River Discharge Estimation Using Multisource Remote Sensing: Application in the Lancang Mekong River Basin

by

xxxxxxxxxx

A Thesis Progress Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Water Engineering and Management

Examination Committee: Prof. xxxxxxx (Chairperson)
Dr. xxxxxxxxx (Member)
Dr. xxxxxxxx (Member)
Dr. xxxxxxxx (Member)

Nationality: xxxxxxxxxx
Previous Degree: Bachelor of Engineering in Agricultural Engineering
Institute of Engineering
xxxxxxxxxxxx

Scholarship Donor: ADB-JSP Scholarship

Asian Institute of Technology
School of Engineering and Technology
Thailand

March 2023

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
AMR	Advanced Microwave Radiometer
AMSR-E	Advanced Microwave Scanning Radiometer for the Earth Observing System
ANN	Artificial Neural Network
CNES	Centre National d'Etudes Spatiales
CTOH	Center for Topographic Studies of the Ocean and Hydrosphere
DEM	Digital Elevation Model
DORIS	Doppler Orbitography and Radio-positioning Integrated by Satellite
ENVISAT	Environmental Satellite
EOS	Earth Observing System
ERS	European Remote Sensing satellite
ESA	European Space Agency
EUMETSAT	Exploitation of Meteorological Satellites
GDR	Geophysical Data Record
GEE	Google Earth Engine
GEOS-3	Geodynamics Experimental Ocean Satellite-3
GFO-RA	GEOSat Follow-On Radar Altimeter
GHz	Gigahertz
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GRDC	Global Runoff Data Centre
HY-2A	Hai Yang 2A
JPL	Jet Propulsion Laboratory
km	Kilometer
Ku band	Kurtz-under band
LRA	Laser Retroflector Array
LRM	Low Resolution Mode
LSTM	Long Short-Term Memory
MLP-ANN	Multilayer perceptron Artificial Neural Network
mm	millimeter
MODIS	Moderate Resolution Imaging Spectroradiometer

MRC	Mekong River Commission
MWR	Microwave Radiometer
NASA	National Aeronautics and Space Administration
NIR	Near Infrared
NOAA	National Oceanic and Atmospheric Administration
NSE	Nash Sutcliffe efficiency
OSTM	Ocean Surface Topography Mission
PBIAS	percent bias
PLRM	Pseudo Low Resolution Mode
POD	Probability of Detection
R^2	coefficient of determination
RMSE	root mean square error
RNN	Recurrent Neural Network
RS	Remote Sensing
SAR	Synthetic Aperture Radar
SARAL	Satellite with ARgos and ALtiKa
SDR	Stage Discharge Rating
SRAL	SAR Radar Altimeter
SWOT	Surface Water and Ocean Topography
TOPEX	The Ocean Topography Experiment
US	United States
VS	Virtual Station

CHAPTER 1
INTRODUCTION

1.1 Background of the Study

Water is nature's most valuable resource, and due to rising demand, fresh water supply is in risk of becoming non-renewable (Kim et al., 2008). Water covers over 71 percent of the Earth's surface. However, only 3% of this is fresh water, with the remaining 2.5% trapped in ice and glaciers. As a result, individuals must rely on 0.5% fresh water for all of their needs (Kashid & Pardeshi, 2014). Inland surface water bodies, such as lakes, reservoirs, creeks, streams, and rivers, are essential to our daily lives. Inland surface water sources provided the majority of the water needed for thermoelectric generation, public supply, agriculture, mining, and industrial purposes (Kenny et al., 2009).

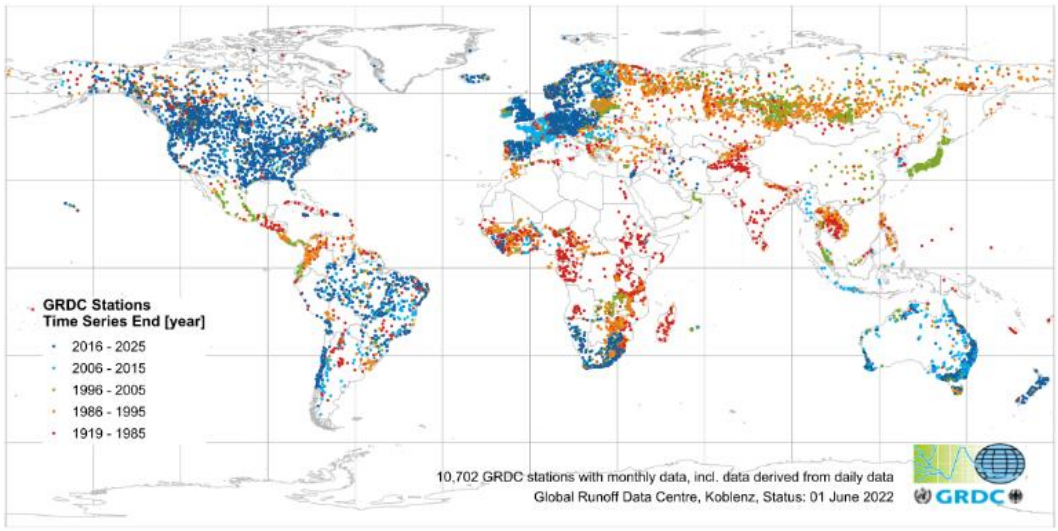
The geographical and temporal patterns of water resource distribution are essential knowledge for water resource management and assessments of water vulnerability. River discharge is an important variable in hydrological studies and applications such as global change monitoring, flood risk assessments, water supply management, dam design, irrigation projects and many more (Van Dijk et al., 2016). River discharge monitoring is also important for establishing a better knowledge of changes in the water cycle and hydrological processes at river-basin scales and globally because of natural disasters and anthropogenic activities (Robert Brakenridge et al., 2012). The level of water bodies is a major concern when it comes to managing and monitoring inland water resources.

Mostly, estimation of river discharge has relied on network of gauging stations measuring water levels which capture data on water level at precise time intervals (Shi et al., 2020). In-situ networks of gauges that record water level at fixed points in rivers and lakes limits number of measurements of inland water bodies. The spatial distribution of gauge stations, on the other hand, is significantly uneven. In-situ gauge network implementation varies by location accessibility, and data availability is determined by national policies. It is also difficult and expensive to set up and maintain these networks in remote areas. Furthermore, the precision of readings is largely reliant on the processing method used and the current state of water bodies. During flood seasons and other extreme events, gauge stations are particularly vulnerable

(Biancamaria et al., 2010). Figure 1-1 depicts the global distribution of hydrologic gauges and their insufficient monitoring capability.

Figure 1-1

Global Distribution of Hydrologic Gauges (*GRDC Data Portal*)



Due to a lack of reliable measurements and insufficient in-situ gauge stations, an accessible and repeatable approach is a must for monitoring river discharge variations. Estimating river stage and discharge based on satellite-measured water level and inundation area is a reasonably easy and alternative method to overcome this issue (Huang et al., 2018). Broadly, use of remote sensing in discharge estimation are grouped into three categories: first category also called rating curve method uses stage-discharge relationship developed from remotely sensed water level or inundated areas with simultaneously collected in-situ data for discharge estimation, second category uses river water level or width from remote sensing data as input parameter or reference for calibration in hydrological or hydraulic models and third category also called remote sensing driven method uses remote sensing data as major input for estimation of river discharge (Shi et al., 2020).

Various studies have been carried out demonstrating applicability of satellite altimetry water level to estimate river discharge (Birkinshaw et al., 2010; Bogning et al., 2018; Calmant & Seyler, 2006; Jiang et al., 2017). However, the temporal resolution of water level and discharge is limited to 10 days (Topex/Poseidon, Jason 1/2/3) or 35 days (EnviSat, SARAL/ALTIKA) due to repeat period of satellite orbit (Tarpanelli, Barbetta,

et al., 2013). Moreover, use of optical sensors and passive microwave sensors in discharge estimation cannot be discarded since they have less revisit time almost daily and large coverage area. The reflectance ratio of land and water body (C/M) from Near Infrared (NIR) band of Moderate Resolution Imaging Spectroradiometer (MODIS), Landsat, Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and Global Flood Detection System (GFDS) has been successfully used as proxy for river discharge estimation (Brakenridge et al., 2005; Hou et al., 2018; Tarpanelli et al., 2019). Cloudy sky hinders the usability of optical sensors data which may result in data outages in measurements and the possible way to overcome is to use multi-mission satellite products to increase sample size (Tarpanelli et al., 2019).

Thus, optical sensor data can be merged with altimetry data to catch information affected by cloudy sky for optical sensors and to improve temporal sampling. Artificial Neuron Networks (ANNs) are used largely in remote sensing applications since they offer an easy but effective possibility of merging input data from different sources into the same retrieval algorithm.

1.2 Statement of the Problem

Water supply planning and management, reservoir management and control, hydropower generation, flood prediction and control, understanding the global water cycle, and other hydrological applications all require accurate measurements of river discharge. To get an accurate measurement of discharge, most rivers require a series of flow velocity measurements in each sector of the cross section. Also, river discharge is frequently obtained from river stage measurements using the stage-discharge rating (SDR) curve to save expense, effort, and time. At several points along the river, in-situ gauge station would offer continuous and dependable river stage measurements (Liu et al., 2016).

However, erosion and sediment deposition processes in river channels and riverbanks constantly alter the geometry of channels, the stage-discharge relationship changes over time. As a result, a periodic comparison of the SDR curves to direct observations is required and hence river gauging stations are expensive, labor-intensive, and time-consuming to maintain (Pan, 2013). On other hand, the number of gauging stations has dropped on a global basis, which is especially problematic in the face of hydrological

regimes that have altered dramatically in recent years or will do so in the future due to climate change (Arnell & Gosling, 2013; Fekete & Vörösmarty, 2007.). Mekong is one of the such basin whose hydrology is altered due to impact of climate change and hydropower developments (Hoang et al., 2016).

Due to the length and remoteness of the Mekong River, installing and maintaining adequate gauge stations to monitor the whole river would be difficult. Other concerns in such transboundary river basin related to economic/political restrictions/data delay, for example, would result in data inaccessibility in certain locations and hence the practicality of obtaining timely and continuous observations at many points along the river continues to be a problem (Liu et al., 2016). Thus, spaceborne sensors can be utilized to address this situation because of the expanding number of earth observation satellites and their improved temporal and geographical resolution. Moreover, a regional model of Lancang Mekong River if developed utilizing the available in-situ data in different river reach and data obtained from satellite sensors will help to estimate discharge in ungauged reaches of the river.

1.3 Research Questions

1. What will be the performance of Remote Sensing in river discharge estimation?
2. What is the spatial and temporal variation of discharge in ungauged reach of Lancang Mekong River?

1.4 Objectives of the Study

The main objective of this research is to utilize multisource remote sensing datasets for river discharge estimation in the Lancang Mekong River Basin (LMRB).

To achieve this main objective, following specific objectives are set:

i. To produce time series water level coping with the limitations of temporal and spatial resolution of satellite altimetry.
ii. To use machine learning approach for daily discharge estimation by merging satellite optical sensor data and altimetry data.
iii. To predict the river discharge at ungauged locations and analyze its spatial and temporal variation in LMRB.

1.5 Scope and limitations of the Study

i. This study uses eight in-situ station data for developing a regional model for estimation of river discharge using remote sensing.
ii. The artificial neural network Long Short-Term Memory (LSTM) will be used in the study.
iii. Daily water level obtained from satellite altimetry products Jason-1/2, Envisat, Saral and reflectance ratio of dry and wet pixel obtained from MODIS will be used as input to LSTM model.
iv. The results of the study will rely on quality and quantity of secondary data.

CHAPTER 2
LITERATURE REVIEW

2.1 Importance of Monitoring River Flow

Monitoring and quantifying river flow is crucial for future forecasting as well as for the sustainable management of this valuable resource (E. Zakharova et al., 2020). Demand for more accurate, high frequency, and accessible water data is increasing because of growing populations and rival water priorities, including the preservation and restoration of aquatic ecosystems.(Hirsch & Costa, 2004; Arnell & Gosling, 2013). Numerous significant applications across a variety of scales necessitate an understanding of river discharge, including global water balances, engineering design, flood predictions, reservoir operations, navigation, water supply, recreation, and environmental management (Gravelle, 2015). For freshwater ecosystems, river flow regimes, such as long-term average flows, seasonality, low flows, high flows, and other forms of flow variability, are crucial (Arnell & Gosling, 2013).

Around the final quarter of the 19th century, technology developed hydroelectricity, and as a result, regular measurements of river flow started to spread globally to advantageous locations to gather data that would help with river discharge forecasting, ensuring continuous power production and water volumes in line with demand (Depetris, 2021). However, the present environmental/ climate catastrophe has introduced new difficulties to the initial uncomplicated objectives. In addition to altering precipitation patterns and the frequency of extreme weather events, climate change also introduces non-stationary characteristics to river flow series (Arnell & Gosling, 2013). Therefore, building solid hydrological data bases is even more crucial so that river management for flood control, water shortage, and water quality maintenance may be based on careful observations of the frequency and intensity of river flows (Derecki & Quinn, 1987).

It is important emphasizing here that, Oceania and Africa, for example, account for more than half of the rivers for which Milliman & Farnsworth, (2011) in their well-known synthesis, could find no hydrological data. Maintaining global river discharge databases is a challenging task and despite major improvements since river flow measurement began, the extensive gauging network is still far from perfect (Milliman & Farnsworth, 2011). For instance, the Global Runoff Data Centre (GRDC) only

includes a limited number of stations that are mostly determined by the participating nations. Additionally, hydrological data is often insufficient or partial, and changes that occur over time in gauging networks could be undetected (Depetris, 2021).

2.2 Measurement Methods of Stream Discharge

Discharge measurements are carried out in natural watercourses to calculate the surface outflow of a basin, its periodic variability, and its outflow characteristics. Data on stream flow must be collected in a cohesive way, for a consistent period, with an assessment of the accuracy and uncertainty involved (Tazioli, 2011). The various methods of discharge estimation includes stream gauging using stage and rating curves, velocity point measurements and dilution gauging and use of current meters (Gravelle, 2015). Most hydrometric stations have a staff gauge, which continually records data during predetermined time intervals. The water level data may be transformed into data on river discharge using a stage-discharge relationship. While analytical formulae may be used to explain these relationships, it is preferable to collect experimental data and calibrate it using actual river discharge measurements (Perumal et al., 2007).

Obtaining mean flow velocity of the section, the traditional methods is carried out by submerging a current meter in several spots of a river cross-section (Tazioli, 2011). However, some problem arises in very high discharge, when water depth is insufficient for immersion of the equipment or flow velocity is lower than the minimum required, in case of turbulence. An alternative for these conditions is the artificial tracing method which is the use of radioactive substances or chemicals like sodium chloride (Florkowski et al., 1969).

To address the gap of on-site measurements, alternative methods of streamflow prediction and monitoring; modeling and satellite observations, have been rapidly developed. While satellite observation may give global views, hydrological modeling is a powerful tool for providing insights at regional or basin scales (E. Zakharova et al., 2020). Since river discharge cannot be monitored directly due to its nature, observations of other hydraulic variables, such as water level, flow velocity, water extent, and slope, both satellite and conventional monitoring methods is used to calculate river discharge (Bjerklie et al., 2005).

2.3 Applicability of Remote Sensing in Stream Discharge Estimation

Remote sensing has developed as a reliable source of observations during the recent years, especially in regions of the world with scarce in situ networks (Bjerklie et al., 2018). There is currently no satellite-based technique that can measure river discharge directly. In order to estimate stream discharge, the use of remotely sensed hydraulic characteristics such water-surface width, gradient, and elevation has been investigated by many researchers (Bjerklie et al., 2005; Dingman & Bjerklie, 2005; Leon et al., 2006; Zakharova et al., 2020). Thus, satellite remote sensing is used to estimate a variety of hydrological status variables and fluxes. Though satellite remote sensing systems may make continuous and up-to-date measurements with broad area coverage depending on the orbital features of the platform, they depend on in-situ observations for algorithm development and validation (Tang et al., 2009).

Contrary to popular belief, remote sensing does not aim to completely replace gauges in discharge estimation (Gleason & Durand, 2020). Remote sensing for discharge estimation is therefore used for substitute of different purpose:

- At the mercy of politics and economics, gauges that exist now might not exist tomorrow.
- Remote Sensing signals are able to expand point gauge measurements in space and time due to high gauge calibration data.
- Water resources may be documented using Remote Sensing for Discharge in a variety of channel configurations and during times of flooding, which are challenges for gauges.
- Another area where Remote Sensing for Discharge may excel is with discontinued gauges, as these gauges can be calibrated and then utilized to parameterize models as they advance into the future.
- When creating a primary data source for usage in locations that are not well monitored, remote sensing might be utilized to prepare inaccurate data.

Bjerklie et al. (2003) assessed the possibilities of the various satellite data sources, described a variety of methods for using remote sensing to calculate river flow, and thought about the possibility of computing river discharge exclusively from remotely sensed data sources. In comparison to other models that merely contain width and slope

or width, slope and velocity, the authors contend that models based on these parameters are often more accurate, especially for big rivers.

Lin et al. (2019) created a cogent worldwide reanalysis of daily river discharge at approximately three million river reaches using a considerable amount of ground and remote sensing data (mostly for precipitation and evapotranspiration) along with the most recent advances in remote sensing hydrography. The power of remote sensing for global hydrologic modeling is illustrated by the level of temporal and spatial precision that has never before been attained. Lin et al. (2019) employed a calibration strategy that takes uncertainty into account, calibrating with gauges when available, remote sensing products when gauges are not, and reanalysis data when remote sensing and gauges are both accessible. This makes it feasible to produce discharge that wouldn't be possible without the use of remote sensing by making the best use possible of both in situ and remote sensing data.

Recent studies have shown that remote sensing has a strong spatial coverage and a longer monitoring period, both of which have driven an increase in interest in using remote sensing to predict discharge (Garkoti & Kundapura, 2021; Huang et al., 2018; Mengen et al., 2020; Sichangi et al., 2016).

2.4 Estimation of Water Level and Discharge Using Satellite Altimetry

Water levels in big rivers, lakes, and floodplains have been continuously monitored by satellite altimetry, and a time series spanning more than 25 years is now available (Papa et al., 2010). Several studies have demonstrated the capability of using satellite altimetry for estimating river discharge in medium to large rivers (with a width of few kilometers), including the Po River, Italy (Tarpanelli, Barbetta, et al., 2013), Niger River (Tourian et al., 2017), Ob River (Kouraev et al., 2004) and several location is Amazon river (E. A. Zakharova et al., 2006) with the use of rating curve developed by correlating altimetry river water level with in-situ measurements of river discharge or development of models.

Thus, the monitoring of river water levels has benefited from recent advancements in radar altimetry technology, and the increasing accuracy of the sensors supports their use as a validation tool for a variety of applications, from basic routing strategies to complex hydraulic models. Though the altimetry mission's spatial-temporal sampling

is a limitation (Tarpanelli et al., 2019). When the satellite ground track repeats (10-day TOPEX/Poseidon and Jason; 35-day ERS-2 and ENVISAT), this may be viewed as a space-borne virtual gauge that measures temporally distinct river channel stages. The upcoming Surface Water and Ocean Topography (SWOT) satellite mission will enhance the temporal and geographical coverage in the future (Birkinshaw et al., 2014).

Chen et al. (1998) demonstrated how models may be integrated with remote sensing to produce discharge using data from the Topex/Poseidon satellite to observe sea level changes and comprehend anomalies in sea surface heights. According to their argument, monitoring ocean anomalies can help us comprehend the necessary adjustments to the global hydrologic cycle that led to such anomalies if the ocean is the ultimate repository for all terrestrial water.

Birkinshaw et al. (2010) highlighted the possibility of employing remote sensing, such as SAR, to give information on the channel cross sections by combining altimetry data with in-situ observed channel cross sections to predict discharge at an ungauged location. The approach makes use of the cross-sectional regions upstream and downstream; altimetry provides the height variation and in situ or remote sensing data provides the river geomorphology. Remote sensing data in particular may be utilized to provide a time series of river width, which when combined with altimetry results in the cross-sectional area at the sub-satellite sites that changes over time.

In the research by Bogning et al., (2018), the performance of the five altimetry missions was evaluated by comparing the results to records from gauge station measurements. A long-term and vastly better water-level time series was produced by combining the data from all of the radar altimetry missions. An improvement in water level peak to peak characterization and, consequently, a more precise annual discharge throughout the shared observation period between the altimetry-based and the in situ mean annual discharge are both produced by the increased data sampling in the river basin.

This technique has however several limitations which includes dependence on quality of altimetry data over continental water bodies, unavailability of in-situ observations to develop rating curves, assumption of static rating curve while extending time series of discharge and temporal sampling rate (Papa et al., 2010).

2.4.1 Satellite Altimetry: Introduction and Working Principle

At the end of the 1960s, satellite radar altimetry first appeared. Based on active microwave observations (or radar) methods, radar altimetry measures distance. The electromagnetic (EM) pulse is transmitted in the nadir direction by the radar sensor, which then accurately calculates the signal's two-way travel time (Δt). In order to record the radar echo or waveform, pulse compression and de-ramping methods are used. The waveform's amplitude and shape, which are connected to the backscattering coefficient, reveal details about the surface's makeup. When the received power reached the center of the leading edge at mid-height, the distance between the satellite and the surface, or altimeter range (R), corresponded to that epoch. It is estimated as

$$R = \frac{c\Delta t}{2} \qquad \text{Equation 2-1}$$

where c is the velocity of light in vacuum.

The range must be adjusted for atmospheric propagation delays, instrument corrections, and surface geophysical adjustments in order to calculate the surface topography accurately (Frappart et al., 2017). Using an ellipsoid to represent the sea-surface height (SSH) over the ocean as shown in Figure 2-1 and the height of the reflecting surface (h) is then determined as;

$$h = H - R - \Delta R_{ion} - \Delta R_{dry} - \Delta R_{wet}(-\Delta R_{ssb}) \qquad \text{Equation 2-2}$$

where H is the height of the center of mass of the satellite above the ellipsoid estimated using precise orbit determination techniques

R is the nadir altimeter range from the center of mass of the satellite to the surface considering instrumental corrections

ΔR_i are the corrections applied to the range and applied over all types of surfaces

ΔR_{ion} is the atmospheric correction which is the atmospheric refraction range delay caused by the ionosphere's dielectric characteristics and free electron content

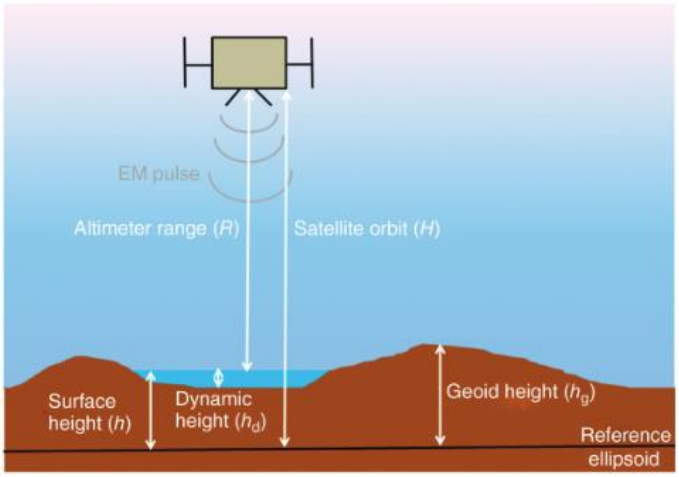
ΔR_{dry} is the atmospheric refraction range delay caused by the troposphere's dry gas component

ΔR_{wet} is the atmospheric refraction range delay caused by the tropospheric water content in clouds and water vapor

ΔR_{ssb} is the range correction over oceans and great lakes which is the interaction of the altimeter's electromagnetic pulse with the scatterers within the footprint (e.g., wave and surface roughness effects)

Figure 2-1

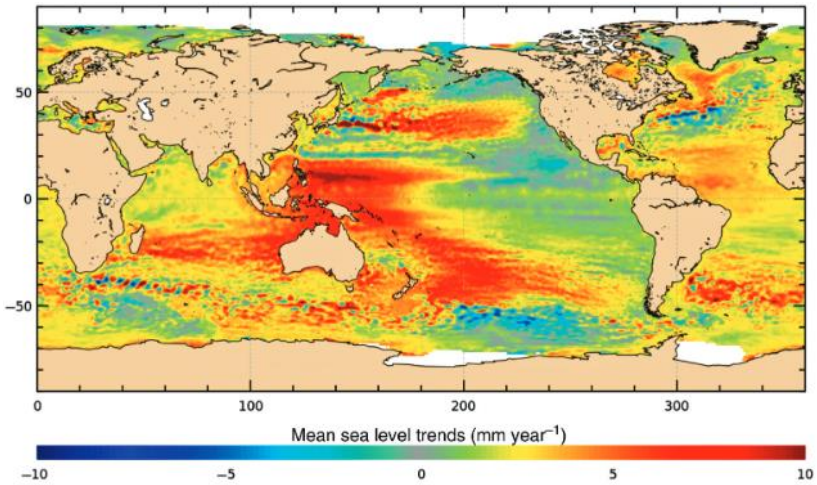
The Principle of Radar Altimetry Measurement (Frappart et al., 2017)



Originally, altimeter products have been used to monitor the surface elevation of oceans and study shows the rise in sea level is not constant. The pace of sea level rise in some areas (such as the western Pacific) is up to three times higher than the average rate worldwide which is shown in Figure 2-2. Later these altimetry satellite missions are being used for observation of large rivers and lakes for long-term observations of water level variations (Frappart et al., 2012).

