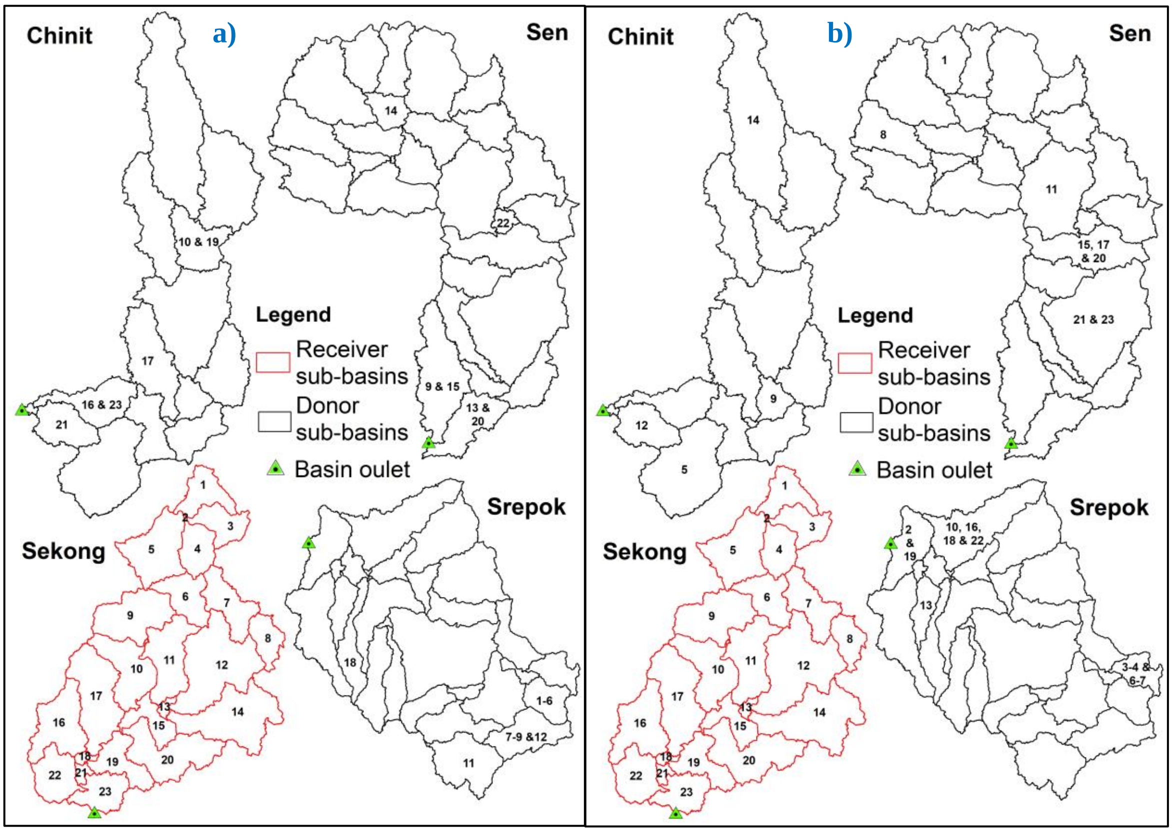


Figure A13. Same as Figure A12 but for the Sekong catchment.

