A Scalable Open-source Web-analytic Framework to Improve Satellite-based Operational Water Management in Developing Countries

Nishan Kumar Biswas

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science in Civil Engineering

University of Washington

2017

Reading Committee:

Faisal Hossain

Bart Nijjssen

Program Authorized to Offer Degree:

Department of Civil and Environmental Engineering

University of Washington

Abstract

A Scalable Open-source Web-analytic Framework to Improve Satellite –based Operational Water Management in Developing Countries

Nishan Kumar Biswas

Chair of the Supervisory Committee:
Associate Professor Faisal Hossain
Department of Civil and Environmental Engineering

Satellite based observations are becoming more widespread for operational and decision making activities in water resources due to their near real-time availability as well as minimization of difficulties related to ground based hydrological measurements. In this study, two crucial hurdles to advance real-world operationalization of satellite datasets are addressed. First, a simple, easy-to-access, easy-to-build, robust and consistent open-source web portal is developed connecting complex models to facilitate satellite based observations by developing nations. Second, to enhance the predictability skill of satellite-based estimations, an innovative and dynamic web analytics-based correction system is developed to reduce the uncertainty related to indirect method of measurement by the satellites. The correction system comprised of precipitation

correction and streamflow correction. In precipitation correction, web crawled precipitation from different sites of government agencies of South Asia region are used to estimate bias in satellite based precipitation. This dynamically bias adjusted precipitation is then used to simulate hydrological models to obtain streamflow. Simulated streamflow is corrected using discharge climatology and gauged discharge observations. These corrected datasets are finally shared through South Asian Surface Water Modelling System (SASWMS) web portal. On an average, these dynamic correction techniques reduced RMSE in streamflow by 80-90%. The take-home message is that with the growth of open-source tools, it is now possible for any resource-constrained water management agency to build robust and cost-effective operational web portals based on satellite data and non-proprietary software. With dynamic correction features that can also be built from open-source tools, the well-known limitations of satellite-based operationalization can be overcome for water management agencies of developing countries.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I wish to express my gratitude and sincerest appreciation to my supervisor Dr. Faisal Hossain for allowing me to do my research under his supervision, expertise, kindness, patience and most of all for intellectual support in all respect. I am also thankful for the support of Dr. Bart Nijssen and his help as a committee member.

I would like to thank Department of Civil and Environmental Engineering (CEE) for supporting me throughout my graduate studies. I am also thankful to professionals of Flood Management Division of Institute of Water Modelling, Bangladesh for their help during this study. I am grateful to the Sustainability, Satellites, Water and Environment (SASWE) Research Group members for their co-operation in every aspect of my research work.

This acknowledgement is not complete without the mention of my parents, other family members and friends for their encouragement. Last, but not least, I would like to thank my wife, Maumita Sharma for her patience and support throughout the study-period.

DEDICATION

This thesis is dedicated to my wife, Maumita Sharma and my son, Arindam Biswas

Chapter 1. INTRODUCTION

1.1 INTRODUCTION

Transboundary water bodies create hydrological, social and economic interdependencies between nations. Thus management of water resources in an integrated way is vital for sustainable development and reducing poverty. There are more than 260 rivers and lakes which are shared by countries and holds 60 percent of the global fresh water (Hossain et al., 2013; Wolf et al., 1999). The Ganges-Brahmaputra (GB) river basin along with Meghna Basin forms Ganges-Brahmaputra-Meghna (GBM) Basin in South Asia. The GBM Basin is considered as the third largest freshwater outlet which is exceeded only by Amazon and Congo river basins. Countries sharing the Ganges-Brahmaputra basin are Bangladesh, India, Nepal, Bhutan with a combined population of more than 630 million. Total area of the basin is over 1.6 million km^2 and catchment area in each country is provided in Table 1.1.

Table 1.1. Contribution of Ganges Brahmaputra Basin in South Asian Countries

Basin	Area (km^2)	Country	Area of Country in Basin (km^2)	Percentage of Total Area of the Basin	Percentage of Total Area of the country
Ganges	1,087,300	India	860000	79	26
		China	33500	3	0.3
		Nepal	147500	14	100
		Bangladesh	46300	4	32
Brahmaputra	543,400	India	195000	36	6
		China	270900	50	3
		Bangladesh	39100	7	27
		Bhutan	38400	7	100

The GBM region is endowed with considerable natural resources that can be used to accelerate sustainable growth. In the recent decades, population increase and urbanization has put more stress on water to fulfill the requirement for agricultural, municipal and industrial uses. Moreover, change in climate pattern has added pressures on many transboundary water resources and affected downstream water availability. An integrated modeling framework is required that can be used as a tool to manage water resources and promote sustainable development in this region.

1.2 BACKGROUND

There are differences in social and economic aspects in managing water resources infrastructure both in political and legal contexts between the nations of transboundary rivers. Water scarcity is growing day-by-day as a result of expanding population demand, environmental concerns and climate change effects and thus cooperation in water resources in large and transboundary water management is necessary.

The most reliable method of estimating amount and availability of water is ground measurements. However, for a feasible water resources research and management approach in a large river basin, estimation of the key components of water cycle is impossible by a purely ground-based observational approach. The ground observation network to monitor water fluxes is also not sufficiently dense in most parts of the world (Shiklomanov et al., 2002; Vörösmarty et al., 2002). Another difficulty is to collect in-situ measurements in real time for making decision in an operational setting. It is impossible to measure all the variables and their changes (e.g. surface runoff, evaporation) in spatial and temporal scale. It is particularly difficult in the developing world due to budgetary constraints as well as due to poor institutional capacity between the riparian nations to share the data. Hydrological modelling driven by insufficient ground or proxy observations can be utilized as an alternative approach for better understanding of the physical processes of the water cycle (Siddique-E-Akbor et al., 2014).

Hydrological models help to understand the past and current state of water resources in the basin and provide a way to explore the implications of management decisions and imposed changes (such as climate change or human impacts of water diversions). The purpose of hydrological modeling at this scale is primarily to support decision making for water resources management planning. For example, it includes resource assessment (resource condition and trend, including spatial and temporal variation in availability, sufficiency, equity and sustainability of supply), vulnerability assessment (response of the system to climate change, demographic change), impact assessment (effectiveness and impacts of proposed development), early warning and flood risk assessment (Johnston and Smakhtin, 2014). Hydrological parameters that can be simulated using hydrological models include surface and subsurface fluxes (e.g. runoff, infiltration, base flow, soil moisture etc.) and streamflow. These parameters are important for sustainable water resources planning, seasonal and extreme event forecasting, crop or drought management etc. (María et al., 2016; Siddique-E-Akbor et al., 2014). Thus hydrologic model can bridge gaps in in-situ measurement as well as keep track of the terrestrial component of the dynamic water cycle to aid the decision making in water management.

Although models and updated computational facilities can minimize the dependence on routine in-situ measurements, there are some key meteorological parameters continuously needed by the model known as forcing parameters. No matter the complexity of the model, every hydrological model aggregate at some level of complexity and requires spatially and temporally meteorological inputs, vegetation and subsurface properties (Vrugt et al., 2007). The major inputs are invariably precipitation and often include other parameters like surface temperature, humidity and wind speed. As stated earlier, due to lack of effective relationships between the countries, it is difficult to obtain in situ meteorological data for basin scale simulation to build a decision making system in transboundary setting.

To overcome the situation of lack of in situ measurements and the dependency of downstream riparian nations on the upstream nations, satellite remote sensing platforms now provide an opportunity to create a level playing field in water resources management through global scale hydrological observations and their near real-time availability. Satellite observations have brought an innovative way of solving water resources conundrum with the availability of remotely-sensed data from different sensors of various wavelengths and wide range of spatiotemporal, radiometric and spectral resolutions. For over a decade, satellite observations have been used for various weather and climate prediction studies and applications at operational scales (Kansakar and Hossain, 2016; Khan et al., 2012; Woldemichael et al., 2012; Gebregiorgis and Hossain, 2011; Nijssen and Lettenmaier, 2004). Several integrated hydrological and water resources modelling systems have been developed based on the satellite data products to enable hydro-meteorological studies and applications (e. g. Global Land Data Assimilation System (GLDAS) - Rodell et al., 2004; Brown et al., 2014).

Despite these advancements, there are some issues that limit the application and predictive capability of remote sensing in basin scale water resources management. In figure 1.1 and 1.2, an example of the unacceptable quality of satellite observations (precipitation) is shown. Average ground measured in situ precipitation over Ganges Basin is compared with the average precipitation obtained from GPM mission and known as IMERG-RT precipitation (Hou et al., 2014; Huffman et al., 2015) from June to August 2016. Ground measured rainfall is collected from National Climatic Data Center – Global Surface Summary of the Day (NCDC-GSOD) of World

Meteorological Organization. The figure shows the amount of overestimation in IMERG product compared to the in-situ observed precipitation.

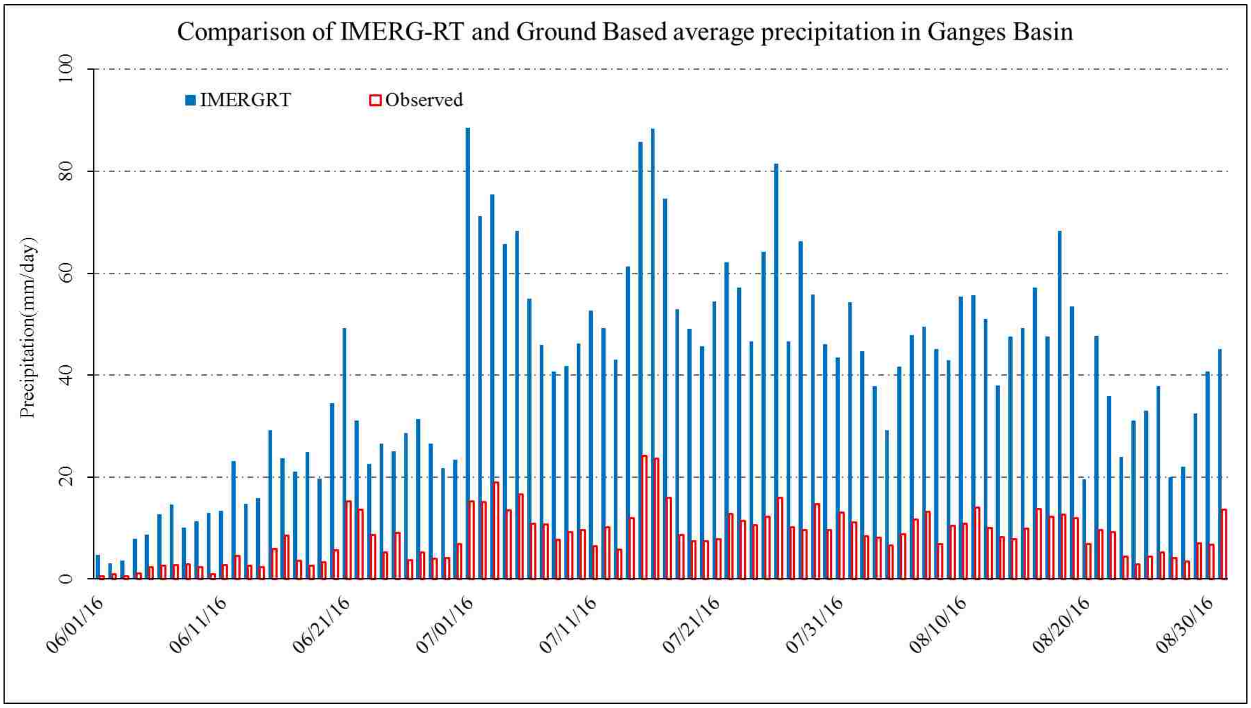


Figure 1.1. Comparison of National Climatic Data Center-Global Surface Summary of the Day (NCDC-GSOD) precipitation with the Latest satellite estimated precipitation product, IMERG-RT estimated precipitation in Ganges Basin

Another example is shown in figure 1.2 where streamflow in the Brahmaputra Basin is simulated using Variable Infiltration Capacity (VIC) Hydrological Model (Liang et al., 1994) and IMERG-RT precipitation, then compared with the observed discharge (rated from observed water level) at Bahadurabad location (considered as the outlet of the Brahmaputra Basin).

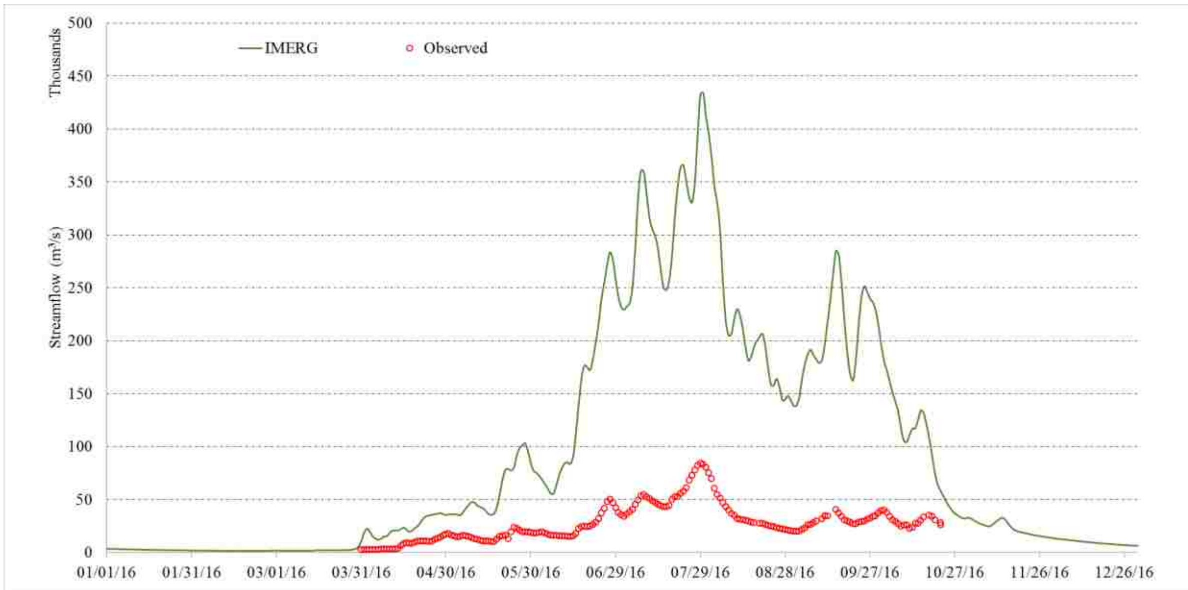


Figure 1.2. Comparison of simulated streamflow using IMERG RT precipitation with the rated discharge at the outlet of Brahmaputra River Basin.

1.3 PROBLEM STATEMENT

Two critical solutions are addressed in the study which are needed to empower stakeholder agencies to become independent users of satellite data for operational water management. These are: 1) an open source and integrated framework that connects complex backend models with front end user needs (which can be built with almost no cost and the users and stakeholder community can built themselves with the power of information technology); 2) an automated correction system that can harness in-situ data availability on the public domain to improve accuracy of satellite data; such a system should be able to take advantage of power of internet and avoid unphysical/unrealistic simulations (as shown in Figure 1.1 and 1.2) due to satellite's indirect method of estimating water cycle variables.

Development of these two solutions is necessary as information technology (IT) development has progressed significantly in the realm of the open-source/non-proprietary community. There are now powerful non-proprietary tools available to empower end users and stakeholder agencies in the developing world and bypass cost-prohibit proprietary software that most developing nations cannot afford (Castronova et al., 2013; Horsburgh, 2009). At present, there are several open source tools and software in the internet that can be used to build customized environment without any hurdles. However, there is no consistent template or methodology for taking advantage of such open-source interface building approach for operationalization of satellite data for water management.

The open-source community now needs to formalize a framework for the water management community. The internet is now being progressively upgraded and updated with vast amounts of in-situ and proxy observations on the water cycle and meteorological conditions that are publicly available. These datasets are often posted online in 'nowcast' mode and remain heavily 'untapped' for dynamic adjustment of satellite observations that otherwise yield unrealistic model simulations. For example, in South Asia, more than a dozen of agencies (see Appendix A.1), to the best of our knowledge, that post only the most current day's measured rainfall on their website for several hundred locations in GBM. This online availability, although limited in record as being only a 'nowcast', provides an opportunity to pursue simple adjustment techniques on the fly and explore if such publicly available data can improve the skill of operational satellite-based hydrologic simulation.

1.4 Objective of the Study

Two key objectives of the study are addressed. Using the Ganges-Brahmaputra-Meghna river basins as an example, these objectives are:

1) To develop an open-source web interface building system that is simple and easy to implement for agencies of the developing world as a 'build it yourself' template for water management;

2) To explore the effectiveness of online and dynamic data quality improvement techniques that leverage the public domain in-situ data posted on the internet to correct satellite data on the fly through web-analytics.

1.5 Literature review

The pressing water management issues that lie in the Ganges-Brahmaputra basin are presented in this section. This section also overviews the hydrological modelling application in large river basin along with satellite based hydrological observations to address potential solutions for water management. The review is broken down into three topics: transboundary issues in Ganges-Brahmaputra; potential of satellite based hydrological observations; and current limitations of application of satellite data.

1.5.1 *Transboundary Issues*

International river basins, including lakes and shallow groundwater, shared by more than one country cover almost half of Earth's land surface (Wolf et al., 1999, Hossain et al., 2013). According to Katiyar and Hossain (2007), about thirty-three countries situated at the most downstream of IRBs in the developing world are heavily depended on hydrologic information from the upstream riparian nations and are challenged in basin-wide hydrologic modeling due to institutional and cost issues. In this era of high frequency computational facilities along with more real-time observational power and engagement in management of water resources, we have very limited knowledge on spatial and temporal dynamics of stream flow and changes in storage globally (Alsdorf et al., 2007). There are several recent success stories on how integrated water resources management can be implemented in International River Basins. Cambodia, Laos, Thailand and Vietnam, have been started to cooperate within the framework of the Mekong River Commission, a framework for the Nile River Basin which is shared among 10 countries and 160 million people was agreed in February 1999 in order to fight poverty and spur economic development in the region by promoting equitable use of, and benefits from, common water resources (UNDP, 2006).

In the GBM basin, most of the problems arise due to the conflicting interests of upstream and downstream riparian nations. Rapidly growing population, expanding agricultural and industrial activities besides the impacts of climate change have resulted in fresh water resources that are stressed. The development and management of the GBM basins have been subject to a number of geopolitical constraints in spite of having huge potential for being a great example of regional cooperation (Brichieri-Colombi and Bradnock, 2003).

A map of GBM Basin is shown in figure 1.3 which shows the basin along with its tributaries and distributaries. Headwaters of both of Ganges and Brahmaputra river originate from the Himalayan Range where glacier melt water is an important source for the headwaters of the river. The Ganges river flows southwest into India and then turns southeast, being joined by many tributaries. The Brahmaputra river (known as Yalung Zangbo in China) flows east through the southern area of China, then flows south into eastern India, turns southwest, then enters Bangladesh (where it is also called Jamuna) before merging with the Ganges. After flowing into Bangladesh, the Ganges, Brahmaputra joined together at the Goalanda Ghat and named Padma and then after about 70 km further down the combined discharge joins the Meghna River at

Chandpur. The combined stream is called the Meghna river, which 90 km further downstream discharges into the Bay of Bengal (McEwen, 2008).

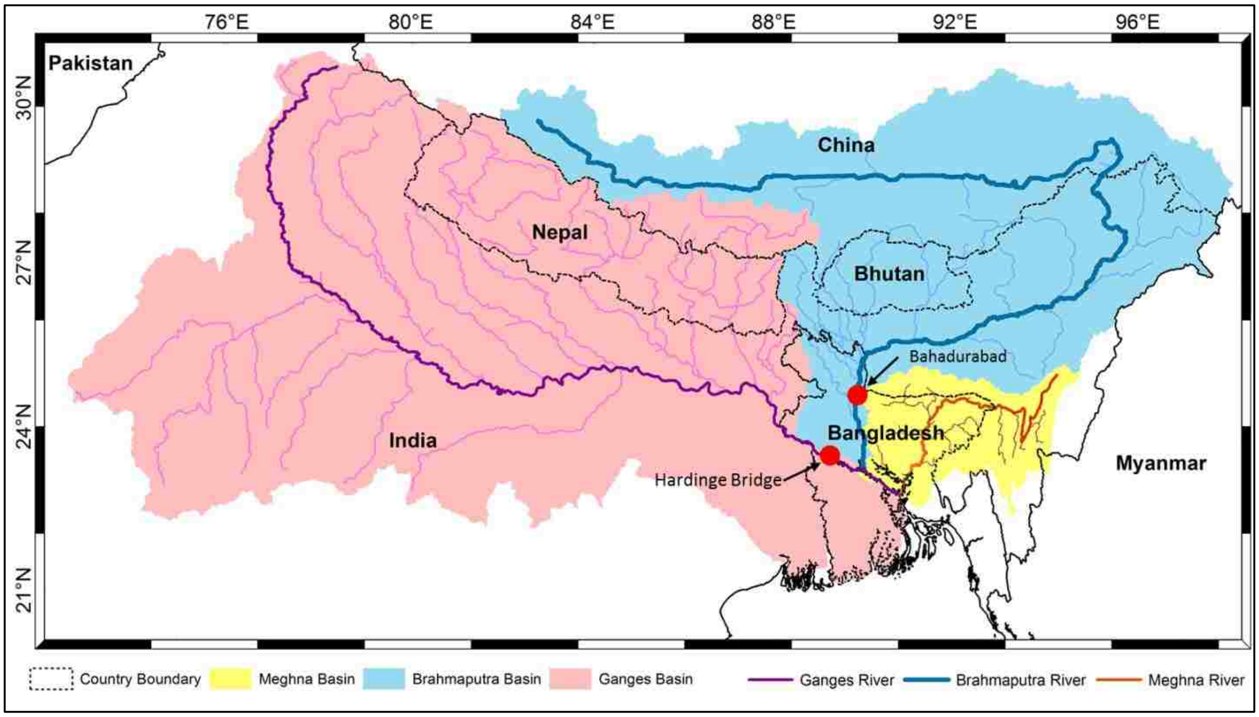


Figure 1.3. Ganges-Brahmaputra-Meghna Basin

There are a number of common water resource management issues for all countries in the GBM basins. There is extreme variation between the peak flows and lean flows in the river systems from monsoon to dry season where 80 to 90 percent of annual rainfall is concentrated in 4–5 months of the monsoon (Hossain et. al., 2014). This significant amount of precipitation is responsible for recurrent floods during the monsoon that cause damage to life, property, and infrastructure. This uneven distribution of rainfall creates the problem of low water availability during the dry season. The spatial variability creates water-stressed conditions in some parts of the basin that often triggers water sharing problems between the state boundary as well as the international boundary (Ahmad and Ahmed, 2003; Biswas, 2008).