Figure 2-2

Spatial trend patterns in sea level over January 1993-2014 based on multi-mission satellite altimetry (https://climate.esa.int/en/projects/sea-level/)



Satellite altimetry can help with hydrological research since it keeps track of surface water heights (Calmant & Seyler, 2006). Using specialized software like the Multi-mission Altimetry Processing Software, a more accurate selection of valid data may be made in order to increase the accuracy of the water stages generated from altimetry. The accuracy and bias of water stages estimated from radar altimetry were evaluated in several studies and found to have good accuracy compared with in-situ observations (Biancamaria et al., 2017).

2.4.2 Satellite Altimetry Products

The time series of measurements of inland water levels that began with the launch of ERS-1 in 1991 and TOPEX/Poseidon in 1992 is continuously being extended by the availability of satellite altimetry from ERS-2 and ENVISAT as well as from Jason-1 and Jason-2 (Birkinshaw et al., 2010).

GEOS-3, SeaSat, Geosat, TOPEX/Poseidon, Geosat Follow-on, and Jason-1/2/3 from the National Aeronautics and Space Administration are among the previous and present satellite altimetry missions listed (NASA). These satellites include ERS-1/2, ENVISAT, CryoSat-2, and Sentinel-3 from the European Space Agency (ESA). Other missions, like the Chinese-planned HY-2A or the combined French-Indian SARAL/AltiKa project, are also in operation. The summary of satellite altimetry missions is shown in Table 2-1.

Table 2-1

Summary of Satellite altimetry missions (Jiang et al., 2017)

Satellite	**Agency**	**Period**	**Altitude (km)**	**Altimeter**	**Frequency Used**	**Repetitivity (Day)**	**Equatorial Inter-Track Distance (km)**
Skylab	NASA	May1973–February1974	435	S193	Ku-band		
GEOS3	NASA	April1975–July1979	845	ALT	Ku and C-band		
SeaSat	NASA	July–October1978	800	ALT	Ku-band	17	
Geosat	US Navy	October1985–January1990	800		Ku-band	17	

Satellite	Agency	Period	Altitude (km)	Altimeter	Frequency Used	Repetitivity (Day)	Equatorial Inter-Track Distance (km)
ERS-1	ESA	July1991–March2000	785	RA	Ku-band	35	80
Topex/ Poseidon	NASA/ CNES	September1992 – October2005	1336	Poseidon	Ku and C-band	10	315
ERS-2	ESA	April1995-July2011	785	RA	Ku-band	35	80
GFO	US Navy/ NOAA	February1998–October2008	800	GFO-RA	Ku-band	17	165
Jason-1	CNES/ NASA	December2001–June2013	1336	Poseidon-2	Ku and C-band	10	315
Envisat	ESA	March2002–April2012	800	RA-2	Ku and S-band	35	80
OSTM/ Jason-2	CNES/ NASA/ Eumetsat/ NOAA	Jun2008–present	1336	Poseidon-3	Ku and C-band	10	315
CryoSat-2	ESA	April2010–present	720	SIRAL	Ku-band	369	7.5
HY-2	China	August2011–present	971		Ku and C-band	14168	
Saral	ISRO/ CNES	February2013–present	800	AltiKa	Ka-band	35	80
Jason-3	CNES/ NASA/ Eumetsat/ NOAA	January2016–present	1336	Poseidon-3B	Ku and C-band	10	315
Sentinel-3A	ESA	February2016–present	814	SRAL	Ku and C-band	27	104

2.5 Multispectral Satellite Data for Monitoring River Discharge

Recent developments in satellite sensors stimulated its use for river flow estimation due to their frequent revisit time and large spatial coverage, taking into account the daily river discharge measurement (Filippucci et al., 2022). Several studies have shown the possibility of discharge estimation from the space with their feasibility to monitor difficult river sites in medium-sized to large catchments from various sensors, including

MODIS, MERIS, Landsat and OLCI datasets (Brakenridge & Anderson, 2006; Shi et al., 2020; Tarpanelli, Brocca, et al., 2013; Tarpanelli et al., 2020). Since discharge cannot be measured directly, hydraulic variables like river channel widths and the extent of surface water are measured using remote sensing data. This information can be gathered by synthetic aperture radar (SAR), visible spectrum digital imagery and they have been used to monitor the extent of a floodplain's inundation, cross section geometry and variation of water level with river extent (L. Smith & Pavelsky, 2008; Sun et al., 2010).

There are two main approaches for monitoring river discharge via multispectral images. One method is to develop regression relationship between in-situ measured discharge and inundated channel reflectivity (including inundated and potentially inundated area of the river channel) and other is to develop regression relationship between river width and in-situ discharge. Both of these methods have been used successfully in various river for discharge estimation (Bjerklie et al., 2003; Li et al., 2019; L. C. Smith & Pavelsky, 2008; Tarpanelli, Brocca, et al., 2013)

Brakenridge & Anderson (2006) proved the practical hydrological uses of the Moderate resolution Imaging Spectroradiometer (MODIS) sensor, including not only flood detection and characterization but also discharge estimates. They only provided one early example where MODIS approximated the discharge using a very short time period (about 30 data).

Sichangi et al. (2018) computed temporal river width data from the MODIS images to derived river discharge. Only remote sensing data with NSE values larger than 0.5 at the analyzed sites were used to estimate discharge. The calculated velocity's values, which range from 0.63 to 1.2 m/s and 0.81 to 1.35 m/s, were found to be within the bounds of the expected velocity 0.96 m/s. The methodology for calculating discharge is a viable strategy for usage on rivers all across the world because it exclusively relies on global satellite databases.

Sahoo et al. (2020) presented an integrated MODIS-Landsat fusion methodology and tested the method in eastern India's Brahmani River Basin to estimate high-frequency discharge. To increase the temporal resolution of the single datasets, a copula-based model was suggested, and the improved results were then shown. This study on the use

of many satellites for different missions demonstrated their value in improving flow monitoring, particularly in highly forested areas.

2.5.1 NIR Band Reflectance as Proxy of Discharge

Optical remote sensing sensors are the essential tools that measure various spectral signatures, according to wavelengths, that each sensor monitors reflected or emitted energy. However, optical remote sensing images are impacted by clouds, haze, and cloud shadows, making it difficult to distinguish between them and other dark objects like water and shadows (Zhu & Woodcock, 2012). In near infrared (NIR) wavelengths, water absorbs more energy (has a low reflectance), whereas non-water reflects more energy (high reflectance) because of the strong water absorption in NIR (Ahn & Park, 2020). Any location that experiences flooding will have an increase in water surface, which will lower the area's NIR reflectance value as seen by the satellite. This different spectral characteristics of water bodies and other objects in near-infrared (NIR) images can be used to monitor variations in the river discharge (Shanlong et al., 2010).

The Calibration/Measurement (C/M) technique makes use of the relationship between surface water extent dynamics and river flow. This method is based on the principle that the ratio of reflectivity from non-inundated area and the inundated area is positively correlated with river discharge. First, Brakenridge et al. (2007) used AMSR-E data at 37 GHz to estimate river flows using the difference between the brightness temperature of a pixel over the river and a pixel over an unaffected pixel. Subsequently, MODIS sensor had been used successfully exploiting different behavior of land and water in NIR band for discharge estimation (Sahoo et al., 2020a; Van Dijk et al., 2016). In contrast, if light, atmospheric scattering, or vegetation characteristics change and recorded radiances change simultaneously, the ratio stays fairly constant. Utilizing MODIS optical data to detect surface water changes economically and effectively is the paired-measurement method (Brakenridge et al., 2005).

While many traditional sensors can readily capture surface water dynamics, the C/M technique has the benefit of requiring less data. The C/M technique may be broadly divided into two categories: one uses optical satellite data, such as the near-infrared (NIR) band of the Moderate Resolution Imaging Spectroradiometer (MODIS). The other uses microwave satellite data such as that from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E), to estimate discharge

and identify floods, even in rivers that are not gauged and are inaccessible. The AMSR-E method of performing C/M is a more complicated procedure that necessitates other auxiliary data, such a precipitation dataset, at the same time (Temini et al., 2011).

Tarpanelli, Brocca, et al. (2013) used this method and successfully predicted discharge for four sections of the Po River in northern Italy using M and C pixels from MODIS NIR images. This work has demonstrated that MODIS data may be used for ungauged river sites as well as medium-sized basins (<10,000 km^2) with considerable temporal variability to provide reliable flow estimates.

Shi et al. (2020) applied linear regression between C/M with observed discharge series using Harmonized Landsat and Sentinel-2 (HLS) surface reflectance product on relatively small rivers with 30~100 m widths in Murray Darling Basin, Australia to estimate river discharge. The results showed high consistency with the observed discharge.

2.5.2 Satellite Optical Data Sets

The multispectral image data can be obtained for different satellite missions. These data availability depends upon the orbital period of satellite and their spatial and temporal resolution. Terra and Aqua satellites both have the MODIS instrument in operation. Every one to two days, it observes the whole surface of the Earth within a 2330-kilometer viewing swath. With three different spatial resolutions of 250 m, 500 m, and 1000 m, its detectors measure 36 spectral bands between 0.405 and 14.385 m (MODIS Web). Likewise the details of other satellite missions are provided below.

S.N.	Data	Agency	Equatorial cross-track separation (km)	Spatial Resolution	Temporal Resolution	Period	Source
1	MODIS AQUA	NASA	2330	250 m, 500 m, 1000 m	1 – 2 days	May 2002 - Present	https://lpdaac.usgs.gov/
2	MODIS TERRA	NASA	2330			Dec 1999 - Present	
3	Landsat ETM +	NASA and USGS	185	30m	16 days	Apr 1999 – Sep 2021	
4	Sentinel 2	ESA	290	10 - 60 m	5 days	June 2015 - Present	https://sentinels.copernicus.eu/
5	Meris	ESA	1150	300 m	3 days	2002 - 2012	https://www.esa.int/

2.6 Multi-Mission Satellite Approach for River Discharge Estimation

The estimations of discharge generated by combining the river stage level from satellite altimetry data with other space-based parameters, such as river width and river bed velocity from optical sensors and Synthetic Aperture Radar (SAR), ought to be superior to those based on a single parameter (Sichangi et al., 2016). Satellite altimetry has a relatively poor spatial sampling that corresponds to its ground track pattern ranging from 10 days for Jason series, 35 days for ERS, Envisat and Saral/AltiKa to 369 days for CryoSat-2 missions. Thus, multi-mission approach has been widely accepted and used in many studies to cope with the limitation of spatiotemporal resolution (Table 2-2) of satellite altimetry and optical images.

The approach involving use of altimetry and optical sensors for river discharge estimation has shown better performance in Niger and Po rivers with NSE of 0.98 and 0.83 (Tarpanelli et al., 2019).

Table 2-2

Previous Studies using multi-mission approach for discharge estimation

Author	Study Area	Data Used	Method	Key Results
Sichangi et al. (2016)	Eight of the world's major rivers	Envisat, Jason 2, MODIS and DEM	River stage is incorporated with the effective river width and compared with empirical equation proposed by (Bjerklie et al., 2003)	Improved prediction accuracy in most of the river. NSE varied between 0.6 and 0.97 with $R^2 > 0.90$.
Huang et al. (2018)	Upper Brahmaputra River, China	Jason 2/3, Saral AltiKa Landsat 5/7/8 and Sentinel 2 images	Modified Manning's equation and Rating Curve	The NSE vary between 0.65 to 0.97 and RMSE vary between 32.97 to 695 m^3/s.
Bjerklie et al. (2018)	Yukon River, North America	Jason 2 Landsat	Manning's equation and Prandtl-Von Karman equation	The calibrated discharge estimate showed the accuracy of +/- 2%
Garkoti & Kundapura, (2021)	Krishna River, India	Jason 3, Sentinel 3A, Sentinel 3B, Sentinel 1 and Sentinel 2 images	Modified Manning's equation	The NSE vary between 0.53 to 0.62 and R^2 vary between 0.83 to 0.97.
Scherer et al. (2020)	Lower Mississippi River	Envisat, Jason-2/3, Sentinel-3A/3B, Landsat-4/5/7/8, Sentinel 2A/2B	Manning's equation	Estimate 18 years discharge time series using satellite altimetry and remote sensing data with NRMSE 7 – 35 %

2.7 Application of Artificial Intelligence in Hydrology

John McCarthy first used the term "artificial intelligence" in 1956. The field of computer science known as artificial intelligence (AI) is concerned with the research and development of intelligent agents that can comprehend their surroundings and take actions that increase their chances of success. The definition of artificial intelligence (AI) is "the capacity to simultaneously hold two different ideas while retaining the capacity to function." However, AI must also have the capacity for inference, quick response, and learning from prior experience (Singh et al., 2013).

Machine learning is an application to AI. It is the practice of assisting a computer in learning without direct instruction by applying mathematical models of data. As a result, a computer system can keep picking up new skills and getting better on its own. Using a neural network, which is a collection of algorithms based after the human brain, is one method for teaching a computer to imitate human reasoning. Machine learning has been effectively used in remote sensing for source separation, classification, regression, clustering, and coding (Camps-Valls, 2009).

In order to predict different hydrological parameters, artificial neural networks (ANNs) are often and accurately employed; yet they are typically only created in one or two hidden layers. In order to successfully address complicated issues, deep learning networks have recently been enlarged with multi-layered design. A recent method of using ANNs is deep learning, which is a subfield of machine learning. It is used to represent complicated ideas by learning at various depths and levels. The unpredictable, complex, and nonlinear character of streamflow makes it viable to utilize AI-based modeling techniques. ANNs can mimic streamflow in particular using additional hydrological factors (Bengio, 2009; Ghumman et al., 2011).

Finding a suitable deep-learning network to use is challenging. However, this variety and adaptability could occasionally result in choosing the incorrect type of network from among the many options. Recurrent neural network (RNN) types of neural networks are an excellent substitute for continuous (time-series) data, such as streamflow, according to recent research (Elumalai et al., 2017). Long-term dependencies between the network's inputs and outputs could be discovered through Long Short-Term Memory (LSTM). Better model performance has been achieved using

this method, highlighting the LSTM's potential for application in hydrological modeling applications (Kratzert et al., 2018).

Several research have been done to address the application of machine learning in the field of hydrology. Some of the researches are discussed below:

Esmaeilzadeh et al. (2017) used data-mining techniques such as ANNs, the M5 tree method, support vector regression, and hybrid Wavelet-ANN methods to estimate the daily flow to the Sattarkhan Dam in Iran. The wavelet artificial neural network (WANN) method outperformed other methods, according to the results, in estimating flow.

According to Kratzert et al. (2018), who used the publicly available CAMELS dataset to model daily streamflow in 291 catchments, the LSTMs as individual catchment models performed better primarily in snow-driven catchments and worse in arid basins because long-term dependencies are more significant for snow-driven processes (e.g., snow accumulation and snowmelt). Additionally, the LSTM model outperformed the benchmark conceptual model by a small margin.

Zhang et al. (2018) calculated water tables in agricultural areas using LSTM-RNN. They contrasted the simulation produced by LSTM-RNN with that produced by Multilayer Perceptron (MLP) and discovered that the former performed better than the latter. So, they suggested model can be used as a substitute for existing methods of estimating water table depth, particularly in locations where hydrogeological data are challenging to come by.

Chiang et al. (2018) combined ensemble approaches into artificial neural networks in order to lower the level of model uncertainty in hourly streamflow projections in the Chinese watersheds of Longquan Creek and Jinhua River. Results showed that as compared to a single neural network, the accuracy of streamflow predictions was enhanced by roughly 19–37% by ensemble neural networks.

According to Kao et al. (2020) the Shihmen Reservoir watershed in Taiwan experienced multi-step-ahead flood forecasting using the LSTM-based Encoder-Decoder (LSTM-ED) model,. Results indicated that the suggested model may boost the interpretability

of the model internals and raise the reliability of flood forecasting by converting and connecting the rainfall sequence with the runoff sequence.

Cheng et al. (2021) developed three machine learning models using ANN, Support Vector Regression (SVR), and LSTM to anticipate discharge variation in North China and adequate accuracy was attained. The results showed that MLP performed marginally better than LSTM-RNN and much better than SVR. Additionally, it was shown that ANNs were more effective than other machine learning techniques for modeling and forecasting karst spring outflow.

2.7.1 Long Short-Term Memory (LSTM)

A unique variety of RNN, known as the LSTM architecture, a type of model or framework for sequential data, was created to address the typical RNN's inability to learn long-term dependencies. It implements gates that control "memory cells" using a unique combination of hidden units, elementwise products, and sums between units. These cells are made to store data unaltered over extended periods of time (Hochreiter & Schmidhuber, 1997). The key advantage of LSTM is that it can learn long-term dependence, which is not achievable with straightforward RNNs.

Runoff was modelled for Fen River basin, China by Hu et al. (2018) using traditional ANN and LSTM models by and results showed LSTM performed better in terms of time series discharge simulation. The LSTM model is more cognitive than the ANN model, and the ANN model is more sensitive when compared to simulations of flooding events.

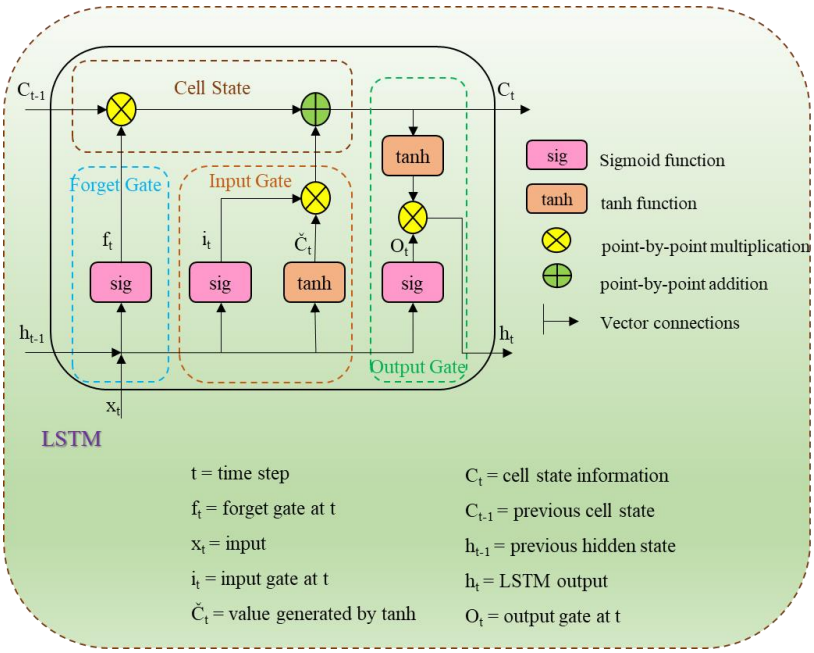
Nogueira Filho et al. (2022) studied the application of LSTM as a regional method against traditional neural network (FFNN) and conceptual models in a data scare catchment. The LSTM model outperformed the Feedforward model in the specific catchment, demonstrating its ability to explain the hydrological dynamics of streamflow in semiarid areas. The performance of the neural networks trained with the regionalization data were better than the neural networks built for single catchments and the neural network approaches also exhibited the capacity to aggregate process understanding from diverse watersheds.

LSTM Architecture

Figure 2-3 depicts the basic LSTM unit, which consists of a cell with an input gate, output gate, and forget gate. To handle the disappearing or ballooning gradient problem, LSTMs employ the idea of gating (Greff et al., 2017). Each of the three gates can be thought of as a typical artificial neuron, computing an activation (using an activation function) of a weighted sum of the current data x_t, a hidden state h_{t-1} from the previous time step, and any bias b. The cell is in charge of remembering values over arbitrary time intervals (Ordóñez & Roggen, 2016).

Figure 2-3

Illustration of a Long-Short Term Memory unit



Forget gate determines what information must be remembered and what can be forgotten. The sigmoid function receives data from the current input X_t and the hidden state h_{t-1}. The values that Sigmoid produces range from 0 to 1. It draws a conclusion on the necessity of the old output's portion (by giving the output closer to 1). The cell will eventually utilize this value of f_t for point-by-point multiplication.

$$f_t = \sigma(W_f . [h_{t-1}, x_t] + b_f)$$

Input gate performs the following operations to update the cell status.

First, the second sigmoid function receives two arguments: the current state X_t and the previously hidden state h_{t-1}. Transformed values range from 0 (important) to 1. (not-important).

The tanh function will then get the identical data from the hidden state and current state. The tanh operator will build a vector (C_t) containing every possible value between -1 and 1 in order to control the network. The output values produced by the activation functions are prepared for multiplication on a point-by-point basis.

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i)$$

$$C_t = tanh(W_c.[h_{t-1}, x_t] + b_c)$$

Output gate determines the value of the next hidden state. Information about prior inputs is contained in this state.

First, the third sigmoid function receives the values of the current state and the prior hidden state. The tanh function is then applied to the new cell state that was created from the original cell state. These two results are multiplied one by one. The network determines which information the hidden state should carry based on the final value. For prediction, this concealed state is employed.

The new hidden state and the new cell state are then carried over to the following time step.

$$o_t = \sigma(W_o.[h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

At conclusion, the forget gate selects whatever pertinent information from the earlier processes is required to conclude. The output gates complete the next concealed state, while the input gate determines what pertinent information may be supplied from the current stage.

2.8 Relevant studies in Mekong River Basin

Several studies have been conducted in Lancang Mekong River Basin for the estimation of river discharge using remote sensing. Some of them are discussed.

Multi mission radar altimetry has been used to develop multiple rating curve and combined in an Ensemble Learning Regression method to estimate discharge (ELQ) at three locations: Stung Treng, Kratie, and Tan Chau in Lower Mekong River (Kim et al., 2019). Since the ELQ approach corrects for deterioration in the performance for Q estimation caused by the poor rating curve with virtual stations far from in-situ Q stations, various H derived from Jason-2 altimetry were employed in this study regardless of distances from in-situ Q stations. Compared to data obtained from a single rating curve, the ELQ estimated discharge revealed more accurate results.

Birkinshaw et al. (2014) estimated discharge for ten years at Nakhon Phanom and Vientiane on the Mekong River using (Bjerklie et al., 2005) equation. Landsat satellite imagery was utilized to offer a variety of channel widths over a stretch of river, while ERS-2, ENVISAT, and satellite altimetry data were used to produce a time series of river channel water levels and the channel slope. The performance evaluation gave Nash–Sutcliffe efficiency value of 0.90 for Nakhon Phanom and 0.86 for Vientiane to obtain river discharge.

A novel decile thresholding method was developed by Mengen et al. (2020) using the concept of at-many-stations hydraulic geometry (AMHG) in Mekong River. The peak flows are virtually always underestimated even though the decile thresholding method produces good results in the low-flow range. The decile thresholding discharge estimation outperformed the Gleason and Wang's optimized AMHG technique, with an RRMSE of 19.5 percent for the overall examined period and 16 percent for just the dry seasons.

However, the discharge estimated by above and other studies are site specific only and no studies have been carried out to develop regional model for whole river. This study is intended to develop a single regional model which can be used to predict discharge at ungauged sites of the river.

CHAPTER 3
STUDY AREA AND DATA

3.1 Study Area

Large rivers including the Yangtze, Salween, Irrawaddy, Red River, and Mekong originate on the Tibetan Plateau. The Mekong drains an area of around 795,000 km^2 from its source and travels south for another 5000 km to the sea. It rises on the Tibetan Plateau and flows through the Yunnan Province of China, Myanmar, Laos, Thailand, Cambodia, and Vietnam. The country wise areas in the Mekong River Basin is shown in Table 3-1 . The Greater Mekong River can be divided into two basins: the Upper Basin, where the river is known as the Lancang and is located in Tibet and China; and the Lower Mekong Basin, which runs from Yunnan to the Sea (Figure 3-1).