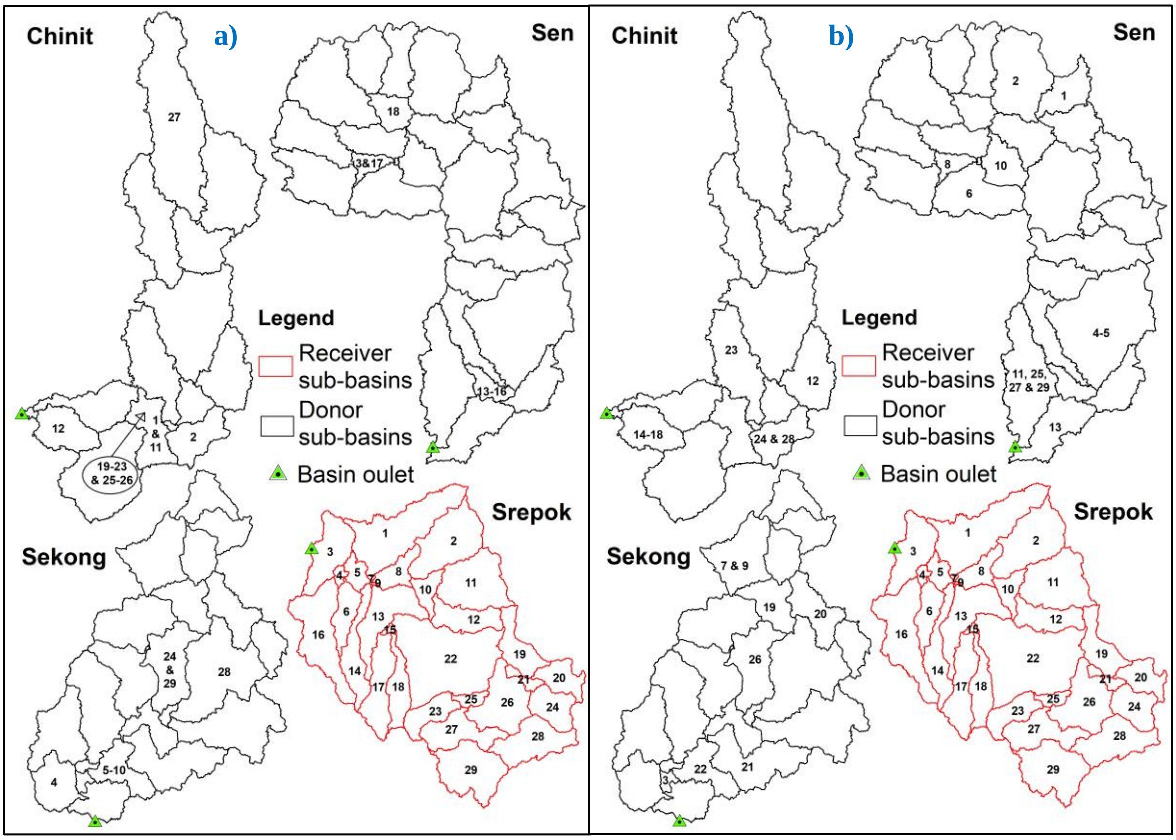


Figure A14. Same as Figure A12 but for the Srepok catchment.

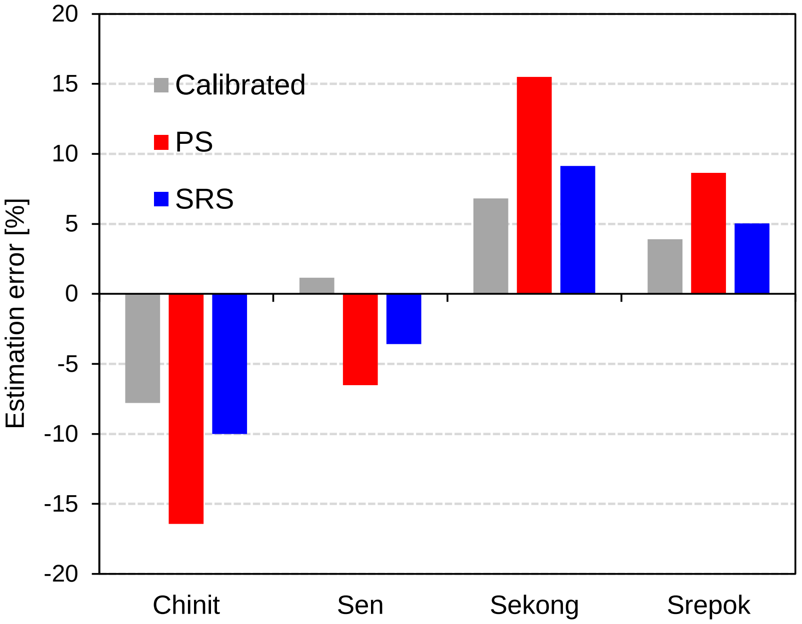


Figure A15. The estimated error of sediment simulation compared with observed data at the gauging station of each catchment.