In the case of GBM, Bangladesh is the largest delta in the world situated at the most downstream end of the Ganges-Brahmaputra-Meghna Basin (Paudyal, 2002) shown in the figure 1.3. Only 8.5 % of the river basin lies within this country, and the rest 91.5% lies in four different countries: India, China, Nepal, and Bhutan (Hossain and Alsdorf, 2010). Annual average renewable fresh water resources in the country is around 1210.6 BM^3 out of which only (9%) is locally generated and the rest of flood flows enter into the country from upstream catchments in India through 54 border rivers (Rashid 1991).

Moreover, about 80% of the area of Bangladesh is covered with the flood plains of Ganges-Brahmaputra-Meghna Basin (Hofer 1998). During the peak period, the entire volume of the GBM river systems (about 142,000 m^3) discharges every second into the Bay of Bengal through a single outlet at the Lower Meghna in Bangladesh (Rahman et al., 1990). Approximately 91% of flood flows enter into Bangladesh from upstream catchments in India through 54 border rivers (Rashid, 1991). In the northern northeast region, the country is overwhelmed by flash floods, and tidal floods and occasional cyclonic storm-surge floods in the coastal region. The construction of Farakka Barrage caused by India increased salinity intrusion and threatened agriculture and ecosystem including the Sundarbans; river erosion and widespread arsenic contamination of groundwater (Chowdhury et al., 1997; Brichiere-Colombi and Bradnock, 2003).

Therefore, incorporation of satellite remote sensing–derived hydrologic information is worthy of exploration for water management in nations of international river basins faced with similar transboundary hurdles. By using an integrated modeling framework derived on the basis on the satellite observations, one can achieve better decision making as well as improve basin-wide monitoring of the dynamics of storage, water flow and availability. Forecasting of extreme events using upstream water availability can help farmers with sufficient lead time (say 7-10 days), then it can be much more helpful for agricultural (sowing or early reaping of crops) decision support to the farmers (Hossain et al., 2017).

1.5.2 *Potential of Satellite Observations*

Water resources management can benefit from the application of remote sensing. Satellite and airborne remote sensing observations can provide valuable information and can be used to support the operational water resources community for management practices and decisions (Bolten et al., 2015). For over a decade, satellite observations have been used for various weather and climate prediction studies and applications at operational scales (Kansakar and Hossain, 2016; Khan et al., 2012; Gebregiorgis and Hossain, 2011; Nijssen and Lettenmaier, 2004). For example, satellite altimetry technique for measuring height of water bodies form space has been proven to be a viable method for river height extraction (Alsdorf et al., 2007).

Satellite observations can indirectly estimate several variables of the water cycle (such as soil moisture, river height, stream flow, vegetation cover etc.) that can be used to force, calibrate or validate hydrological models and allow decision making for water management in challenging situations such as IRBs in the developing world (Musa et al., 2015; Gebregiorgis and Hossain, 2014). A comprehensive review of the application of satellite based estimation of extent of flood inundation, stream flow and water stage is presented in several studies (Smith, 1997; Schumann et al., 2009). Several integrated hydrological and water resources modelling systems have been developed based on the satellite data products to enable hydro-meteorological studies and applications (e. g. Global Land Data Assimilation System (GLDAS) - Rodell et al., 2004; Brown et al., 2014).

Specific examples of satellite earth observation applications are broadly discussed in Kansakar and Hossain (2016). They explored some of the most successful applications in land cover/land use change, carbon biomass assessment, disaster management, floods, landslide and water resources. The use of satellite remote sensing data is expected to increase further for International River Basins (IRBs) with the proposed missions that are planned for routine estimation of certain hydrological variables, such as river discharge (Surface Water and Ocean Topography, Alsdorf et al., 2006), soil moisture (Hydrospheric State Mapper; Entekhabi et al., 2005). Maswood and Hossain (2016) discussed the emerging role that satellite precipitation can potentially have in improving the flood forecasting limitations of downstream nations in IRBs of flood-prone developing nations.

However, several challenges remain for effective application these data, such as uncertainty, limited sampling and skill. Other challenges include identifying and categorizing near-term priority observations and finding opportunities to address near-term information gaps or limitations for improving access to data, information, and decision making. Some of these challenges are discussed next.

1.5.3 *Limitation of Satellite Observations*

In-situ rainfall measurement covers a very few centimeter area and very much limited in spatial characterization. Whereas satellite rainfall estimates are indirect method of measurement (from infrared, passive microwave, or radar sensors). Their spatial resolution typically varies from 0.04° to 0.25° with a precipitation value representative of an area 16 and 625 square kilometers respectively (Ebert et al., 2007). Time steps of measurement ranging from half an hour to three hours or a day. The difference in footprints and algorithms used causes the uncertainties in amount of precipitation from the in-situ measured observations. The uncertainties can be summarized in different groups. It may be due to, 1) temporal sampling scale from sub daily to annual, 2) Spatial Scale ranges from cell to basin and 3) Latency, which sometime causes the data invaluable for operational decision making.

Quality of satellite data can often become unacceptable, resulting in simulations that are found limited in skill or unrealistic for decision making. A good example is satellite precipitation estimation, where the uncertainties at smaller space-time scales are known to be complex and often the limiting factor to its operational use for hydrological applications (Hossain and Huffman, 2008). International precipitation working group has taken an initiative to assess the performance of precipitation algorithms over several parts of the world and then published in their websites. Brown (2006) assessed suitability of the precipitation estimates by simulating hydrological models to get the discharge. Ebert et al. (2006) examined 12 precipitation products either from satellite or NWP models and they validated the datasets over Australia, United States and Northwestern Europe. They studied the relative performance in seasonal variation and also spatial variation and got that in the summer and in low latitudes, satellite estimates work better than in the winter with deep convection.

In addition to data quality issues, satellite observations also suffer from delayed transmission (i.e., latency) and various data formatting issues whose awareness is mostly limited to scientific community but not to the application world. When these issues are considered in sum, the increasing observational capability of satellites will not have an equivalent impact on increasing societal applications until creative and cost-effective solutions are devised to improve the utility of satellite data for decision makers (Hossain, 2015, Bulatewicz, 2014).

1.6 Thesis Outline

The chapters of the thesis are outlined as follows. Chapter 2 is about materials and methods. It presents the hydrological models and datasets that are used in the study. Using these model and datasets, a simple and robust “build it yourself” type portal development methodology is discussed elaborately under ‘methods’. For quality assurance of the satellite based hydrological model simulated results and to increase the predictability of remote sensing applications, a dynamic and on-the-fly correction method based on the power of information technology and internet is proposed. The developed portal with backend models and web based real time correction system is applied to Ganges Brahmaputra Basin of South Asia region. In the Results and Analysis chapter (Chapter 3), performance of the integrated framework and the correction system is assessed. Finally, findings and conclusions are discussed in the Conclusion and Recommendation chapter (Chapter 4).

Chapter 2. MATERIALS AND METHODS

2.1 HYDROLOGICAL MODEL

In this study, Variable Infiltration Capacity (VIC) hydrological model is used for simulating water fluxes. VIC model is a semi-distributed macro scale research grade hydrological model and first developed by Liang et al. (1994). It is capable of solving full water balance and energy balances. The basic structure of the VIC model is described in detail by Liang et al. (1994). Cherkauer et al. (2003) provides updates for cold land process and Bowling and Lettenmaier (2010) provided updates for lakes and wetlands to the basic structure of this model. The model has been widely applied for purposes such as seasonal hydrological forecasting (Nijssen et al., 2001; Yuan 2016), climate change impacts studies (Wu et al., 2012, Zhao et al., 2012), and water and energy budget studies (Haddeland et al., 2007, Gao et al., 2010), among various other applications. VIC's distinguishing hydrologic features are its representation of the role of sub-grid variability as a control on soil water storage and in turn runoff generation and its parameterization of base flow, which occurs from a lower soil moisture zone as a nonlinear recession (Dumenil and Todini 1992).

The basic model features of VIC are as follows: 1) the land surface is modeled as a (lumped) grid of large (0.1 km), flat, uniform cells; 2) inputs to the model are time series of daily or sub-daily meteorological drivers (e.g., rainfall, snow, air temperature, wind speed); 3) land–atmosphere fluxes and the water and energy balances at the land surface are simulated at a daily or sub-daily time step, and water can only enter a grid cell via the atmosphere; 4) grid cells are simulated independently of each other, and the entire simulation is run for each grid cell separately, one grid cell at a time, rather than for each time step, looping over all grid cells. A schematic view of VIC Land cover tiles along with the soil column and major fluxes are shown in figure 2.1. The VIC model used in this study is collected from a previous application on the GBM basins by Siddique-E-Akbor et. al. (2014).

To generate the streamflow from the simulated water cycle variables from VIC Model, a separate model named as Horizontal Routing Model is used which is developed by Lohmann et al. (1998).

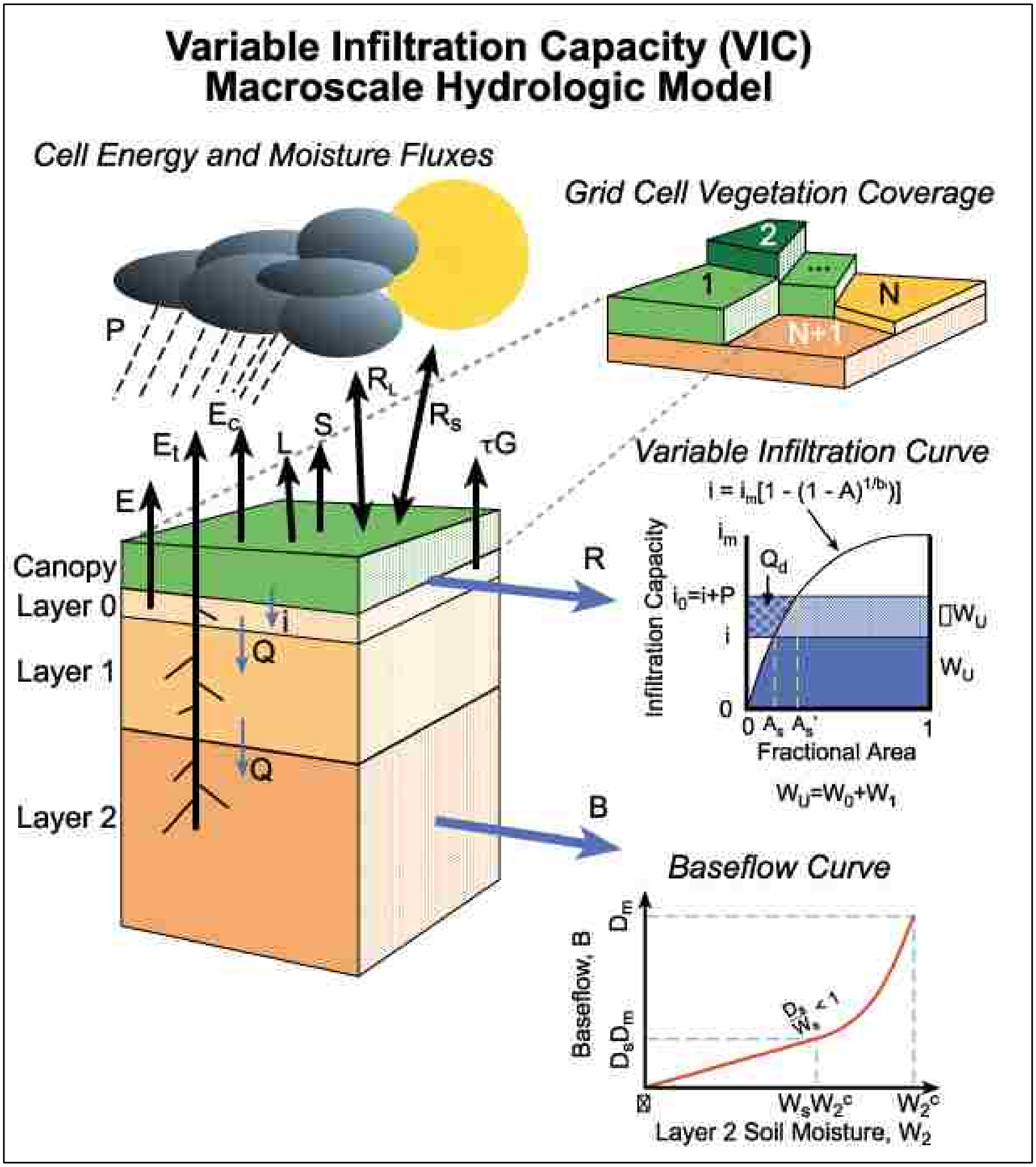


Figure 2.1. Schematic diagram of Variable Infiltration Capacity (VIC) Hydrological Model. (Image Source: Figure 1 of http://vic.readthedocs.io/en/master/Overview/ModelOverview/)

2.2 DATASETS

2.2.1 *Satellite Estimated Datasets*

Four types of satellite based meteorological parameters (precipitation, maximum and minimum temperature and average wind speed at 2-meter height) are used in this study to demonstrate the applicability of satellite estimates for operational water management. NASA Global Precipitation Measurement (GPM) estimated Integrated Multi-satellitE Retrievals for GPM (IMERG) early run (Hou et al., 2014 and Huffman et al., 2015) dataset is used to estimate amount of precipitation. Other meteorological forcing parameters are retrieved from NCEP (National Center for Environmental Prediction) final analysis server.

The Global Precipitation Measurement (GPM) mission is an international constellation of one core observatory satellite and approximately ten partner satellites. With joint effort of NASA and Japan Aerospace Exploration Agency (JAXA), the GPM Core observatory was deployed on February 28, 2014 to the TRMM satellite which was launched at November 27, 1997. A dual-frequency precipitation radar and a conical-scanning multichannel GPM Microwave Imager is carried by the GPM Core observatory. In figure 2.2, the GPM core observatory is shown. Due to embedding of improved instruments with TRMM Microwave imager, the GPM can detect light and solid precipitation more accurately than all of the previous precipitation sensors (Hou et al., 2014; Huffman et al., 2015). IMERG precipitation derivation algorithm inter-calibrates, merges and interpolates all satellite passive microwave precipitation estimates together with microwave calibrated infrared satellite estimates, monthly precipitation gauge analyses other precipitation estimators at fine time and space scales (Huffman et al., 2015). The IMERG products are characterized by high temporal and spatial resolutions (half-hour and 0.1° × 0.1°).

Figure 2.2. Global Precipitation Mission Core Observatory

(Image Source: https://eospso.gsfc.nasa.gov/sites/default/files/publications/What%27s%20Up%20With%20Precip_508.pdf, Page 6)

There are several reasons to select GPM derived IMERG product for this study. Besides the enhanced measurement and sampling capabilities of GPM, it also offers many advanced science contributions and societal benefits. It provides improved knowledge of the Earth's water cycle and its link to climate change. Also new insights into storm structures and large-scale atmospheric processes, precipitation microphysics are some of the advanced feature of IMERG product. The amount of precipitation improved forecasting abilities for natural hazards, including floods, droughts and landslides and also agricultural crop forecasting and monitoring of freshwater resources (https://pmm.nasa.gov/GPM).

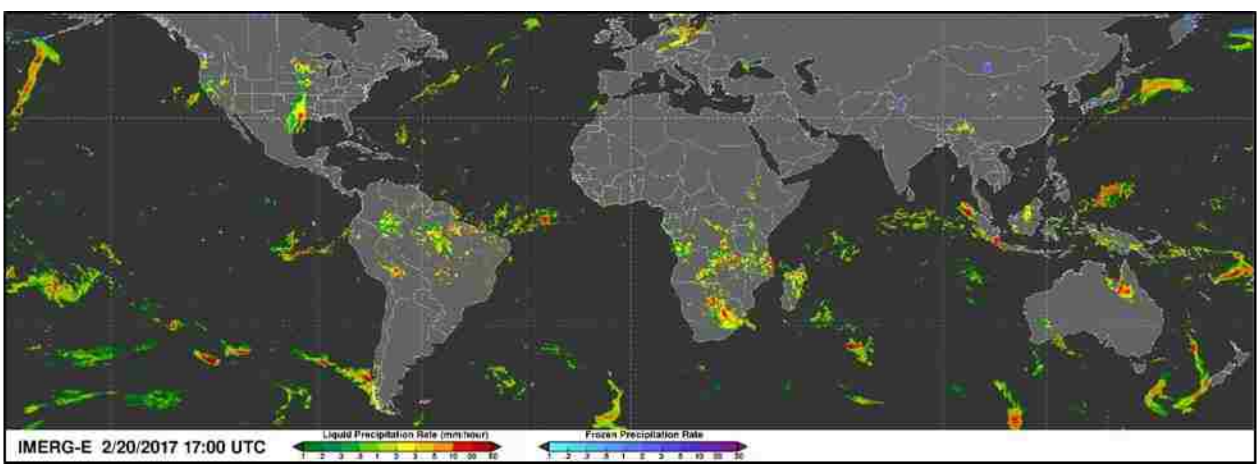


Figure 2.3. Half Hourly Global precipitation rate estimated by IMERG for 2017-02-20 17:00

(Image Source:
https://trmm.gsfc.nasa.gov/trmm_rain/Events/ATLA/latest_big_half_hourly_gridded.jpg)

Daily maximum and minimum temperature and average wind speed datasets are collected from NCEP FNL Operational Model Global Tropospheric Analyses (National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce, 2000). These datasets downloaded from https://rda.ucar.edu/data/ and they are derived from the NCEP FNL Server comprising temperature and wind speed. These NCEP Final Operational Global Analysis data are on 1-degree by 1-degree grids and prepared operationally every six hours. This product is from the Global Data Assimilation System (GDAS), which continuously collects observational data from the Global Telecommunications System (GTS), and other sources. The final products are prepared about one hour after the Global forecasting datasets are initialized so that more observational data can be utilized.

2.2.2 *NCDC-GSOD Datasets*

Global Surface Summary of the Day (GSOD) is an in situ observational data of National Climatic Data Center (NCDC) and derived from Integrated Surface Hourly dataset. ISH dataset includes global data obtained from the USAF Climatology Center. Latency of the latest daily summary data is 1-2 days and over 9000 stations data are normally available. The data are reported and summarized based on Greenwich Mean Time (GMT). A quality control checking is done before open the dataset to the public domain (e.g. precipitation amount will only appear if the station reports the data sufficiently to provide a valid value). Several hydrological parameters available in GSOD datasets. Of them, precipitation amount, mean wind speed, maximum temperature, minimum temperature is used for the performance analysis of satellite observations (https://www7.ncdc.noaa.gov/CDO/cdoselect.cmd?datasetabbv=GSOD). Number of stations found in NCDC-GSOD Server in Ganges-Brahmaputra Basin and their spatial distribution are shown in figure 2.4.

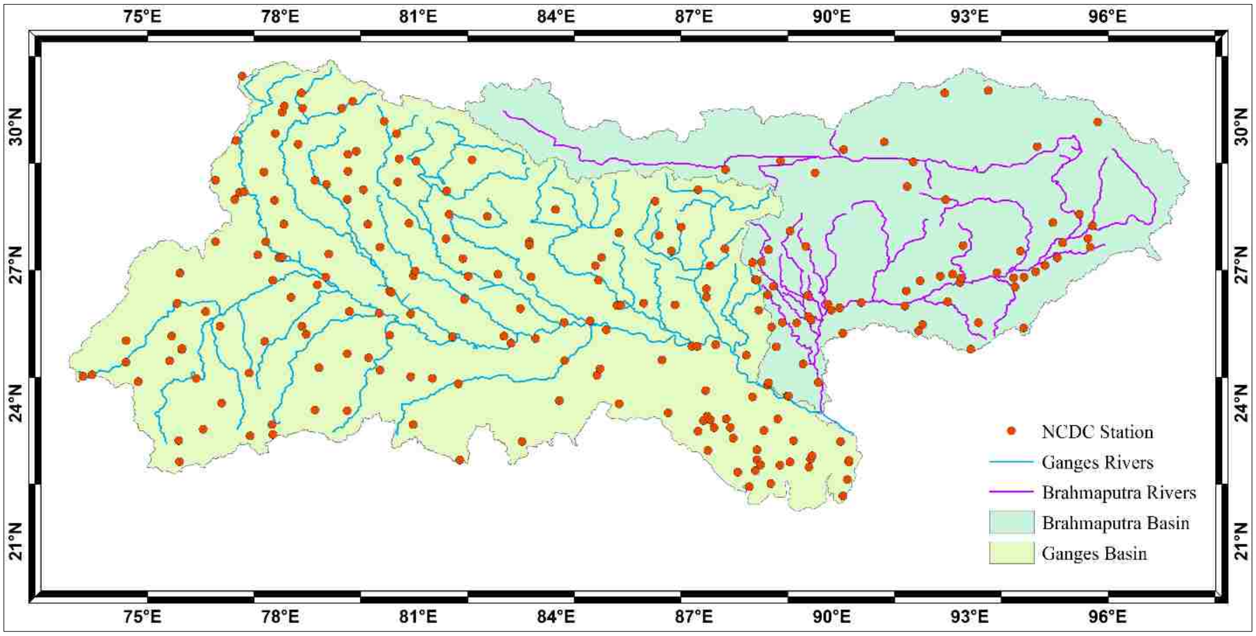


Figure 2.4. Spatial distribution of NCDC-GSOD Stations in Ganges-Brahmaputra Basin

2.3 OPEN-SOURCE "BUILD-IT-YOURSELF" WEB PORTAL DEVELOPMENT

To refresh, the study had two specific objectives: 1) to develop an open source web portal that connects complex models and can be used as an interface template for operational water management by stakeholder agencies, 2) to develop a web-based real-time correction system for dynamically increasing the accuracy of satellite based observations.

To fulfil the first study objective, a very simple, well organized, easy to navigate and consistent web portal is developed with the facility of data visualization and downloading. Because the web portal is implemented over the GBM basin, it will be called the South Asian Surface Water Modeling System (SASWMS) hereafter (http://depts.washington.edu/saswe). A fast loading and consistent layout enabled template is downloaded, necessary *HTML (Hypertext Markup Language)*, *Cascading Style Sheet (CSS)* codes were modified, and *JavaScript* codes are added in the webpages to make it more dynamic. Facilities are provided to visualize and download necessary observations and simulation results from the portal. Figure 2 illustrates how the portal is developed using free-of-cost online resources.