Figure 3-1

Map of Study Area

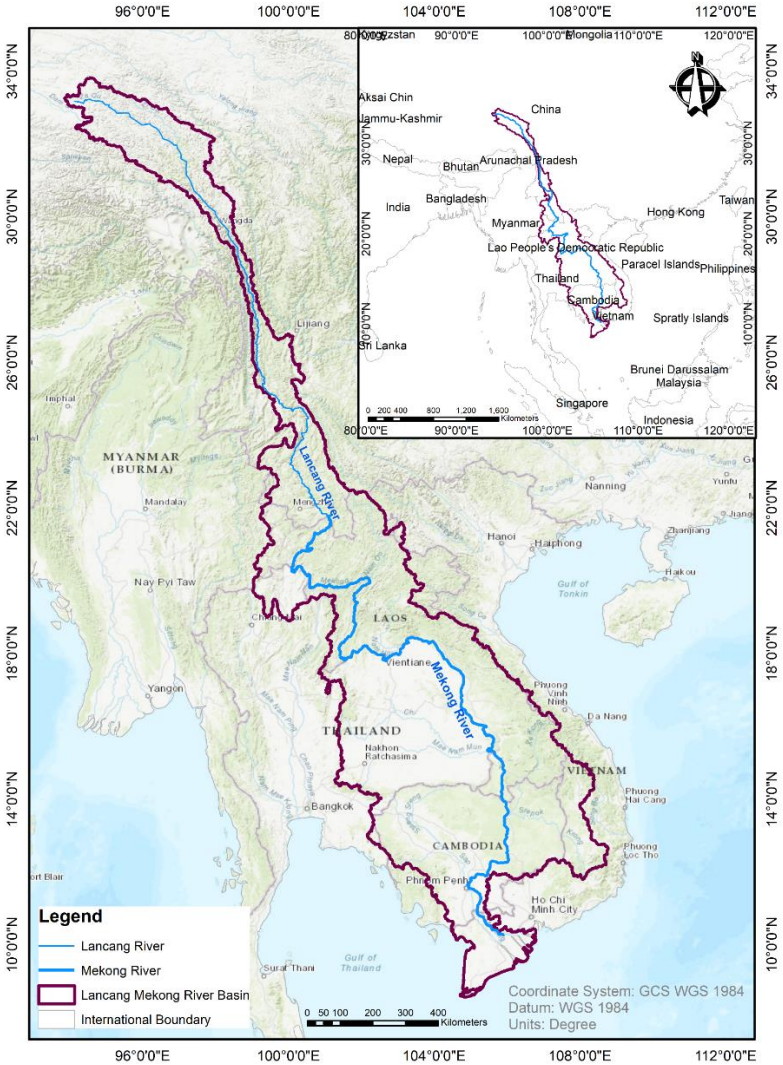


Table 3-1

Country wise area in the Mekong River Basin

Basin	Area (km^2)	Countries	Area of country in basin (km^2)	% of the total area of the basin	% of the total area of the country
Mekong	795,000	China	165,000	21	2
		Myanmar	24,000	3	4
		Lao PDR	202,000	25	85
		Thailand	184,000	23	36
		Cambodia	155,000	20	86
		Vietnam	65,000	8	20

The Upper Basin (China and Myanmar), which contributes around 15% of the flow, accounts for 24% of the entire catchment. There are no large tributaries, therefore all future water resource development will follow the main Mekong. On the other hand, the Lower Mekong Basin is also nourished by sizable tributaries, which account for around 85% of the annual flow.

3.1.1 Climate

In the region upstream, the climate is cold, whereas in the region downstream, the climate is tropical. Yunnan province in southern China's Upper Basin also experiences monsoons, though there are significant topographic variations. Yunnan has a variety of climates, from subtropical and tropical monsoons in the south to temperate monsoons in the north. However, in Yunnan province, there is a significantly broader variance from year to year in the timing of the start of the southwest monsoon, which affects the pattern of rainfall in the Upper Basin of China. Although yearly levels drop to as little as 600 mm in the north, the seasonal pattern of rainfall is the same as for the Lower Basin. The variations in climatic zones of the Lancang Mekong Basin is shown in Figure 3-2.

The southwest monsoon, which typically lasts from May through late September or early October and coincides with the Lower Basin's flood season, dominates the region's climate. In most areas of the basin, there are frequent periods of intense rainfall lasting one or two days. The wettest months of the year are August, September, and in the delta even October due to tropical cyclones that affect most of the region later in the season. Less than 1,500 mm of rain fall each year averages throughout the floodplain of Cambodia and the Mekong Delta, while more than twice that amount falls in the

Central Highlands of the Lao People's Democratic Republic and inside the major valley at Pakse (MRC 2005). The precipitation between upper Mekong and Lower Mekong is quite different and presented in Table 3-2.

Figure 3-2

The Climatic zones of the Lancang Mekong Basin (He et al., 2009)

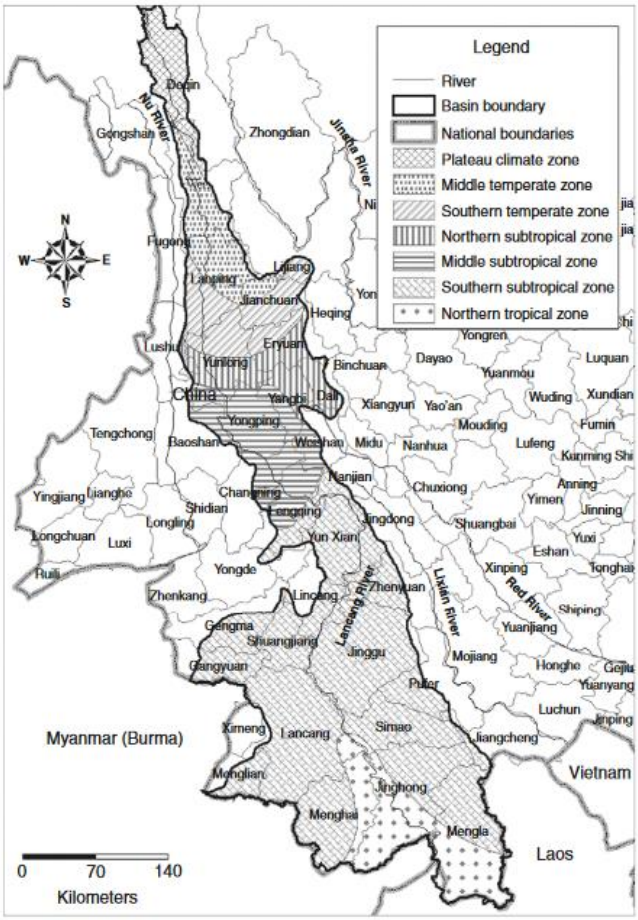


Table 3-2

Monthly Average Precipitations in the Mekong River Basin (in mm) (He et al., 2009)

Basin	Month	January	February	March	April	May	June	July	August	September	October	November	December	Annual
Upper Mekong	Rainfall	17	19	27	41	98	174	224	228	132	102	43	19	1125
	%	1.52	1.65	2.43	3.67	8.70	15.5	19.9	20.3	11.7	9.10	3.81	1.72	100
Lower Mekong	Rainfall	8	15	40	77	198	241	269	292	299	165	54	14	1672
	%	0.5	0.9	2.4	4.6	11.8	14.4	16.1	17.5	17.9	9.9	3.2	0.8	100

3.1.2 Hydrology

The Mekong River ranks eighth in the global basins with a mean annual flow into the South China Sea of around 14,500 m^3/s. Although the flow from the Upper Basin makes up just 13% of the normal annual flow, it may make up to 30% of the flow during the

dry season (MRC 2010). A large network of tributaries makes up the Mekong River basin, creating several sub-basins. In the Lower Basin, significant tributary networks emerge. These systems may be split into two categories: those that contribute to the main rainy season flow and those that drain low relief areas with lesser rainfall. Tropical monsoonal regions are typically characterized by an abundance of water and a relatively consistent flow pattern. However, the flow continues to vary dramatically from year to year (MRC, 2018). The estimated mean annual flow of the basin is almost 460 km^3 and of the total annual flow, in an average year about 75 per cent occurs within just four months between July and October.

The mean annual runoff in the basin is shown in Figure 3-3. In the Lancang Mekong watershed, river discharge rises sharply from north to south, with significant regional variations which is shown in Figure 3-4.

Figure 3-3

Mean Annual Runoff in the Mekong River Basin (MRC, 2010)

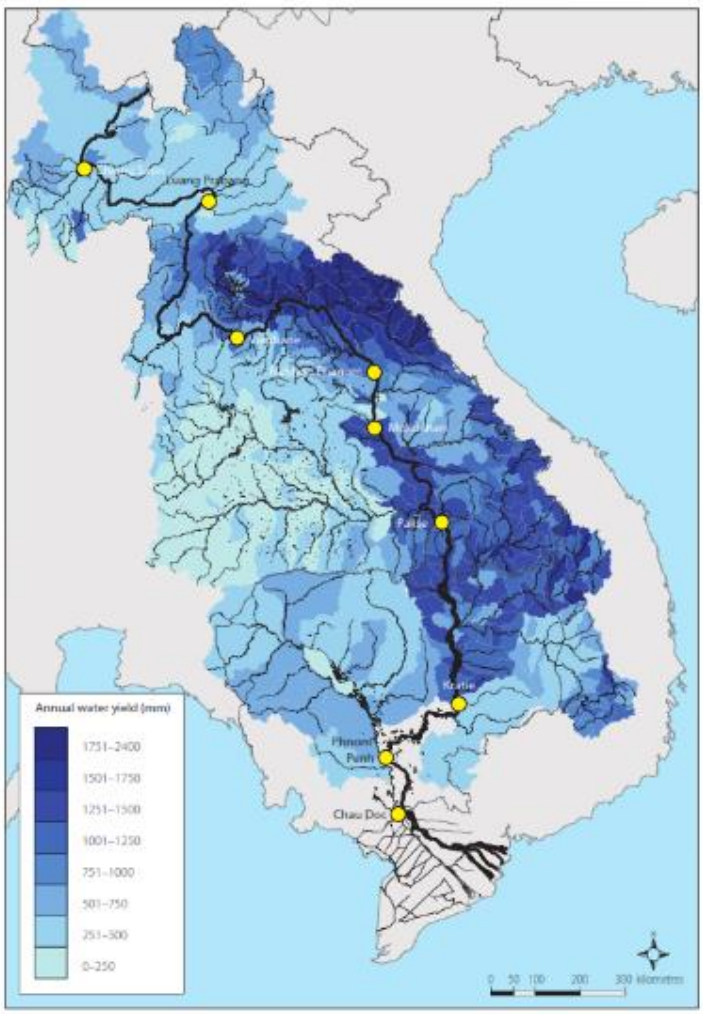
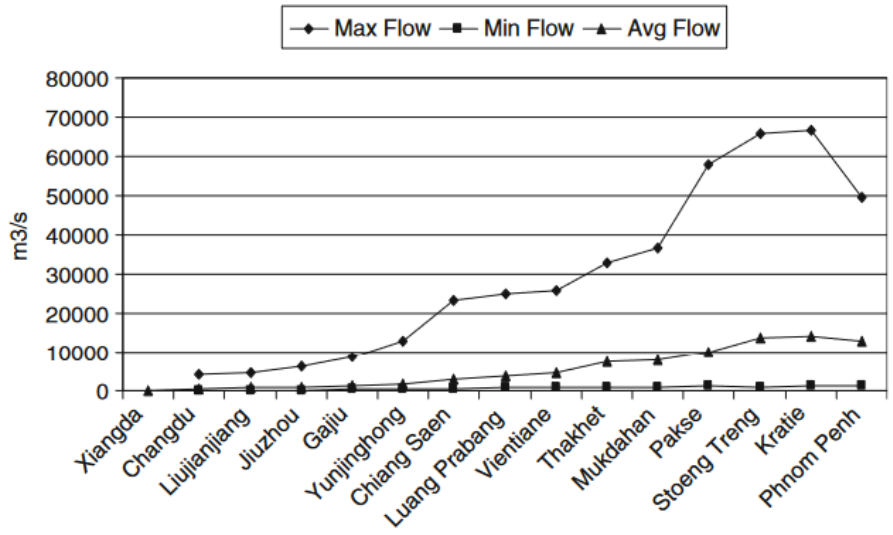


Figure 3-4

Maximum, minimum and mean annual flows at sites along the Lancang Mekong River (He et al., 2009)



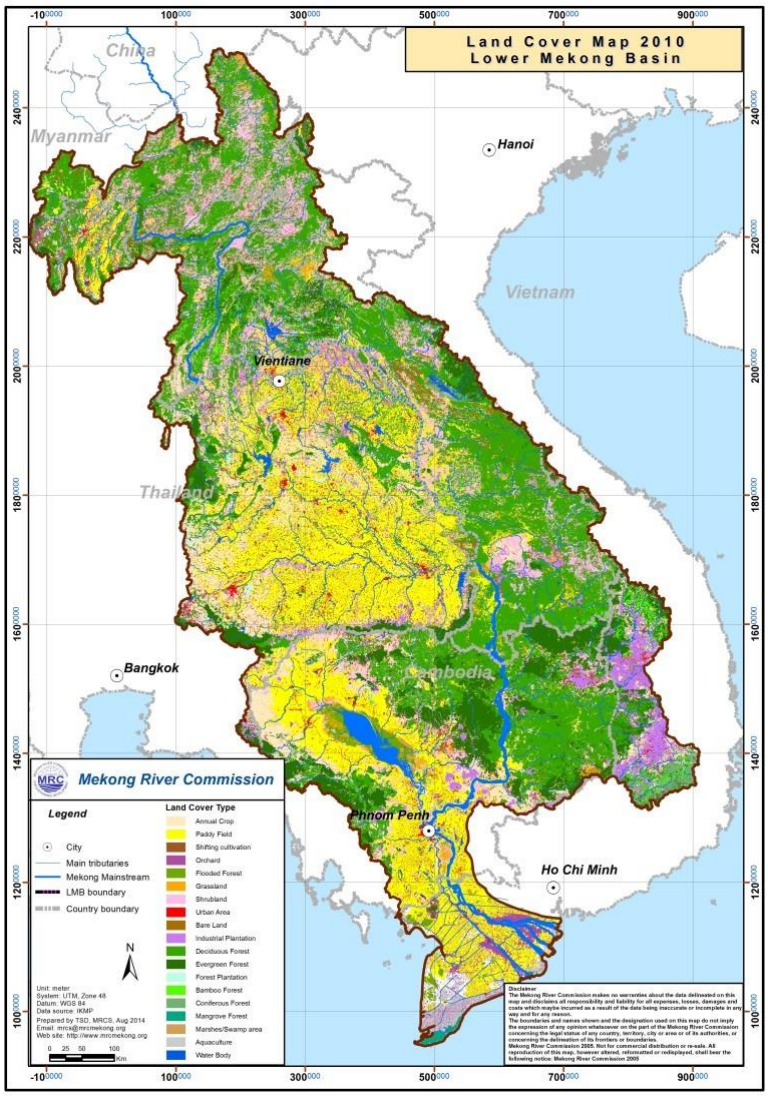
3.1.3 Land use and Land cover

The river and its tributaries are constrained by narrow, steep gorges in the upper Mekong River Basin. In this region of the basin, there are few tributary river systems. The river alters as the floodplain widens, the valley expands, and the river widens and slows. Soil erosion is the main issue here. The primary determinants of the river's hydrology are the climate, geography, and land use of the lower Mekong Basin. The soils recover after upland shifting agriculture (slash and burn), but the vegetation takes much longer. Similar to other areas of the basin, shifting and permanent cultivation have gradually decreased the amount of forest cover over the past three decades. Over the past 50 years, Thailand's Lower Basin regions have had the greatest rate of forest cover loss among all the countries along the Lower Mekong. The land use and land cover map of the Mekong Basin is presented in Figure 3-5.

The most prevalent soil types are salty and sandy, making a large portion of the area unsuitable for growing wet rice. However, agriculture is intense despite the low fertility. The main crops are cassava, maize, and glutinous rice. The main hydrological risk in this area is drought.

Figure 3-5

Land Use and Land Cover Map of Mekong River Basin (MRC, 2010)



3.2 Data

3.2.1 In-situ River Discharge

In this study, in-situ data of daily discharge is obtained from Mekong River Commission (MRC) (https://mrcmekong.org/) both for training and validation purposes. Nine locations on Mekong River is used whose details is shown in Table 3-3 and is sorted according to the river's flow direction (1-9). These hydrological stations represent the different hydrological characteristics in the basin. These data are used to develop a regional model. Further, last two-gauge data (10-11) is used to validate the predictive accuracy of regional model. The location of these stations is indicated in Figure 3-1.

Table 3-3

List of in Situ Gauge Stations in the Mekong Mainstream used this Study

S.N.	Name of Station	Station Code	Basin Coverage	Location		Streamflow Record
			km^2	Latitude (°N)	Longitude (°E)	
1	Chiang Saen	TH_010501	213027	20.273	100.083	Jan 1, 2002 to Dec 31, 2018*
2	Luang Prabang	LA_011201	237938	19.892	102.137	
3	Nong Khai	TH_012001	304990	17.887	102.739	
4	Nakhnon Phanom	TH_013101	353753	17.40	104.800	
5	Mukdahan	TH_013402	399746	16.540	104.737	
6	Pakse	LA_013901	510889	15.117	105.800	
7	Stung Treng	KH_014501	624016	13.545	106.017	
8	Kratie	KH_014901	726763	12.240	105.987	
9	Tan Chau	VN_019803	768492	10.803	105.243	
10	Jinghong	CN_092600	153386	22.00	100.770	
11	Chiang Khan	TH_011903	295193	17.884	101.612	

**The period varies based on data availability*

3.2.2 Satellite Altimetry Data for Water Level

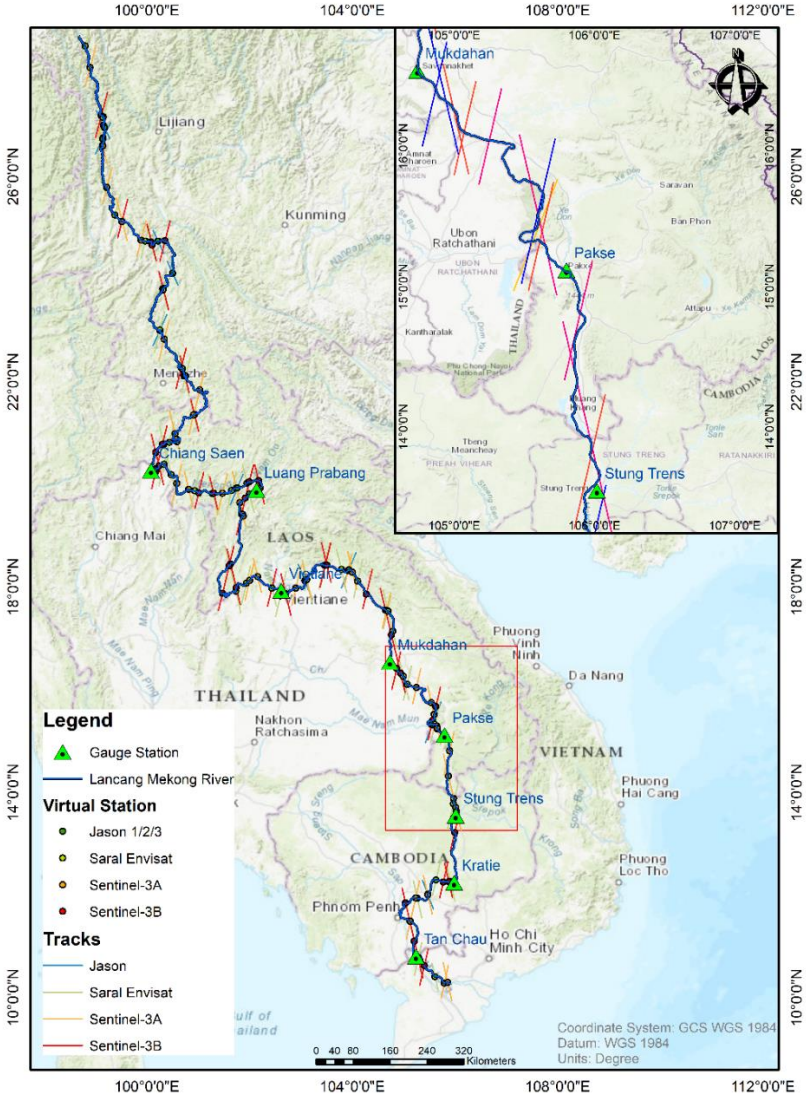
This study uses Jason-1/2, Envisat and Saral altimetry products to derive time series water level from 2002 to 2018 since these datasets are freely available by the service provider. Also, from these altimetry products, sufficient virtual stations is located along the Lancang Mekong River. Figure 3-6 shows the altimeter tracks crossing Lancang Mekong River and Virtual Stations. Envisat provides data from 2002 to 2010 and has an orbital period of 35 days. The 18-Hz along-track range data in the Geophysical Data Record (GDR), which is publicly available from the Center for Topographic Studies of the Ocean and Hydrosphere (CTOH; http://ctoh.legos.obs-mip.fr/data/alongtrack-data/datarequest) is used for Envisat.

Jason-1 and Jason-2 orbit at an altitude of 1336 km, with an inclination of 66°, on a 10-day repeat cycle, providing observations of the Earth surface (ocean and land) from 66° latitude North to 66° latitude South, with an equatorial ground-track spacing of about 315 km.

A collaboration between Centre National d'Etudes Spatiales (CNES) and the National Aeronautics and Space Administration (NASA) launched the Jason-1 mission on December 7th, 2001. The Jason-1 mission's sensors are based on the Poseidon-2 altimeter, a two-frequency altimeter with C (5.3 GHz) and Ku (13.575 GHz) bands that was part of the previous Topex/Poseidon missions. Along with the Jason Microwave Radiometer from NASA, the Doppler Orbitography and Radio-positioning Integrated by Satellite (DORIS) instrument from CNES, the Black Jack Global Positioning System receiver from NASA, and a Laser Retroflector Array (LRA) from NASA/Jet Propulsion Laboratory (JPL) make up its payload, which allows for precise orbit determination. Jason-1 was decommissioned on June 21, 2013, and it stayed in its intended orbit until 26 January 2009.

Figure 3-6

Virtual Stations Track along the Lancang Mekong River



A collaboration between CNES, NASA, Exploitation of Meteorological Satellites (EUMETSAT), National Oceanic and Atmospheric Administration (NOAA) launched Jason-2 mission on June 20, 2008. The Advanced Microwave Radiometer (AMR) from JPL/NASA, the Global Navigation Satellite System (GNSS), and an LRA from JPL/NASA make up most of its payload, which also includes the real-time monitoring system DIODE of DORIS device from CNES, a GNSS transceiver from NASA/JPL, and the Poseidon-3 radar altimeter from CNES. Jason-2 remained in its nominal orbit until 3 July 2016.

SARAL (Satellite with ARgos and ALtiKa) is a mission jointly undertaken by the Indian Space Research Organisation (ISRO) and Centre National d'Études Spatiales (CNES) to study ocean circulation and sea surface elevation through altimetric measurements. The mission was launched on 25 February 2013, and its prime payload, the ALtiKa altimeter, is the first spaceborne altimeter to operate at Ka-band. The ALtiKa altimeter was designed and built by CNES, and is a key component of the SARAL mission. Table 3-4 provides the major characteristics of altimetry missions used in this study.

Table 3-4

Major Features of the Radar Altimetry Missions used in this Study

Mission	**Jason-1/2/3**	**ENVISAT**	**Saral**
Instrument	Poseidon-2 Poseidon-3 Poseidon-3B	Radar Altimeter (RA-2)	Altika
Space agency	Centre National d'Etudes Spatiales (CNES) National Aeronautics and Space Administration (NASA)	European Space Agency (ESA)	Centre National d'Etudes Spatiales (CNES) Indian Space Research Organization (ISRO)
Operation	2001–2013 Since 2008 Since 2016	2002–2012	Since 2013

Mission	Jason-1/2/3	ENVISAT	Saral
Acquisition mode	Low Resolution Mode (LRM)	LRM	Pulse Limited Mode
Acquisition	Along-track	Along-track	Along-track
Frequency (GHz)	13.575 (Ku) 5.3 (C)	13.8 (Ku) 13.575 (Ku) 3.2 (S)	35.75 (Ka)
Altitude (km)	1315	800	800
Orbit inclination (°)	66	98.55	98.55
Repetitively (days)	9.9156	35	35
Equatorial cross-track separation (km)	315	80	80

3.2.3 Satellite Optical Data Sets

The different satellite missions used in this study for the optical images are moderate-resolution imaging spectroradiometer (MODIS); MODIS AQUA and TERRA. Due to the coarse spatial resolution, MODIS has provided better results in large rivers with a width of more than 100 m for C/M estimation and thus selected for Mekong River. The optimal location for selecting wet pixel is the river area and inundation sensitive area during flood events ensuring minimum vegetation coverage while for dry pixel, area should be located far from river area preferably urban areas (Shi et al., 2020; Tarpanelli, Brocca, et al., 2013).

The MODIS is among the sensors on board the Earth Observing System (EOS) Terra (since 1999) and Aqua (since 2002) satellites. Due to its high temporal resolution of 1-2 days with occasionally two passes during a day at mid-latitude (3 h apart from each other), its moderate spatial resolution (2 channels at 250 m, 5 at 500 m, and 29 at 1 km), and its high spectral resolution of 36 bands ranging in wavelength from 0.4 m to 14.4 m, it is widely used for monitoring a variety of terrestrial, atmospheric, and ocean phenomena. This study uses level-2 products MOD09GQ and MYD09GQ from TERRA and AQUA, respectively, at daily resolution.

Table 3-5 provides the major characteristics of optical satellite sensors used in this study.