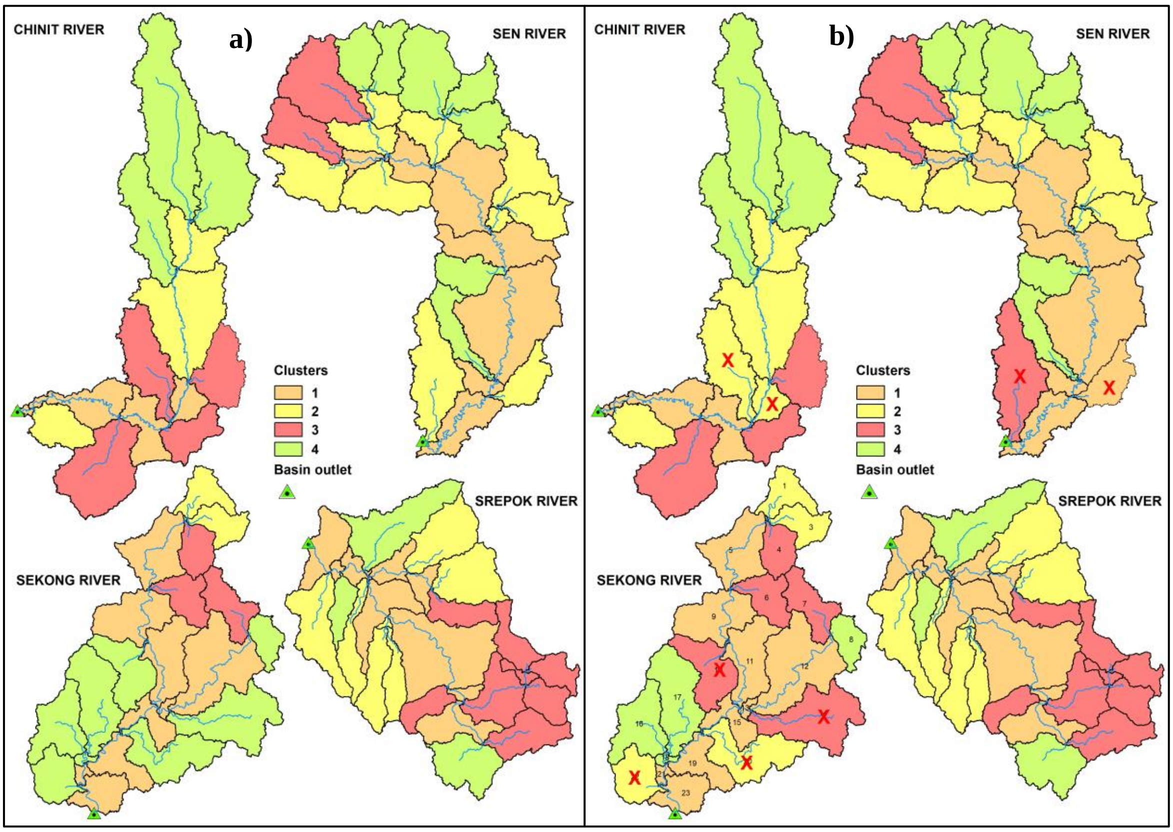


Figure A16. a) Sub-catchment clusters based on rainfall and sediment response, using the tentative estimate of model parameters by the arithmetic mean method in the first phase of the SRS procedure. b) Sub-catchment clusters based on rainfall and sediment response after updating the model parameters in the second phase of the SRS procedure. See phases 1 and 2 of the SRS procedure in Figure 3.1. In b), the red cross-marks shown on the clustering maps denote a different cluster number of corresponding sub-catchments compared with that in a).