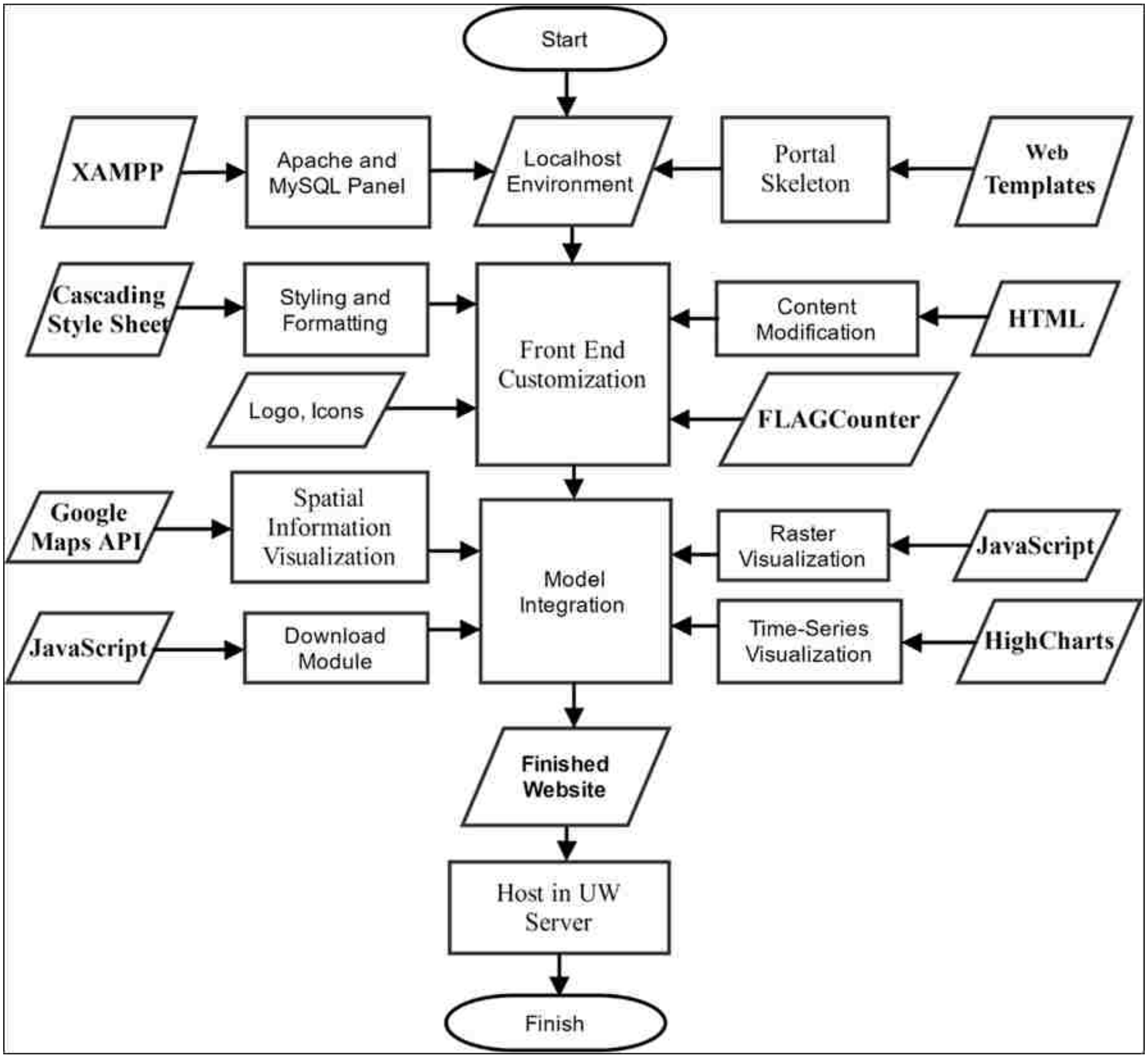


Figure 2.5. SASWMS Portal Development Methodology

2.3.1 *Localhost Environment Development*

For the purpose of the development of the portal, first step was to develop a localhost environment. The localhost is the default name describing the local computer address also known as the loopback address. It is considered as the most convenient way for network developers and programmers to test their programs and checks close at hand, rather than have everything go onto the internet, every time they want to test something they are working on. There are several tools available in the internet to make localhost. In building SASWMS portal, *XAMPP* is chosen to make the localhost environment. *XAMPP* stands for *Cross-Platform (X), Apache (A), MariaDB*

(M), PHP (P) and Perl (P). It is a completely free, easy to install apache distribution under the terms of GNU Public License. The main use of *XAMPP* is it facilitates the developers to create a local web server for testing and development purposes. *XAMPP* is downloaded from the link: https://www.apachefriends.org/index.html and installed to make localhost environment for the development of the front end as well as to check the backend connection with the complex models. A free template was downloaded from https://www.templated.co/transit which will fit the purpose of building simple interface of this study. After installing *XAMPP* and placing the downloaded template contents in htdocs folder (default *XAMPP* folder to run localhost), *Apache Module* and *MySQL* module is started so that http://localhost can work on the user computer and any modification in the default template can be visualized without web publishing. In Figure 2.6, *XAMPP* control panel is shown where *Apache* and *MySQL* module is turned on.

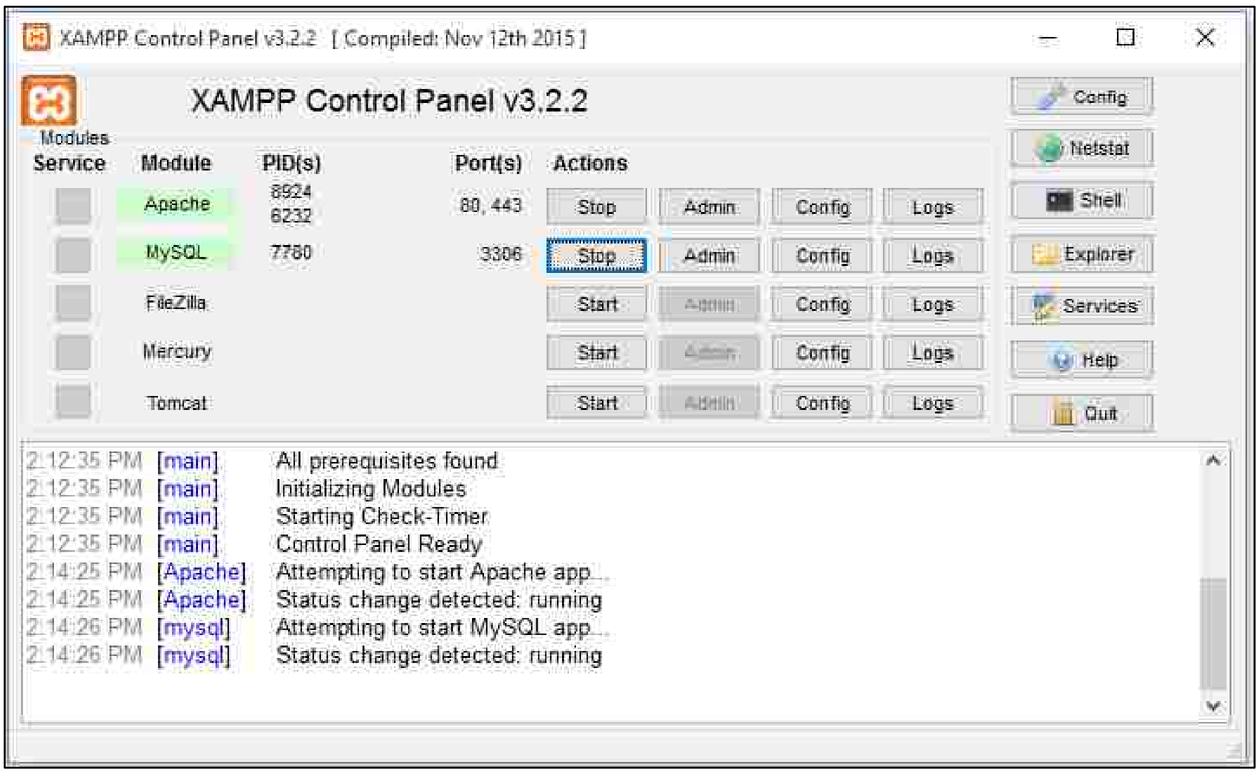


Figure 2.6. Cross-Platform, Apache, MariaDB, PHP and Perl (XAMPP) Localhost Environment Control Panel

2.3.2 *Front End Customization*

During frontend customization for the web portal (user interface), latest version of *HTML*, *HTML5* version is used so that the portal can be viewed in most of the browsers including smart device compatibility. After enabling *XAMPP* control panel, the situation makes it easier to modify/edit the html codes, *CSS* styling behavior to customize the front-end. Unnecessary styles, formats, background images and sections were completely removed to keep it simple so that the users can use this template to develop their own with minimum efforts. Home tab is designed as the index page/homepage of the portal which is shown in figure 2.7 and sample *HTML* code behind this page is shown in figure A.10 of appendix. Latest updates tab is (see figure A.1 in the appendix) added for providing the information on latest updates and major modification made in the portal. Visualization menu (see figure A.2 of appendix) is responsible to visualize the model results in different ways. Runoff, base flow, evaporation and soil moisture is visualized through raster gridded visualization (see figure A.3 of appendix) and streamflow in different prominent location is visualized in Timeseries visualization. Download page (figure A.4 of appendix) is developed to provide scope to the users to download specific parameter in specific temporal range. Publication tab is provided mainly to retrieve relevant publications of SASWE Research Group. Acknowledgements (figure A.5 of appendix), Contact Us (figure A.6 of appendix) and Terms of use (figure A.7 of appendix) is added to inform users about the portal and functionality and policies behind it.

Figure 2.7. Frontend Homepage of South Asian Surface Water Modelling System

2.3.3 *Backend Development*

Most of the backend connectivity purposes is solved by using *JavaScript* (https://www.javascript.com/) in the portal as it runs on visitor's computer and doesn't require any special programs or any downloads from any website. *JavaScript* is known as an object-oriented computer programming language commonly used to create interactive effects within web browsers. The primary focus of using *JavaScript* is to facilitate the connection between the results and the front end of the portal. In both of the raster gridded surface viewer and Date download page, initially date is taken previous local date of the user, *JavaScript* code is shown in figure A.11 of the appendix. In the raster gridded surface visualization, *JavaScript* is used to take arguments from the user selected Basin, Dataset, Temporal accumulation and selected date and then it constructs the path to the map and then returns the particular map. A screenshot of the JavaScript code is added in the figure A.12 of appendix.

Google Maps JavaScript Application Programming Interface - API (https://developers.google.com/maps/documentation/javascript/) is used to enable google map in the time series viewer page. The portal is registered in the google maps *API* to get the unique user key. For interactive visualization, basin boundary and river network is embedded on the google map as *KML* (*Keyhole Markup Language* by the Open Geospatial Consortium) layer and called through *JavaScript* which is shown in figure A.9 of appendix. Locations of the streamflow gauging are added as a form of coordinates on the map and hovering option is added so that the name of station can be viewed by dragging mouse over the location without clicking on the icon. Google map's *addListener* feature is used to visualize the streamflow of the station in an iframe (an example of *iframe* is shown in figure A.8 of appendix). Extended screen option is provided in the iframe for the convenience of the users. To visualize the timeseries chart in the iframe and extended screen, *Highcharts* (http://www.highcharts.com/) is added and linked with streamflow comma separated files. In the same way of raster visualization, *JavaScript* is used in the Download Page to capture user query and then retrieving text files according to it.

2.3.4 *Vertical Integration*

From the dataset download to the model results posting to the SASWMS portal, an integrated framework design is completed. Near real-time precipitation data is downloaded regularly from the IMERG precipitation server. Other meteorological forcing datasets (e.g. Maximum and Minimum Temperature, Average Wind Speed is downloaded from GFS-final analysis server - https://rda.ucar.edu/data/) which are in 6-hour interval and then they are converted into daily temporal accumulation.

After getting all the necessary datasets, precipitation correction system is triggered to correct IMERG estimated daily precipitation using the above described precipitation correction algorithm. After completing precipitation correction, corrected precipitation is feed into Variable Infiltration Capacity Model and the model is simulated for last 2 years to minimize the errors related to model spin up and uncertainties due to model initialization and instability.

Following completion of VIC Model simulation, Horizontal Routing Model (as stated in Hydrological Model section) is initiated to get the streamflow at predefined stations. Another task is to post process of VIC simulation results to extract spatial water fluxes. The simulation results from VIC Model is post processed and from the Route model simulated streamflow is updated in the SASWMS portal. Immediately after posting all these results into the portal, raster maps are also prepared to facilitate the users to visualize. Python integrated ArcGIS module named as *arcpy* (http://desktop.arcgis.com/en/arcmap/10.3/analyze/arcpy/a-quick-tour-of-arcpy.htm) is used to prepare the raster gridded surface of different parameters (e.g. baseflow, runoff, evaporation and soil moisture). Daily operational procedure of SASWMS is illustrated in figure 2.8.

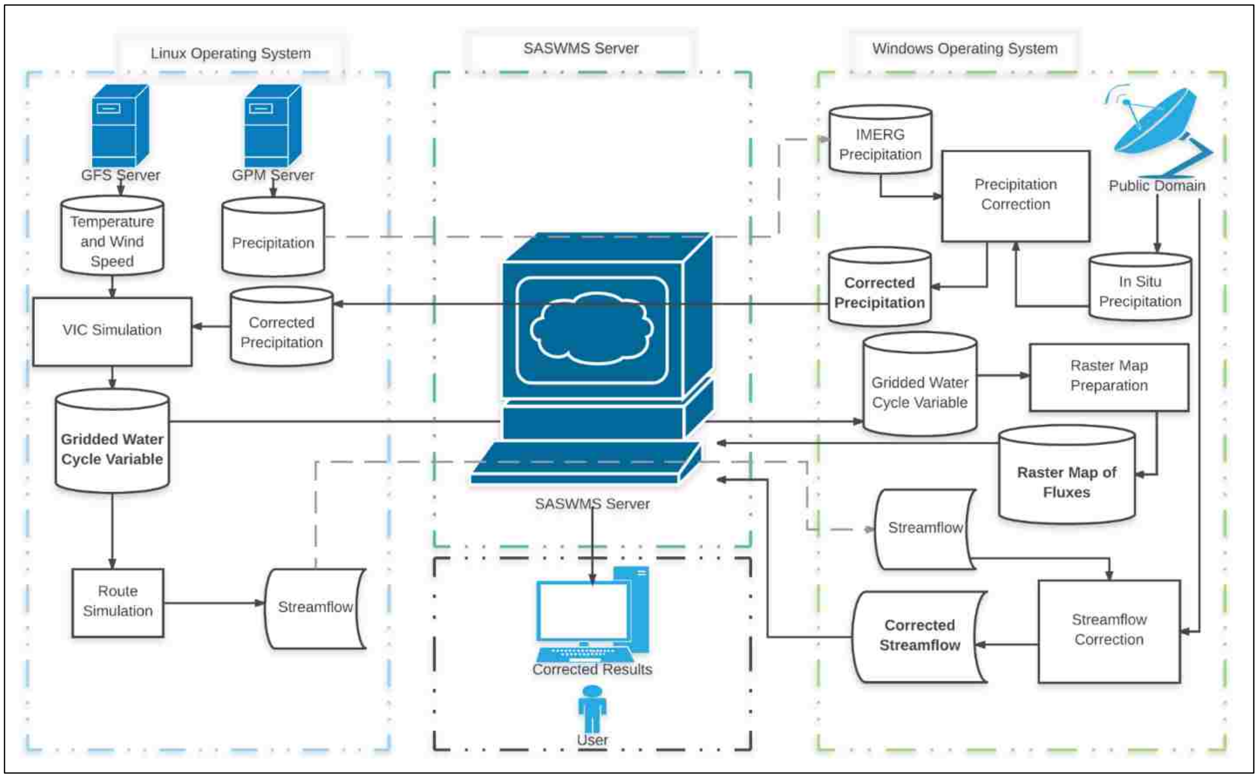


Figure 2.8: Regular operational procedure of SASWMS

To utilize most prominent modules of different operating system, multiple platforms are used in the SASWMS System development. Forcing data preprocessing and Model simulation is completed in the *Ubuntu Operating System (OS).* Whereas the precipitation and streamflow correction system is developed in the Windows OS. Also the raster maps preparation is done in the Windows OS. Multiple programming language (e.g. Visual Studio is for Web Crawler development, Python for mapping and raster processing) is used to automate the steps and sequential running. Data transfer between the server and processing computers using *WinSCP* (https://winscp.net/eng/index.php) *API* is used.

2.3.5 *System Monitoring and Garbage Collection*

To ensure real-time posting of model results to the server and for convenience of the users of the portal, system is monitored in each step carefully from the satellite data download to the web posting of all the results. During performing individual steps, a two way monitoring system is maintained, first option is an on screen error/warning console messaging system and second option is log maintain to collect all the success and failure notices. An example of the on-screen notification system is shown in figure 2.9. Different color combinations are used in this system for monitoring convenience, cyan colored notification means new steps are initialized, green color notification is for successful completion of the step and red color notification means failure of any step.

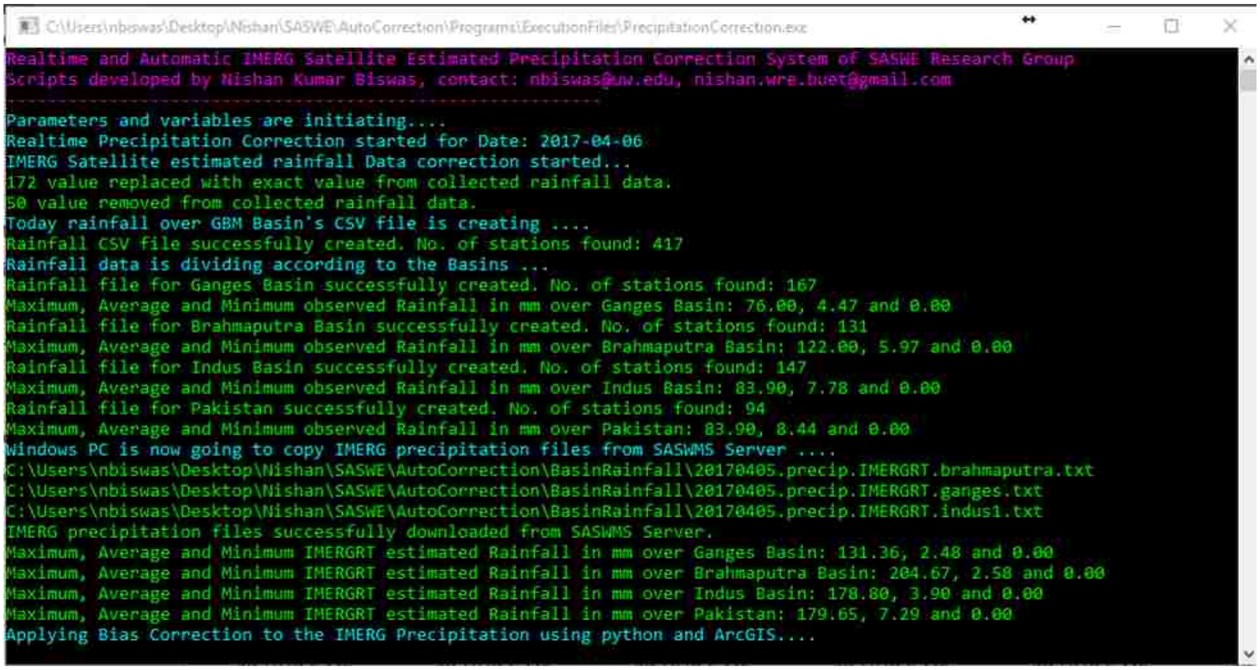


Figure 2.9: Example of on screen console notification system of SASWMS

In the log mention method, all the screen notifications along with the message of failure are saved in texts. Also in some steps, basic statistical figures (e.g. average observed precipitation, no. of stations used in the correction) are recorded for future uses. Log file contents of web crawler is shown in figure A.13 of appendix. In the download log file, number of stations found in each of the agency website is gathered and saved. In the precipitation correction log, number of in-situ crawled precipitation in each of the basin and average observed precipitation in each basin is saved. Average IMERG estimated precipitation and the amount of Bias which is applied in the IMERG precipitation calculated and recorded in this log. An example is shown in figure A.14 of appendix.

All the steps involved in uploading the model findings to the server (ASCII datasets, timeseries streamflow and rasters) are tracked. The list of files uploaded to the server are saved to another log file, contents of which is shown in figure A.15 of appendix.

2.4 Development of Web-analytic Correction System

For attaining the second objective of the study, a web analytic based correction system is developed. This correction method uses the free online resources to capture in in situ information from stakeholder agencies of South Asia. Using web crawled information, satellite estimated datasets are corrected and then posted in the South Asian Surface Water Modelling System portal. There are 3 components of the correction system, a web crawler to crawl the in situ water information in different agency websites, a precipitation correction system to correct the IMERG satellite estimated precipitation and a streamflow correction system to ensure the quality of estimated streamflow. The ultimate goal is that decision makers can utilize the untapped information posted on the internet and can get valuable information from the satellite observations in real-time basis. The components of the Web Analytic based correction system are discussed below.

2.4.1 *Web Crawler Development*

A real-time web based data crawler was developed using *Microsoft Visual C# (Sharp). Microsoft Visual C#* (pronounced C sharp) is a new programming language designed for building a wide range of enterprise applications that run on the *.NET Framework*. An evolution of Microsoft C and Microsoft C++, C# is simple, modern, type safe, and object oriented. C# code is compiled as managed code, which means it benefits from the services of the common language runtime. These services include language interoperability, garbage collection, enhanced security, and improved versioning support.

The developed crawler crawls a pre-defined list of websites at a particular time and extracts ground measured rainfall data from bona-fide government water management agencies. The agency websites included in the SASWMS portal are from Part of Tibet, India, Nepal, Bangladesh and Bhutan to cover the Ganges-Brahmaputra Basin. Location map of the stations included in the crawling system is shown in figure 2.10. A list of the websites and the index or location of the table of corresponding websites where rainfall information is posted was predefined in the web crawler. The captured precipitation data are posted in the portals in different ways varies from static *HTML* table (which updates regularly) to dynamic posting to the websites directly from the database.