Table 3-5

Major Features of the Optical Satellite Sensors used in this Study

Optical Sensor	Product	Band (Spectral range in nm)	Temporal Resolution (days)	Spatial Resolution (m)
MODIS AQUA	MYD09GQ	2 (841-876)	1-2	250
MODIS TERRA	MOD09GQ	2 (841-876)	1-2	250

3.2.4 Basin Characteristics

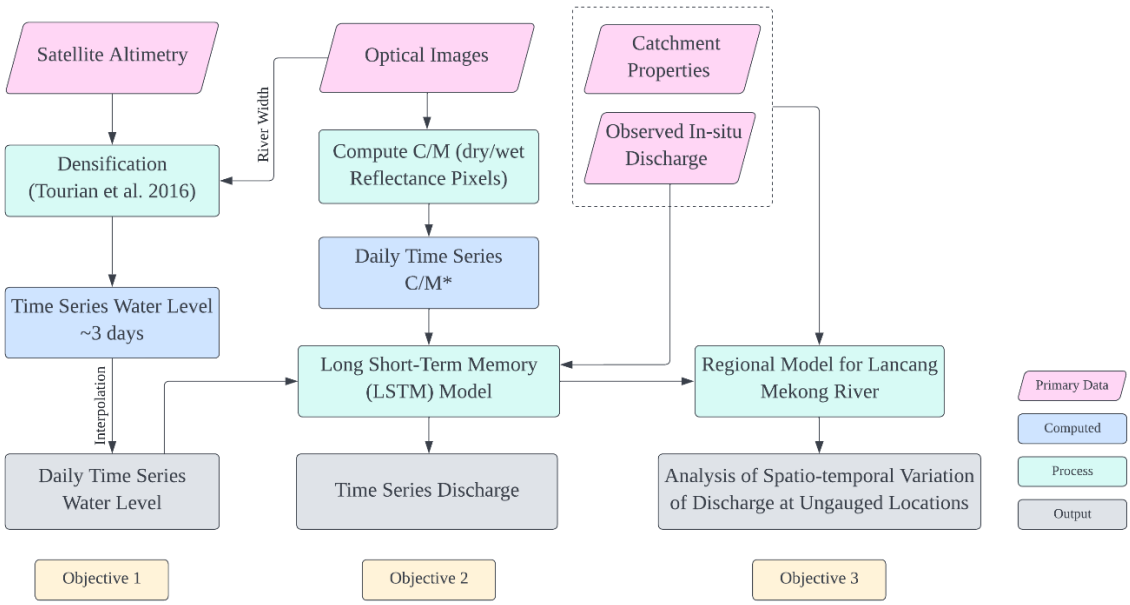
The characteristics of basin draining to each of the hydrological stations used in this study are used as an input to LSTM model while developing regional model. The parameters that are used are catchment area, catchment centroid latitude, catchment centroid longitude, average catchment elevation, difference between highest and lowest elevated points in a catchment, average catchment slope, curve number and average mainstream bed slope gradient. The parameters are computed from available DEM and land use data.

CHAPTER 4
METHODOLOGY

The general procedure for river discharge estimation is shown in Figure 4-1. The details of the methodology are explained in this chapter. Long Short-Term Memory machine learning algorithm is employed for predicting river discharge and is evaluated. The input to the model is water level data obtained from satellite altimetry, reflectance ratio of dry and wet pixels from optical images and other catchment parameters. The performance of the model to simulate river discharge is evaluated using different statistical indicators.

Figure 4-1

Overall Methodological Framework for the study

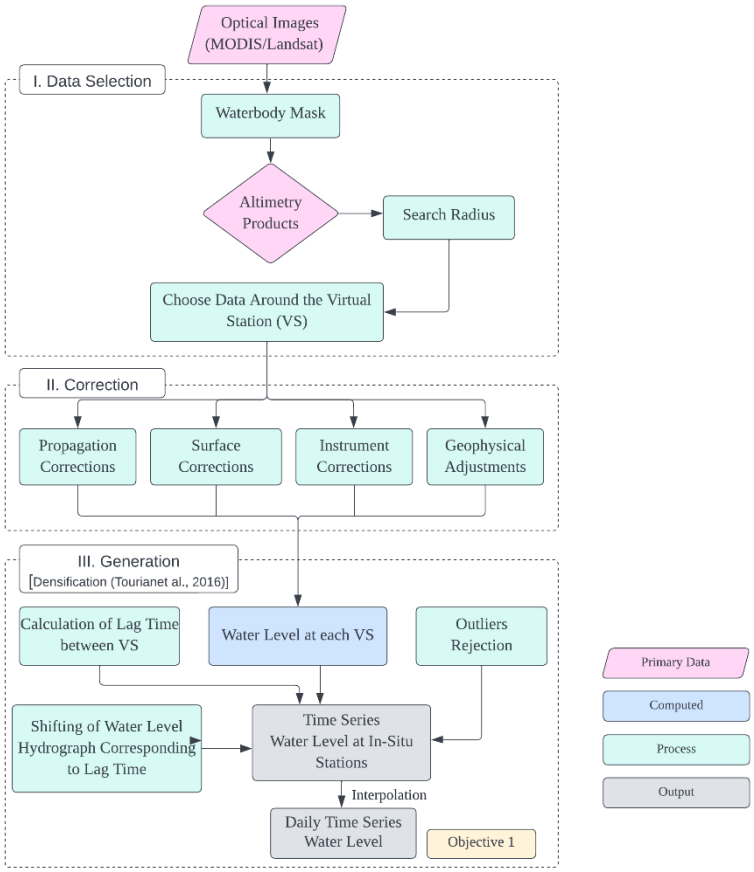


4.1 Estimation of water level at different stations in Mekong River

The methodological approach to estimate the water level at different reach of river using multi-mission satellite altimetry is shown in Figure 4-2.

Figure 4-2

Methodological Approach to Estimate Water Level using Satellite Altimetry (Objective 1)



4.1.1 Time Series Water Level from Satellite Altimetry

Particularly in inaccessible and complex areas, satellite remote sensing offers the exceptional opportunity to acquire freely accessible water level heights. Numerous altimetric missions, including TOPEX/Poseidon, Jason-1/2/3, ERS-1/2, Environmental Satellite (Envisat), CryoSat-2, SARAL/AltiKa, and Sentinel 3A, 3B, have contributed data over the past two decades that can be utilized to track the behavior of water bodies on Earth (Garkoti & Kundapura, 2021). The increased accuracy of the altimetry sensors encourages its use as a validation tool for many applications, from straightforward routing approaches to complex hydraulic models, and recent advancements in radar altimetry technology provide crucial information for river water level monitoring (Domeneghetti et al., 2015; Schneider et al., 2017).

The water surface level is calculated from multiple satellite altimetry data at the corresponding time. Altimeter measurements of surface topography are distorted, and thus corrections are applied. For example, atmospheric propagation effects in the

troposphere and the ionosphere, electromagnetic bias, residual geoid errors, and inverse barometer effects can all distort measurements (Sichangi et al., 2016). Considering propagation delays from the interactions of electromagnetic waves in the atmosphere and geophysical corrections, the height of the reflecting surface (H) with respect to a reference ellipsoid is estimated. Thus, the types of corrections to be applied are geophysical corrections for wet troposphere, dry troposphere, ionosphere, solid earth tides and pole tides. The accuracy of obtained water level is evaluated comparing with in-situ water level for available locations.

4.1.2 Densification of Altimetry derived Time Series Water Level at In-situ Stations

Multi-mission altimetry products are used to increase the temporal resolution of water level in this study. The time series water level is derived at various Virtual Stations (VS) from different satellite missions. The water level obtained from different satellite mission have bias within due to mission's geophysical data and these intersatellite bias is estimated and removed and other errors of the altimetry data is corrected as mentioned. Then, the Virtual Stations (VS) are connected hydraulically and statistically using an algorithm developed by (Tourian et al., 2016, 2017) using the time series water level data, river width and slope. For this, time lag between virtual stations is estimated.

The time lag between VSs is determined using the anticipated river width and the slope obtained from satellite altimetry using the Equation 4-1.

$$T_L = \frac{L}{C} = \frac{L}{bV} \qquad \text{Equation 4-1}$$

Where, T_L is travel time, L is the distance between the virtual stations, c is celerity, b = constant=5/3 and V is the velocity. The velocity is computed by the Equation 4-2 developed by (Tourian et al., 2016).

$$V = 1.48W^{-0.8}S^{0.6} \qquad \text{Equation 4-2}$$

where, W refers to river width and S is the slope of the river.

Then, at any location along the river, integration of all the altimetric readings is done by stacking all the obtained altimetric measurements at a reference location by shifting the water level hydrographs of all virtual stations according to corresponding time lag. The measurement is merged by normalizing the time series in accordance with their

statistical properties. The readings are rescaled down to their actual water level values after an outlier identification procedure.

River Width Estimation:

Google Earth Engine (GEE) based river width algorithm RivWidthCloud developed by Yang et al. (2020) is used to extract river centerline and width from remotely sensed images. The simplicity of usage based on a well-liked cloud computing environment GEE and the flagging power to automatically reduce negative impact from cloud and shadows are two key advantages of RivWidthCloud over previous river width measuring techniques. These benefits spare users the expense of downloading, storing, and locally processing remote sensing data, enabling them to easily extract common width statistics and time series.

Firstly, extraction of river mask from satellite image is done. Then, river centerline from the river mask is derived and river width along the centerline is obtained at each river section. To calculate the river width, for each centerline pixel the direction orthogonal to the local centerline is computed at first and then river width along these orthogonal directions is computed. The extracted centerline is also used to define the along-river coordinate distance between Virtual Station (VS)s and to compute the slope.

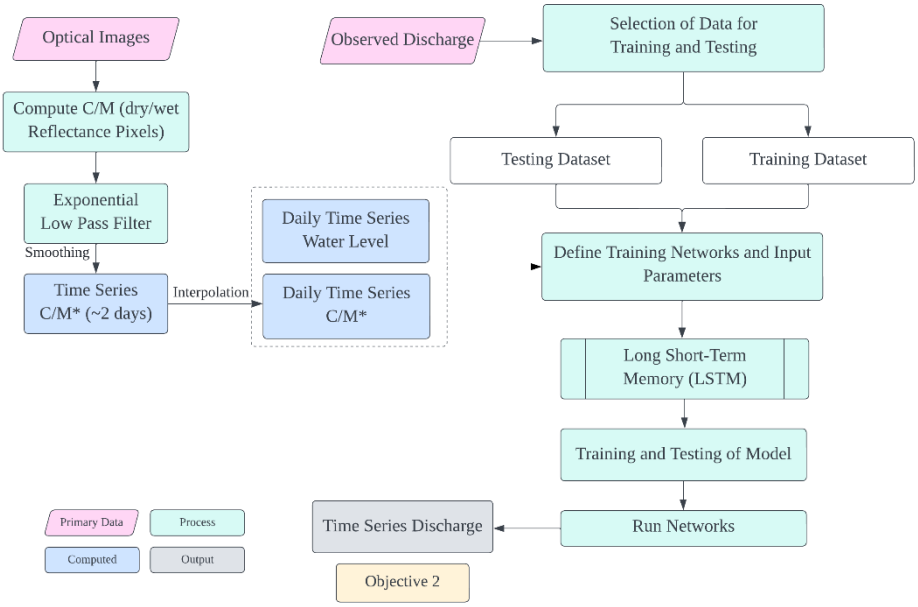
4.2 Daily discharge estimation by merging satellite optical and altimetry data

The methodological framework to accomplish this objective is shown in Figure 4-3.

The surface reflectance between water and land from Near Infra-Red (NIR) band of optical sensors will have different pixel values. The water pixel (M) has lower reflectance than land pixel (C) in which reflectance value of water pixel decreases with flooding events while land pixel doesn't vary much and assumed to be constant (Tarpanelli et al., 2017). Thus, measuring time series variation in reflectance of water pixel with respect to land pixel helps to provide variation in river discharge indirectly. The ratio between the spectral reflectance of the C and M pixels should theoretically allow a minimizing of the atmospheric effects because disturbances derived from the atmosphere influence both C and M in the same way (Tarpanelli, Brocca, et al., 2013). The values of C/M increase with the presence of water and, hence, with discharge. Hence, ratio of C and M is taken as the proxy for measurement of river discharge.

Figure 4-3

Methodological Framework for Discharge Estimation Merging Optical Sensor Data and Altimetry Data (Objective 2)



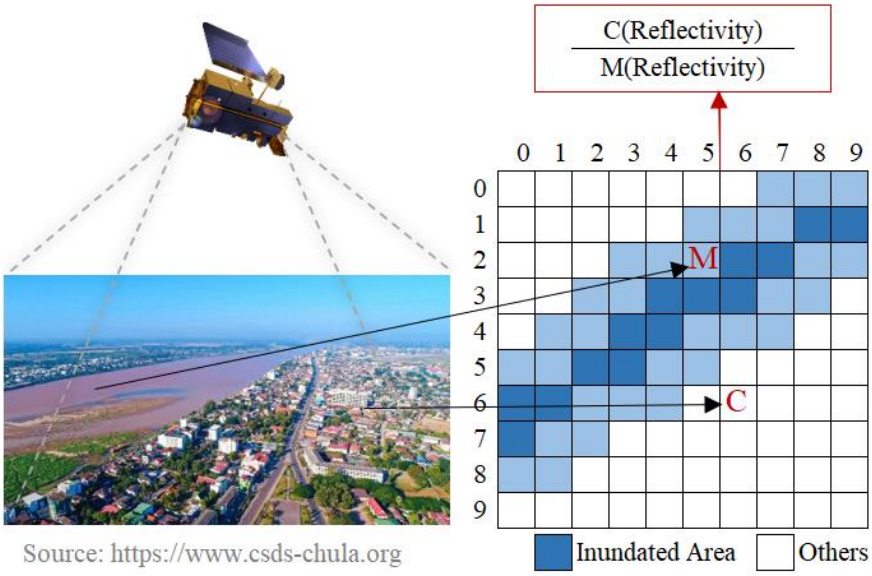
A cloudy sky, on the other hand, renders the optical pictures useless and causes data outages in the measurements. Also, the selected location of C and M pixels may also affect the accuracy of discharge estimation. The data from more satellite missions of the same kind might be gathered and combined to provide the necessary information as a feasible solution in these circumstances. Thus, this study used Long Short Term Memory (LSTM) approach to combine observations from optical sensors with altimetry data for estimation of river discharge to overcome these issues.

4.2.1 C/M from Satellite Optical Data Sets

The ratio of the surface reflectance of a land pixel named C (calibration), situated close to the river in an area free of surface water even during high flooding, and of a water pixel named M (measurement), situated within the river with the permanent presence of water, is computed. Figure 4-4 shows the typical representation of wet and dry pixel reflectivity. First, the cloud-affected pixels are located and eliminated. Then, the location for selecting M and C pixels is chosen. The optimal site for wet pixel M is next to a river in a region that is both entirely submerged in water and sensitive to changes in the flooded area during flood occurrences. The calibration's dry pixel C is situated outside the river, above urban areas that are not impacted by the seasonal cycle of vegetation, and in places that are not surrounded by water.

Figure 4-4

A representation of wet and dry pixel reflectivity



To create a time series of C/M, multiple images for the study period are evaluated. Then, to reduce the noise effects due to atmospheric contribution, time series of C/M is smoothed with a low pass filter (averaging moving window) to get C/M*.

4.2.2 Merging Multiple Satellite Datasets

All the datasets should have common time series for successful merging of different data from various sources. While the water level altimetry time series is restricted to the passage of satellites over the river, the C/M* ratio retrieved from optical satellite sensors is impacted by cloudy images. These circumstances result in missing data, which breaks up the time series. In addition, the accuracy of the altimetry multi-mission satellites might vary based on the altimeter utilized and the number of satellites available at the same time. Due to the satellites' repetition cycle, the time series generated by optical sensors have different temporal resolutions. When compared to altimetric time series, MODIS-derived time series have a better temporal resolution. The expected temporal resolution of MODIS-derived data is approximately 2 to 3 days.

Therefore, to overcome this issue, all satellite data sets is interpolated at the daily scale to provide time series with the same frequency that are consistent with the temporal resolution of ground observations. Daily data gaps are linearly filled up by interpolating values between the closest previous and next values.

The LSTM methodology is used to carry out the merging process. With this method, data from many sources is combined into a single retrieval strategy by simply adding or removing inputs from the LSTM configuration and updating the training accordingly. The daily data sets function the inputs to the LSTM: river water level and C/M*. The number of hidden layers and hidden neurons is established by trial-and-error procedure. The training set and validation set is sampled randomly.

Long Short-Term Memory (LSTM):

LSTM networks are specifically designed to learn long-term dependencies and can overcome the problems of vanishing and exploding gradients. LSTM networks are composed of an input layer, one or more memory cells, and an output layer (Zebin et al., 2018). The number of neurons in the input layer is equal to the number of explanatory variables. The main characteristic of LSTM networks is contained in the hidden layer consisting of so-called memory cells. Each of the memory cells has three gates maintaining and adjusting its cell state s_t: a forget gate (f_t), an input gate (i_t), and an output gate (o_t).

At every time-step t, each of the three gates is presented with the input x_t (one element of the) as well as the output h_{t-1} of the memory cells at the previous time-step $t-1$. Hereby, the gates act as filters, each fulfilling a different purpose:

- The forget gate defines what information is removed from the cell state.
- The input gate specifies what information is added to the cell state.
- The output gate specifies what information from the cell state is used

Each of the gates has parameters for its weights and biases, giving many parameters for deep networks with many units' present. The weights of these connections are learned or updated during the training of the network.

4.2.3 Evaluation of the performance of Long Short-Term Memory (LSTM) Model

The performances of the LSTM model to simulate discharge at various sites of Lancang Mekong River is evaluated based on the correlation coefficient (R), coefficient of determination (R^2), Nash-Sutcliffe efficiency (NSE) and percent bias (PBIAS), and visual interpretation using a line diagram and scatter diagram. This is done by comparing the observed discharge at in-situ stations with simulated discharge.

$$R^2 = \frac{n\sum Q_i^{obs} Q_i^{pre} - \sum Q_i^{obs} \sum Q_i^{pre}}{\left(\sqrt{n(\sum Q_i^{obs})^2 - (\sum Q_i^{obs})^2}\right) * \left(\sqrt{n(\sum Q_i^{pre})^2 - (\sum Q_i^{pre})^2}\right)}$$ Equation 4-3

$$NSE = 1 - \frac{\sum_{i=1}^{n} (Q_i^{obs} - Q_i^{sim})^2}{\sum_{i=1}^{n} (Q_i^{obs} - \overline{Q_i}^{obs})^2}$$ Equation 4-4

$$PBIAS = \frac{\sum_{i=1}^{n} (Q_i^{obs} - Q_i^{pre}) * 100}{\sum_{i=1}^{n} (Q_i^{obs})}$$ Equation 4-5

where, Q_i^{obs} is observed discharge, Q_i^{pre} is predicted discharge and $\overline{Q_i}^{obs}$ is the average observed discharge.

Coefficient of Determination (R^2):

The coefficient of determination describes the statistical relationship between the variables and helps to show the nature of association among the predicted and observed data. R^2 is the ratio of explained variation compared to the total variation. It ranges from 0 to 1; its higher value indicates less error variance, and generally, a value greater than 0.5 is considered acceptable. This statistical tool is highly sensitive to outliers and insensitive to additive and proportional differences between observed and predicted data.

Nash-Sutcliffe Efficiency (NSE):

The Nash-Sutcliffe Efficiency (NSE) is a metric used to evaluate the agreement between observed and simulated data. The NSE value ranges from negative infinity to 1, with a value of 1 indicating a perfect match between the simulated and observed data. In contrast, if the NSE value is negative, it implies that the average observed value provides a better estimate than the model, indicating poor predictions (Nash & Sutcliffe, 1970).

Percent Bias (PBIAS):

Percent bias measures the relationship between the observed data and its predicted data; it measures the average tendency of observed data to be larger or smaller than the predicted data. Percent bias describes whether the simulated model is overestimated or underestimated. A low PBIAS value or a value that tends to zero indicates the optimal model. A negative value indicates the overestimation of the model. In contrast, a

positive PBIAS value indicates an underestimation of the model. When the data are evaluated, PBIAS reveals any deviation of the data as a percentage.

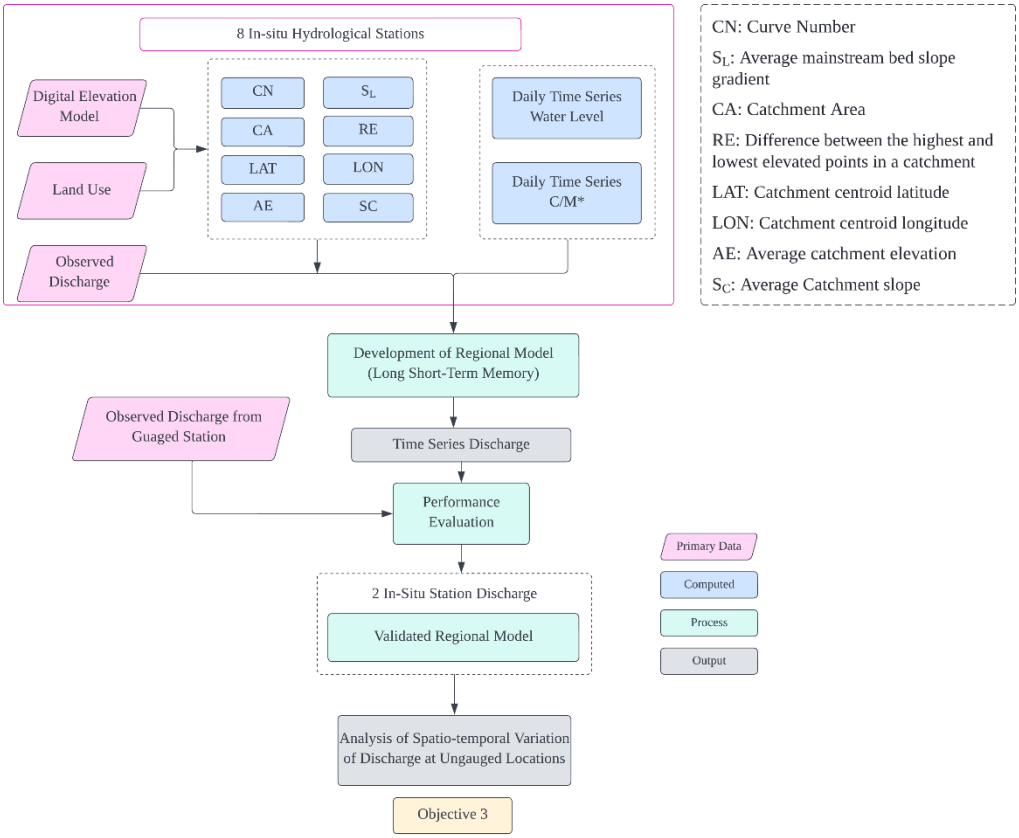
A model with higher R^2 and NSE values, and PBAIS values, decrees a relatively better model for the simulation of discharge.

4.3 Development of a regional model for estimation of the river discharge

The methodological framework to achieve this objective is presented in Figure 4-5.

Figure 4-5

Methodological Framework for Development of Regional Model for Discharge Estimation (Objective 3)



There are limited number of hydrological gauging stations in Mekong River, and it is not always feasible to measure discharge at every location in the river reach. On other hand, it is not always easy to obtain gauge data due to various reasons. However, developing regional model considering parameters obtained from remote sensing and catchment characteristics for entire basin will help to predict discharge at any required location without physically being present. The initial lists of catchment parameters that

are expected to have impact on discharge are provided in Table 4-1. Sensitivity analysis of these parameters is carried out to obtain most sensitive parameters that will have profound effect on discharge and final catchment parameters are obtained.

Then, a regional model will be established by integrating the catchment parameters along with time series water level and C/M* for nine in-situ stations within the river reach. This model could then be applied simply but with strong predictive power to estimate catchment-scale discharge in ungauged reach of Lancang Mekong River. Physical catchment descriptors and model parameters are known to be interdependent. Also, it is well known that their interactions are highly nonlinear; therefore, LSTM can be a suitable tool for the regionalization of model parameters (Soni et al., 2021). Thus, the integration will be done using the best selected machine learning approach from Section 4.2.3. The performance of regional model will be tested against other two gauged hydrological stations to check the predictive accuracy using the statistical indicators.

Table 4-1

Basin attributes and parameters used for regionalization

S.N.	Parameters	Acronym	Remarks
1	Catchment Area	CA	
2	Catchment centroid latitude	LAT	
3	Catchment centroid longitude	LON	
4	Average catchment elevation	AE	
5	Difference between the highest and lowest elevated points in a catchment	RE	
6	Average Catchment slope	S_C	
7	Curve Number	CN	
8	Average mainstream bed slope gradient	S_L	

CHAPTER 5
RESULTS AND DISCUSSIONS

5.1 Water Level Estimation from Satellite Altimetry

Satellite altimetry offers near-global profile data along tracks, confined solely by the satellite's orbital parameters, yet restricted to its ground track by the nadir-only measurement capability of the sensors. Thus, initially location where the path of satellites intersected with Lancang Mekong river were focused. These points were designated as Virtual Stations (VS) and water level data was collected for each of these virtual stations using satellite altimeter missions from ENVISAT, Jason-1, Jason-2, and SARAL. This data was then processed and analyzed to gain insight into the water levels at 54 different Virtual Stations. The location of virtual stations for different altimetry are presented in Figure 5-1 and details are presented in Table 5-1. Radar altimeters do not determine the range value at each point by direct measurement. Instead, they rely on the echo of a radar signal to extract the range value. The data rate for each of these missions differ, with ENVISAT having 18 Hz, Jason-2 having 20 Hz, and SARAL/AltiKa having 40 Hz, resulting in along-track distances of 380 m, 294 m, and 173 m respectively between altimeter measurements.