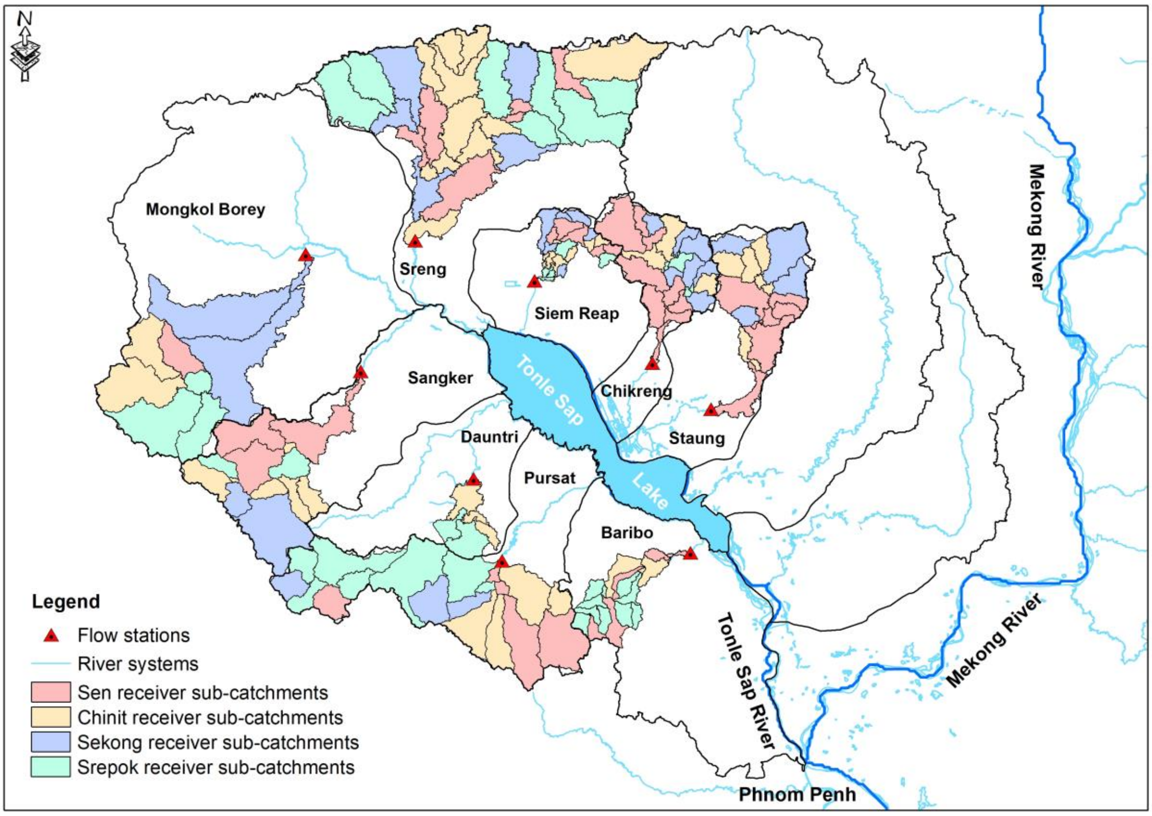


Figure A17. Receiver sub-catchments of the actual ungauged catchment of the TSL basin. Refer to Figure 3.3 for the location of Chinit, Sen, Sekong and Srepok catchments.