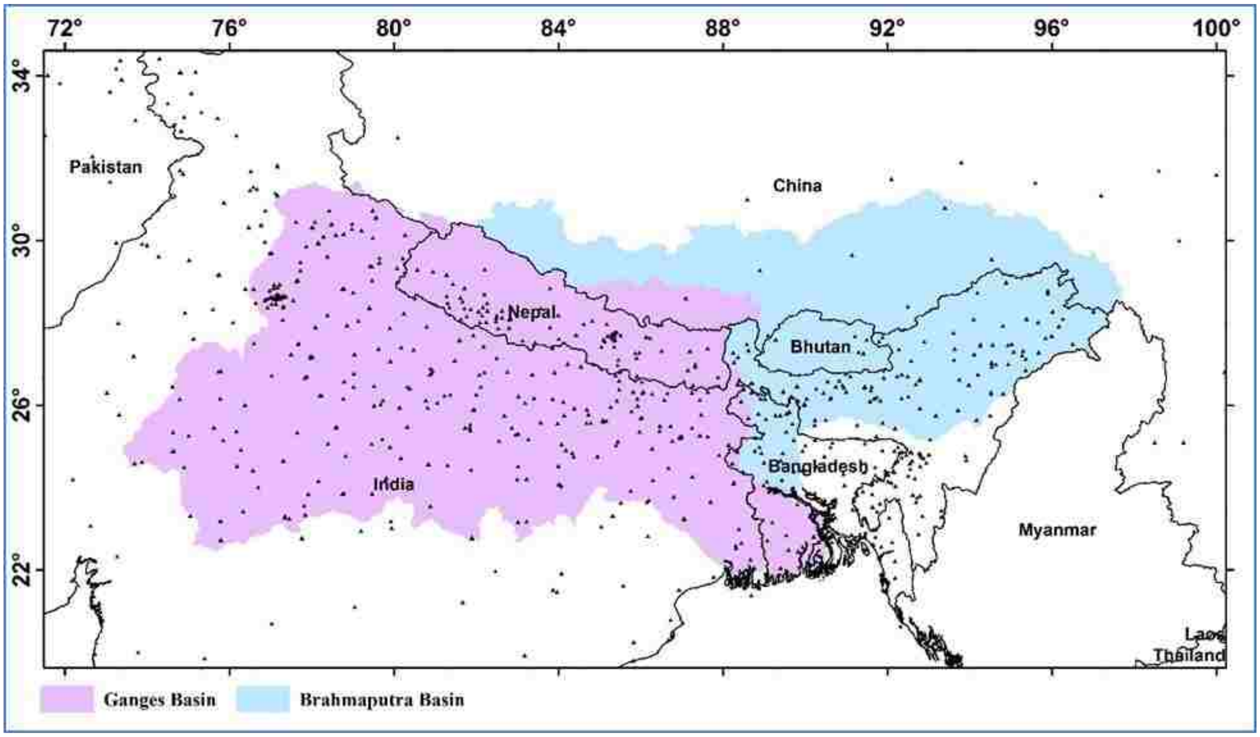


Figure 2.10: Stations included in the Web Crawler of SASWMS to crawl in situ rainfall information from public domains

2.4.1.1 HTML table crawling

In the major portion of the websites, rainfall data is posted using static html table where the location of the table, number of stations and format of the table is fixed. The crawler iterates through the websites where the table is static and grab rainfall data table according to the specified index number. To process the html table and datasets to get the amount of rainfall, a standard library of Microsoft visual studio named as *Html Agility Pack* (https://htmlagilitypack.codeplex.com/) is used. Using the library, the rainfall posting date is also captured from the websites. After extracting rainfall data from table and acquiring date of rainfall, it downloads datasets in text format where rainfall date, station name and rainfall amount (in mm) is saved. A sample code snippet of processing *HTML* table is provided in the figure A.16 of appendix.

2.4.1.2 PDF File Processing

Some of the websites of the public domains share rainfall information via portable document format (PDF). To read data in shared PDF, a *.NET PDF* library named *iTextSharp* (https://sourceforge.net/projects/itextsharp/) is used. The *iText* is a PDF library that allows users to read, create, modify, maintain PDF files programmatically. To read the agency provided PDF, the target PDF is downloaded and then the page index of rainfall information is provided. Using the library, all texts are extracted and reorganized. Finally, the obtained (i.e., 'crawled') stations, rainfall amount and date of rainfall is generated and saved. Sample chunk of code which is responsible to process online PDF is added in the figure A.17 of appendix.

2.4.1.3 HTTP Webpage Posting

For dynamic webpages (e.g. www.cwc.gov.in), *HTTP (Hyper Text Transfer Protocol)* post web request class is used which is a two-way communication system between the server and the local machine. To enable this class, a list of station code (a web based string ID) is copied from the target server. Then through the web client feature of C-Sharp, the initial response from the website is captured. The html code content from the response is modified and the station code is placed from collected station code and iterates for each station. After completing modification in the *HTML*, the regenerated html code is sent to the server to like a pseudo *HTML*. The server then returns the information of the station which comes with real-time rainfall value, date and time of rainfall and station name. This information is saved in the same way of the static *HTML* processing using *Html Agility Pack.* Sample code to process the request and obtaining rainfall is provided in figure A.18 of appendix.

2.4.1.4 Data Quality Assurance

After completing download of all the web information, a quality check is done by calculating the number of stations, maximum and average rainfall amount found in each of the basins. Rainfall with only the current date is taken into consideration. Any unwanted information is excluded from the downloaded rainfall. Representative characters are replaced with the actual amount of rainfall. Characters such as "999", "9999", "999.99", "N/A", "*", "**", "NA", "" is tagged as no information available."NIL", "NILL", "N" is replaced as zero rainfall and "TRACE", "Traces", "TR", "TR." and "T" is used as trace rainfall and represented as 0.01 mm. There are 913 stations currently included in the download program of SASWMS. Of these, on an average 650-800 station's data were found to be posted regularly by the agency websites as 'nowcast' for that day.

2.4.2 *Precipitation Correction System*

Precipitation correction system is developed as one of the primary data quality issue associated with Satellite precipitation data (such as IMERG) was related to excessive bias (Prakash et. al., 2016) that often renders the data unusable or results in physically unrealistic simulation of water cycle variables. There are different methods of precipitation bias correction which are suitable for real-time satellite estimated precipitation. Some of the most popular methods of bias correction of satellite estimated precipitation using in situ measurements are mean bias correction (Briedenbach et. al., 1999), use of a regression equation (Immerzeel, 2009; Cheema and Bastiaanssen, 2010), distribution transformation (Bouwer, 2004) and spatial bias method (Cheema and Bastiaanssen, 2010).

Mean bias correction is a straightforward correction methodology that determines the average bias for all stations on a specific day and this mean bias is subsequently used to correct the estimated rainfall amount. It is not suitable for large basins where topography varies and also stations density is not same in all over the basin. Since the spatial variation of rainfall in Ganges-Brahmaputra basin is large, this method will not be a good choice.

In the regression equation method, a regression equation is derived for each station using historical time series between the observed and satellite estimated precipitation. Then the coefficients (slope and constants in linear relationship) are spatially interpolated in the whole basin to estimate the correction factor of satellite estimated precipitation. A good correlation between observed precipitation and satellite estimated precipitation is necessary to apply this method. As there is lack of in-situ data in the study region, a long historical time-series is not available for most of the stations.

Distribution transformation method is specially developed for the statistical downscaling of climate model data (Bouwer, 2004). This method calculates statistical distribution of all station data on a particular day and of satellite estimated rainfall data at the same locations. The mean and standard deviation is used to correct estimated satellite precipitation. This method has been tested in case of Mekong Basin for flood forecasting purposes. Spatial pattern of satellite estimated precipitation does not change in this method, only change in the magnitude of rainfall is made. One of the major limitations in this method is short term convective rainfall cannot be represented or corrected by using this way of bias correction.

In this study, as the spatial distribution of stations included in the web crawler are not very dense and also the variation of bias is heterogeneous and also there is a lack of long term observed rainfall, spatial bias method was found most suitable method among all the methods for applying real-time bias correction.

In the spatial bias correction method, bias amount between observed and satellite estimated precipitation is calculated in all the observed station location of the basin in a particular day. The bias amount at each station is then used to form a bias raster by completing spatial interpolation of the bias amount using a suitable interpolation technique. In this study, Inverse Distance Weightage (IDW) and Spline interpolation techniques are used to assess their relative performance and suitability. To apply the bias in most accurate way, the parameters of chosen interpolation method is tuned. After completing bias interpolation, bias amount is subtracted from the IMERG estimated precipitation. All the precipitation amount less than zero is discarded after applying bias to the satellite estimated precipitation. In this way, the corrected precipitation datasets are prepared and it is uploaded to SASWMS server along with the raster map, also it is used by the hydrological model.

To apply spatial bias method for correcting the satellite rainfall estimates, the pixels of IMERG rainfall are detected where web crawled rainfall data is available. The bias amount in the available observed locations are calculated using the following formula

$$Bias\ in\ SRE = SRE_O - WCR \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \quad (1)$$

Where SRE_O is uncorrected satellite rainfall estimates and WCR is the web crawled rainfall. This bias amount is interpolated over the whole basin to get a bias raster surface for the particular date of correction. This Interpolated Bias Rainfall (IBR) is then used in the following equation to get the corrected satellite rainfall estimates, SRE_C.

$$SRE_C = SRE_O - IBR \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \quad (2)$$

Finally, a control check is completed to avoid negative amount of precipitation in corrected IMERG precipitation. During interpolation of the bias amount of rainfall over the basin, two different methods of interpolation techniques (Inverse Distance Weightage known as IDW method and Spline method) is used with the default parameterization. For the purpose of optimizing the quality of corrected dataset, the interpolation parameters of selected method of

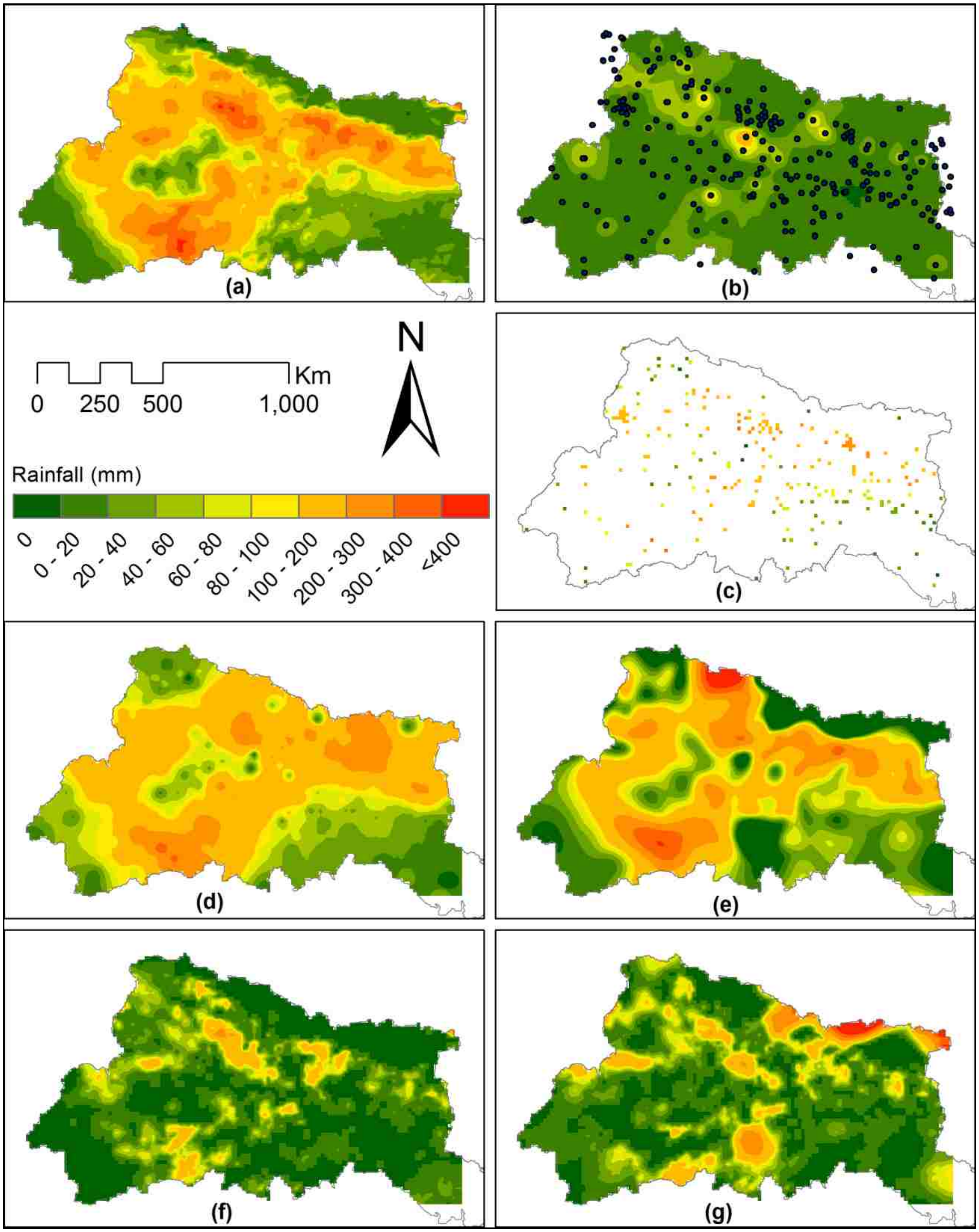


Figure 2.11. Example of Precipitation Correction Methodology for Ganges Basin (Date: 01 July 2016, the rainiest day in Ganges Basin). (a) IMERG-RT estimated precipitation, (b) In-situ

precipitation and web-crawled station locations, (c) Calculated Bias between IMERG and Web crawl precipitation, (d) IDW method of interpolation of Bias over Ganges Basin, (e) Spline interpolation of Bias over the Basin, (f) Corrected IMERG precipitation using IDW bias interpolation, (g) Corrected IMERG precipitation using spline bias interpolation

interpolation is optimized. As an example of precipitation bias correction methodology, IMERG data correction over Ganges Basin in 21th July of 2016 is shown in figure 2.11.

In part (a) of the figure, IMERG precipitation over the basin is shown. In (b), in–situ observed precipitation along with the location of the stations is shown. By subtracting IMERG and web-crawled precipitation, bias amount is calculated in available web-crawled locations which are shown in (c). The estimated bias amount is then interpolated using IDW interpolation technique and shown in (d) and spline interpolation technique is shown in (e). Finally, the bias amount found in (d) and (e) is subtracted from IMERG precipitation of (a) to get the corrected IMERG precipitation. Bias corrected precipitation using IDW and spline method of bias interpolation is shown in (f) and (g) accordingly.

2.4.3 *Streamflow Correction System*

In streamflow correction system, the simulated streamflow is corrected by using climatology of discharge (rated) and drainage area contribution and lag time of each correction station. Climatology discharge was derived using data pertaining from 1970-2015 for Bahadurabad station in Brahmaputra and 1910-2015 for Hardinge Bridge station of Ganges River. Streamflow correction methodology is illustrated in Figure 2.12.

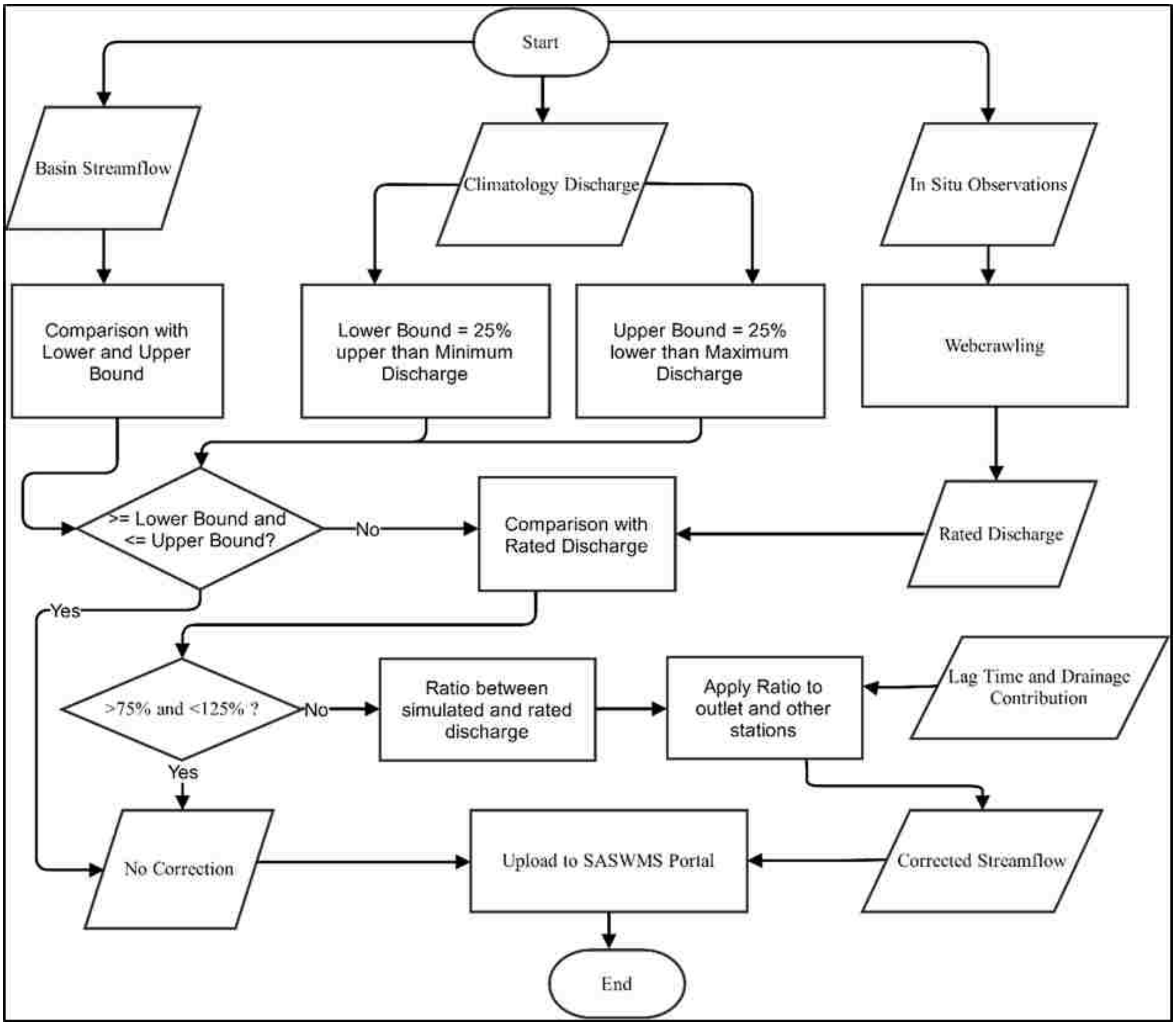


Figure 2.12. Streamflow correction Methodology

2.4.4 *Correction at Basin Outlet*

A "no-correction" envelope of streamflow was developed using climatology datasets for each Julian day. This range covers the region between 25% higher than the climatologically minimum discharge and 25% lower than the climatologically maximum discharge (climatology discharge of Brahmaputra and Ganges Basin are shown in figure 2.13 and 2.14). This range is considered as a safe and physically realistic zone where the system will not trigger any streamflow correction. The idea behind this is that satellite observations of precipitation typically overestimate or underestimate by an order of magnitude more under certain circumstances. It is those circumstances that yield physically unrealistic hydrologic simulation of streamflow that this correction system is designed to avoid and thereby retain the trust of the end users. It should be noted that this type of correction system will not take care of other kinds of corrections where streamflow is inaccurate but physically realistic.

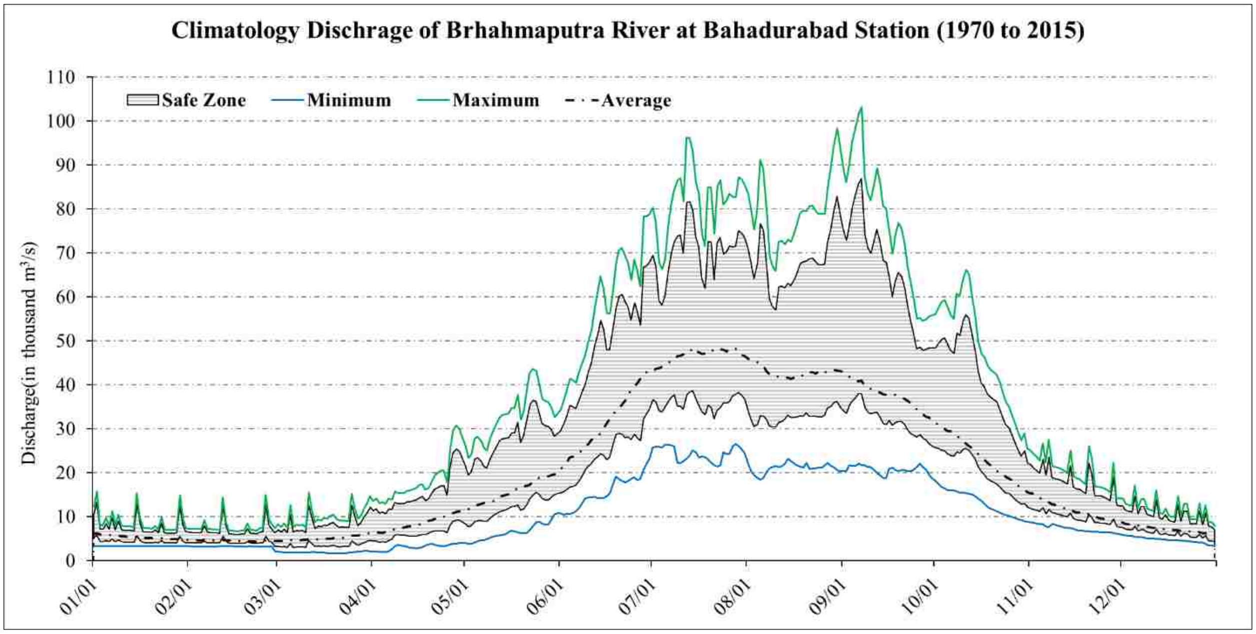


Figure 2.13. Climatology discharge of Brahmaputra Basin at Bahadurabad Station (Safe zone refers to no streamflow correction zone)

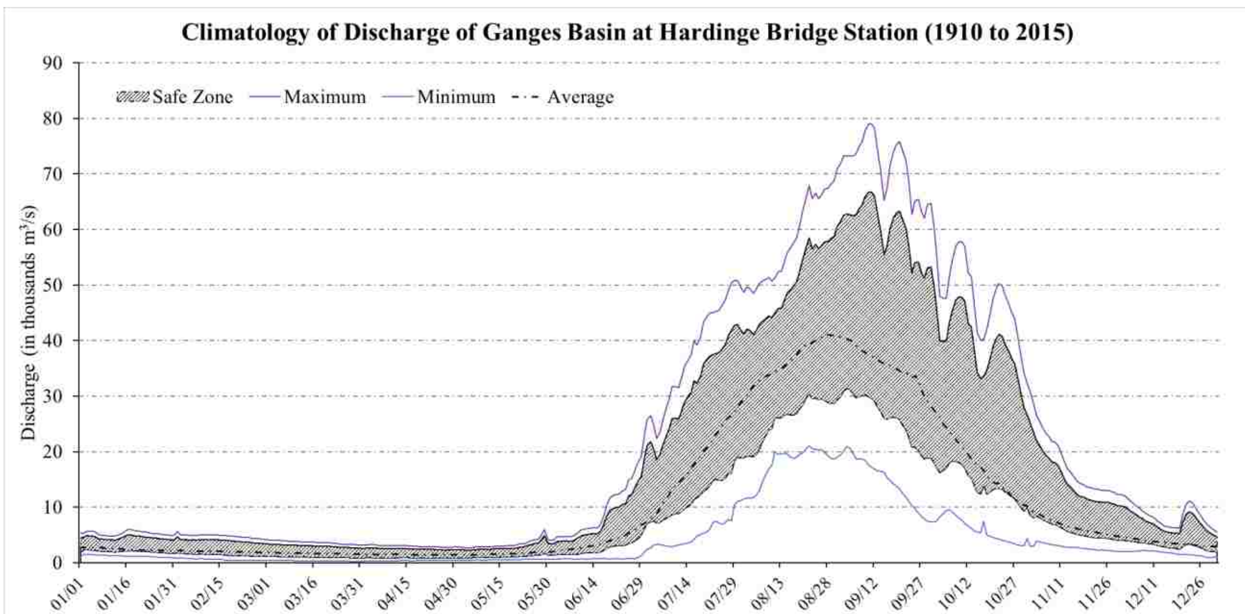


Figure 2.14. Climatology discharge of Ganges Basin at Hardinge Bridge Station (Safe zone refers to no streamflow correction zone)

However, when the simulated streamflow is outside this safe and physical no-correction zone, the system crawls the in-situ discharge of that day derived from observed water level records of basin outlet and compares the values. If the simulated streamflow is lower than 75% of public domain in-situ discharge or higher than 125% of the rated discharge, only then an automatic correction is triggered. During correcting the streamflow at the basin outlet, a ratio is derived between the rated discharge and the model simulated discharge. This ratio is finally applied to the model simulated discharge to obtain the 'corrected' streamflow of the basin.

$$Ratio, R = \frac{Q_0}{Q_s} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (4)$$

Where Q_s is the simulated streamflow and Q_0 is the rated discharge. The simulated discharge, Q_s is then multiplied by the ratio, R to get the corrected streamflow at the outlet.

$$Corrected\ Streamflow, Q_c = Q_s * R \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . (5)$$

This ratio is also used to correct the streamflow in the intermediate stations of the basin.