Figure 5-1

The ground track and the selected virtual stations from Jason-1/2 (left) and ENVISAT, SARAL (right)

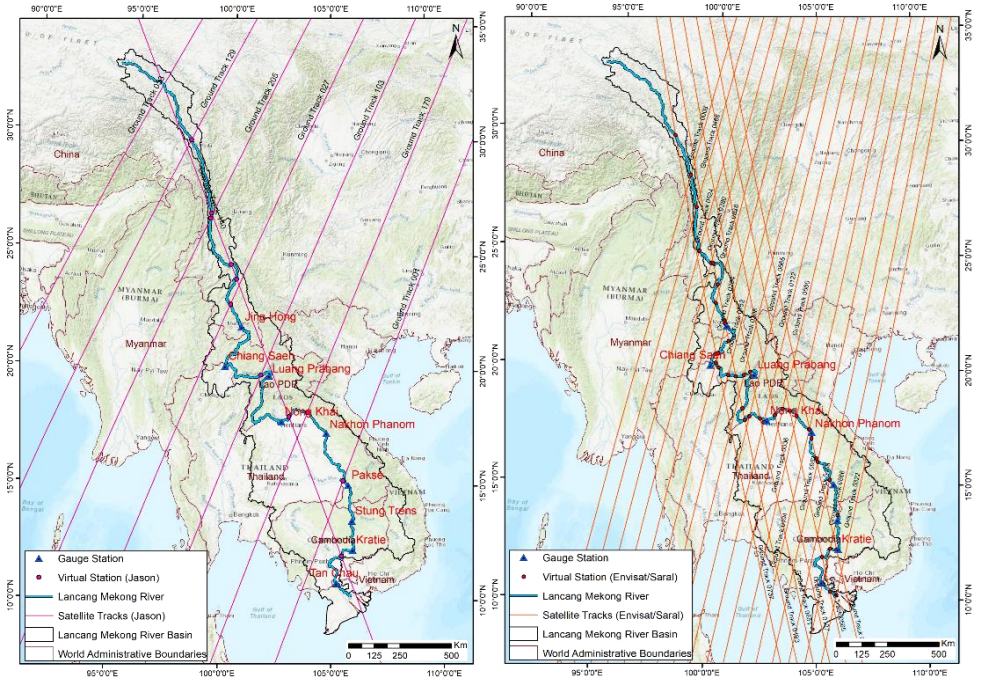


Table 5-1

List of virtual station along with chainage and location along Mekong River

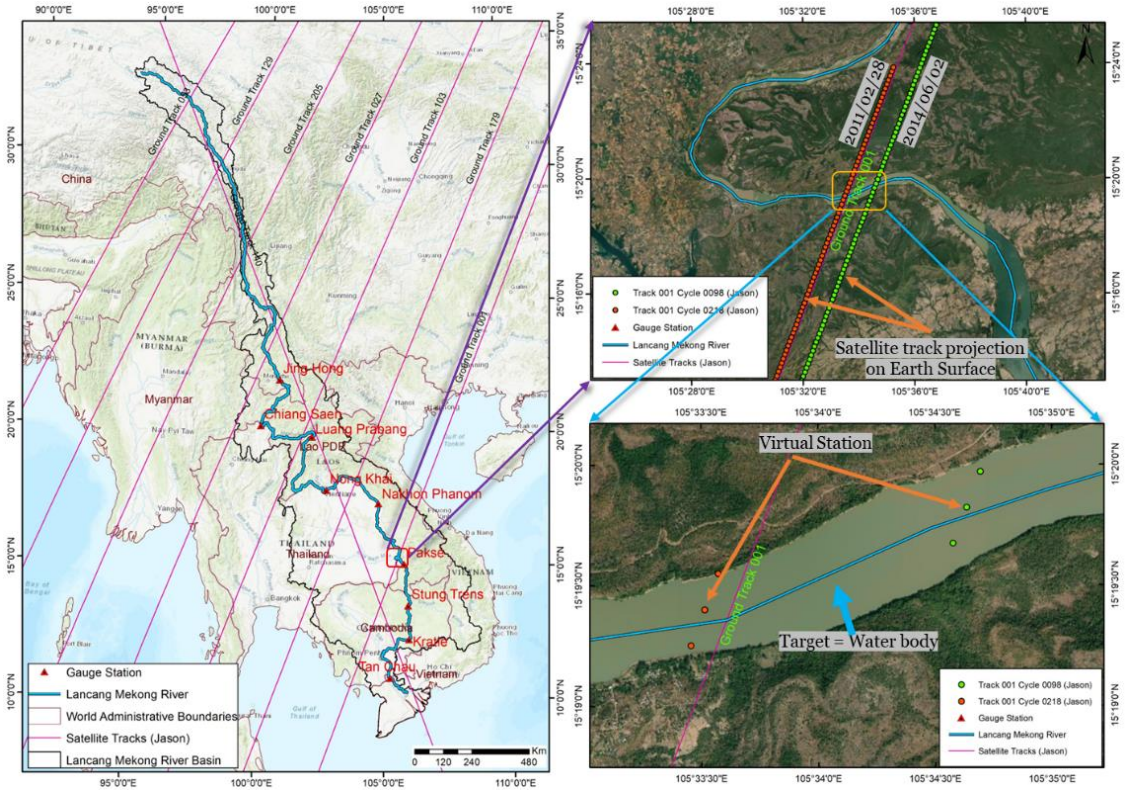
SN	Chainage (km)	Satellite Mission	Track Number	Latitude	Longitude
1	0	Envisat/Saral	0737	30.231	97.922
2	18	Jason	129	30.105	97.992
3	255	Envisat/Saral	0008	28.544	98.797
4	464	Envisat/Saral	0466	27.161	99.152
5	495	Jason	140	26.955	99.176
6	523	Jason	205	26.735	99.140
7	742	Envisat/Saral	0924	25.254	99.379
8	836	Envisat/Saral	0380	24.744	99.970
9	847	Envisat/Saral	0193	24.730	100.078
10	873	Jason	140	24.744	100.218
11	980	Jason	140	24.119	100.504
12	1037	Envisat/Saral	0193	23.812	100.308
13	1159	Jason	027	23.009	100.254
14	1159	Envisat/Saral	0838	23.009	100.255
15	1282	Envisat/Saral	0193	22.256	100.693
16	1300	Envisat/Saral	0193	22.116	100.727
17	1303	Envisat/Saral	0294	22.116	100.755
18	1460	Envisat/Saral	0193	21.356	100.912
19	1586	Envisat/Saral	0294	20.825	100.441
20	1601	Envisat/Saral	0737	20.786	100.331
21	1712	Envisat/Saral	0294	20.393	100.337
22	1742	Envisat/Saral	0737	20.192	100.474
23	1840	Envisat/Saral	0752	19.879	100.931
24	1885	Envisat/Saral	0193	19.826	101.281
25	1933	Envisat/Saral	0208	19.883	101.652
26	1947	Jason	103	19.934	101.762
27	1970	Envisat/Saral	0651	20.027	101.951
28	2045	Envisat/Saral	0651	19.819	102.001
29	2233	Envisat/Saral	0193	18.427	101.614
30	2363	Envisat/Saral	0193	17.923	101.733
31	2403	Envisat/Saral	0666	18.078	101.940
32	2476	Envisat/Saral	0651	17.982	102.437
33	2509	Envisat/Saral	0122	17.833	102.603
34	2590	Jason	140	18.126	103.077

SN	Chainage (km)	Satellite Mission	Track Number	Latitude	Longitude
35	2595	Envisat/Saral	0107	18.151	103.117
36	2662	Envisat/Saral	0580	18.426	103.460
37	2699	Envisat/Saral	0565	18.346	103.788
38	2720	Jason	179	18.334	103.934
39	2753	Envisat/Saral	0036	18.117	104.107
40	2845	Envisat/Saral	0494	17.537	104.688
41	2847	Envisat/Saral	0021	17.531	104.700
42	2894	Envisat/Saral	0021	17.140	104.792
43	3006	Envisat/Saral	0021	16.262	104.996
44	3027	Envisat/Saral	0952	16.105	105.070
45	3190	Jason	001	15.323	105.561
46	3196	Envisat/Saral	0408	15.326	105.610
47	3394	Envisat/Saral	0866	13.782	105.972
48	3592	Envisat/Saral	0021	12.262	105.915
49	3624	Envisat/Saral	0866	12.296	105.634
50	3664	Jason	140	12.013	105.473
51	3698	Envisat/Saral	0565	11.938	105.268
52	3886	Envisat/Saral	0866	10.783	105.292
53	3947	Envisat/Saral	0565	10.409	105.613
54	3986	Envisat/Saral	0322	10.274	105.898

Each VS is in fact formed as a box sized with corresponding river width (Figure 5-2). However, since ground tracks cross the river at non perpendicular angles at most VSs, the size of VS is sometimes different from the river width. The water level time series at the VS locations are then obtained by averaging only the water related measurements inside the VS boxes.

Figure 5-2

The ground track of two cycles of Jason-2 representing water related measurement around virtual station



The missions ENVISAT and SARAL succeeded the previous mission ERS-2, employing the same orbital configuration with a cross-track resolution below 80 km and a repeat cycle of 35 days. ENVISAT was equipped with a dual frequency radar altimeter that operated in Ku and S bands. However, in 2010, its orbital height was lowered, resulting in a different repeat cycle and ground tracks. The study utilized only the high frequency (18 Hz) Ku-band data collected from 2002 to 2010 during regular operation mode. The altimetry data were stored in the Sensor Geophysical Data Records (SGDRs) format and contained parameters such as satellite position and timing, distance between satellite and surface, and corrections. The ENVISAT had four different ranges calculated by standard waveform retracking algorithms, including Ocean, ICE-1, ICE-2, and SEAICE. The ICE-1 retracker was initially developed to observe ice sheets but was used in this study to calculate range measurements for inland waterbodies due to their similar waveform.

Jason-2, launched in June 2008, is the successor to the previous missions Topex/Poseidon and Jason-1. It employs the same orbital configuration with a cross-

track resolution of less than 315 km and a repeat cycle of 10 days. To estimate water levels, high-frequency ranges from Sensor Geophysical Data Records (SGDR) altimeter products are used. However, it is known that switching retracking algorithms along a single satellite track results in height offsets. To avoid these offsets, ICE1 retracker is used to obtain all altimeter range measurements of a specific inland water body under investigation. By accounting for the propagation delays of electromagnetic waves due to their interaction with the atmosphere and geophysical entities, a formula is employed to estimate the height of the reflecting water surface relative to an ellipsoid or a geoid.

$$h = H - R - C_{ionosphere} - C_{drytroposphere} - C_{wettroposphere} - Cs_{olidEarthtide} - C_{poletide}$$

where h is the height of reflecting water surface (water body or river section), H is the altitude of satellite with reference to an ellipsoid, R is the altimeter range, $C_{ionosphere}$ is the correction for delayed propagation through the ionosphere, $C_{drytroposphere}$ and $C_{wettroposphere}$ are corrections for delayed propagation in the troposphere from pressure and humidity variations, respectively, and $C_{solidEarthtide}$ and $C_{poletide}$ are corrections that account for crustal vertical motions from the solid and polar tides, respectively.

One of the primary obstacles encountered was the challenge of comparing water level measurements obtained from different satellite missions, taken at various locations and with different characteristics. The issue of intersatellite biases made it challenging to combine water level measurements from different altimetry missions. Furthermore, errors in atmospheric corrections and data interpolation along the track can lead to significant inaccuracies, especially in inland water bodies and rivers where the topography is non-uniform (Tourian et al., 2016). To address the issue of intersatellite bias, relevant studies were consulted to obtain the necessary information, which was then applied accordingly.

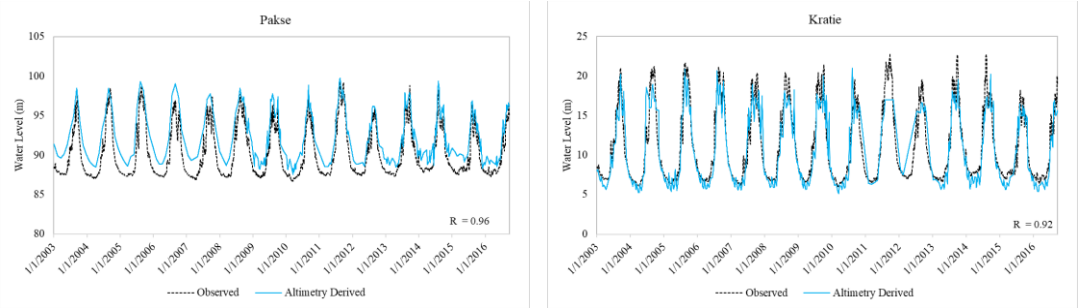
With this, water level is obtained for 54 virtual stations along the Mekong mainstream. The temporal resolution of obtained water level is 10 days for Jason and 35 days for ENVISAT and SARAL. Then, these water levels are transferred to the location of in-situ gauging stations following the densification procedure described by Tourian et al., 2016 where data from multiple satellite altimeters at VSs are connected and more detailed water level time series are obtained.

This approach utilizes ancillary measurements of average river width and slope, which are fed into a simple empirical hydraulic equation (based on Bjerklie et al. 2003). This equation estimates both the average flow velocity and the time lag caused by streamflow between the altimetric virtual stations and the chosen location along the river. Each measurement is then adjusted by the estimated time lag and shifted at the designated location. The resulting time series is standardized between the 3rd and 85th percentiles and any outliers are removed using Student's t test. Finally, the time series is scaled to match the statistical water level distribution at the selected location.

Figure 5-3 shows the altimetry derived water level and observed water level at two gauging stations: Pakse and Kratie. The correlation between observed and derived water level is 0.96 and 0.92 for Pakse and Kratie respectively.

Figure 5-3

The altimetry derived water level compared with observed water level at Pakse and Kratie



5.2 Discharge Modelling using Water Level and Optical Datasets

This study uses ratio of surface reflectance from dry and wet pixel (C/M) and altimetry derived water level as major inputs for river discharge estimation. The performance of discharge estimation with C/M ratio from AQUA and TERRA satellite and combined with altimetry water level is carried out by comparing with observed discharge as well.

5.2.1 Time Series C/M from Optical Datasets*

The ratio of reflectance from more stable land pixel (C) and variable water pixel (M) in Near Infrared (NIR) band of optical images serves as an indicator of discharge variation. This principle is adapted to train LSTM model and predict discharge based on remote sensing data. For each optical dataset AQUA and TERRA, all images were processed and analyzed using Google Earth Engine (GEE) platform and Python 3.9. Location of

C and M pixel plays pivotal role in discharge estimation and this needs to be carefully determined. Here, the locations of C and M pixels were determined adopting the method described by Tarpanelli, Brocca, et al., 2013.

An area with 38 x 38 pixel surrounding the gauging station is selected for each investigating stations. At first, cloudy images were filtered; the pixel with a value of band 1 higher than 0.2 were excluded since optical images are contaminated by clouds and their shadows which impacts surface reflectance values whereas band 2 provides the river discharge signal. Also, visual inspection was done to analysis of cloud free images. The total number of images after the cloud filtering is reduced to almost 50%. The mean temporal resolution was decreased to ~2 days from ~1 day. So, linear interpolation was applied to obtain daily time series data. Table 5-2 shows the number of images selected for analysis for each station.

Table 5-2

The number of images selected for analysis for each station.

Station	Period	Total Number of Images	Selected Images			
			Aqua		Terra	
			M Pixel	C Pixel	M Pixel	C Pixel
Chiang Saen	2003 - 2021	6939	4272	4318	3875	3804
Luang Prabang	2003 - 2018	5844	3366	3482	2929	2927
Nong Khai	2003 - 2021	6939	3952	3698	3717	3550
Nakhon Phanom	2003 - 2021	6939	3304	3526	3650	3706
Mukdahan	2003 - 2021	6939	3092	3141	3595	3396
Pakse	2003 - 2018	5844	2950	3158	3071	3261
Stung Treng	2003 - 2021	6939	3594	3743	3536	3683
Kratie	2003 - 2021	6939	3630	3878	3713	3998
Tan Chau	2003 - 2021	6939	3424	3387	3607	3498

To best locate C and M pixel for each gauging stations, C/M ratio was computed for each pixel within the box obtaining [38 x 38] x [(38 x 38) - 1] time series C/M ratios. Here, each pixel of the box was assumed to be M and remaining pixel as C and hence C/M time series were calculated in all the pixels. Similarly, the ratio was calculated for

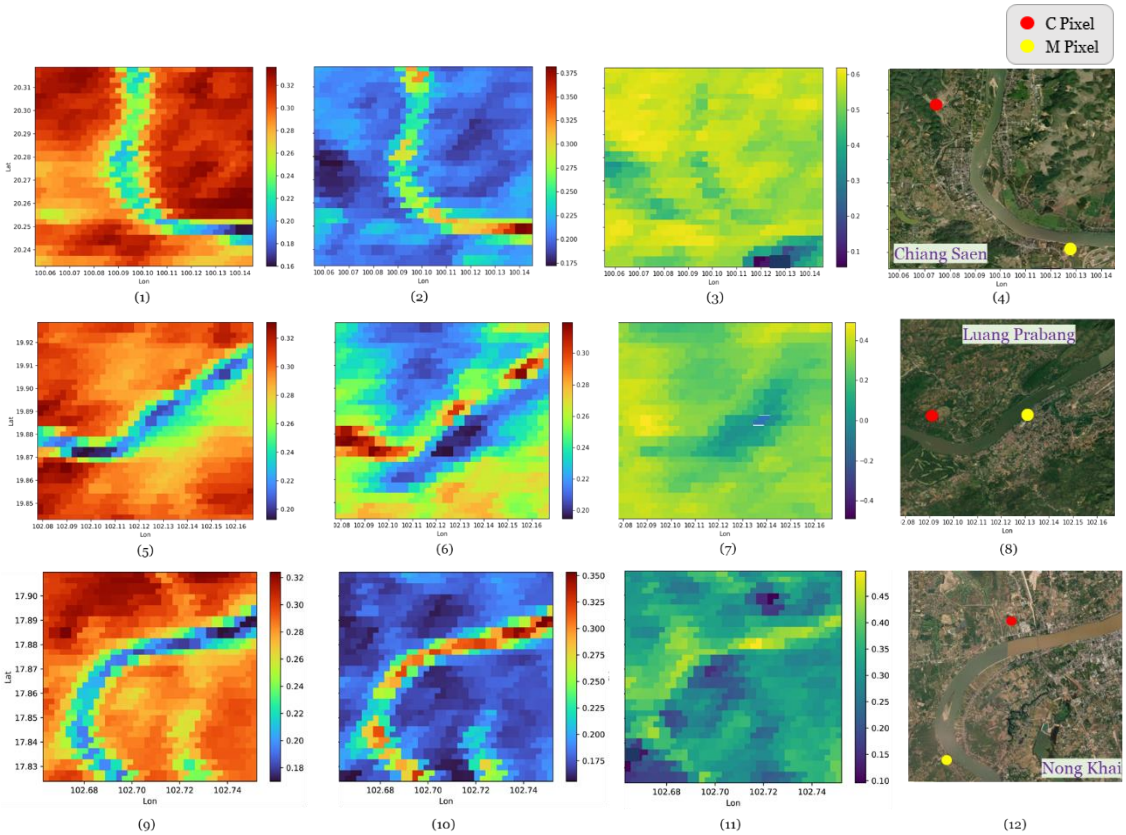
all pixels of the box i.e., 1443, assuming M and varying the C location, following the procedure as above.

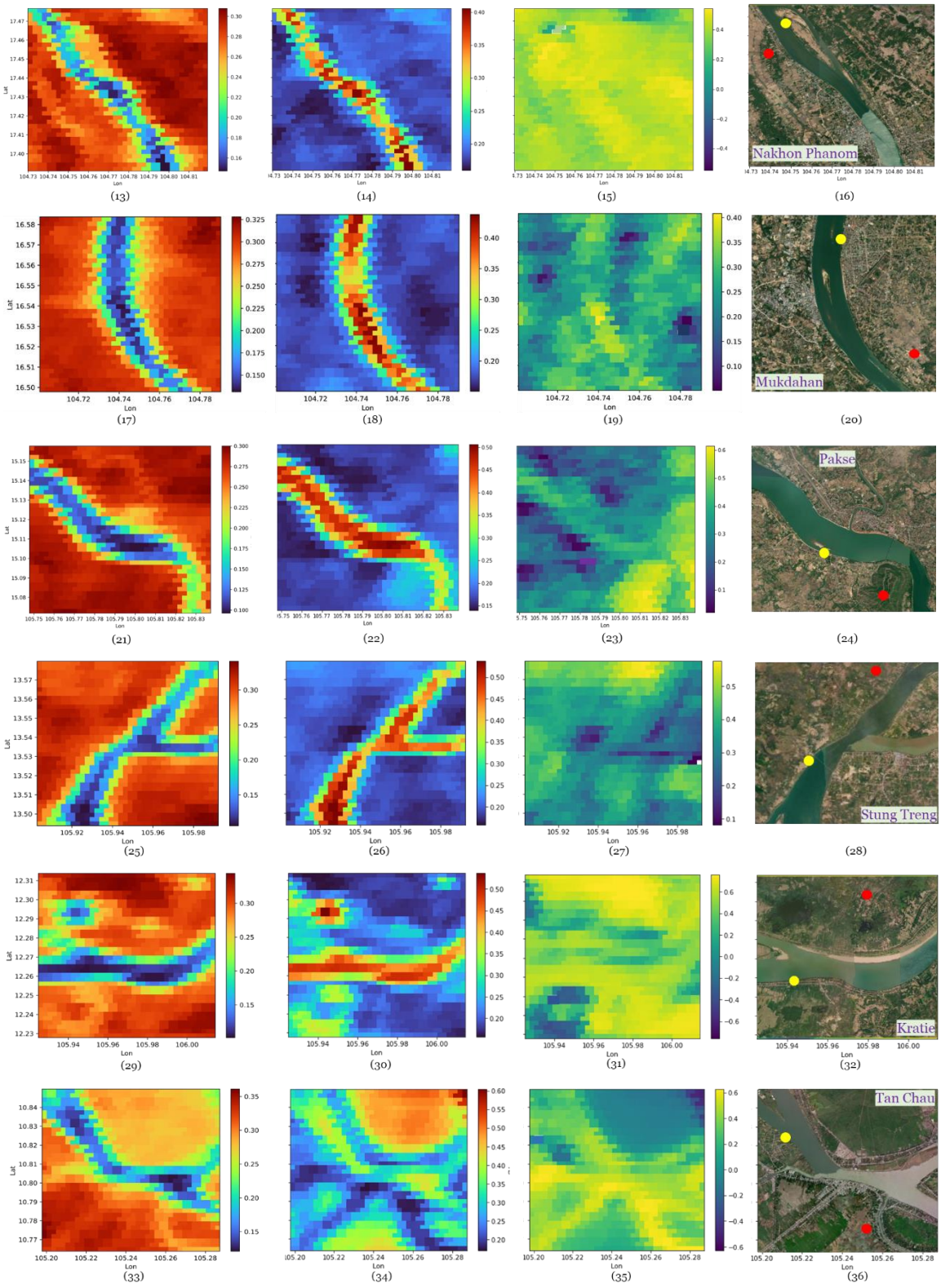
Obtained C/M ratios were correlated with in-situ discharge and maximum coefficient of correlation for each pixel was obtained. Finally, best correlated combination of C and M pixel were selected for discharge estimation. Figure 5-4 shows the best location of C and M pixel for each gauging stations. The coefficient of variation for C pixel is below 0.3 and relatively stable while that for M pixel shows higher variability.

Best location for C Pixels were obtained near to the urban areas or densely vegetated areas with lower coefficient of variations of surface reflectance whereas best location for M Pixel were obtained near the river flood plain portions where higher sensitivity to water variability is expected.

Figure 5-4

Maps of median value of Reflectance of band 2 (first column), coefficient of variation of Reflectance (second column), coefficient of correlation between C/M and discharge (third column) and the location of the C and M pixel having maximum correlation is shown with the satellite image corresponding to the selected box (fourth column)





(13) (14) (15) (16)

(17) (18) (19) (20)

(21) (22) (23) (24)

(25) (26) (27) (28)

(29) (30) (31) (32)

(33) (34) (35) (36)

Although the C/M ratio helps to reduce the noise in the surface reflectance measurements, C/M fluctuates quickly over time and seems to be unstable. An exponentially smoothing filter was added to the C/M ratio to reduce the effects of short-term and observation noises and C/M* was computed for each gauging stations. The

correlation between satellite data and observed discharge for each gauging stations is improved after filtering of C/M ratio. Figure 5-5 shows the comparison of correlation between observed data and C/M* with C/M at seven reaches of the Mekong River. It can be observed the correlation coefficient is improved from 0.44 – 0.63 to 0.64 – 0.76 for AQUA products. However, for TERRA products, the coefficient of correlation improved from 0.55 – 0.76 to 0.63 – 0.83. It is notable to observe that TERRA products are highly correlated with observed discharge in the Lancang Mekong Basin.