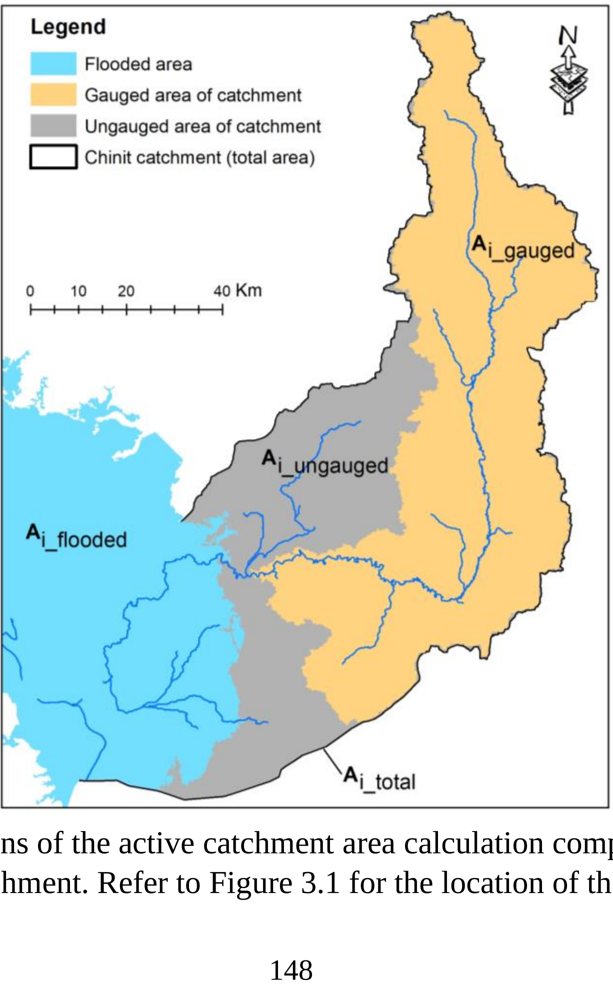


Figure A18. Definitions of the active catchment area calculation components and flooded area in the Chinit catchment. Refer to Figure 3.1 for the location of the Chinit catchment.

Table A14. Equations for calculating sediment load in the ungauged catchments of the TSL Basin

Nº	Accounting catchment area	Equation*
1	Total area	$SED_{i\_total} = (A_{i\_total} / A_{i\_gauged}) * SED_{i\_gauged}$
2	Total area – flooded area	$SED_{i\_TRIB} = [(A_{i\_total} - A_{i\_flooded}) / A_{i\_total}] * SED_{i\_total}$

* SED_{i_total} is the total sediment load of tributary *i* [metric tons], A_{i_total} is the total catchment area of tributary *i* [km^2], A_{i_gauged} is gauged catchment area of tributary *i* [km^2], SED_{i_gauged} is sediment load of the gauged catchment area of tributary *i* [metric tons], SED_{i_TRIB} is final sediment load of tributary *i* [metric tons] and A_{i_flood} is the flooded area of tributary *i* [km^2]

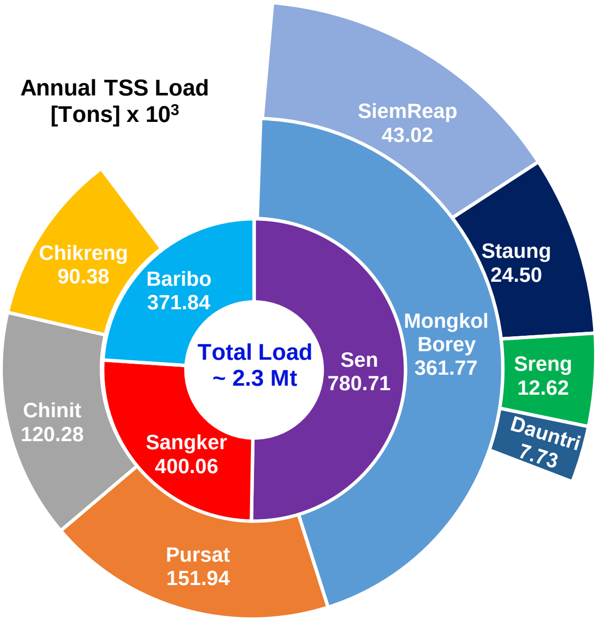


Figure A19. Long-term mean annual sediment load of each catchment excluding the flooded areas from 2001–2011.