2.4.5 *Propagation of Correction in Intermediate Stations*

Streamflow correction applied to the outlet of the basin has been applied to some prominent locations of both basins. The main focus of this correction is that downstream decision makers can utilize the time of travel of water from the upstream location to the downstream. Streamflow correction locations in both basins are shown in figure 2.15. To propagate the correction, two important parameters are considered. These are relative drainage area of the station and lag time between the station and basin outlet.

The relative drainage area of a station is calculated by dividing the catchment area upstream of that station by total area of the basin.

$$Relative\ Drainage\ Area, R = \frac{Contributing\ drainage\ area\ of\ the\ station}{Total\ drainage\ area\ of\ the\ Basin} \dots \dots \dots \dots \dots (6)$$

Lag time of the station is calculated using velocity of water and the distance of the station from the outlet. According to Akhter et al. (2009), velocity of water in Ganges-Brahmaputra basin can be assumed as 1 ms^{-1} and 0.1 ms^{-1} for channel and overland flow respectively. By using this velocity, Lag time of station is calculated using following formula,

$$Lag\ time\ of\ station, n\ (in\ days) = \frac{Distance\ between\ station\ and\ outlet}{Flow\ velocity\ of\ water} \dots \dots \dots \dots \dots \dots (7)$$

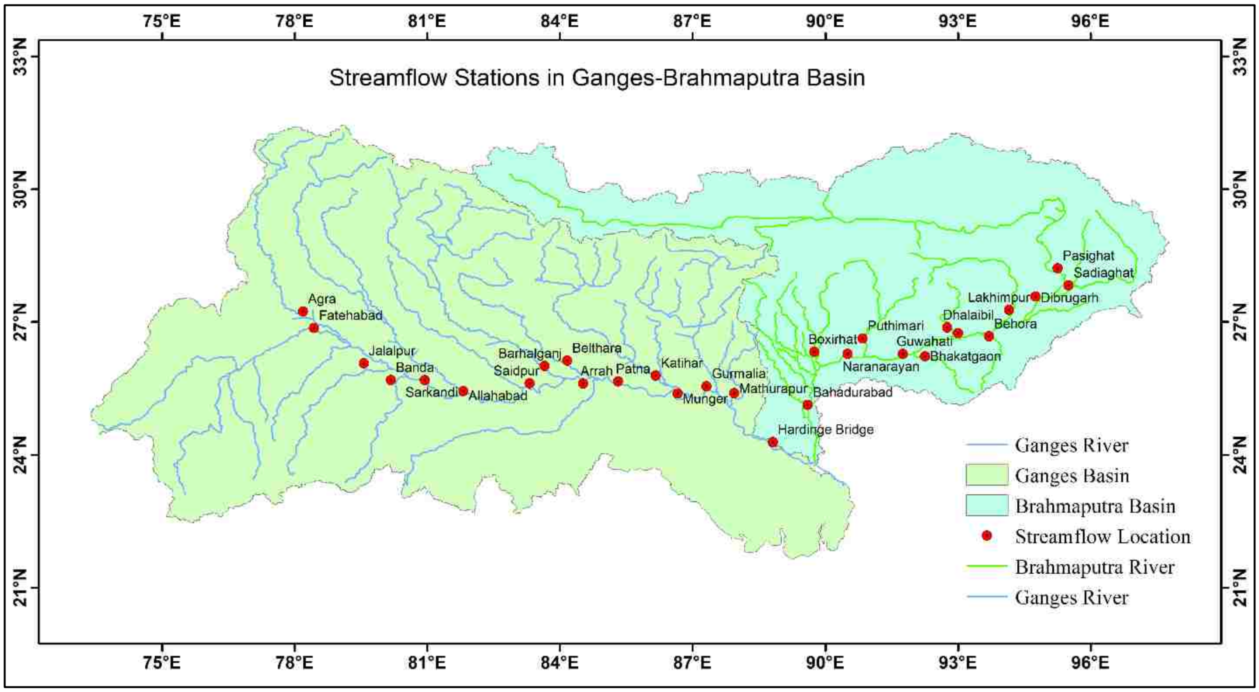


Figure 2.15. Streamflow correction locations in Ganges-Brahmaputra Basin

Using the equation 7, lag time map of both basins are calculated and shown in figure 2.16.

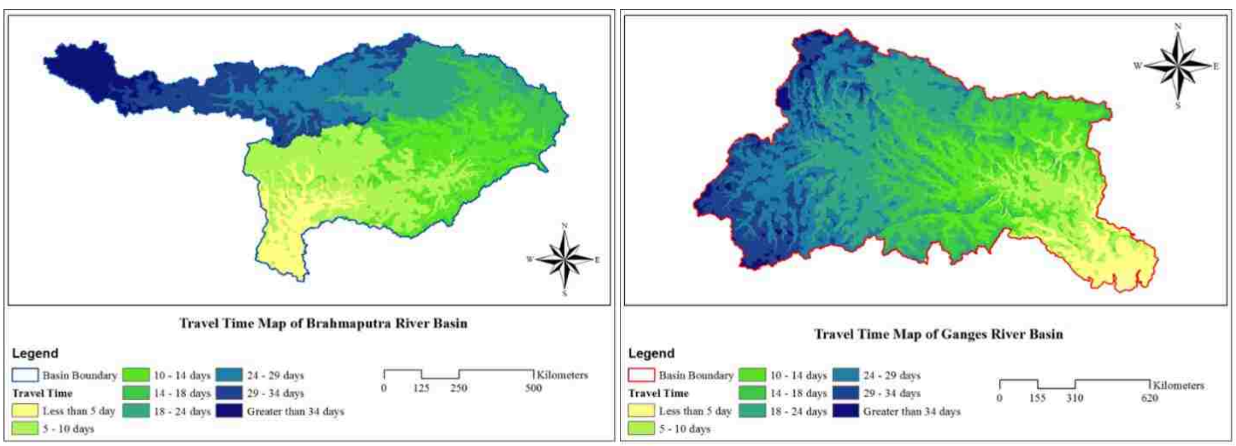


Figure 2.16. Lag time (in no. of days) map of Brahmaputra (left) and Ganges (right) Basin

From this map, lag time for each of the stations is selected. Lists of the streamflow stations of Brahmaputra and Ganges Basin along with their relative drainage area and lag time (in days) is shown in Table 2.1 and table 2.2.

Table 2.1. List of intermediate streamflow correction stations of Brahmaputra Basin

Station	River	Latitude	Longitude	Lag time (Days)	Relative Drainage Area
Bahadurabad	Brahmaputra	25.125	89.6	0	1
Naranarayan Bridge	Brahmaputra	26.275	90.5	1	0.876

Guwahati	Brahmaputra	26.275	91.75	2	0.793
Biswasnath	Brahmaputra	26.745	93	4	0.691
Dibrugarh	Brahmaputra	27.575	94.75	6	0.555
Pasighat	Brahmaputra	28.215	95.25	7	0.458
Sadiaghat	Lohit	27.825	95.5	7	0.094
Behora	Dhansiri	26.675	93.7	5	0.053
North Lakhimpur	Subansiri	27.275	94.15	5	0.072
Dhalaibil	Kameng	26.875	92.75	4	0.041
Puthimari	Manas	26.625	90.84	2	0.076
Boxirhat	Dharla	26.325	89.75	1	0.107
Bhakatgaon	Kopili	26.225	92.25	3	0.03

Table 2.2. List of intermediate streamflow correction stations of Ganges Basin

Station	River	Latitude (Degree)	Longitude (Degree)	Lag time *n* (Days)	Relative Drainage Area, *R*
Hardinge Bridge	Ganges	24.2875	88.812	0	1
Gurmalia	Ganges	25.5475	87.312	2	0.957
Patna	Ganges	25.6575	85.312	4	0.826
Saidpur	Ganges	25.6125	83.312	6	0.464
Sarkandi	Ganges	25.6875	80.937	9	0.348
Fatehabad	Chambal	26.8625	78.437	12	0.161
Agra	Yamuna	27.2375	78.187	12	0.074
Jalalpur	Betwa	26.0625	79.562	11	0.067
Banda	Ken	25.6875	80.172	10	0.036
Arrah	Son	25.6075	84.522	5	0.107
Mathurapur	Mahananda	25.3875	87.937	1	0.035
Katihar	Saptakoshi	25.7875	86.162	3	0.099
Munger	Ganlt	25.3825	86.662	2	0.025
Barhalganj	Ghaghara	26.0075	83.647	6	0.185
Belthara	Gandaki	26.1275	84.162	5	0.062
Allahabad	Ganges	25.4375	81.812	8	0.109

Using the derived relative drainage area, R and the lag time n, streamflow at any station is calculated using the following equation,

$$Corrected\ Streamflow\ at\ Intermediate\ Station, Q_{CI,T-n} = Q_{CO,T} * R \dots\dots\dots\dots\dots\dots\dots\dots \quad (8)$$

Where, T indicates the time step of streamflow correction at the outlet, $Q_{CO,T}$ is the corrected streamflow at the basin outlet (e.g. Bahadurabad Station of Brahmaputra Basin) at time step T and $Q_{CI,T-n}$ is the corrected streamflow at the intermediate station at T-n time step.

Chapter 3. RESULTS AND ANALYSIS

3.1 SASWMS Portal Development

In South Asian Surface Water Modelling System Portal, several hydrological parameters are shared regularly. Parameters are precipitation, reference evapotranspiration, streamflow, runoff, base flow, evaporation and soil moisture. Through the vertically integrated system, IMERG Satellite estimated precipitation is corrected regularly using the web based dynamic correction system. Corrected precipitation datasets are uploaded to the SASWMS server in raster map and ASCII formatted dataset. Using this precipitation dataset along with other variables as forcings, VIC and Route Model are simulated daily. Simulated water cycle fluxes are uploaded in the SASWMS server in map and data mode. Route simulated streamflow is corrected using dynamic streamflow correction system and then it also uploaded to the server. Users can visualize any datasets in spatial raster format maps and streamflow can be viewed in time series format from the visualization tab.

3.1.1 *Timeseries Visualization*

Simulated streamflow from the Route Model can be viewed in time series format. Any station's streamflow can be viewed for last one year and it is added in the streamflow time series visualization button of Visualization parent menu. In the Streamflow Visualization page, after selecting the desired basin, "View Time series" button need to be clicked. Immediately after clicking, selected basin with the stations located within it are viewed and laid on google map. When user drags mouse over the locations, station name appears. Single click on the station location produces time series of the station and comes in the interface. From there, zooming in option for viewing the time series plot and downloading facilities are provided. In the time series option, streamflow of any stations can be found for last one year.

Station's information (river, basin, latitude, longitude) are also provided in the streamflow visualization interface. Information can also be retrieved from http://depts.washington.edu/saswe/show_stns.html. In figure 3.1, Brahmaputra Basin along with the location of stations shown in the interface is shown.

Figure 3.1. Typical outlook of Streamflow timeseries visualization page where Brahmaputra Basin along with the location of available streamflow timeseries data is shown

3.1.2 *Raster Gridded Map Visualization*

Spatial water fluxes are added under Raster Gridded Surface Viewer. Here, available datasets and their temporal accumulation are depended on the user choice of the basin. When a particular basin is selected in the "Basin Name" option, available datasets in that particular basin are shown in the datasets. Available "Temporal accumulation" options are shown after selecting a particular

Dataset. In the bottom of the option box elements, a tracking element collects the user feedbacks and redirects to the next pending job to visualize a particular raster map. After completing minimum required queries, the tracker shows metadata of the raster map (Basin, Dataset, Temporal Accumulation and Date) and the raster map box shows the available image raster if the image is available in the system. A typical user query, tracker information along with the raster maps is shown in figure 3.2. Date and temporal accumulation can be changed to detect the change in the selected parameter. One mentionable point in here is that date and temporal accumulation preserves user queries, whereas changing the basin or dataset will remove all previous queries as the available datasets and their accumulation in the basins are not same.

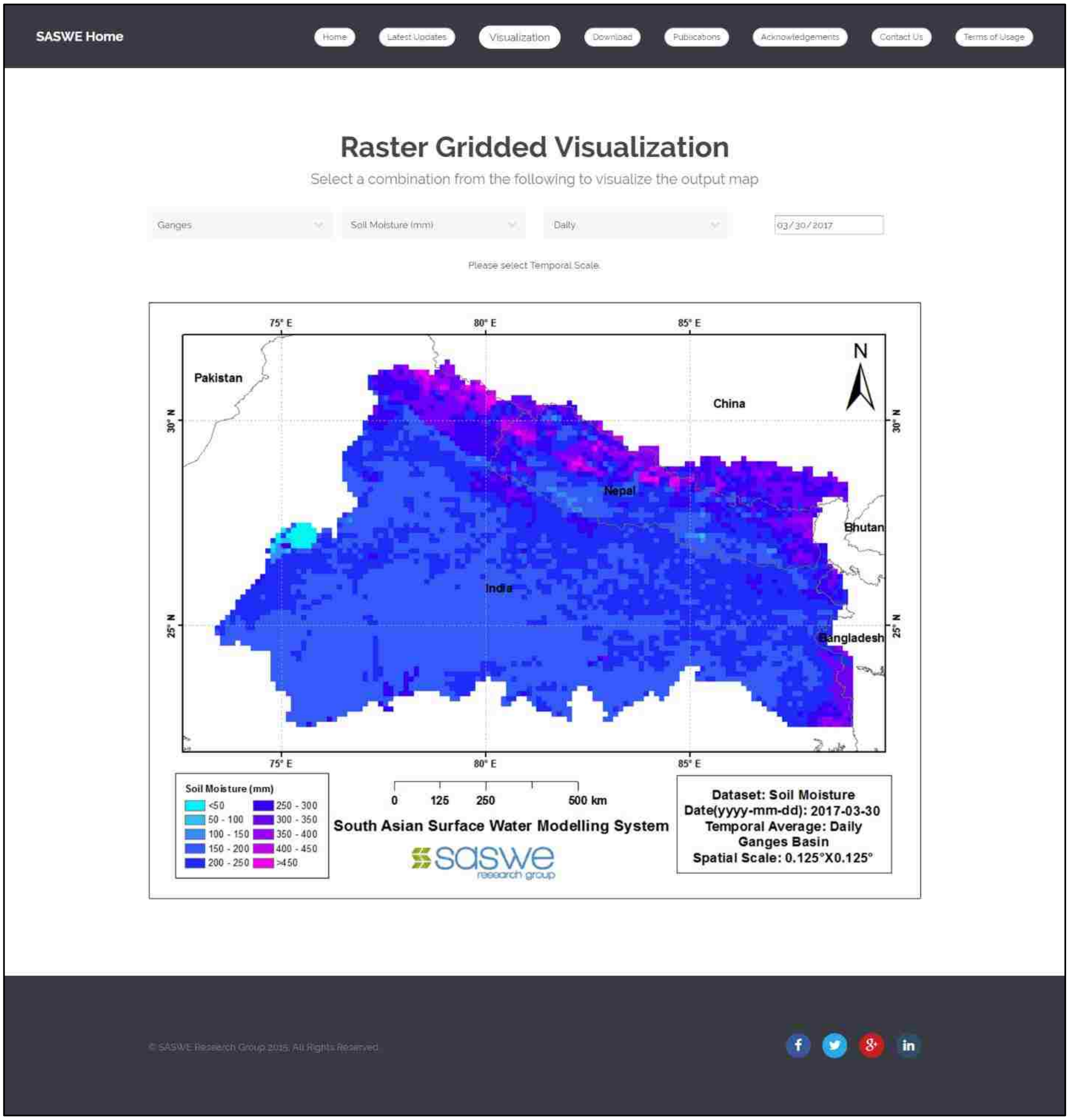


Figure 3.2. Outlook of Raster Map visualization page where daily soil moisture of Ganges Basin of 30th March 2017 is shown.

3.1.3 *Data Download*

All datasets of each parameter are available for downloading in the download parent menu of the server. During the initialization of this page, an option row consists of basin name, dataset, temporal accumulation, year, month and download button. Like the raster visualization page, available dataset menu is depended on the basin selected and all other option can be selected simultaneously. Here, there is an option check follower which collects user selection queries and shows below the option row as same as in the raster visualization. It also redirects the user for next available option that needs to be selected to complete the minimum required queries for downloading a particular dataset.

The option check follower is embedded in the download button so that when the download button is clicked, first the options are checked and then, if there is any incorrect or missed selection, an error/warning message is popped from the server and no data is retrieved. After each successful query following by clicking on the download button, available number of ESRI (Environmental System Research Institute) ASCII formatted raster dataset is downloaded to the predefined download location of the user PC.

File naming convention of the ASCII files are very important for users. The different parts are divided by using a single dot (.). First part is the date in ISO date/time strings (e.g. 20170303), second part is the basin name (e.g. brahmaputra), third part is the parameter name (e.g. runoff). Weekly and monthly files naming conventions end differently from the daily files. For weekly files, “weekly” is added before the file extension. For monthly file, first part is only year and month (e.g. 201608). All files are maintained as text file extension (.txt). A typical ASCII formatted data file along with the header is shown in figure

To make the download process more convenient for the users, all of the files can be retrieved through FTP (file transfer protocol) or SSH (Secure Shell) terminal. In that users need prior permission from the system which can be obtained by sending request to saswe@uw.edu.

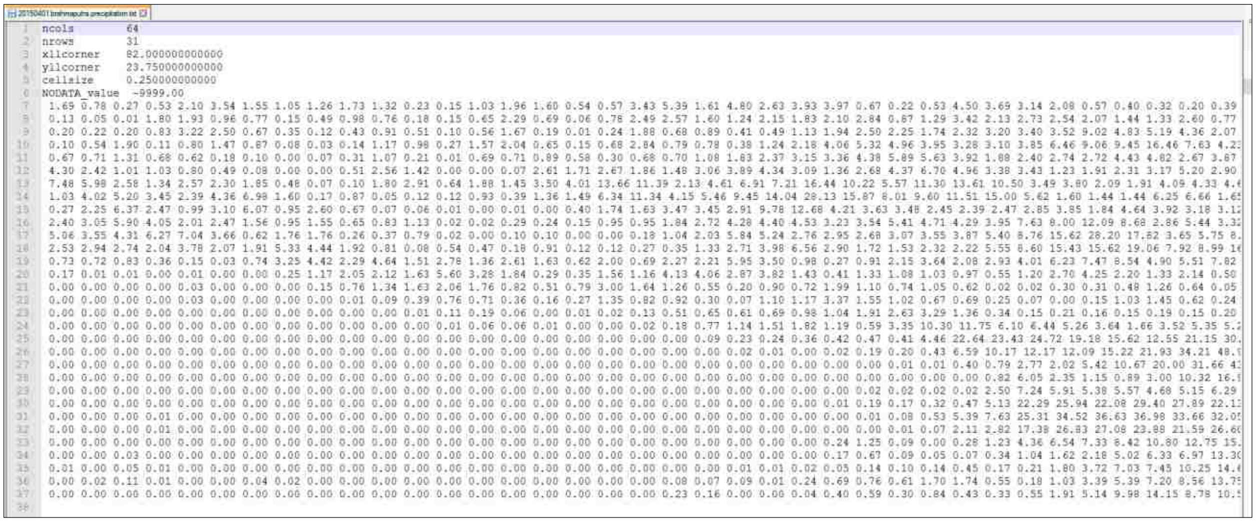

```
ncols        64
nrows        31
xllcorner    82.000000000000
yllcorner    23.750000000000
cellsize     0.250000000000
NODATA_value -9999.00
```

Figure 3.3. Contents along with the header rows of a typical ESRI Formatted file

3.2 Performance of Web Analytics Based Correction System

Web analytics based correction system is applied in Ganges and Brahmaputra River Basin. Here, performance of precipitation correction is compared with the reference in situ precipitation of NCDC-GSOD rainfall. Change in precipitation and simulated streamflow using this method of correction system is studied. To obtain the best estimation of streamflow, performance of only precipitation correction, only streamflow correction and combined correction of precipitation and streamflow is assessed for Ganges and Brahmaputra Basin. In case of precipitation correction, data was used from 1st January 2016 to 31st December 2016. As there are limitations in sharing in situ water level data in the dry season, water level data from 27th March and up to 30th August is used in here for assessing performance of streamflow correction.