Figure 5-5

Coefficient of correlation between C/M, C/M and observed discharge at seven gauging stations within the Mekong River for AQUA and TERRA satellite data*

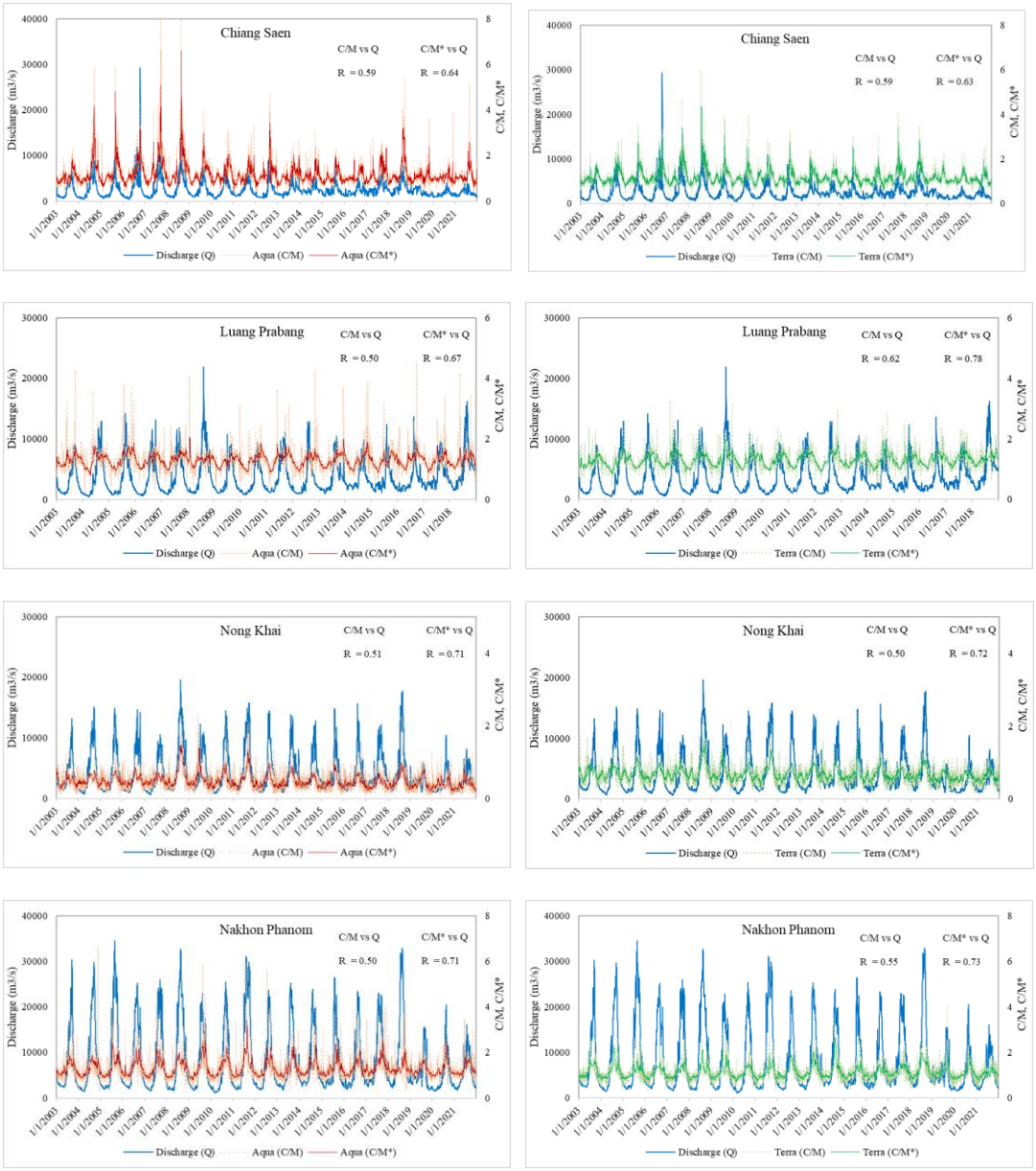


Table 5-3 summarizes the coefficient of correlation between observed discharge and reflectance ratio for each gauging stations.

Table 5-3

Coefficient of Correlation between C/M and C/M with Observed Discharge.*

Station	Aqua		Terra	
	C/M	C/M*	C/M	C/M*
Chiang Saen	0.59	0.64	0.59	0.63
Luang Prabang	0.29	0.43	0.49	0.66
Nong Khai	0.51	0.71	0.50	0.72
Nakhon Phanom	0.50	0.71	0.55	0.73
Mukdahan	0.36	0.64	0.42	0.69
Pakse	0.50	0.67	0.62	0.78
Stung Treng	0.44	0.66	0.58	0.71
Kratie	0.69	0.79	0.76	0.83
Tan Chau	0.63	0.76	0.64	0.72

Thus, time series C/M* is obtained for each gauging station on daily resolution for each AQUA and TERRA satellite products. Then, this time series C/M* are used as input for developing LSTM model for each of the gauging stations.

5.2.2 Discharge prediction with C/M using LSTM model*

5.2.2.1 Data Preprocessing:

In this study, three LSTM model are developed for seven gauging stations using AQUA, TERRA and combined AQUA-TERRA products as input features for LSTM model. Before using data with an LSTM model, it is important to preprocess it to make it suitable for the model to work with. This included cleaning and normalizing the data, breaking it down into smaller units, making sure all data has the same length, separating it into training and test sets, and creating vector representation of the input features.

Data Scaling:

Scaling helps to standardize the range of independent variables and improves the performance and interpretability of ML algorithms. Additionally, it can speed up the training process and improve the numerical stability of optimization algorithms. The commonly used techniques in data scaling are Standardization and Normalization or Min-Max scaling. Standardization assumes that the data will have a bell-shaped distribution and changes the values to have a mean of zero and standard deviation of

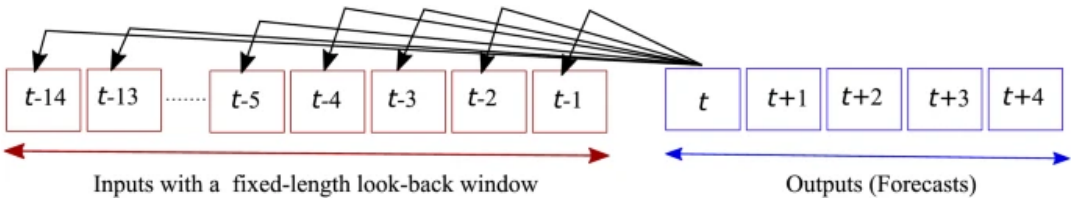
one. Min-Max scaling, on the other hand, transforms the values ranging between 0 and 1. In this research, input features were Min-Max scaled.

Windowing of input data:

The next step included preparing input data for predictions/forecasts using a sliding window technique where past observations were grouped to predict/forecast future discharge. Since the current state of a watershed is dependent on the past behavior of the system, the input features are reorganized using window size (Figure 5-6). This creates fixed-length vectors that machine learning model can use to find patterns and relationships in past observations to make future discharge predictions. Various configuration of window size was checked for discharge prediction during hyperparameter tuning and look back window size of 5 days was adopted in this study for discharge prediction.

Figure 5-6

Schematic representation of the sliding window approach used in this study



Data Splitting

Data splitting in an LSTM model is the process of dividing a dataset into multiple subsets that are used for training, validation, and testing. The training dataset represents the portion of the data that is utilized to train the ML algorithm by learning from historical data to estimate the parameters of the model, thus allowing for effective predictions when the machine learning model is subsequently exposed to previously unseen data. The validation dataset, on the other hand, is employed to optimize the model's hyperparameters. The test dataset, in contrast, is utilized solely for the evaluation of the algorithm after the model has been fully trained and optimized.

The most common data splitting method for LSTM models is to use a train-test split, where a portion of the data is used for training the model and the remaining portion is used for testing. The train-test split is typically done in a random manner, so that the

model is trained and tested on different data each time. When splitting data for an LSTM model, it is important to ensure that the split preserves the temporal order of the data if the data is a time series. Thus, in this research training set and test set were prepared from each alternate days which accounted 50% of total data as training set and remaining as test set. Alternate days was selected to preserve all characteristics of input feature during the training process. A separate validation set was used taking 10% of training data to evaluate the model's performance during the training process before testing.

5.2.2.2 Model Development:

The LSTM model was implemented in Keras with TensorFlow backend and computational code was implemented in Python 3.9. The LSTM architecture consists of three main components: the input gate, the forget gate, and the output gate. Each component is made up of a sigmoid layer and a dot product operation. The input gate controls the amount of information that is allowed to flow into the cell state, the forget gate controls the amount of information that is kept in the cell state, and the output gate controls the amount of information that is output from the cell state. The cell state is a memory unit that can retain information over a longer period. The output of the LSTM cell is passed through a fully connected layer with a linear activation function to produce the final output.

After preprocessing the input feature, LSTM network architecture was defined. Defining the architecture of an LSTM model involves several important decisions. The first included determining the number of layers to use in the model. LSTM models can have one or more layers, with each layer containing multiple LSTM cells. Each layer can extract different levels of abstraction from the input data. More layers can capture more complex patterns, but it also increases the risk of overfitting. In this study, two layers were implemented.

The next step was determining the number of neurons in each layer. The number of neurons in each layer will affect the model's capacity to learn and represent the underlying patterns in the data. Determining the number of neurons in each layer of an LSTM model is crucial for optimal performance. The complexity of the problem and the dimensionality of the input are key factors that should be considered when deciding on the number of neurons. Additionally, a larger number of neurons in a layer can

increase the capacity of the model to learn more complex representations, however, it can also lead to increased computation and the risk of overfitting. Therefore, the number of neurons was carefully chosen to strike a balance between model performance and computational efficiency.

Another important decision is choosing the type of activation function to use in the LSTM cells. The activation function controls the output of each neuron, and it is used to introduce non-linearity in the model. The activation functions used in this study is Leaky ReLu. This activation function is selected because Leaky ReLU mitigates the "dying ReLU" problem by allowing a small negative slope for negative input values, thus preventing the gradients from becoming negligible. This facilitates the training process by enabling the network to learn more complex representations, whilst remaining computationally efficient. Additionally, it is a simple and straightforward modification to implement.

Dropout regularization can be added to the LSTM model to prevent overfitting by randomly dropping out some neurons during the training process. Overfitting occurs when a model becomes too complex and starts to memorize the training data instead of generalizing to new data. Dropout works by randomly dropping out (setting to zero) a certain percentage of the neurons during training. This forces the remaining neurons to adapt and learn more robust features, making the model less sensitive to the specific training data. Additionally, dropout also increases the diversity of the models, which can lead to better generalization performance. By adding dropout to LSTM models, it helps to prevent overfitting and improve the overall performance on unseen data.

Hyperparameter tuning is done in LSTM to find the optimal combination of hyperparameters that results in the best performance on the validation set. It is crucial for LSTM model because it can significantly impact the performance of the model. Hyperparameter tuning is the process of adjusting the parameters of a machine learning model that are not learned from the data during training in LSTM. It involves training the LSTM model multiple times with different combinations of hyperparameters and evaluating the model's performance on a validation set to find the optimal combination. There are different methods for hyperparameter tuning such as Grid Search, Random Search and Bayesian Optimization.

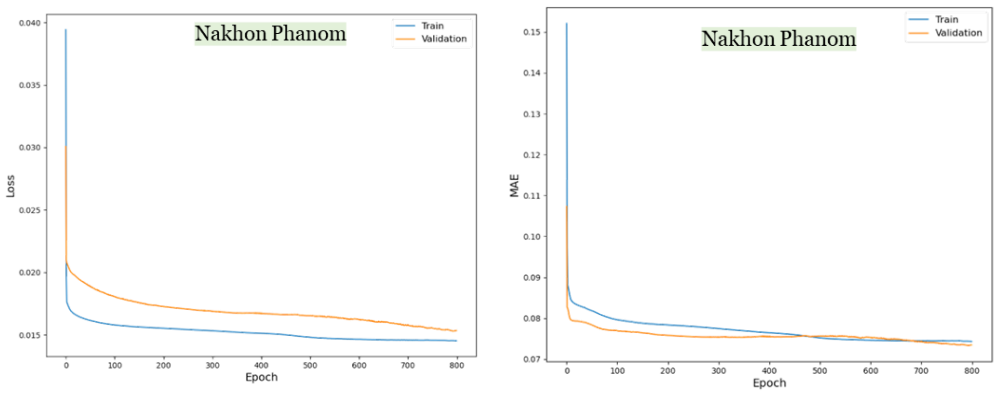
Hyperparameter tuning was performed in this research to optimize LSTM model performance. The hyperparameters used for tuning include the number of neurons per layer, the learning rate, the batch size, the dropout rate, number of epochs, window size and the number of layers. To perform the tuning, grid search method and manually trial and error method was explored for the different combinations of the hyperparameters. Grid search is a simple and exhaustive method of searching through a predefined range of hyperparameters by specifying a set of values for each hyperparameter and training the model for each combination of those values. This process allowed to test a wide range of combinations and determine the optimal set of hyperparameters that resulted in the best performance on the validation set.

After the optimal combination of hyperparameters was determined, the final LSTM model was trained using these hyperparameters on the entire training set and evaluated on the test set to report the final performance of the model. It's worth noting that the process of hyperparameter tuning can be computationally time-consuming, but it is crucial for LSTM model to achieve the best performance. Table 5-4 summarizes the optimal parameters of LSTM model used in this research.

Figure 5-7 shows the performance of LSTM model in terms of loss and mean squared error with increasing number of epochs. It is observed the model performance has improved with increasing number of epochs and stabilizes with epochs about 800. Thus the number of epoch considered for training of LSTM model in this study is taken as 800.

Figure 5-7

Model Loss and Mean Squared Error during training and Validation



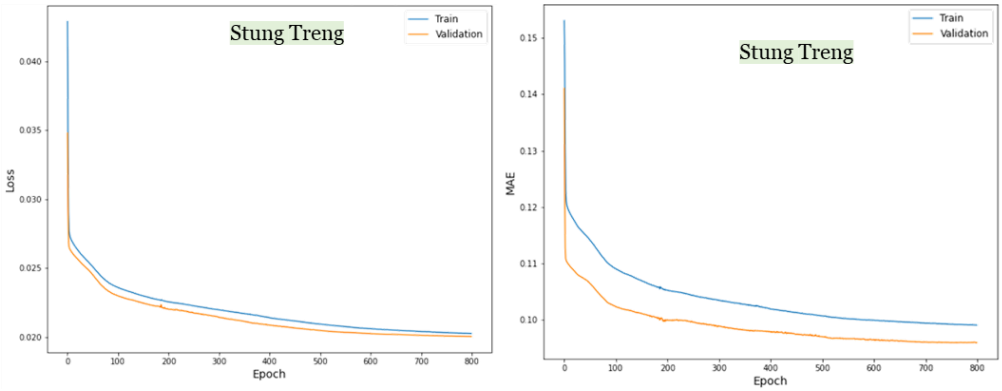


Table 5-4

Optimized Hyperparameters for LSTM Model Development.

Station	No. of Hidden Layer	No. of Units	Window Size	Batch Size	Optimizer / Activation Function
Chiang Saen	1	16	5	5	Adam / Leaky ReLu
Luang Prabang	1	16	4	5	
Nong Khai	1	16	5	5	
Nakhon Phanom	1	16	1	1	
Mukdahan	1	16	5	5	
Pakse	1	16	1	1	
Stung Treng	1	16	5	5	
Kratie	1	16	1	1	
Tan Chau	1	16	1	1	

5.2.2.3 Discharges Prediction using Optical Datasets

The LSTM model was able to identify patterns and relationships in the data that were not immediately apparent, which has the potential to lead to new insights. For analyzing the performance of discharge simulated by single product and combined product, model was trained using Aqua and Terra products separately at first and then was trained using combined input from Aqua and Terra products. Then, model was tested for next set of data which were not used during the training period. To include all pattern and seasonality of the data, model was trained using data for alternate days and tested for next alternate days. Overall, the results from the model are promising and indicate robustness of using optical data for discharge estimation. Figure 5-8 shows the comparison between the observed and simulated discharge using LSTM for input features from MODIS AQUA, TERRA and combination of both.

The Table 5-5 provides the performance of different products (Aqua, Terra and combination of Aqua and Terra) in different stations (Chiang Saen, Luang Prabang, Nong Khai, Nakhon Phanom, Mukdahan, Pakse, Stung Treng, Kratie and Tan Chau) for simulated discharge compared with the observed discharge. The three input features used are Aqua, Terra, and combined Aqua and Terra products. The results for each station have been tabulated separately for both the training and testing periods. The evaluation has been done in terms of Coefficient of Correlation (R), Coefficient of Determination (R^2), Nash–Sutcliffe Efficiency (NSE), and Percent Bias.

In terms of the correlation coefficient, the results show moderate to strong positive correlations between simulated and observed discharge for most of the stations. The results indicate that combined input feature of Aqua and Terra product shows the highest correlation coefficient values between simulated and observed discharge, ranging from 0.70 to 0.88, for all the nine stations. For the Aqua product, the correlation coefficients range from 0.49 to 0.78, and for the Terra product, the correlation coefficients range from 0.63 to 0.86. The highest correlation coefficient of 0.88 was observed for the Kratie station with the Aqua+Terra product, while the lowest correlation coefficient of 0.49 was observed for the Luang Prabang station with the Aqua product.

The coefficient of determination values is positive for all stations and products, indicating that the simulated discharge data explains some portion of the variance in the observed discharge data. The coefficient of determination values ranges from 0.24 to 0.77 for Aqua, 0.39 to 0.73 for Terra, and 0.46 to 0.77 for combined Aqua and Terra. The combined Aqua and Terra product generally showed higher R-squared values than the individual Aqua and Terra products, indicating that the combination of both products leads to a better fit between the simulated and observed discharge data. However, the R-squared values are generally lower than the corresponding correlation coefficients, indicating that the simulated discharge data does not explain all the variance in the observed discharge data.

The Nash–Sutcliffe Efficiency (NSE) values for combined Aqua and Terra are also the highest among the three input features for most of the stations, indicating better performance in simulating the observed discharge. The value of NSE across all stations and products ranges from 0.23 to 0.76. The highest NSE value of 0.76 is achieved by

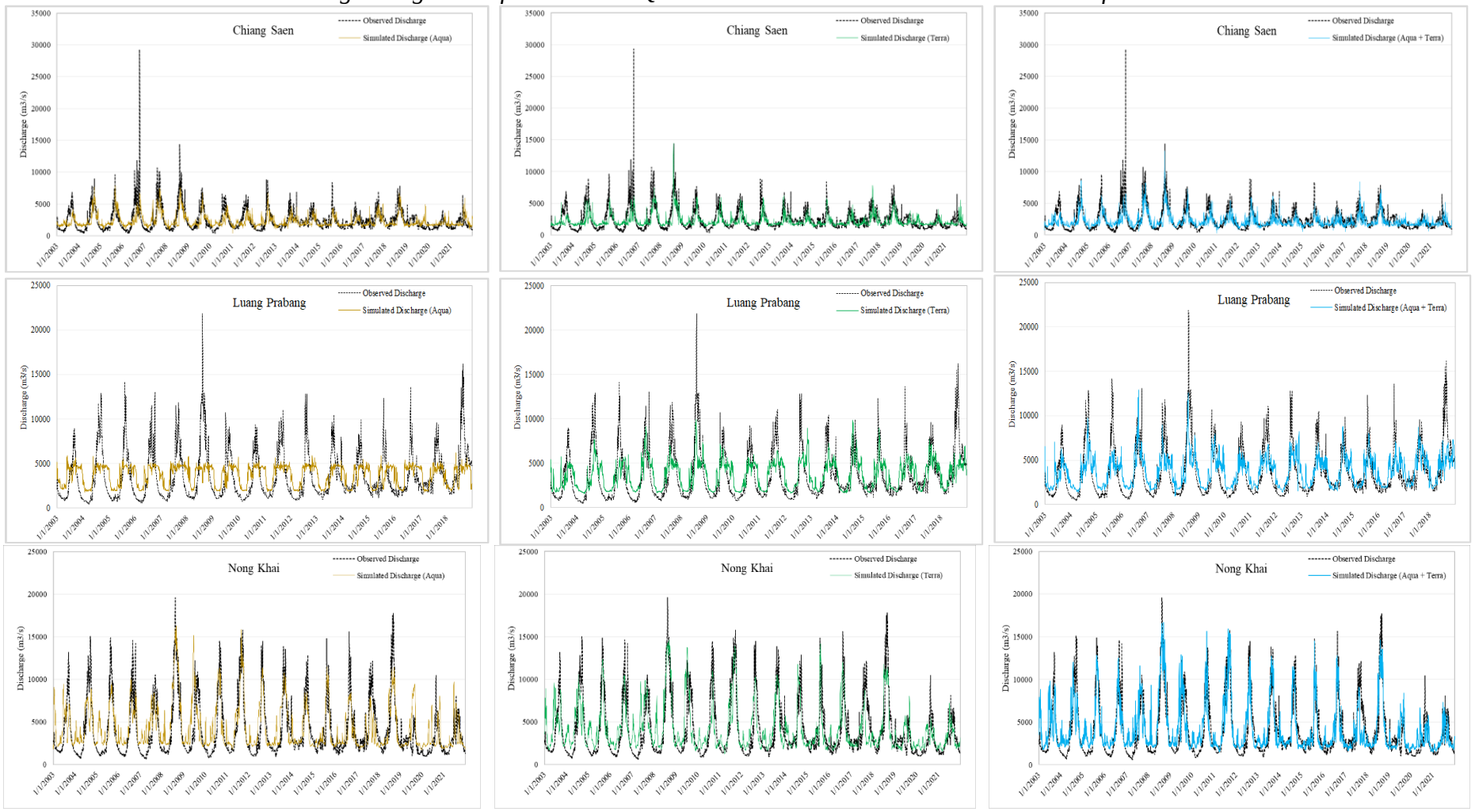
combined Aqua and Terra products at Kratie station, while the lowest NSE value of 0.23 is achieved by Aqua at Luang Prabang station during both calibration and validation period. In general, the combined Aqua and Terra product outperforms both Aqua and Terra products separately in terms of NSE values.

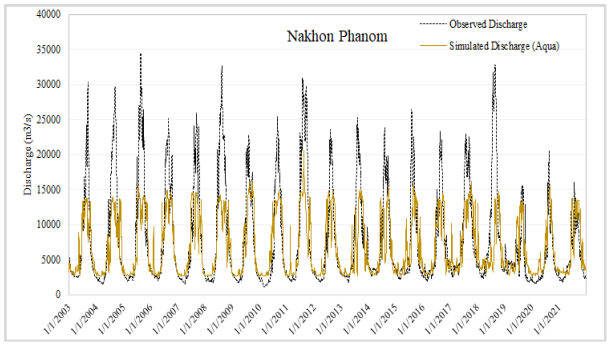
The Percent Bias (PBIAS) values for all the three products are negative, indicating underestimation of the simulated discharge in Chiang Saen, Nakhon Phanom, Stung Treng and Kratie stations. However, for remaining stations, PBIAS values are positive indicating overestimation of simulated discharge compared to observed discharge. The magnitude of the PBIAS values varies greatly across stations and products. The PBIAS varies across all stations and products range from -1.21 % to 25.12%. The highest PBIAS value of 25.12% is found for Aqua product at Mukdahan station, while the lowest PBIAS value of -1.21 % is found for Aqua product at Stung Treng station.

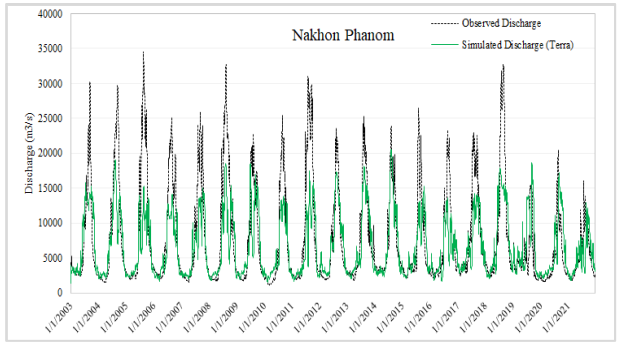
Thus, the results highlight the importance of using multiple products for simulating discharge and the benefits of combining them. The combined Aqua and Terra product shows the highest correlation coefficient, coefficient of determination, and Nash–Sutcliffe Efficiency values for all the nine stations. It is worth noting that, the performance of simulated discharge is found to be better in the downstream reaches compared to upstream reaches Chiang Saen and Luang Prabang. Further this provides insight on exploring the potential of using other combinations of satellite products or incorporating other data sources to improve the accuracy of discharge simulation.

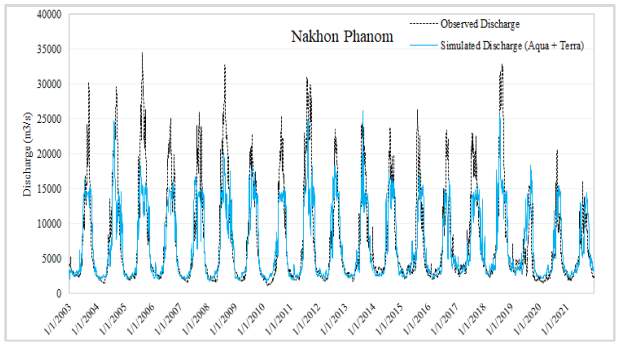
Figure 5-8

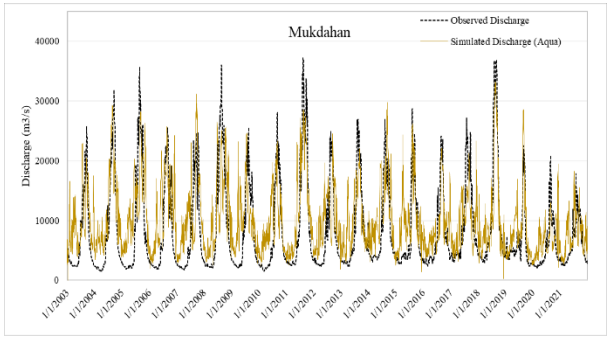
Observed and Simulated Discharge using LSTM for MODIS-AQUA and MODIS-TERRA and combination of both.