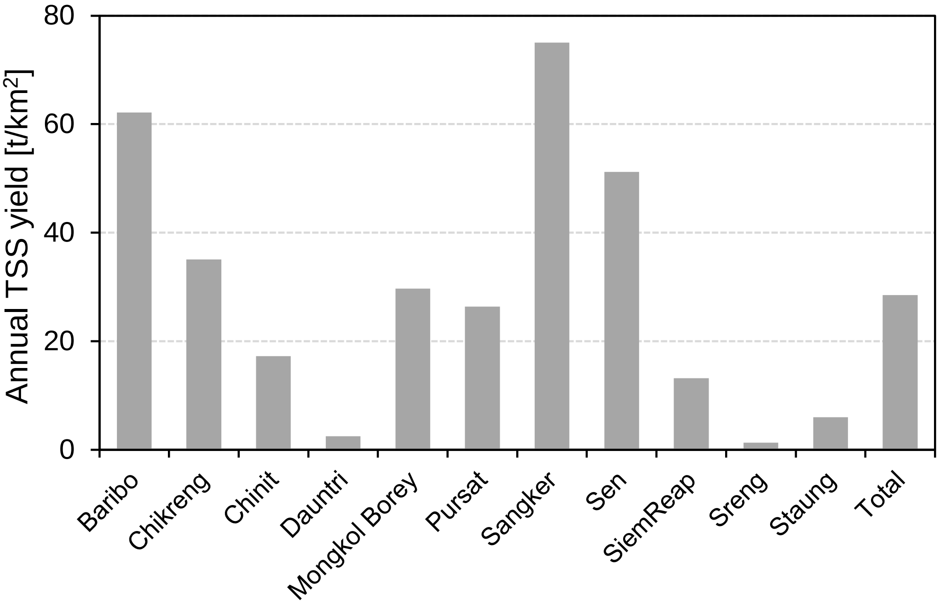


Figure A20. Long-term mean annual sediment yield of each catchment of the TSL basin excluding the flooded areas from 2001–2011.