3.2.1 *Performance of Precipitation Correction System*

IMERG satellite estimated precipitation without any correction and Bias corrected (IDW and Spline bias interpolation) precipitation compared with NCDC-GSOD in-situ observed precipitation. During IDW method of interpolation of bias, exponent of distance is taken as 2 and 12 number (variable) of nearest input sample points is used to perform interpolation. For spline interpolation, regularized spline is used where weight of the third derivatives of the surface in the curvature minimization expression was 0.1 and 12 number of points per region is used for local approximation.

3.2.1.1 Brahmaputra Basin

Firstly, the error in satellite precipitation estimates is studied in daily accumulation scale. In figure 3.4, a comparison of daily average precipitation is shown. In the figure, high amount of overestimation in daily average precipitation over Brahmaputra Basin can be observed. It can also be illustrated from the figure that IDW bias interpolated precipitation pattern is closer to the NCDC-GSOD and web crawled precipitation pattern. From the figure, another conclusion that can be made is the abrupt behavior of Spline bias interpolation. The overestimation in IMERG can be further explained by the scatter plot of different daily average precipitation methods with NCDC-GSOD precipitation over Brahmaputra Basin which is shown in figure 3.5. From there, we can see that IDW and Web crawl precipitation represents better compared to the IMERG and spline method.

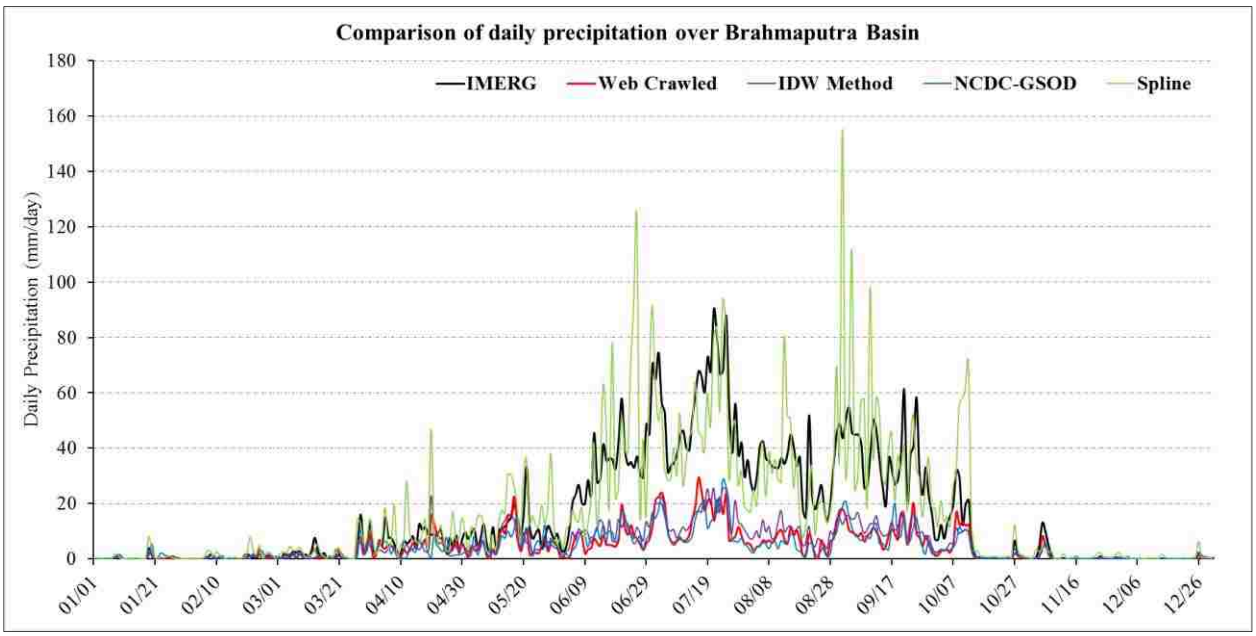


Figure 3.4. Comparison of daily average precipitation over Brahmaputra Basin for the year of 2016

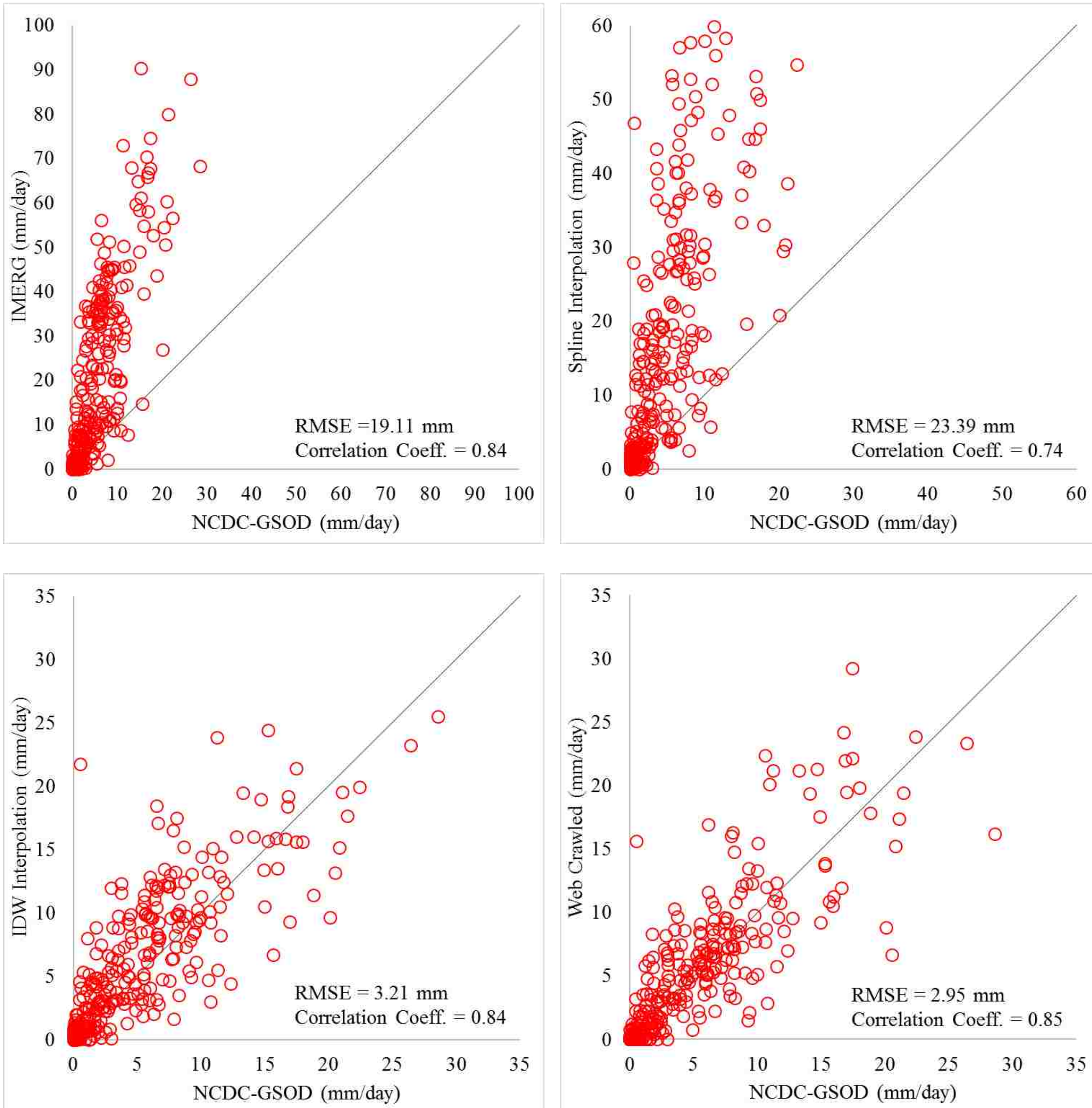


Figure 3.5. Scatter plot of daily average precipitation over Brahmaputra Basin for the year of 2016, NCDC-GSOD precipitation and (upper left) IMERG RT precipitation, (upper right) Spline bias corrected precipitation, (lower left) IDW bias corrected precipitation and (lower right) Web Crawled precipitation

Daily average precipitation amount found from different method is compared with NCDC-GSOD precipitation numerically and shown in Table 3.1. Root Mean Square (RMSE), Correlation coefficient and Mean Absolute Error (MAE) is calculated for each of the precipitation by using NCDC-GSOD source. From the error metrics shown, we can see that Root Mean Square Error

(RMSE) of IMERG-RT is reduced from 19.11mm to 3.21 mm when IDW bias interpolation method is used to correct IMERG product. Mean absolute Error (MAE) also decreased significantly due to use of this dynamic correction method.

Table 3.1. Error comparison metrics for precipitation correction in Brahmaputra Basin

Precipitation Product	Average Precipitation (mm)	RMSE (mm)	MAE (mm)	Correlation Coefficient
IMERG Precipitation	15.42	19.11	11.48	0.84
Web Crawl Precipitation	4.45	2.95	1.66	0.85
IDW Bias corrected precipitation	5.00	3.21	1.97	0.84
Spline bias corrected precipitation	16.93	23.39	12.95	0.74

Monthly comparison of these precipitation methods are shown in figure 3.6. It can be illustrated from here that overestimation by IMERG product is higher during the monsoon season than during season. That means IMERG precipitation behaves poorly during high rainfall events.

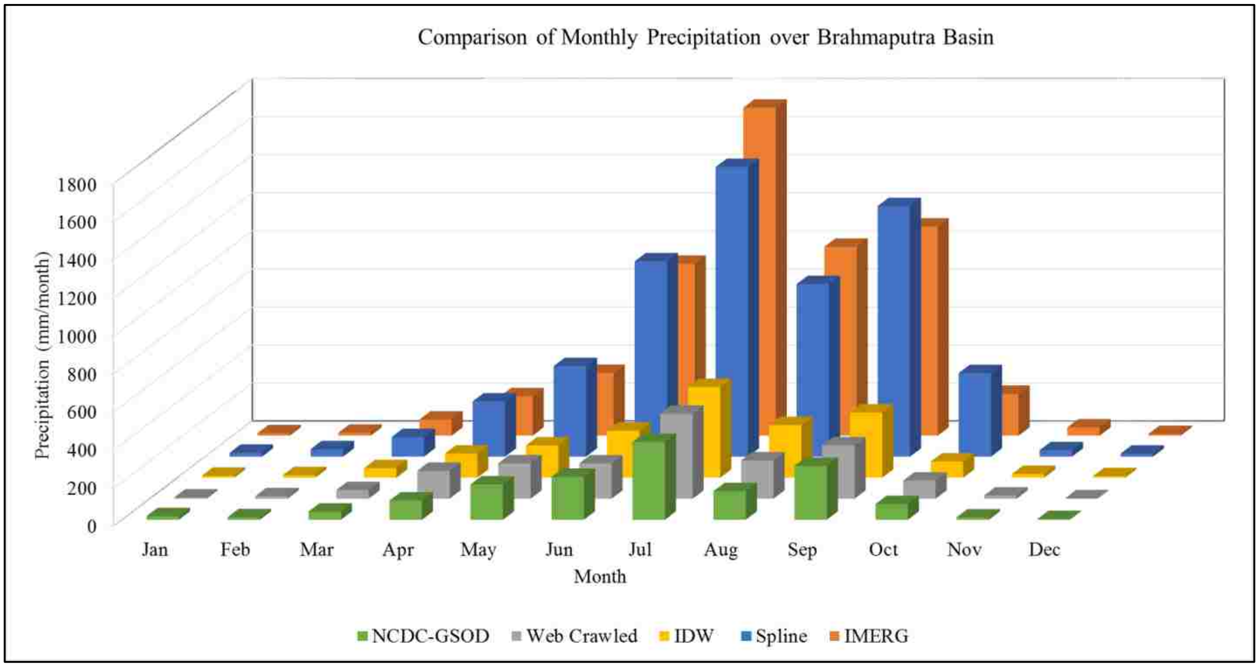


Figure 3.6. Comparison of monthly average precipitation over Brahmaputra Basin for 2016

Spatial variation of the bias in the basin is studied by diving the whole year into four seasons. The seasons are: dry season (January-March), pre-monsoon (April-June), monsoon (July-September) and post-monsoon (October-December). During each season of the year, IMERG, Web crawled, IDW interpolated and spline interpolated precipitation is compared spatially with

NCDC-GSOD estimated precipitation. Bias in each of the product is calculated by comparing with NCDC-GSOD precipitation. In figure 3.7, monsoon precipitation amount (from July to September) is shown for NCDC-GSOD, IMERG, IDW, Spline and Web crawled sources. Comparison can be made in amount reduction of bias from uncorrected IMERG to the Spline, IDW and Web-crawled precipitation. In most of the cells, IMERG estimated precipitation is more than 5 times larger than the NCDC-GSOD estimated precipitation. Even in a considerable number of cells, bias is more than 1000%. In middle part of the basin, negative bias can also be seen, which is comparatively small. In appendix A.19, appendix A.20 and appendix A.21, other season maps are shown.

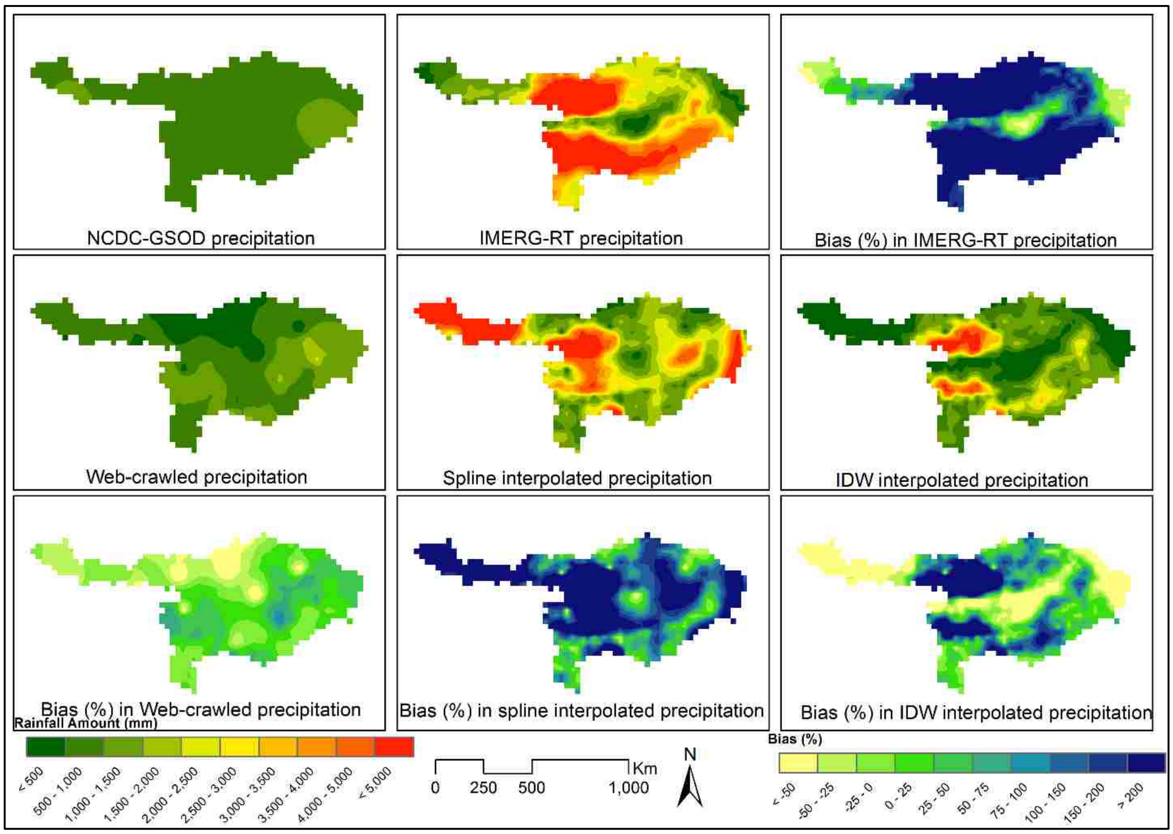


Figure 3.7. Comparison of precipitation from IMERG, Web crawled, IDW bias interpolated and spline bias interpolated precipitation along with bias amount (in percentage with respect to NCDC-GSOD precipitation) in every product over Brahmaputra Basin for monsoon season.

The impact of the precipitation correction system on streamflow prediction can be seen by simulating VIC Hydrological Model using IMERG-RT precipitation, IDW bias corrected precipitation, spline bias corrected precipitation and NCDC-GSOD precipitation data. The result of VIC Model simulated streamflow is shown in figure 3.8. From the figure, uncorrected IMERG-

RT simulated streamflow is shown as highly unrealistic. Although spline correction is used for correcting precipitation, it decreases peak flow to a certain extent from the uncorrected precipitation, but still it is also showing an unrealistic simulation of streamflow (as there was also some unrealistic correction in IMERG-RT precipitation). IDW method of bias correction shows more significant improvement in the streamflow, it decreases the highly unphysical flows and also captured the pattern of in situ rated discharge. Although both methods of correction system improved the quality of simulated streamflow from the IMERG-RT precipitation, overall performance of IDW method is found to be superior to the spline method. The modest but systematic overestimation in streamflow prediction that remains can be taken care of through agency-based adjustment factors.

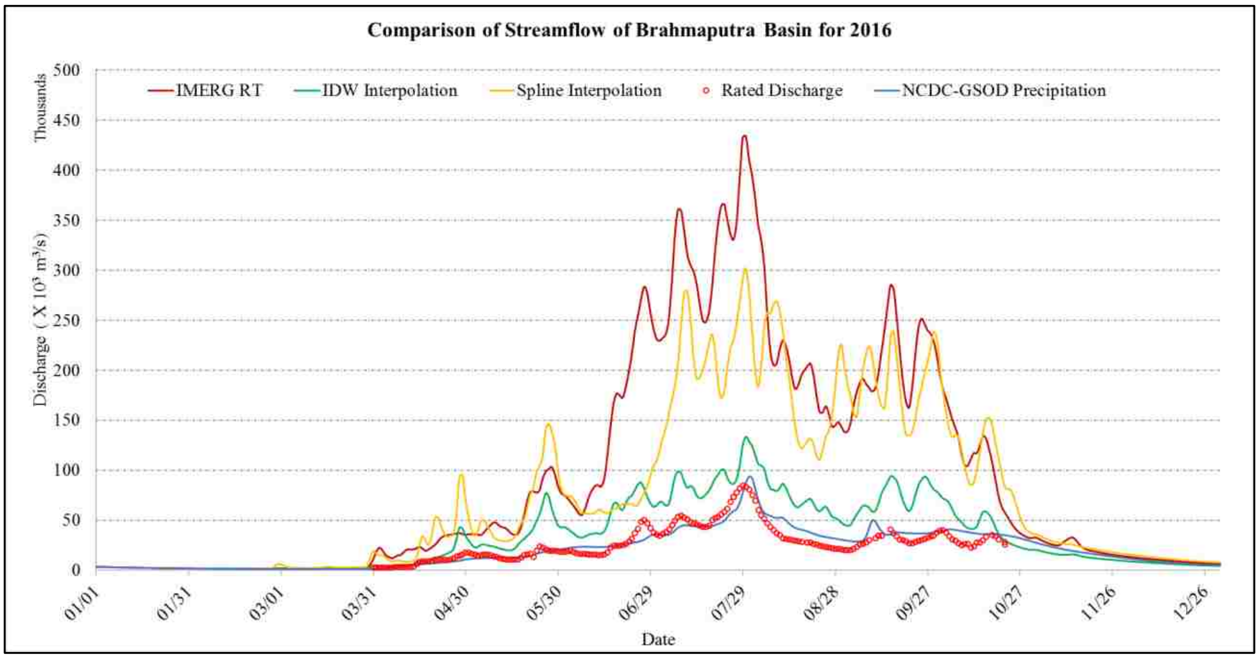


Figure 3.8. Comparison of simulated streamflow of Brahmaputra Basin using IMERG-RT precipitation, Bias corrected precipitation and NCDC-GSOD precipitation with rated discharge at the Bahadurabad station

3.2.1.2 Ganges Basin

For comparison of performance, daily average precipitation found from different methods is compared like that done in Brahmaputra Basin, which is shown in Figure 3.9. Here also, IMERG-RT highly overestimated the NCDC-GSOD precipitation and Spline performs poor compared to IDW method of bias interpolation.