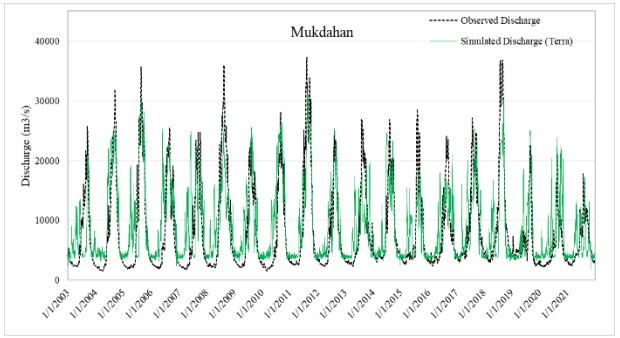


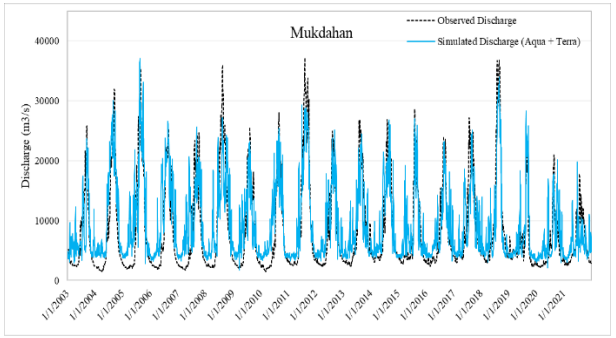

Nakhon Phanom
Observed Discharge
Simulated Discharge (Aqua)
Discharge (m3/s)

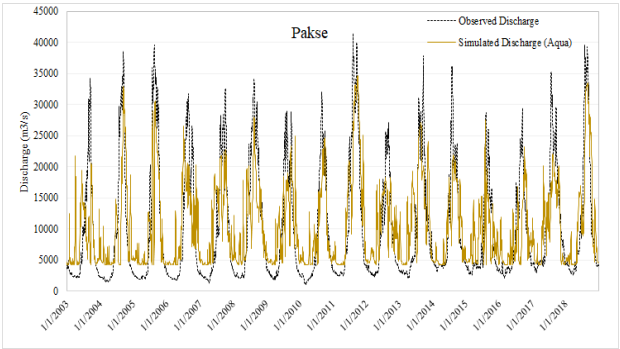

Nakhon Phanom
Observed Discharge
Simulated Discharge (Terra)
Discharge (m3/s)

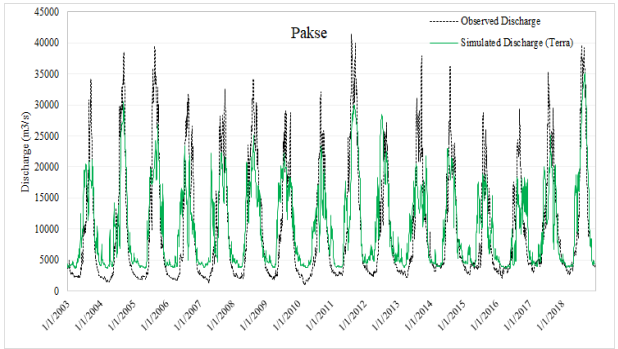

Nakhon Phanom
Observed Discharge
Simulated Discharge (Aqua + Terra)
Discharge (m3/s)

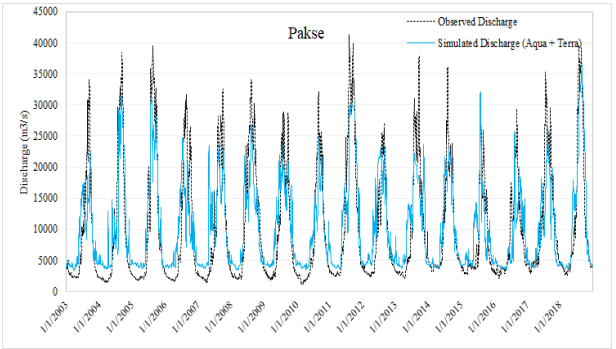

Mukdahan
Observed Discharge
Simulated Discharge (Aqua)
Discharge (m3/s)

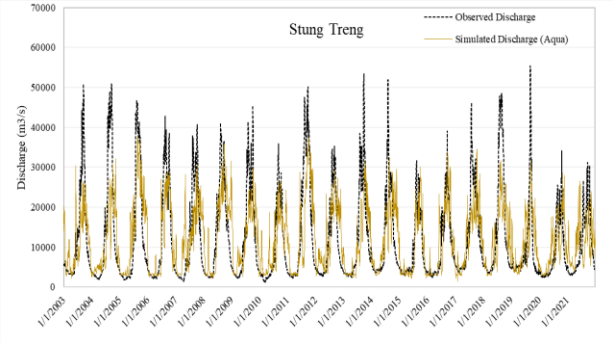

Mukdahan
Observed Discharge
Simulated Discharge (Terra)
Discharge (m3/s)

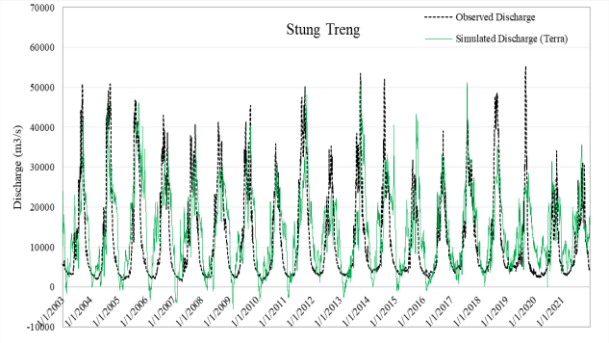

Mukdahan
Observed Discharge
Simulated Discharge (Aqua + Terra)
Discharge (m3/s)

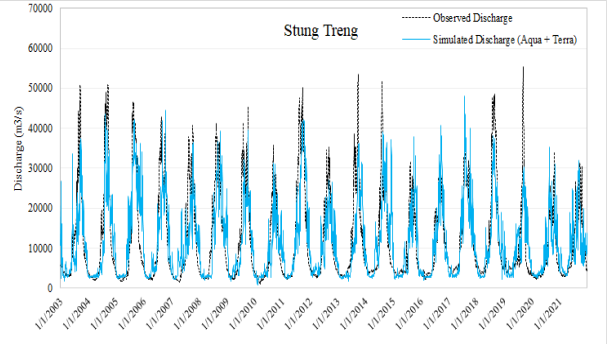

Pakse
Observed Discharge
Simulated Discharge (Aqua)
Discharge (m3/s)

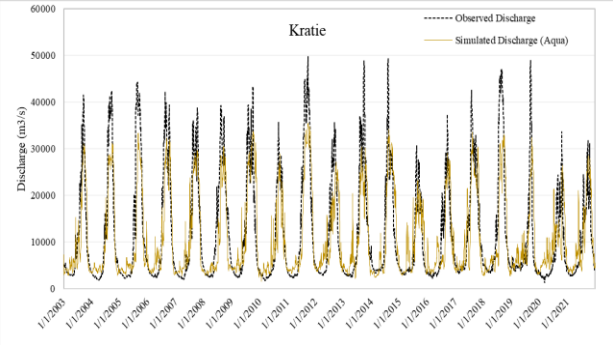

Pakse
Observed Discharge
Simulated Discharge (Terra)
Discharge (m3/s)

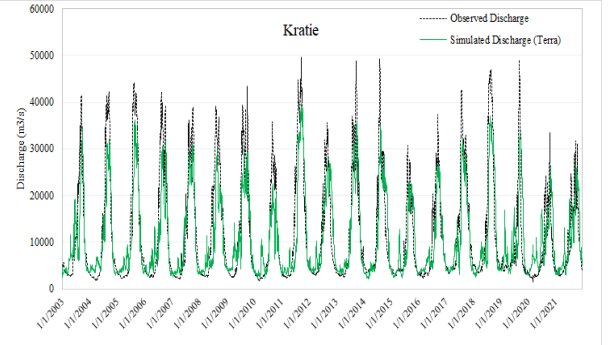

Pakse
Observed Discharge
Simulated Discharge (Aqua + Terra)
Discharge (m3/s)

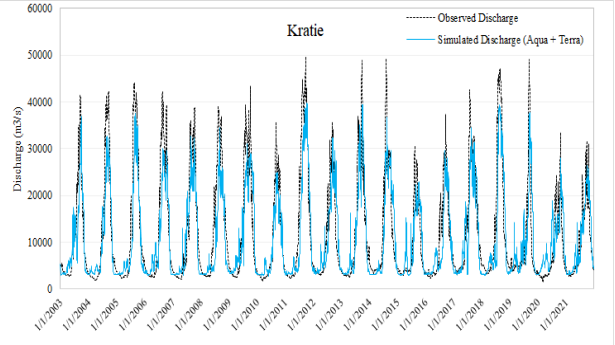

Stung Treng
Observed Discharge
Simulated Discharge (Aqua)
Discharge (m3/s)

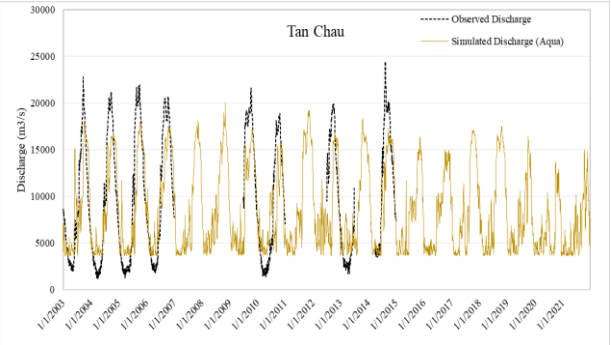

Stung Treng
Observed Discharge
Simulated Discharge (Terra)
Discharge (m3/s)

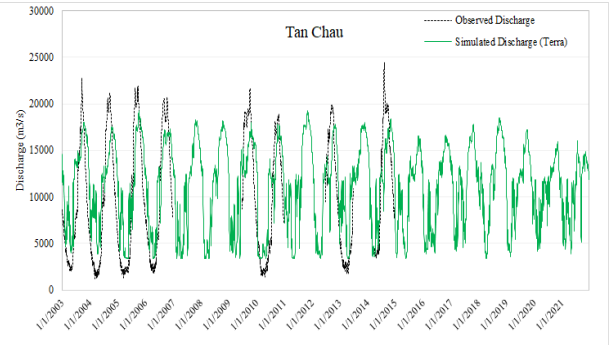

Stung Treng
Observed Discharge
Simulated Discharge (Aqua + Terra)
Discharge (m3/s)


Kratie
Observed Discharge
Simulated Discharge (Aqua)
Discharge (m3/s)


Kratie
Observed Discharge
Simulated Discharge (Terra)
Discharge (m3/s)


Kratie
Observed Discharge
Simulated Discharge (Aqua + Terra)
Discharge (m3/s)


Tan Chau
Observed Discharge
Simulated Discharge (Aqua)
Discharge (m3/s)


Tan Chau
Observed Discharge
Simulated Discharge (Terra)
Discharge (m3/s)

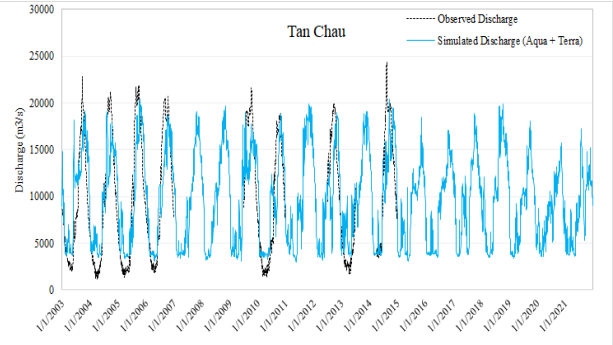

Tan Chau
Observed Discharge
Simulated Discharge (Aqua + Terra)
Discharge (m3/s)

Table 5-5

Performances of simulated discharge with observed discharge using MODIS Aqua, Terra and combination of both.

Station	Product	Training				Testing			
		R	R^2	NSE	PBIAS	R	R^2	NSE	PBIAS
Chiang Saen	Aqua	0.65	0.42	0.41	-5.97	0.64	0.41	0.40	-5.76
	Terra	0.63	0.40	0.38	-9.11	0.63	0.39	0.37	-9.33
	Aqua + Terra	0.70	0.49	0.47	-9.40	0.70	0.48	0.46	-9.53
Luang Prabang	Aqua	0.49	0.24	0.23	4.98	0.49	0.24	0.23	4.96
	Terra	0.67	0.45	0.44	4.48	0.67	0.45	0.44	4.22
	Aqua + Terra	0.71	0.50	0.48	6.37	0.71	0.51	0.49	5.96
Nong Khai	Aqua	0.75	0.56	0.55	6.95	0.75	0.56	0.55	7.03
	Terra	0.75	0.57	0.56	4.23	0.75	0.57	0.56	4.43
	Aqua + Terra	0.81	0.66	0.65	4.89	0.81	0.65	0.65	5.04
Nakhon Phanom	Aqua	0.73	0.54	0.50	-12.56	0.73	0.54	0.50	-12.64
	Terra	0.75	0.56	0.52	-13.74	0.75	0.56	0.52	-13.75
	Aqua + Terra	0.83	0.69	0.67	-9.96	0.83	0.69	0.67	-10.03
Mukdahan	Aqua	0.71	0.50	0.43	24.40	0.71	0.50	0.42	25.12
	Terra	0.75	0.57	0.55	10.03	0.76	0.57	0.55	9.95
	Aqua + Terra	0.82	0.68	0.66	13.29	0.82	0.68	0.65	13.61
Pakse	Aqua	0.74	0.55	0.55	2.56	0.74	0.55	0.55	2.63
	Terra	0.81	0.66	0.65	3.56	0.81	0.66	0.65	3.57
	Aqua + Terra	0.85	0.72	0.71	2.45	0.85	0.72	0.71	2.45
Stung Treng	Aqua	0.70	0.48	0.48	-1.21	0.70	0.49	0.49	-1.26
	Terra	0.70	0.49	0.43	17.78	0.70	0.49	0.43	17.75
	Aqua + Terra	0.78	0.61	0.61	-3.90	0.78	0.61	0.61	-3.84
Kratie	Aqua	0.84	0.71	0.70	-7.18	0.84	0.71	0.70	-7.25
	Terra	0.86	0.73	0.71	-9.18	0.86	0.73	0.71	-9.19
	Aqua + Terra	0.88	0.77	0.76	-9.41	0.88	0.77	0.76	-9.42
Tan Chau	Aqua	0.78	0.61	0.59	-8.30	0.78	0.61	0.59	-8.30
	Terra	0.77	0.60	0.54	15.26	0.77	0.60	0.54	15.22
	Aqua + Terra	0.80	0.64	0.63	1.92	0.80	0.64	0.63	2.04

5.2.3 Discharge prediction with Water Level and C/M using LSTM model*

The Table 5-6 shows the results of simulated discharge compared with observed discharge for two different stations, Pakse and Kratie. The comparison was done for two different cases: using water level only as input feature and using a combination of Aqua, Terra, and water level as input features. Four metrics were used to evaluate the accuracy of the simulated discharge: Coefficient of Correlation (R), Coefficient of Determination (R2), Nash–Sutcliffe Efficiency (NSE), and Percent Bias.

For the Pakse station, both cases produced similar results for training and testing periods, with R values of 0.96 and R^2 values of 0.92 for first case and R values of 0.97 and R^2 values of 0.94 for second case. The NSE values are above 0.90 for both cases, indicating good agreement between simulated and observed discharge. However, there was a difference in the PBIAS values, with values of 9.26 and 12.64 during testing period and 9.37 and 12.81 for testing period for the water level only and the combined case, respectively. This suggests that the combined case may slightly overestimate the discharge. However, in terms of other performance indicators, combination of Aqua, Terra and water level outperformed use of single water level only for discharge estimation.

For the Kratie station, trend of results is similar to Pakse in terms of correlation coefficient, coefficient of determination and Nash Sutcliff Efficiency but the magnitude of performance indicators are lower than Pakse. However, in terms of percent bias, the simulated discharge is slightly underestimated by about 0.25 % which is far better than Pakse for combined input features during both training and testing period. This suggests that the combined case may slightly underestimate the discharge in Kratie station.

Overall, the results indicate that the combination of Aqua, Terra, and water level data produced better results than the single water level, as evidenced by higher R, R^2, and NSE values for both stations.

Figure 5-9 shows the graphical comparison between simulated discharge and observed discharge at Pakse and Kratie for both cases. It can be noted, water level only has underestimated high flows but with combined input of multiple products, these peak

flows have been better predicted. This can be due to missing water level attributed by temporal resolution of satellite altimetry and fulfilling this gap by use of reflectance ratio which have better temporal resolution. This also highlights the importance of using multisource remote sensing data for discharge prediction.

Figure 5-9

Observed and Simulated Discharge using LSTM for Water Level and combination of Water Level, MODIS Aqua and Terra.

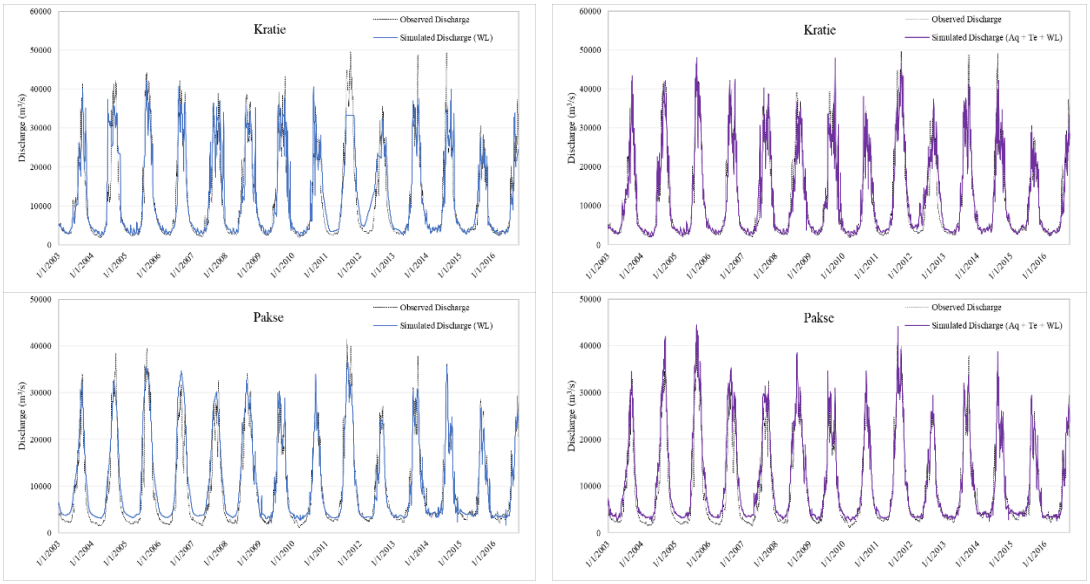


Table 5-6

Performances of Simulated Discharge with observed discharge using Water level and combination of Water Level, MODIS Aqua and Terra.

Station	Product	Training				Testing			
		R	R^2	NSE	PBIAS	R	R^2	NSE	PBIAS
Pakse	WL	0.96	0.92	0.91	9.26	0.96	0.92	0.91	9.37
	WL + Aqua + Terra	0.97	0.94	0.92	12.64	0.97	0.94	0.92	12.81
Kratie	WL	0.92	0.84	0.84	2.78	0.92	0.84	0.84	2.89
	WL + Aqua + Terra	0.95	0.90	0.90	-0.26	0.95	0.90	0.90	-0.25

CHAPTER 6
CONCLUSIONS

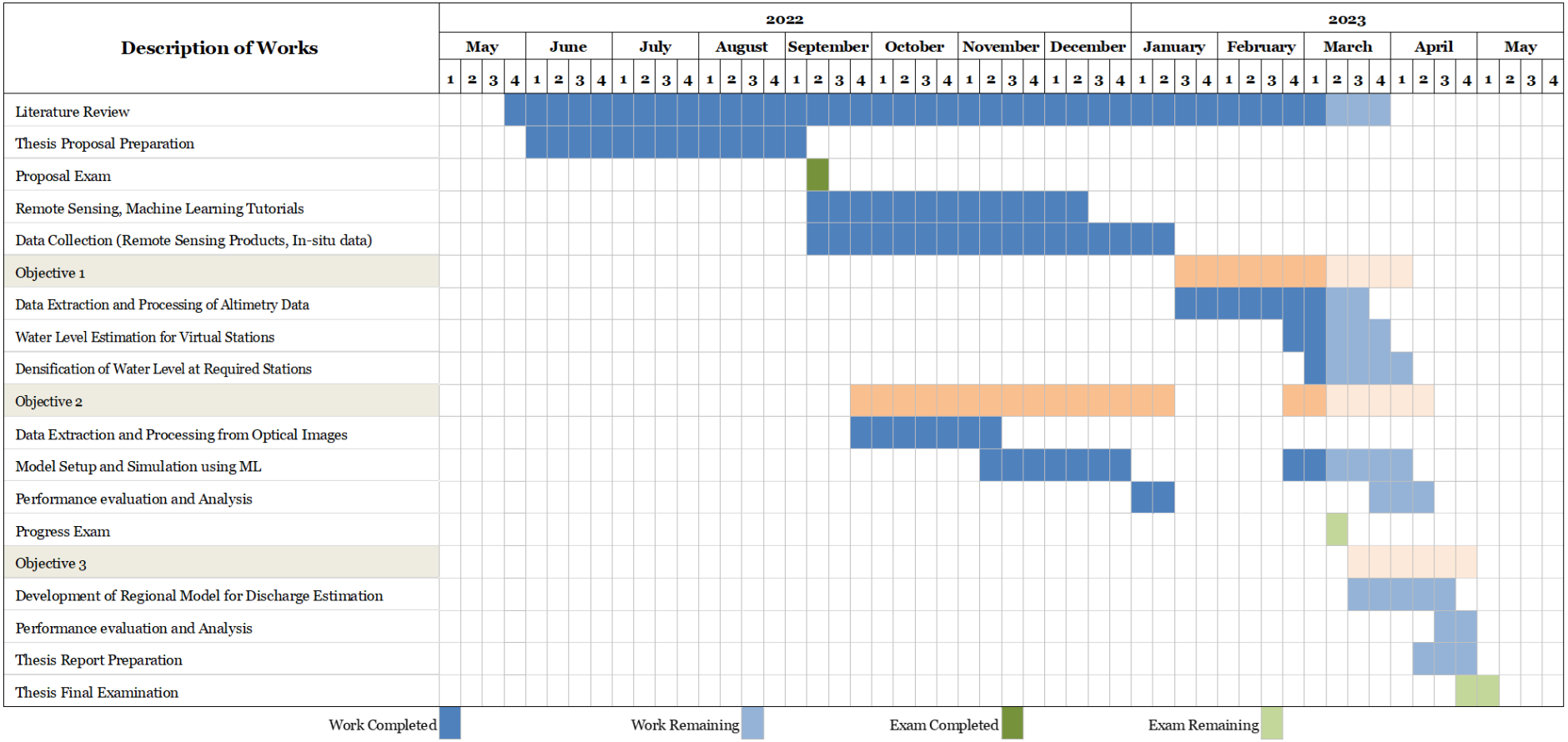
The following outputs are anticipated after the successful completion of proposed study:

1. Time series water level can be estimated from the satellite observations with good accuracy and use of multiple altimetry missions helps to increase the temporal resolution of water level in Mekong River.
2. The combination of multiple optical sensors provided better performance in simulating discharge compared to use of single sensor. This is demonstrated by combined use of MODIS Aqua and Terra products which have predicted discharge with higher performance indices at all reaches of Mekong River compared to individual Aqua and Terra products.
3. The multi mission approach which uses optical and radar altimetry within LSTM model demonstrated its effectiveness in estimating river discharge at two locations along the Mekong River.

CHAPTER 7

TIME SCHEDULE

The overall timeline of the study is presented below:



REFERENCES

Ahn, J.-H., & Park, Y.-J. (2020). Estimating Water Reflectance at Near-Infrared Wavelengths for Turbid Water Atmospheric Correction: A Preliminary Study for GOCI-II. *Remote Sensing*, *12*(22), Article 22. https://doi.org/10.3390/rs12223791

Arnell, N. W., & Gosling, S. N. (2013). The impacts of climate change on river flow regimes at the global scale. *Journal of Hydrology*, *486*, 351–364. https://doi.org/10.1016/j.jhydrol.2013.02.010

Bengio, Y. (2009). Learning Deep Architectures for AI. *Foundations and Trends® in Machine Learning*, *2*(1), 1–127. https://doi.org/10.1561/2200000006

Biancamaria, S., Andreadis, K. M., Durand, M., Clark, E. A., Rodriguez, E., Mognard, N. M., Alsdorf, D. E., Lettenmaier, D. P., & Oudin, Y. (2010). Preliminary Characterization of SWOT Hydrology Error Budget and Global Capabilities. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(1), 6–19. https://doi.org/10.1109/JSTARS.2009.2034614

Biancamaria, S., Frappart, F., Leleu, A.-S., Marieu, V., Blumstein, D., Desjonquères, J.-D., Boy, F., Sottolichio, A., & Valle-Levinson, A. (2017). Satellite radar altimetry water elevations performance over a 200m wide river: Evaluation over the Garonne River. *Advances in Space Research*, *59*(1), 128–146. https://doi.org/10.1016/j.asr.2016.10.008

Birkinshaw, S. J., Moore, P., Kilsby, C. G., O'Donnell, G. M., Hardy, A. J., & Berry, P. A. M. (2014a). Daily discharge estimation at ungauged river sites using remote sensing: DAILY DISCHARGE ESTIMATION USING REMOTE SENSING. *Hydrological Processes*, *28*(3), 1043–1054. https://doi.org/10.1002/hyp.9647

Birkinshaw, S. J., Moore, P., Kilsby, C. G., O'Donnell, G. M., Hardy, A. J., & Berry, P. A. M. (2014b). Daily discharge estimation at ungauged river sites using remote sensing: DAILY DISCHARGE ESTIMATION USING REMOTE SENSING. *Hydrological Processes*, *28*(3), Article 3. https://doi.org/10.1002/hyp.9647

Birkinshaw, S. J., O'Donnell, G. M., Moore, P., Kilsby, C. G., Fowler, H. J., & Berry, P. A. M. (2010). Using satellite altimetry data to augment flow estimation techniques on the Mekong River. *Hydrological Processes*, *24*(26), 3811–3825. https://doi.org/10.1002/hyp.7811

Bjerklie, D. M., Birkett, C. M., Jones, J. W., Carabajal, C., Rover, J. A., Fulton, J. W., & Garambois, P.-A. (2018). Satellite remote sensing estimation of river discharge: Application to the Yukon River Alaska. *Journal of Hydrology*, *561*, 1000–1018. https://doi.org/10.1016/j.jhydrol.2018.04.005

Bjerklie, D. M., Lawrence Dingman, S., Vorosmarty, C. J., Bolster, C. H., & Congalton, R. G. (2003). Evaluating the potential for measuring river discharge from space.