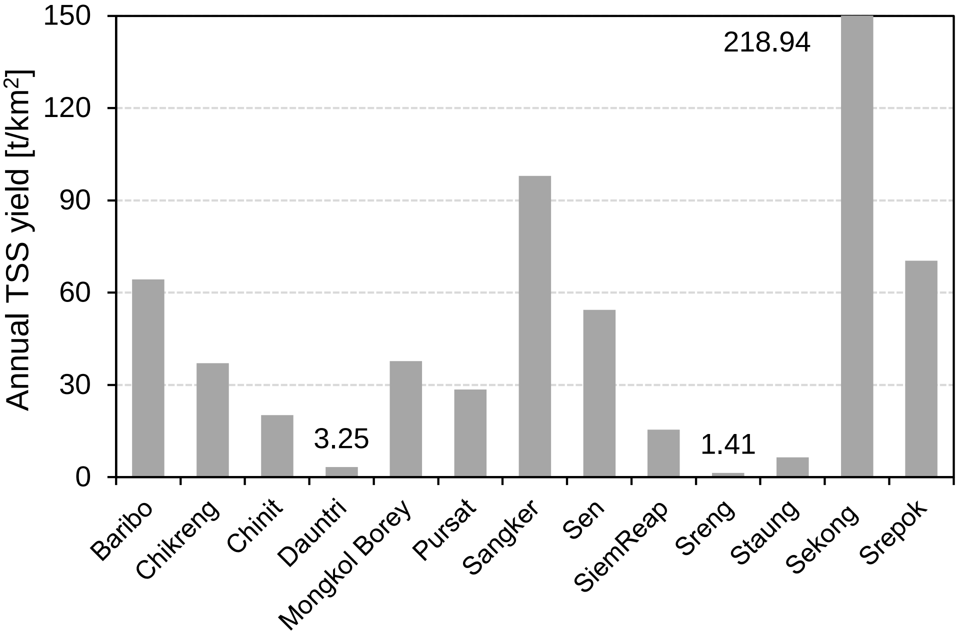


Figure A21. Comparison of long-term mean annual sediment yield of Sekong and Srepok catchments with all catchments of the TSL basin. The annual sediment yield was calculated from the gauged area of each catchment from 2001–2011.

Table A15. The performance of statistical indices in the calibration and validation periods for study in chapter 4.

Tributaries	**Calibration**			**Validation**		
	R^2	NSE	PBIAS [%]	R^2	NSE	PBIAS [%]
Baribo	0.56	0.54	-4.50	0.50	0.35	-22.50
Chikreng	0.41	0.40	-11.10	0.59	0.24	48.30
Chinit	0.75	0.74	-1.80	0.67	0.65	10.80
Dauntri	0.50	0.47	-32.10	0.17	0.15	-23.20
Mongkol Borey	0.70	0.70	1.70	0.63	0.63	-0.90
Pursat	0.31	0.31	-1.20	0.30	0.29	-5.40
Sangker	0.42	0.40	-19.90	0.53	0.46	-37.00
Sen	0.85	0.85	-8.90	0.79	0.79	2.50
Siem Reap	0.65	0.65	-6.40	0.65	0.63	-10.10
Sreng	0.84	0.83	-6.60	0.57	0.55	22.60
Staung	0.66	0.65	-6.10	0.65	0.64	15.00
Average	***0.60***	***0.59***	***-8.81***	***0.55***	***0.49***	***0.01***