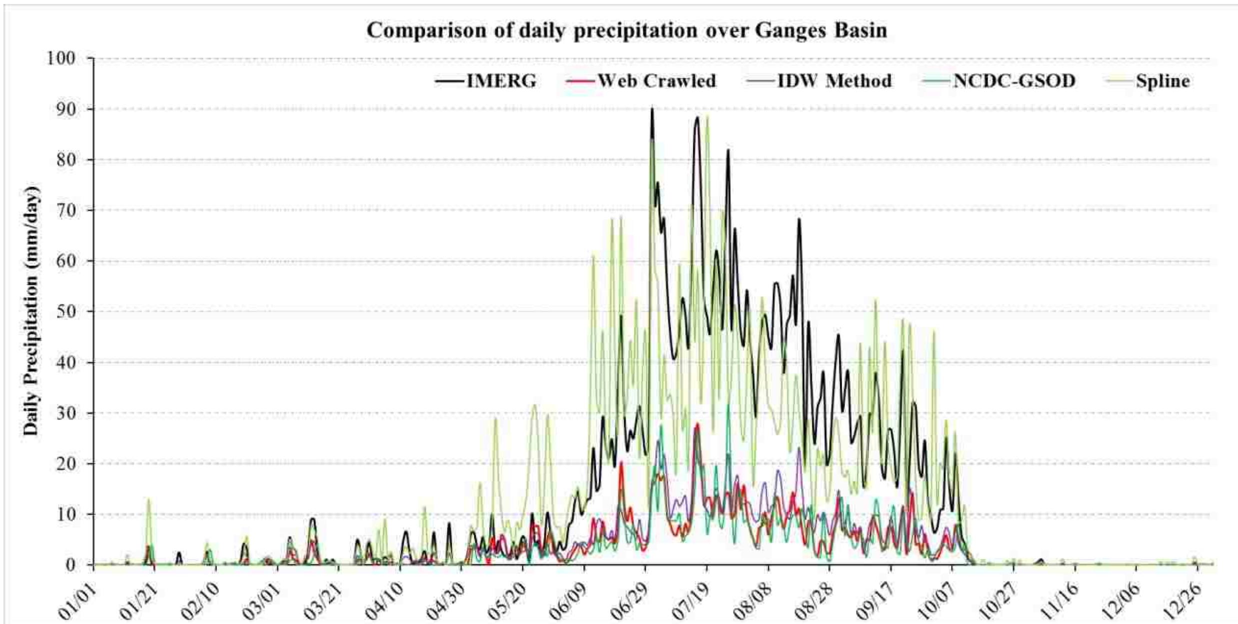
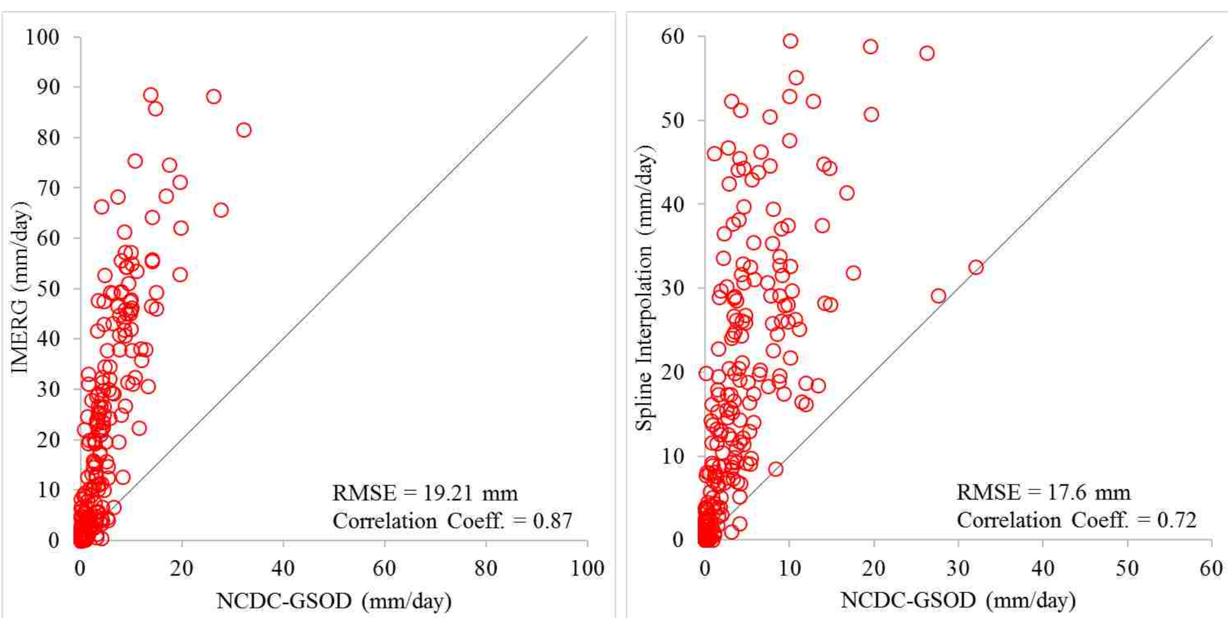


Figure 3.9. Comparison of daily average precipitation over Ganges Basin for the year of 2016

Scatter plot of all the products with NCDC-GSOD precipitation is shown in figure 3.10 to further illustrate the scenario of overestimation of IMERG-RT precipitation. Performance of the precipitation comparison is shown in table 3.2. Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and correlation coefficient is calculated to compare the performance with respect to NCDC-GSOD precipitation.



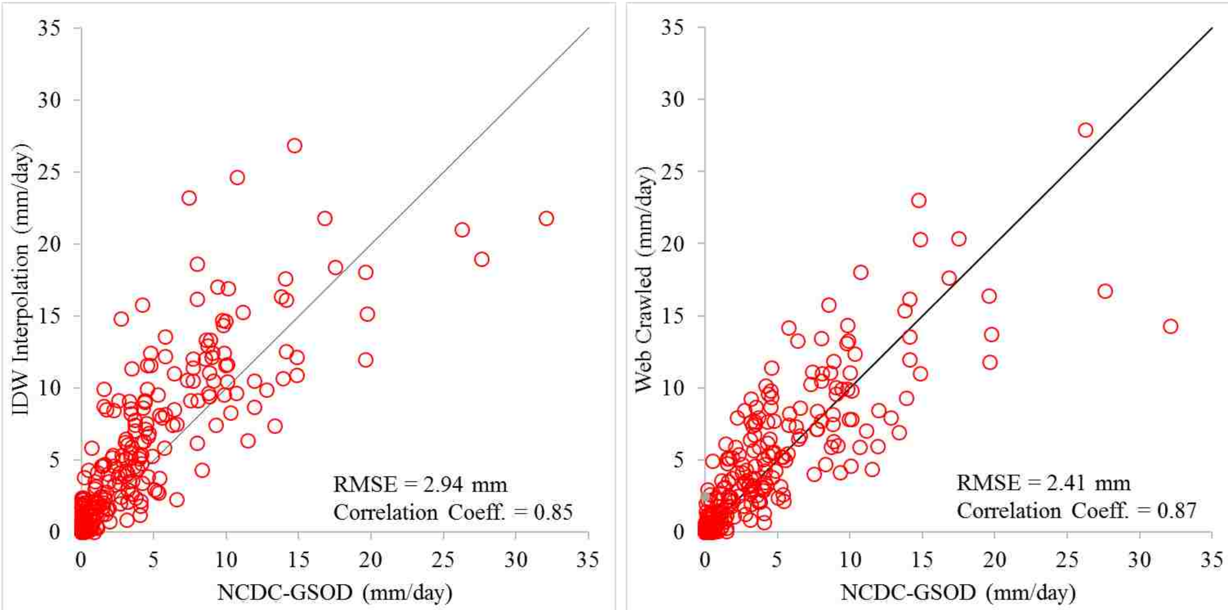


Figure 3.10. Scatter plot of daily average precipitation over Ganges Basin for the year of 2016, NCDC-GSOD precipitation and (a) IMERG RT precipitation, (b) Spline bias corrected precipitation, (c) IDW bias corrected precipitation and (d) Web Crawled precipitation

Table 3.2. Error comparison metrics for precipitation correction in Ganges Basin

Precipitation Product	RMSE (mm)	Mean Average Error (mm)	Average Precipitation (mm)	Correlation Coefficient
IMERG Precipitation	19.21	10.54	13.36	0.87
Web Crawl Precipitation	2.41	1.28	3.23	0.87
IDW Bias corrected precipitation	2.94	1.67	3.96	0.85
Spline bias corrected precipitation	17.60	9.89	12.76	0.72

Like monthly comparison of precipitation in Brahmaputra Basin, in Figure 3.11, monthly precipitation is calculated and compared among the products for Ganges Basin. IMERG overestimation amount increases with the increase in amount of monthly rainfall. From the following figure, in July, August and September, IMERG estimated precipitation is abruptly high compared to the other months of the year.

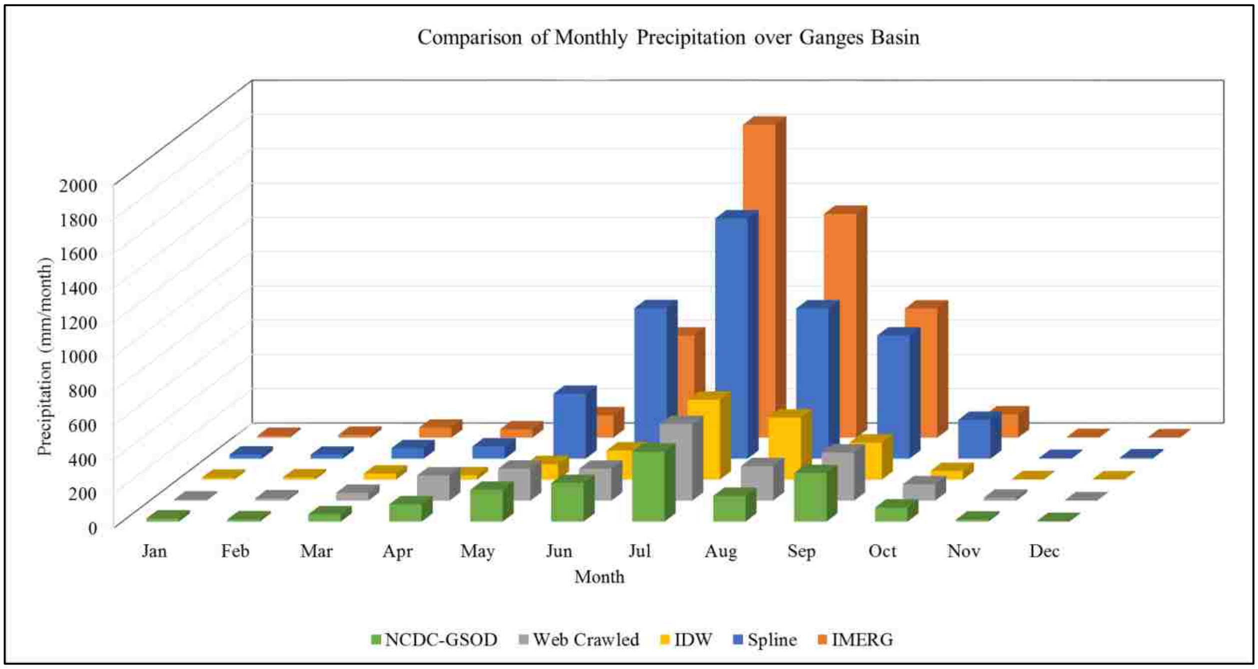


Figure 3.11. Comparison of monthly average precipitation over Ganges Basin for 2016

Seasonal precipitation comparison along with bias estimation as well as spatial distribution of this bias is shown in figure 3.12. The following figure is for monsoon season only; where a high amount of positive bias is revealed in almost all over the basin as like as Brahmaputra Basin. Here positive bias goes up to more than 1500% compared with NCDC-GSOD precipitation. Although there is some negative bias in IMERG precipitation, it is of a very low magnitude compared to the positive bias. Other season's comparison map is shown in appendix A.22-appendix A.25.

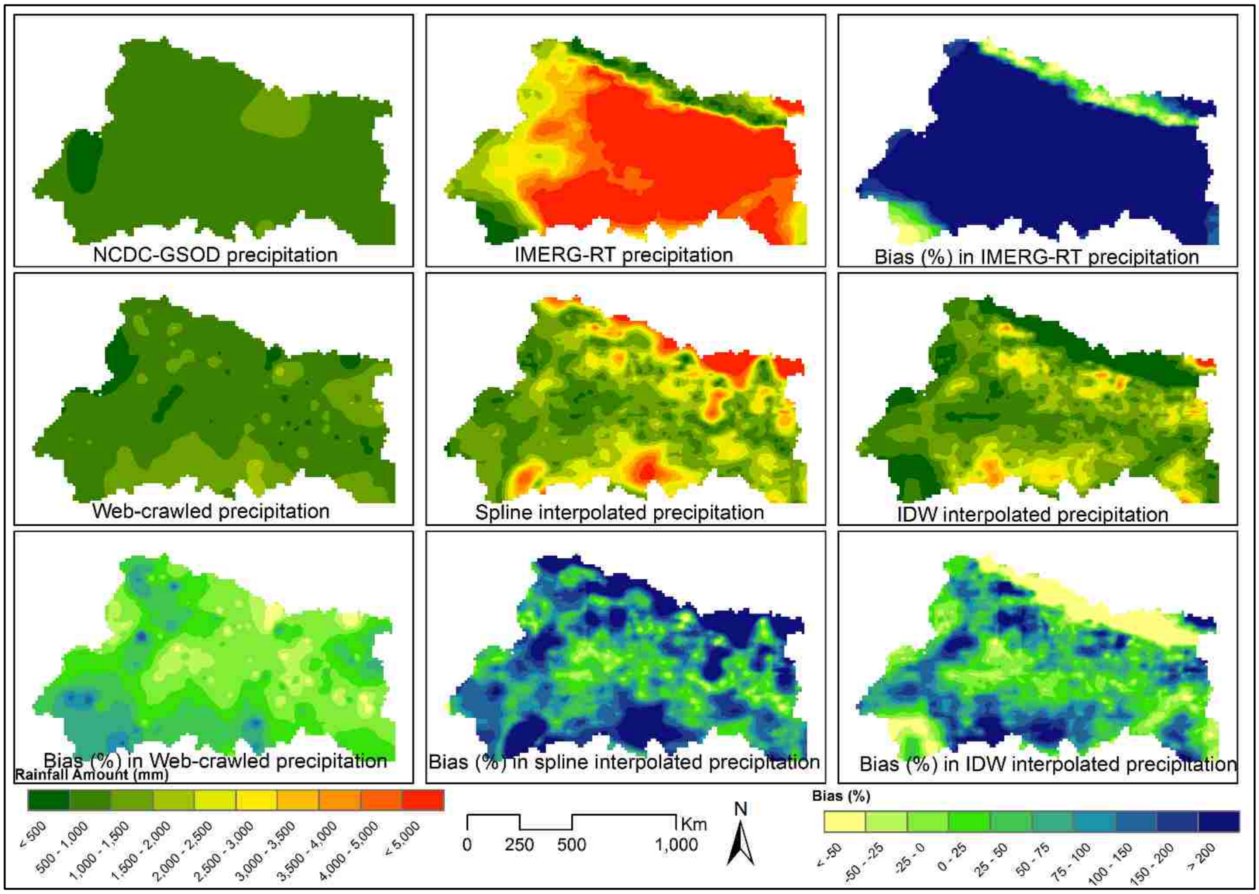


Figure 3.12. Comparison of precipitation from IMERG, Web crawled, IDW bias interpolated and spline bias interpolated precipitation along with bias amount (in percentage with respect to NCDC-GSOD precipitation) in every product over Ganges Basin for monsoon season.

After comparing precipitation products and their bias in daily, monthly and seasonal scale, impact of precipitation correction in simulated streamflow is also studied. Comparison of streamflow simulated by using every precipitation products is shown in figure 3.13.

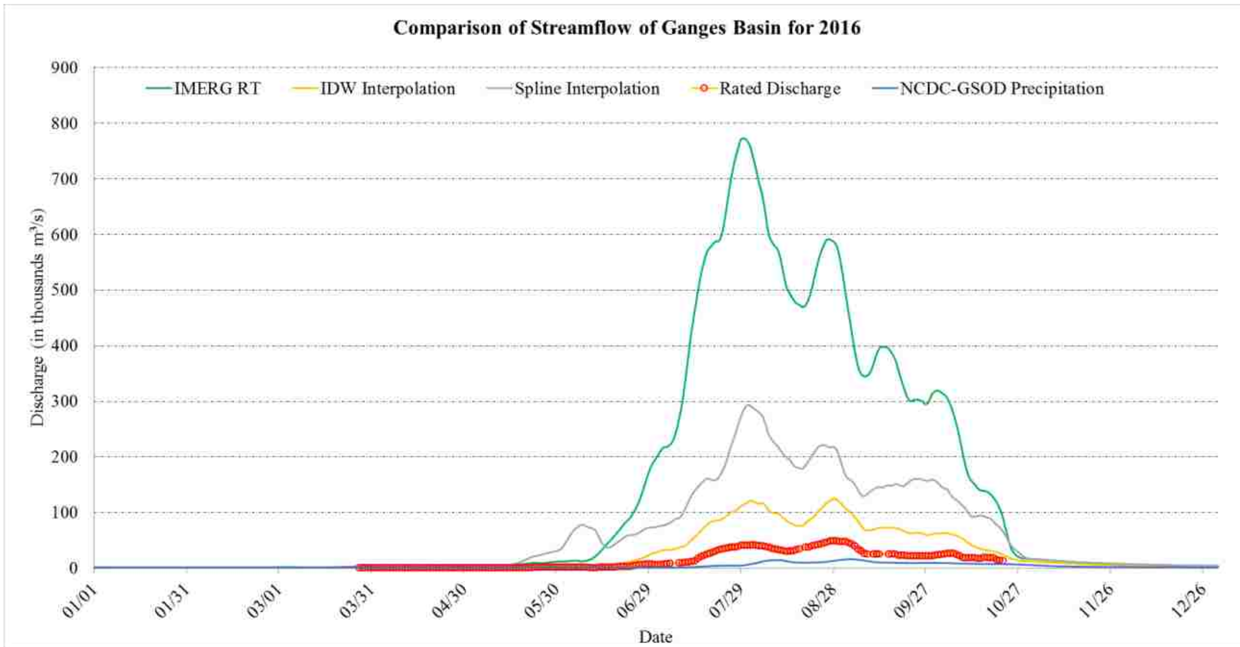


Figure 3.13. Comparison of simulated streamflow of Ganges Basin using IMERG-RT precipitation, Bias corrected precipitation and NCDC-GSOD precipitation with rated discharge at Hardinge Bridge station for the year of 2016

For both Ganges and Brahmaputra Basin, simulated streamflow is compared and their relative performance is reported in Table 3.3. RMSE in streamflow of Brahmaputra Basin decreased to 31306 m^3/s from about 158000 m^3/s and for Ganges Basin, 36188 m^3/s from more than 300,000 m^3/s due to the implementation of IDW method of bias correction.

Table 3.3. Error comparison metrics in simulated streamflow of Ganges and Brahmaputra Basin

Basin	Error Metrics	IMERG-RT	Precipitation Bias Correction	
			IDW Method	Spline Method
Brahmaputra	RMSE (m^3/s)	158568	31306	115655
	Correlation Coefficient	0.94	0.90	0.80
	NSE	-82.03	-2.24	-43.17
Ganges	RMSE (m^3/s)	316566	36188	107239
	Correlation Coefficient	0.93	0.98	0.94
	NSE	-406.9	-4.3	-45.8

3.2.2 *Streamflow Correction System*

In figure 3.14 and 3.15, simulated streamflow by using corrected and non-corrected precipitation along with streamflow correction and rated discharge is shown for Brahmaputra and Ganges Basin respectively. In both cases, streamflow from IMERG-RT product is plotted in the secondary axis as their value is very high compared to the other streamflow values. Table 3.4 and

3.5 summarizes the skill of different combination of correction and their effect in streamflow estimation. In column 2, simulation results from IMERG-RT datasets without any correction is shown. In column 3, results of only streamflow correction techniques are shown. The last two columns are combination of streamflow and precipitation correction for both IDW and spline method of precipitation bias interpolation. During only streamflow correction of both Basins, streamflow from IMERG product always lies outside the climatologic range (min and max) of discharge and thus they are corrected using streamflow correction algorithm to fit with rated discharge. For this reason, their streamflow prediction skill is maximum compared to only precipitation correction and combination of precipitation and streamflow correction. Streamflow from IDW method of bias interpolated precipitation decreases a lot from the IMERG-RT simulated values and thus less amount of outlier found from the no correction region which leads to deviation from rated discharge and their performance seems less than only streamflow correction technique.

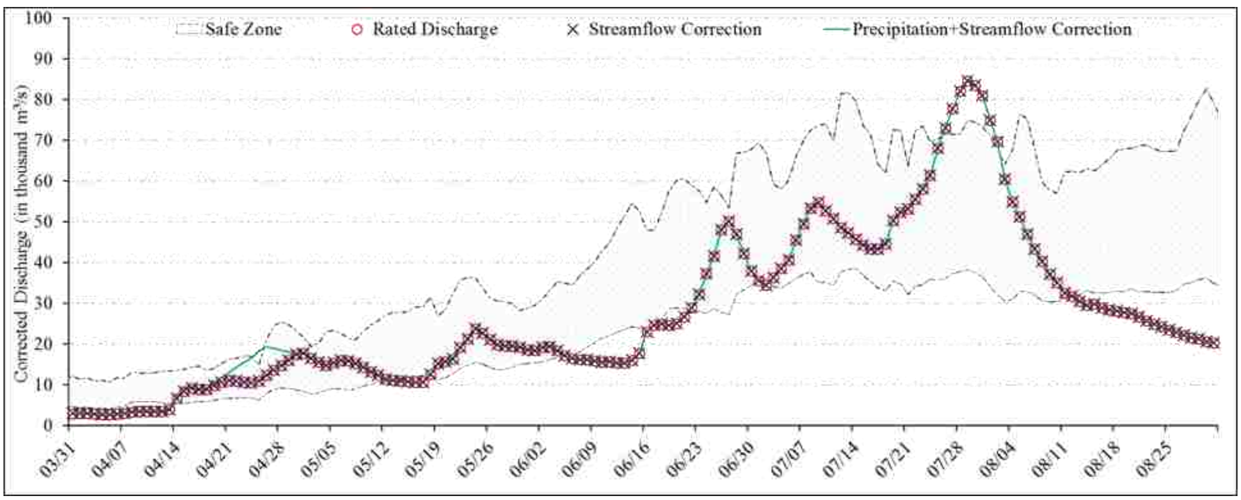


Figure 3.14. Streamflow comparison of Brahmaputra Basin from uncorrected precipitation, by using only streamflow correction and using precipitation and streamflow correction.

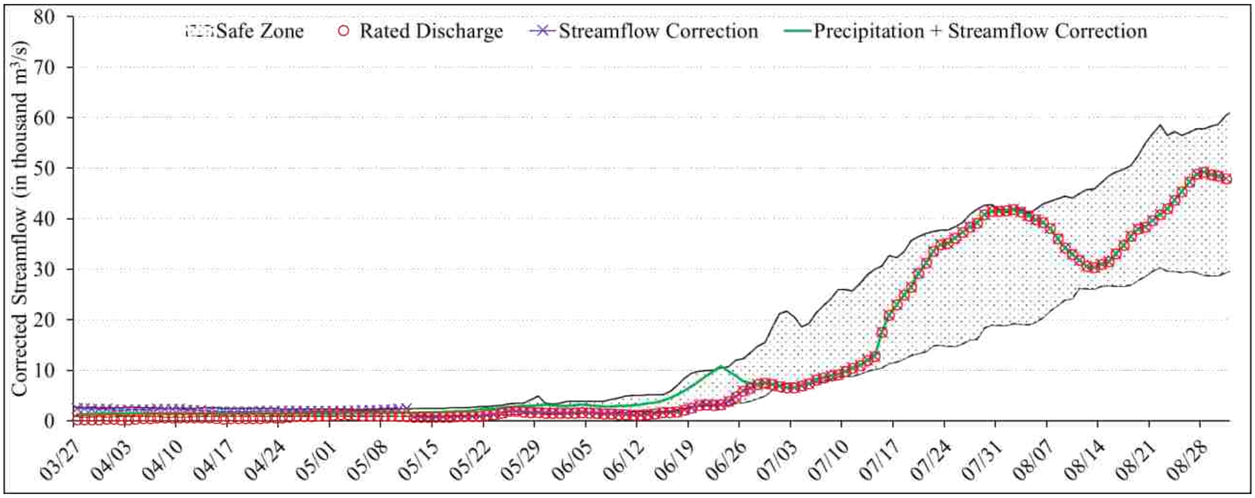


Figure 3.15. Streamflow comparison of Ganges Basin from uncorrected precipitation, by using only streamflow correction and using precipitation and streamflow correction.