Journal of Hydrology, *278*(1–4), 17–38. https://doi.org/10.1016/S0022-1694(03)00129-X

Bjerklie, D. M., Moller, D., Smith, L. C., & Dingman, S. L. (2005a). Estimating discharge in rivers using remotely sensed hydraulic information. *Journal of Hydrology*, *309*(1–4), 191–209. https://doi.org/10.1016/j.jhydrol.2004.11.022

Bjerklie, D. M., Moller, D., Smith, L. C., & Dingman, S. L. (2005b). Estimating discharge in rivers using remotely sensed hydraulic information. *Journal of Hydrology*, *309*(1–4), Article 1–4. https://doi.org/10.1016/j.jhydrol.2004.11.022

Bogning, S., Frappart, F., Blarel, F., Niño, F., Mahé, G., Bricquet, J.-P., Seyler, F., Onguéné, R., Etamé, J., Paiz, M.-C., & Braun, J.-J. (2018). Monitoring Water Levels and Discharges Using Radar Altimetry in an Ungauged River Basin: The Case of the Ogooué. *Remote Sensing*, *10*(2), Article 2. https://doi.org/10.3390/rs10020350

Brakenridge, G. R., Nghiem, S. V., Anderson, E., & Chien, S. (2005). Space-based measurement of river runoff. *Eos, Transactions American Geophysical Union*, *86*(19), 185–188. https://doi.org/10.1029/2005EO190001

Brakenridge, G. R., Nghiem, S. V., Anderson, E., & Mic, R. (2007). Orbital microwave measurement of river discharge and ice status. *Water Resources Research*, *43*(4). https://doi.org/10.1029/2006WR005238

BRAKENRIDGE, R., & ANDERSON, E. (2006). MODIS-BASED FLOOD DETECTION, MAPPING AND MEASUREMENT: THE POTENTIAL FOR OPERATIONAL HYDROLOGICAL APPLICATIONS. In J. Marsalek, G. Stancalie, & G. Balint (Eds.), *Transboundary Floods: Reducing Risks Through Flood Management* (pp. 1–12). Springer Netherlands. https://doi.org/10.1007/1-4020-4902-1_1

Calmant, S., & Seyler, F. (2006). Continental surface waters from satellite altimetry. *Comptes Rendus Geoscience*, *338*(14–15), 1113–1122. https://doi.org/10.1016/j.crte.2006.05.012

Camps-Valls, G. (2009). Machine learning in remote sensing data processing. *2009 IEEE International Workshop on Machine Learning for Signal Processing*, 1–6. https://doi.org/10.1109/MLSP.2009.5306233

Cheng, S., Qiao, X., Shi, Y., & Wang, D. (2021). Machine learning for predicting discharge fluctuation of a karst spring in North China. *Acta Geophysica*, *69*(1), 257–270. https://doi.org/10.1007/s11600-020-00522-0

Chiang, Y.-M., Hao, R.-N., Zhang, J.-Q., Lin, Y.-T., & Tsai, W.-P. (2018). Identifying the Sensitivity of Ensemble Streamflow Prediction by Artificial Intelligence. *Water*, *10*(10), 1341. https://doi.org/10.3390/w10101341

Depetris, P. J. (2021). The Importance of Monitoring River Water Discharge. *Frontiers in Water*, *3*. https://www.frontiersin.org/articles/10.3389/frwa.2021.745912

Derecki, J. A., & Quinn, F. H. (1987). Use of current meters for continuous measurement of flows in large rivers. *Water Resources Research*, *23*(9), 1751–1756. https://doi.org/10.1029/WR023i009p01751

Dingman, S. L., & Bjerklie, D. M. (2005). Estimation of River Discharge. In M. G. Anderson & J. J. McDonnell (Eds.), *Encyclopedia of Hydrological Sciences* (p. hsa069). John Wiley & Sons, Ltd. https://doi.org/10.1002/0470848944.hsa069

Domeneghetti, A., Castellarin, A., Tarpanelli, A., & Moramarco, T. (2015). Investigating the uncertainty of satellite altimetry products for hydrodynamic modelling. *Hydrological Processes*, *29*(23), 4908–4918. https://doi.org/10.1002/hyp.10507

Elumalai, V., Brindha, K., Sithole, B., & Lakshmanan, E. (2017). Spatial interpolation methods and geostatistics for mapping groundwater contamination in a coastal area. *Environmental Science and Pollution Research*, *24*(12), 11601–11617. https://doi.org/10.1007/s11356-017-8681-6

Esmaeilzadeh, B., Sattari, M. T., & Samadianfard, S. (2017). Performance evaluation of ANNs and an M5 model tree in Sattarkhan Reservoir inflow prediction. *ISH Journal of Hydraulic Engineering*, *23*(3), 283–292. https://doi.org/10.1080/09715010.2017.1308277

Fekete, B. M., & Vörösmarty, C. J. (n.d.). *The current status of global river discharge monitoring and potential new technologies complementing traditional discharge measurements*. 8.

Filippucci, P., Brocca, L., Bonafoni, S., & Tarpanelli, A. (2022). *River Discharge estimation from optical satellite data: Latest advances using NIR sensors* [Other]. display. https://doi.org/10.5194/egusphere-egu22-3513

Florkowski, T., Davis, T. G., Wallander, B., & Prabhakar, D. R. L. (1969). The measurement of high discharges in turbulent rivers using tritium tracer. *Journal of Hydrology*, *8*(3), 249–264. https://doi.org/10.1016/0022-1694(69)90001-8

Frappart, F., Blumstein, D., Cazenave, A., Ramillien, G., Birol, F., Morrow, R., & Rémy, F. (2017). Satellite Altimetry: Principles and Applications in Earth Sciences. In J. G. Webster, *Wiley Encyclopedia of Electrical and Electronics Engineering* (pp. 1–25). John Wiley & Sons, Inc. https://doi.org/10.1002/047134608X.W1125.pub2

Frappart, F., Papa, F., Silva, J. S. da, Ramillien, G., Prigent, C., Seyler, F., & Calmant, S. (2012). Surface freshwater storage and dynamics in the Amazon basin during the 2005 exceptional drought. *Environmental Research Letters*, *7*(4), 044010. https://doi.org/10.1088/1748-9326/7/4/044010

Garkoti, A., & Kundapura, S. (2021). Deriving water level and discharge estimation using satellite altimetry for Krishna River, Karnataka. *Remote Sensing Applications: Society and Environment*, *22*, 100487. https://doi.org/10.1016/j.rsase.2021.100487

Ghumman, A. R., Ghazaw, Y. M., Sohail, A. R., & Watanabe, K. (2011). Runoff forecasting by artificial neural network and conventional model. *Alexandria Engineering Journal*, *50*(4), 345–350. https://doi.org/10.1016/j.aej.2012.01.005

Gleason, C., & Durand, M. (2020). Remote Sensing of River Discharge: A Review and a Framing for the Discipline. *Remote Sensing*, *12*(7), 1107. https://doi.org/10.3390/rs12071107

Gravelle, R. (2015). *Discharge Estimation: Techniques and Equipment*.

GRDC Data Portal. (n.d.). Retrieved June 16, 2022, from https://portal.grdc.bafg.de/applications/public.html?publicuser=PublicUser#dataDownload/Home

Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(10), 2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924

He, D., Lu, Y., & Li, Z. (n.d.). *Chapter 14—Watercourse Environmental Change in Upper Mekong*. 28.

Hirsch, R. M., & Costa, J. E. (2004). U.S. stream flow measurement and data dissemination improve. *Eos, Transactions American Geophysical Union*, *85*(20), 197–203. https://doi.org/10.1029/2004EO200002

Hoang, L. P., Lauri, H., Kummu, M., Koponen, J., van Vliet, M. T. H., Supit, I., Leemans, R., Kabat, P., & Ludwig, F. (2016). Mekong River flow and hydrological extremes under climate change. *Hydrology and Earth System Sciences*, *20*(7), 3027–3041. https://doi.org/10.5194/hess-20-3027-2016

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hou, J., van Dijk, A. I. J. M., Renzullo, L. J., & Vertessy, R. A. (2018). Using modelled discharge to develop satellite-based river gauging: A case study for the Amazon Basin. *Hydrology and Earth System Sciences*, *22*(12), 6435–6448. https://doi.org/10.5194/hess-22-6435-2018

Hu, C., Wu, Q., Li, H., Jian, S., Li, N., & Lou, Z. (2018). Deep Learning with a Long Short-Term Memory Networks Approach for Rainfall-Runoff Simulation. *Water*, *10*(11), 1543. https://doi.org/10.3390/w10111543

Huang, Q., Long, D., Du, M., Zeng, C., Qiao, G., Li, X., Hou, A., & Hong, Y. (2018a). Discharge estimation in high-mountain regions with improved methods using multisource remote sensing: A case study of the Upper Brahmaputra River.

Remote Sensing of Environment, *219*, 115–134. https://doi.org/10.1016/j.rse.2018.10.008

Huang, Q., Long, D., Du, M., Zeng, C., Qiao, G., Li, X., Hou, A., & Hong, Y. (2018b). Discharge estimation in high-mountain regions with improved methods using multisource remote sensing: A case study of the Upper Brahmaputra River. *Remote Sensing of Environment*, *219*, 115–134. https://doi.org/10.1016/j.rse.2018.10.008

Jiang, L., Schneider, R., Andersen, O., & Bauer-Gottwein, P. (2017). CryoSat-2 Altimetry Applications over Rivers and Lakes. *Water*, *9*, 211. https://doi.org/10.3390/w9030211

Kao, I.-F., Zhou, Y., Chang, L.-C., & Chang, F.-J. (2020). Exploring a Long Short-Term Memory based Encoder-Decoder framework for multi-step-ahead flood forecasting. *Journal of Hydrology*, *583*, 124631. https://doi.org/10.1016/j.jhydrol.2020.124631

Kashid, S. G., & Pardeshi, S. A. (2014). A survey of water distribution system and new approach to intelligent water distribution system. *2014 First International Conference on Networks & Soft Computing (ICNSC2014)*, 339–344. https://doi.org/10.1109/CNSC.2014.6906645

Kenny, J. F., Barber, N. L., Hutson, S. S., Linsey, K. S., Lovelace, J. K., & Maupin, M. A. (2009). *Estimated use of water in the United States in 2005* (Report No. 1344; Circular, p. 60). USGS Publications Warehouse. https://doi.org/10.3133/cir1344

Kim, Lee, Chang, Bui, Jayasinghe, Basnayake, Chishtie, & Hwang. (2019). Daily River Discharge Estimation Using Multi-Mission Radar Altimetry Data and Ensemble Learning Regression in the Lower Mekong River Basin. *Remote Sensing*, *11*(22), Article 22. https://doi.org/10.3390/rs11222684

Kim, Y., Schmid, T., Charbiwala, Z. M., Friedman, J., & Srivastava, M. B. (2008). NAWMS: Nonintrusive autonomous water monitoring system. *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems - SenSys '08*, 309. https://doi.org/10.1145/1460412.1460443

Kouraev, A. V., Zakharova, E. A., Samain, O., Mognard, N. M., & Cazenave, A. (2004). Ob' river discharge from TOPEX/Poseidon satellite altimetry (1992–2002). *Remote Sensing of Environment*, *93*(1), 238–245. https://doi.org/10.1016/j.rse.2004.07.007

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). *Rainfall-Runoff modelling using Long-Short-Term-Memory (LSTM) networks* [Preprint]. Catchment hydrology/Modelling approaches. https://doi.org/10.5194/hess-2018-247

Leon, J. G., Calmant, S., Seyler, F., Bonnet, M.-P., Cauhopé, M., Frappart, F., Filizola, N., & Fraizy, P. (2006). Rating curves and estimation of average water depth at

the upper Negro River based on satellite altimeter data and modeled discharges. *Journal of Hydrology*, *328*(3), 481–496. https://doi.org/10.1016/j.jhydrol.2005.12.006

Li, H., Li, H., Wang, J., & Hao, X. (2019). Extending the Ability of Near-Infrared Images to Monitor Small River Discharge on the Northeastern Tibetan Plateau. *Water Resources Research*, *55*(11), 8404–8421. https://doi.org/10.1029/2018WR023808

Lin, P., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., David, C. H., Durand, M., Pavelsky, T. M., Allen, G. H., Gleason, C. J., & Wood, E. F. (2019). Global Reconstruction of Naturalized River Flows at 2.94 Million Reaches. *Water Resources Research*, *55*(8), 6499–6516. https://doi.org/10.1029/2019WR025287

Liu, K.-T., Tseng, K.-H., Shum, C., Liu, C.-Y., Kuo, C.-Y., Liu, G., Jia, Y., & Shang, K. (2016). Assessment of the Impact of Reservoirs in the Upper Mekong River Using Satellite Radar Altimetry and Remote Sensing Imageries. *Remote Sensing*, *8*(5), 367. https://doi.org/10.3390/rs8050367

Mengen, D., Ottinger, M., Leinenkugel, P., & Ribbe, L. (2020). Modeling River Discharge Using Automated River Width Measurements Derived from Sentinel-1 Time Series. *Remote Sensing*, *12*(19), Article 19. https://doi.org/10.3390/rs12193236

Milliman, J. D., & Farnsworth, K. L. (2011). *River Discharge to the Coastal Ocean: A Global Synthesis* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511781247

MODIS Web. (n.d.). Retrieved July 30, 2022, from https://modis.gsfc.nasa.gov/data/

Nogueira Filho, F. J. M., Souza Filho, F. de A., Porto, V. C., Vieira Rocha, R., Sousa Estácio, Á. B., & Martins, E. S. P. R. (2022). Deep Learning for Streamflow Regionalization for Ungauged Basins: Application of Long-Short-Term-Memory Cells in Semiarid Regions. *Water*, *14*(9), Article 9. https://doi.org/10.3390/w14091318

Ordóñez, F. J., & Roggen, D. (2016). Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition. *Sensors*, *16*(1), Article 1. https://doi.org/10.3390/s16010115

Pan, F. (2013). Remote sensing of river stage and discharge. *SPIE Newsroom*. https://doi.org/10.1117/2.1201212.004611

Papa, F., Durand, F., Rossow, W. B., Rahman, A., & Bala, S. K. (2010). Satellite altimeter-derived monthly discharge of the Ganga-Brahmaputra River and its seasonal to interannual variations from 1993 to 2008. *Journal of Geophysical Research: Oceans*, *115*(C12). https://doi.org/10.1029/2009JC006075

Perumal, M., Moramarco, T., Sahoo, B., & Barbetta, S. (2007). A methodology for discharge estimation and rating curve development at ungauged river sites: RATING CURVE ESTIMATION AT UNGAUGED RIVER SITES. *Water Resources Research*, *43*(2). https://doi.org/10.1029/2005WR004609

Robert Brakenridge, G., Cohen, S., Kettner, A. J., De Groeve, T., Nghiem, S. V., Syvitski, J. P. M., & Fekete, B. M. (2012). Calibration of satellite measurements of river discharge using a global hydrology model. *Journal of Hydrology*, *475*, 123–136. https://doi.org/10.1016/j.jhydrol.2012.09.035

Sahoo, D. P., Sahoo, B., & Tiwari, M. K. (2020a). *Performance Evaluation of Remote Sensing-based High Frequent Streamflow Estimation Models at the Bramhani River Basin Outlet*. *2020*, H011-0003.

Sahoo, D. P., Sahoo, B., & Tiwari, M. K. (2020b). Copula-based probabilistic spectral algorithms for high-frequent streamflow estimation. *Remote Sensing of Environment*, *251*, 112092. https://doi.org/10.1016/j.rse.2020.112092

Scherer, D., Schwatke, C., & Dettmering, D. (n.d.). *Estimation of River Discharge using Multi-Mission Satellite Altimetry and Optical Remote Sensing Imagery*. 15.

Schneider, R., Godiksen, P., Ranndal, H., Madsen, H., & Bauer-Gottwein, P. (2017). Application of CryoSat-2 altimetry data for river analysis and modelling. *Hydrology and Earth System Sciences*, *21*, 751–764. https://doi.org/10.5194/hess-21-751-2017

Sea Level. (n.d.). ESA Climate Office. Retrieved August 10, 2022, from https://climate.esa.int/en/projects/sea-level/

Shanlong, L., Bingfang, W., Nana, Y., Fapeng, L., Meiping, W., & Jing, W. (2010). Progress in River Runoff Monitoring by Remote Sensing□□. *Advances in Earth Science*, *25*(8), 820. https://doi.org/10.11867/j.issn.1001-8166.2010.08.0820

Shi, Z., Chen, Y., Liu, Q., & Huang, C. (2020). Discharge Estimation Using Harmonized Landsat and Sentinel-2 Product: Case Studies in the Murray Darling Basin. *Remote Sensing*, *12*(17), Article 17. https://doi.org/10.3390/rs12172810

Sichangi, A. W., Wang, L., & Hu, Z. (2018). Estimation of River Discharge Solely from Remote-Sensing Derived Data: An Initial Study Over the Yangtze River. *Remote Sensing*, *10*(9), Article 9. https://doi.org/10.3390/rs10091385

Sichangi, A. W., Wang, L., Yang, K., Chen, D., Wang, Z., Li, X., Zhou, J., Liu, W., & Kuria, D. (2016a). Estimating continental river basin discharges using multiple remote sensing data sets. *Remote Sensing of Environment*, *179*, 36–53. https://doi.org/10.1016/j.rse.2016.03.019

Sichangi, A. W., Wang, L., Yang, K., Chen, D., Wang, Z., Li, X., Zhou, J., Liu, W., & Kuria, D. (2016b). Estimating continental river basin discharges using multiple remote sensing data sets. *Remote Sensing of Environment*, *179*, 36–53. https://doi.org/10.1016/j.rse.2016.03.019

Singh, G., Mishra, A., & Sagar, D. (n.d.). *SBIT JOURNAL OF SCIENCES AND TECHNOLOGY ISSN 2277-8764 VOL-2, ISSUE 1, 2013.* 4.

Smith, L. C., & Pavelsky, T. M. (2008). Estimation of river discharge, propagation speed, and hydraulic geometry from space: Lena River, Siberia: RIVER DISCHARGE AND HYDRAULIC GEOMETRY. *Water Resources Research*, *44*(3), Article 3. https://doi.org/10.1029/2007WR006133

Smith, L., & Pavelsky, T. (2008). Estimation of river discharge, propagation speed, and hydraulic geometry from space: Lena River, Siberia. *Water Resour. Res*, *44*. https://doi.org/10.1029/2007WR006133

Soni, P., Tripathi, S., & Srivastava, R. (2021). A comparison of regionalization methods in monsoon dominated tropical river basins. *Journal of Water and Climate Change*, *12*(5), 1975–1996. https://doi.org/10.2166/wcc.2021.298

Sun, W. C., Ishidaira, H., & Bastola, S. (2010). Towards improving river discharge estimation in ungauged basins: Calibration of rainfall-runoff models based on satellite observations of river flow width at basin outlet. *Hydrology and Earth System Sciences*, *14*(10), 2011–2022. https://doi.org/10.5194/hess-14-2011-2010

Tang, Q., Gao, H., Lu, H., & Lettenmaier, D. P. (2009). Remote sensing: Hydrology. *Progress in Physical Geography: Earth and Environment*, *33*(4), 490–509. https://doi.org/10.1177/0309133309346650

Tarpanelli, A., Amarnath, G., Brocca, L., Massari, C., & Moramarco, T. (2017). Discharge estimation and forecasting by MODIS and altimetry data in Niger-Benue River. *Remote Sensing of Environment*, *195*, 96–106. https://doi.org/10.1016/j.rse.2017.04.015

Tarpanelli, A., Barbetta, S., Brocca, L., & Moramarco, T. (2013). River Discharge Estimation by Using Altimetry Data and Simplified Flood Routing Modeling. *Remote Sensing*, *5*(9), 4145–4162. https://doi.org/10.3390/rs5094145

Tarpanelli, A., Brocca, L., Lacava, T., Melone, F., Moramarco, T., Faruolo, M., Pergola, N., & Tramutoli, V. (2013a). Toward the estimation of river discharge variations using MODIS data in ungauged basins. *Remote Sensing of Environment*, *136*, 47–55. https://doi.org/10.1016/j.rse.2013.04.010

Tarpanelli, A., Brocca, L., Lacava, T., Melone, F., Moramarco, T., Faruolo, M., Pergola, N., & Tramutoli, V. (2013b). Toward the estimation of river discharge variations using MODIS data in ungauged basins. *Remote Sensing of Environment*, *136*, 47–55. https://doi.org/10.1016/j.rse.2013.04.010

Tarpanelli, A., Iodice, F., Brocca, L., Restano, M., & Benveniste, J. (2020). River Flow Monitoring by Sentinel-3 OLCI and MODIS: Comparison and Combination. *Remote Sensing*, *12*(23), Article 23. https://doi.org/10.3390/rs12233867

Tarpanelli, A., Santi, E., Tourian, M. J., Filippucci, P., Amarnath, G., & Brocca, L. (2019). Daily River Discharge Estimates by Merging Satellite Optical Sensors and Radar Altimetry Through Artificial Neural Network. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(1), 329–341. https://doi.org/10.1109/TGRS.2018.2854625

Tazioli, A. (2011). Experimental methods for river discharge measurements: Comparison among tracers and current meter. *Hydrological Sciences Journal*, *56*(7), Article 7. https://doi.org/10.1080/02626667.2011.607822

Temini, M., Lacava, T., Tarendra, L., Tramutoli, V., Ghedira, H., & Riadh, A. (2011). A multi-temporal analysis of AMSR-E data for flood and discharge monitoring during the 2008 flood in Iowa. *Hydrological Processes*, *25*(16), 2623. https://doi.org/10.1002/hyp.8020

Tourian, M. J., Schwatke, C., & Sneeuw, N. (2017). River discharge estimation at daily resolution from satellite altimetry over an entire river basin. *Journal of Hydrology*, *546*, 230–247. https://doi.org/10.1016/j.jhydrol.2017.01.009

Tourian, M. J., Tarpanelli, A., Elmi, O., Qin, T., Brocca, L., Moramarco, T., & Sneeuw, N. (2016). Spatiotemporal densification of river water level time series by multimission satellite altimetry: SPATIOTEMPORAL DENSIFICATION OF ALTIMETRY OVER RIVERS. *Water Resources Research*, *52*(2), 1140–1159. https://doi.org/10.1002/2015WR017654

Van Dijk, A. I. J. M., Brakenridge, G. R., Kettner, A. J., Beck, H. E., De Groeve, T., & Schellekens, J. (2016a). River gauging at global scale using optical and passive microwave remote sensing. *Water Resources Research*, *52*(8), 6404–6418. https://doi.org/10.1002/2015WR018545

Van Dijk, A. I. J. M., Brakenridge, G. R., Kettner, A. J., Beck, H. E., De Groeve, T., & Schellekens, J. (2016b). River gauging at global scale using optical and passive microwave remote sensing. *Water Resources Research*, *52*(8), Article 8. https://doi.org/10.1002/2015WR018545

Yang, X., Pavelsky, T. M., Allen, G. H., & Donchyts, G. (2020). RivWidthCloud: An Automated Google Earth Engine Algorithm for River Width Extraction From Remotely Sensed Imagery. *IEEE Geoscience and Remote Sensing Letters*, *17*(2), 217–221. https://doi.org/10.1109/LGRS.2019.2920225

Zakharova, E. A., Kouraev, A. V., Cazenave, A., & Seyler, F. (2006). Amazon River discharge estimated from TOPEX/Poseidon altimetry. *Comptes Rendus Geoscience*, *338*(3), 188–196. https://doi.org/10.1016/j.crte.2005.10.003

Zakharova, E., Nielsen, K., Kamenev, G., & Kouraev, A. (2020). River discharge estimation from radar altimetry: Assessment of satellite performance, river

scales and methods. *Journal of Hydrology*, *583*, 124561. https://doi.org/10.1016/j.jhydrol.2020.124561

Zebin, T., Sperrin, M., Peek, N., & Casson, A. J. (2018). Human activity recognition from inertial sensor time-series using batch normalized deep LSTM recurrent networks. *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1–4. https://doi.org/10.1109/EMBC.2018.8513115

Zhang, J., Zhu, Y., Zhang, X., Ye, M., & Yang, J. (2018). Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. *Journal of Hydrology*, *561*, 918–929. https://doi.org/10.1016/j.jhydrol.2018.04.065

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, *118*, 83–94. https://doi.org/10.1016/j.rse.2011.10.028