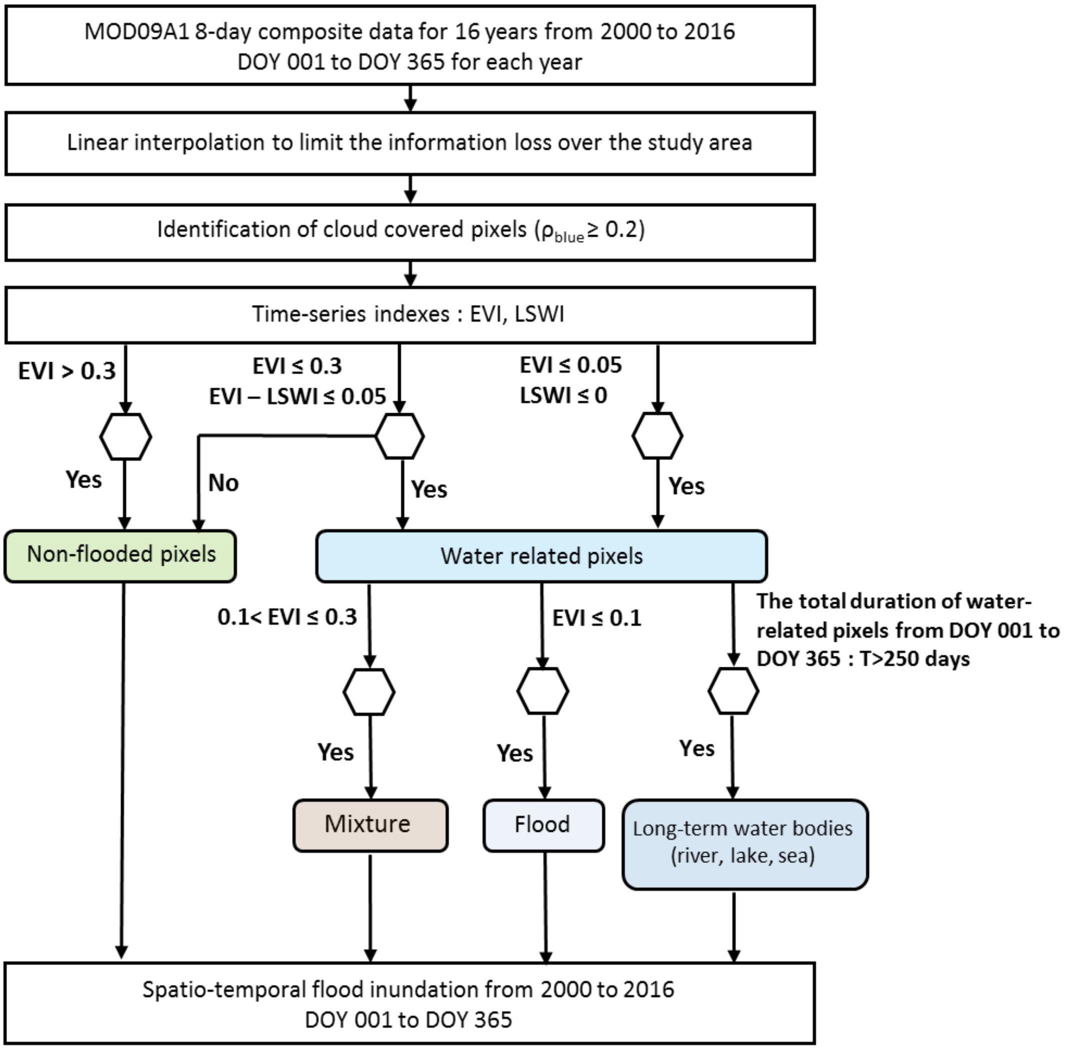


Figure A22. Processing steps of the MODIS images for estimating inundated areas in the TSL Basin adopted from Frappart et al. (2018).

Table A16. Summaries of natural forest cover (i.e., i.e., forest, evergreen broadleaf and mixed forest) in each tributary basin from 1995 to 2018. Area [%] and Change [%] refer to the percentage area shared with the total area of the tributary basin and the relative change compared to the 1995 land cover area, respectively.

Tributary basins	Basin area [km^2]	1995		2002		2010		2018	
		Area [km^2]	Area [%]*	Area [km^2]	Change [%]**	Area [km^2]	Change [%]	Area [km^2]	Change [%]
Baribo	7138.81	1255.14	17.58	839.21	-5.83	777.09	-6.70	729.85	-7.36
Chikreng	2713.15	1194.19	44.01	882.49	-11.49	852.22	-12.60	792.19	-14.82
Chinit	8217.39	3049.42	37.11	963.58	-25.38	588.82	-29.94	531.52	-30.64
Dauntri	3682.20	724.39	19.67	595.74	-3.49	580.54	-3.91	525.18	-5.41
Mongkol Borey	14938.19	2483.88	16.63	1294.75	-7.96	1255.31	-8.22	1165.14	-8.83
Pursat	5945.43	3598.59	60.53	3059.40	-9.07	3019.14	-9.75	2923.11	-11.36
Sangker	6035.80	1506.52	24.96	1217.46	-4.79	1202.71	-5.03	1101.66	-6.71
Sen	16344.33	10282.30	62.91	7282.30	-18.35	7027.76	-19.91	6878.89	-20.82
Siem Reap	3618.88	383.96	10.61	259.18	-3.45	254.80	-3.57	230.75	-4.23
Sreng	9966.05	4217.41	42.32	2668.53	-15.54	2512.53	-17.11	2148.30	-20.76
Staung	4355.05	1279.75	29.39	974.37	-7.01	965.83	-7.21	971.99	-7.07

Note: Flooded forest cover is not included since the flooded area is excluded from SWAT modeling.