Table 3.4. Assessment of precipitation correction, streamflow correction and combined correction system in Brahmaputra Basin

Error Metrics	IMERG-RT	Streamflow	Precipitation + Streamflow	
			IDW	Spline
RMSE (m^3/s)	158274.68	698.56	6929.40	1663.14
Correlation	0.96	1.00	0.95	1.00
NSE	-65.36	1.00	0.87	0.99
Peak Discharge (m^3/s)	433269.19	84674.97	90693.46	84674.97
Error in Peak (m^3/s)	348594.22	0.00	6018.49	0.00
Percentage Error in Peak (%)	411.69	0.00	7.11	0.00
Peak Discharge Ratio	5.12	1.00	1.07	1.00
Total Runoff Volume (Billion M^3)	1996.20	368.13	394.75	372.98
Runoff Ratio with Observed	5.43	1.00	1.07	1.02
Percentage Error in Runoff (%)	443.36	0.20	7.45	1.53

Table 3.5. Skill assessment of precipitation correction, streamflow correction and combined correction system in Ganges Basin

Error Metrics	IMERG-RT	Streamflow	Precipitation + Streamflow	
			IDW	Spline
RMSE (m^3/s)	323246.28	774.53	1409.19	563.62
Correlation	0.95	1.00	1.00	1.00
NSE	-378.74	1.00	0.99	1.00
Peak Discharge (m^3/s)	771669.31	49263.93	49263.93	49263.93
Error in Peak (m^3/s)	722405.37	0.00	0.00	0.00
Percentage Error in Peak (%)	1466.40	0.00	0.00	0.00

Peak Discharge Ratio	15.66	1.00	1.00	1.00
Total Runoff Volume (Billion M^3)	2971.66	188.33	193.57	185.37
Runoff Ratio with Observed	16.84	1.07	1.10	1.05
Percentage Error in Runoff (%)	1583.67	6.70	9.67	5.02

Chapter 4. CONCLUSION AND RECOMMENDATION

4.1 CONCLUSION

The "Built-it-yourself" portal that is developed in this study has numerous applications in water resources management and decision making applications. The growth of open-source and non-proprietary tools has now made it possible for any resource-constrained agencies in the developing world to build robust and cost-effective operational web portals using internal resources. The two key hurdles that have been resolved in this study are:

- An easy-to-access, easy-to-maintain web portal interfaces that connect physical models and can be developed in user agency environment without software complexities and high development and maintenance costs. The SASWMS portal is easily replicable in the stakeholder organization by non-IT professionals with the power of information technology and freely available untapped internet resources.
- The low degree of satellite based hydrological prediction that often results untrustworthy scenarios can be improved by using web based dynamic correction system. This system can be utilized in operational water resources management specially in transboundary setting which lead benefits for the downstream countries by utilizing nowcast mode of upstream information.

4.2 RECOMMENDATIONS

The whole system including the portal and web based dynamic correction has been developed with most updated available resources over the internet. Some of the limitations and future scope of the study are discussed in this section.

- In SASWMS, amount of data transmission and storage capacity can be optimized by using GIS database (e.g. PostgreSQL which is well known for spatial data management, https://www.postgresql.org/). *JavaScript* can be used to prepare the map using any raster processing option. *Google Heatmap* (https://developers.google.com/maps/documentation/javascript/examples/layer-heatmap) is a good option to plot any raster or vector data on the web based background map layer.

- In Raster gridded surface viewer page, static images are linked behind the button and date-time selector. A dynamic way of visualization of these raster maps in animated fashion and automatic date-time slider can be designed to help the users to detect daily/seasonal changes more easily.
- In the Data Download page, specific query can be used for temporal and spatial queries (e.g. start date and end date query) to provide the user for flexibility in accessing data at the required range of selection in both spatial and temporal direction.
- In the web crawler, the maximum possible number of stations are predefined. For this reason, any new stations reported in the websites beyond the prefixed stations cannot be added into interpolation scheme. The list of public domains (total number of 14 websites) are built-in coded which limits its capability to get new agency portals through this scheme of crawling system. More agencies are building new portals to share hydrological information which can be added later to include more station data into the web analytics based correction.
- During designing the web crawler, the location of rainfall information within any webpage is hardcoded into the system. Therefore, the crawler cannot keep track of any format/domain change of the agency webpages. In the dry season, many of the agencies do not share any observed data in their portals and the crawler will not be able to work. Integration of World Meteorological Organization (WMO- GTS) stations will also improve the station coverage of web crawling.
- In the precipitation correction methodology, four methods of real time precipitation correction are described and only one method-spatial bias method is tested in this study. There are three other methods of real time bias correction of satellite estimation which have not assessed during bias correction application. Furthermore, in the spatial bias method, there are different interpolation methods such as Kriging, Nearest Neighbor and other interpolation algorithm can also be tested.
- Spatial distribution of the stations is not homogeneous and currently no option developed to inspect their spatial coverage in this automatic system. Sometimes bias interpolation techniques don't work properly which can worsen the quality of satellite estimated precipitation and thus finally unphysical values in streamflow and other simulated water cycle parameters.

- A simpler version of streamflow correction can also be implemented in the system to reduce the computational steps and complexity. The streamflow quality control filter with the climatology based safe zone discharge can be omitted. Bias in simulated streamflow from in-situ discharge can be adjusted by only comparing with the rated discharge based on the ground measured data availability.

REFERENCES

Ahmad, Q. K., & Ahmed, A. U. (2003). Regional Cooperation in Flood Management in the Ganges-Brahmaputra-Meghna Region: Bangladesh Perspective. Flood Problem and Management in South Asia. 181-198. doi:10.1007/978-94-017-0137-2_9.

Akanda, A. S. (2012). South Asia's water conundrum: hydroclimatic and geopolitical asymmetry, and brewing conflicts in the Eastern Himalayas. International Journal of River Basin Management, 10(4), 307-315. doi:10.1080/15715124.2012.727824.

Akhtar, M. K., Corzo, G. A., J., V. A., & Jonoski, A. (2009). River flow forecasting with Artificial Neural Networks using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges river basin. Retrieved March 25, 2017, from http://adsabs.harvard.edu/abs/2009HESSD...6.3385A.

Castronova, A. M., Goodall, J. L., & Ercan, M. B. (2013). Integrated modeling within a Hydrologic Information System: An OpenMI based approach. Environmental Modelling & Software, 39, 263-273. doi:10.1016/j.envsoft.2012.02.011.

Bakker, M. H. (2009). Transboundary River Floods and Institutional Capacity. JAWRA Journal of the American Water Resources Association, 45(3), 553-566. doi:10.1111/j.1752-1688.2009.00325.x.

Biswas, A. K. (2008). Management of Ganges-Brahmaputra-Meghna System: Way Forward. Water Resources Development and Management of Transboundary Rivers and Lakes, 143-164. doi:10.1007/978-3-540-74928-8_6.

Brichieri-Colombi, S., & Bradnock, R. W. (2003). Geopolitics, water and development in South Asia: cooperative development in the Ganges-Brahmaputra delta. The Geographical Journal, 169(1), 43-64. doi:10.1111/1475-4959.t01-1-00002.

Bolten, J., Lee, C., & Houser, P. (2015). Satellite Data for Water Resources Management. Eos, 96. doi:10.1029/2015eo035971.

Bonnema, M., Sikder, S., Miao, Y., Chen, X., Hossain, F., Pervin, I. A., . . . Lee, H. (2016). Understanding satellite-based monthly-to-seasonal reservoir outflow estimation as a function of hydrologic controls. Water Resources Research, 52(5), 4095-4115. doi:10.1002/2015wr017830.

Bouwer, L.M., J.C.J.H Aerts, G.M Van de Coterlet, N. Van de Giessen, A. Gieske, A. and C. Manaerts. (2004). Evaluating downscaling methods for preparing Global Circulation Model (GCM) data for hydrological impact modelling. Chapter 2, in Aerts, J.C.J.H. & Droogers, P.

(Eds.), Climate Change in Contrasting River Basins: Adaptation Strategies for Water, Food and Environment, 25-47. doi:10.1079/9780851998350.0025.

Seo, D., Breidenbach, J., & Johnson, E. (1999). Real-time estimation of mean field bias in radar rainfall data. Journal of Hydrology, 223(3-4), 131-147. doi:10.1016/s0022-1694(99)00106-7.

Bulatewicz, T., D. Andresen, S. Auvenshine, J. Peterson, D.R. Steward (2014). A distributed data component for the Open Modeling Interface. Environmental Modelling & Software, 57, 138-151. doi:10.1016/j.envsoft.2014.02.017.

Cheema, M. J., & Bastiaanssen, W. G. (2012). Local calibration of remotely sensed rainfall from the TRMM satellite for different periods and spatial scales in the Indus Basin. International Journal of Remote Sensing, 33(8), 2603-2627. doi:10.1080/01431161.2011.617397.

Cherkauer, K. A., L. C. Bowling and D.P. Lettenmaier, (2003). Variable Infiltration Capacity (VIC) cold land process model updates. Global and Planetary Change, 38(1-2), 151-159. doi:10.1016/S0921-8181(03)00025-0.

Dan, L., Ji, J., Xie, Z., Chen, F., Wen, G., & Richey, J. E. (2012). Hydrological projections of climate change scenarios over the 3H region of China: A VIC model assessment. Journal of Geophysical Research: Atmospheres, 117(D11). doi:10.1029/2011jd017131.

Ebert, E. E., Janowiak, J. E., & Kidd, C. (2007). Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models. Bulletin of the American Meteorological Society, 88(1), 47-64. doi:10.1175/bams-88-1-47.

F. Silvestro, S. Gabellani, F. Delogu, R. Rudari, and G. Boni1, (2013). Exploiting remote sensing land surface temperature in distributed hydrological modelling: the example of the Continuum model. Hydrol. Earth Syst. Sci., 17, 39–62. doi:10.5194/hess-17-39-2013.

Gao, H., Tang, Q., Ferguson, C. R., Wood, E. F., & Lettenmaier, D. P. (2010). Estimating the water budget of major US river basins via remote sensing. International Journal of Remote Sensing, 31(14), 3955-3978. doi:10.1080/01431161.2010.483488.

Gebregiorgis, A., & Hossain, F. (2011). How Much Can a Priori Hydrologic Model Predictability Help in Optimal Merging of Satellite Precipitation Products? Journal of Hydrometeorology, 12(6), 1287-1298. doi:10.1175/jhm-d-10-05023.1.

Gebregiorgis, A., & Hossain, F. (2014). Making Satellite Precipitation Data Work for the Developing World. IEEE Geoscience and Remote Sensing Magazine, 2(2), 24-36. doi:10.1109/mgrs.2014.2317561.

Haddeland, I., Skaugen, T., & Lettenmaier, D. P. (2007). Hydrologic effects of land and water management in North America and Asia: 1700–1992. Hydrology and Earth System Sciences, 11(2), 1035-1045. doi:10.5194/hess-11-1035-2007.

Horsburgh, J. S., Tarboton, D. G., Piasecki, M., Maidment, D. R., Zaslavsky, I., Valentine, D., & Whitenack, T. (2009). An integrated system for publishing environmental observations data. Environmental Modelling & Software, 24(8), 879-888. doi:10.1016/j.envsoft.2009.01.002.

Hossain, F., C K, Shum, F.J. Turk, S. Biancamaria, H. Lee, A. Limaye, M. Hossain, S. Shah-Newaz, L.C. Mazumder, T. Ahmed, W. Yigzaw and A.H.M. Siddique-E-Akbor (2014). A Guide for Crossing the Valley of Death: Lessons Learned from Making a Satellite based Flood Forecasting System Operational and Independently Owned by a Stakeholder Agency. Bulletin of the American Meteorological Society, 95(8), 1201-1207. doi:10.1175/bams-d-13-00176.1.

Hossain, F., & Huffman, G. J. (2008). Investigating Error Metrics for Satellite Rainfall Data at Hydrologically Relevant Scales. Journal of Hydrometeorology, 9(3), 563-575. doi:10.1175/2007jhm925.1.

Hossain, F. (2012). Do Satellite Data Portals Today Reach Out to Diverse End Users Around the World? Bulletin of the American Meteorological Society, 93(11), 1633-1634. doi:10.1175/bams-d-12-00035.1.

Hossain, F. (2015). Data for All: Using Satellite Observations for Social Good. Eos, 96. doi:10.1029/2015eo037319.

Hossain, F., Siddique-E-Akbor, A. H., Mazumder, L. C., Shahnewaz, S. M., Biancamaria, S., Lee, H., & Shum, C. K. (2014). Proof of Concept of an Altimeter-Based River Forecasting System for Transboundary Flow Inside Bangladesh. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(2), 587-601. doi:10.1109/jstars.2013.2283402.

Hossain, F. M. Maswood, A.H.M. Siddique-E-Akbor, W. Yigzaw, L.C. Mazumder, T. Ahmed, M. Hossain, S. Shah-Newaz, A. Limaye, H. Lee, S. Pradhan, B. Shrestha, B. Bajracharya, C K Shum, F. J. Turk. (2014). A Promising Radar Altimetry Satellite System for Operational

Flood Forecasting in Flood-Prone Bangladesh. IEEE Geoscience and Remote Sensing Magazine, 2(3), 27-36. doi:10.1109/mgrs.2014.2345414.

Hou, A.Y. et al., (2014). The global precipitation measurement mission. Bulletin of American Meteorological Society, 95, 701–722. doi:10.1175/BAMS-D-13-00164.1.

Huffman, G.J., D.T. Bolvin, E.J. Nelkin, (2015). Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation. NASA/GSFC Code 612, 47.

Human Development Report 2006. Beyond scarcity: Power, poverty and the global water crisis. Chapter 6. UNDP, 2006.

Immerzeel, W., Rutten, M., & Droogers, P. (2009). Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula. Remote Sensing of Environment, 113(2), 362-370. doi:10.1016/j.rse.2008.10.004.

Kansakar, P., & Hossain, F. (2016). A review of applications of satellite earth observation data for global societal benefit and stewardship of planet earth. Space Policy, 36, 46-54. doi:10.1016/j.spacepol.2016.05.005.

Khan, S. I., Yang, h., & Vergara, H.J. (2012). Microwave Satellite Data for Hydrologic Modeling in Ungauged Basins. IEEE Geoscience and Remote Sensing Letters, 9, 663–667. doi:10.1109/ LGRS.2011.2177807.

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research, 99(D7), 14415. doi:10.1029/94jd00483.

Lohmann, D., Raschke, E., Nijssen, B., & Lettenmaier, D. P. (1998). Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model. Hydrological Sciences Journal, 43(1), 131-141. doi:10.1080/02626669809492107.

Pedro-Monzonís, M., Solera, A., Ferrer, J., Andreu, J., & Estrela, T. (2016). Water accounting for stressed river basins based on water resources management models. Science of The Total Environment, 565, 181-190. doi:10.1016/j.scitotenv.2016.04.161.

Maswood, M. and F. Hossain (2016). Advancing River Modeling in Ungauged River Basins using Remote Sensing: The Case of Ganges-Brahmaputra-Meghna Basins. Journal of River Basin Management, 14(1), 103-117. doi:10.1080/15715124.2015.1089250.

McEwen T. (2008), Ganges-Brahmaputra-Meghna/Barak basin. (Online at: http://www.ce.utexas.edu/prof/mckinney/ce397/Topics/Ganges/Ganges(2008).pdf)

Brown, M.E. A.E. Racoviteanu, D.G. Tarboton, A. Sen Guptc, J. Nigro, F. Policelli, S. Habib, M. Tokay, M.S. Shrestha, S. Bajracharya, P. Hummel, M. Gray, P. Duda, B. Zaitchik, V. Mahat, G. Artan, S. Tokar (2014). An integrated modeling system for estimating glacier and snow melt driven streamflow from remote sensing and earth system data products in the Himalayas. Journal of Hydrology, 519, 1859-1869. doi:10.1016/j.jhydrol.2014.09.050.

Musa, Z. N., Popescu, I., & Mynett, A. (2015). A review of applications of satellite SAR, optical, altimetry and DEM data for surface water modelling, mapping and parameter estimation. Hydrology and Earth System Sciences, 19(9), 3755-3769. doi:10.5194/hess-19-3755-2015.

National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce (2000). NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. http://dx.doi.org/10.5065/D6M043C6.

Nijssen, B., G. M. O'Donnell, D. P. Lettenmaier, D. Lohmann, and E. F. Wood (2001). Predicting the discharge of global rivers. Journal of Climate, 14(15), 3307–3323. doi:10.1175/1520-0442(2001)014<3307:PTDOGR>2.0.CO;2.

Nijssen, B. (2004). Effect of precipitation sampling error on simulated hydrological fluxes and states: Anticipating the Global Precipitation Measurement satellites. Journal of Geophysical Research, 109(D2). doi:10.1029/2003jd003497.

Katiyar. N. and F. Hossain, (2007). An Open-book Watershed Model for Prototyping Space-borne Flood Monitoring Systems in International River Basins. Environmental Modeling and Software. 22(12), 1720-1731. doi:10.1016/j.envsoft.2006.12.005.

Eltayeby, O., John, D., Patel, P., & Simmerman, S. (2013). Comparative case study between D3 & Highcharts on Lustre metadata visualization. 2013 IEEE Symposium on Large-Scale Data Analysis and Visualization (LDAV). doi:10.1109/ldav.2013.6675172.

Prakash, S., Mitra, A. K., Aghakouchak, A., Liu, Z., Norouzi, H., & Pai, D. (2016). A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. Journal of Hydrology. doi:10.1016/j.jhydrol.2016.01.029

Rodell, M., P.R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J.K. Entin, J.P. Walker, D. Lohmann, and D. Toll

(2004). The Global Land Data Assimilation System. Bull. Amer. Meteor. Soc., 85, 381–394. doi: 10.1175/BAMS-85-3-381.

Schumann, G., Bates, P. D., Horritt, M. S., Matgen, P., & Pappenberger, F. (2009). Progress in integration of remote sensing–derived flood extent and stage data and hydraulic models. Reviews of Geophysics, 47(4). doi:10.1029/2008rg000274

Siddique-E-Akbor, A.H.M., F. Hossain, C K Shum, F.J. Turk, Steven Tseng, and Yuchan Yi (2014). Satellite Precipitation Data Driven Hydrologic Modeling for Water Resources Management in the Ganges, Brahmaputra and Meghna Basins. Earth Interact., 18, 1–25. doi: 10.1175/EI-D-14-0017.1.

Vrugt, J. A., Braak, C. J., Clark, M. P., Hyman, J. M., & Robinson, B. A. (2008). Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. Water Resources Research, 44(12). doi:10.1029/2007wr006720.

Water Without Borders. Backgrounder. United Nations Department of Public Information, 2004.

Woldemichael, A. T., Hossain, F., Pielke, R., & Beltrán-Przekurat, A. (2012). Understanding the impact of dam-triggered land use/land cover change on the modification of extreme precipitation. Water Resources Research, 48(9). doi:10.1029/2011wr011684.

Wu, H., Kimball, J. S., Elsner, M. M., Mantua, N., Adler, R. F., & Stanford, J. (2012). Projected climate change impacts on the hydrology and temperature of Pacific Northwest rivers. Water Resources Research, 48(11). doi:10.1029/2012wr012082.

Wu, Y., Liang, Z., & Liu, L. (2013). Design and implementation of tourism information system based on Google Maps API. 2013 21st International Conference on Geoinformatics. doi:10.1109/geoinformatics.2013.6626139.

Yuan, X. (2016). An experimental seasonal hydrological forecasting system over the Yellow River basin – Part II: The added value from climate forecast models. Hydrology and Earth System Sciences Discussions, 1-24. doi:10.5194/hess-2016-102.

Zhao, F., Chiew, F. H., Zhang, L., Vaze, J., Perraud, J., & Li, M. (2012). Application of a Macroscale Hydrologic Model to Estimate Streamflow across Southeast Australia. Journal of Hydrometeorology, 13(4), 1233-1250. doi:10.1175/jhm-d-11-0114.1.

Zhu, C., Leung, L. R., Gochis, D., Qian, Y., & Lettenmaier, D. P. (2009). Evaluating the influence of antecedent soil moisture on variability of the North American monsoon precipitation in the Coupled MM5/VIC Modeling System. Journal of Advances in Modeling Earth Systems, 2. doi:10.3894/james.2009.1.13.

APPENDIX A

Table A.1. List of Websites in GBM region and maximum available stations in each websites

Website Title	Website Link	Stations
IMD City Weather, India	http://14.139.247.11/citywx/citywxnew.php	229
City Weather (cityWX), India	http://202.54.31.7/citywx/city_weather.php	33
City WX updated Page, India	http://202.54.31.7/citywx/city_weather1.php	14
Weather Underground Website	http://www.wunderground.com	18
Meteorological Forecasting Division, Nepal	http://www.mfd.gov.np/	19
Flood Forecasting and Warning Center, Bangladesh	http://www.ffwc.gov.bd/	53
Department of Hydrology and Meteorology, Nepal	http://hydrology.gov.np/	64
Regional Meteorological Center (AMSS), Delhi, India	http://amssdelhi.gov.in/	7
Regional Meteorological Center, New Delhi, India	http://121.241.116.157/dynamic/weather/delhiregion.html	8
Regional Meteorological Center, Gangtok, India	http://www.imdsikkim.gov.in/daily_Forecast.pdf	Variable
Weather Report for NCR Delhi, India	http://121.241.116.157/dynamic/weather/Delhi.pdf	Variable
Indian Meteorological Department, Guwahati, India	http://www.imdguwahati.gov.in/dwr.pdf	Variable
Bangladesh Meteorological Department, Bangladesh	http://www.bmd.gov.bd/	39
Central Water Commission, India	http://www.cwc.gov.in/	181
	Maximum Possible Number of Stations	990

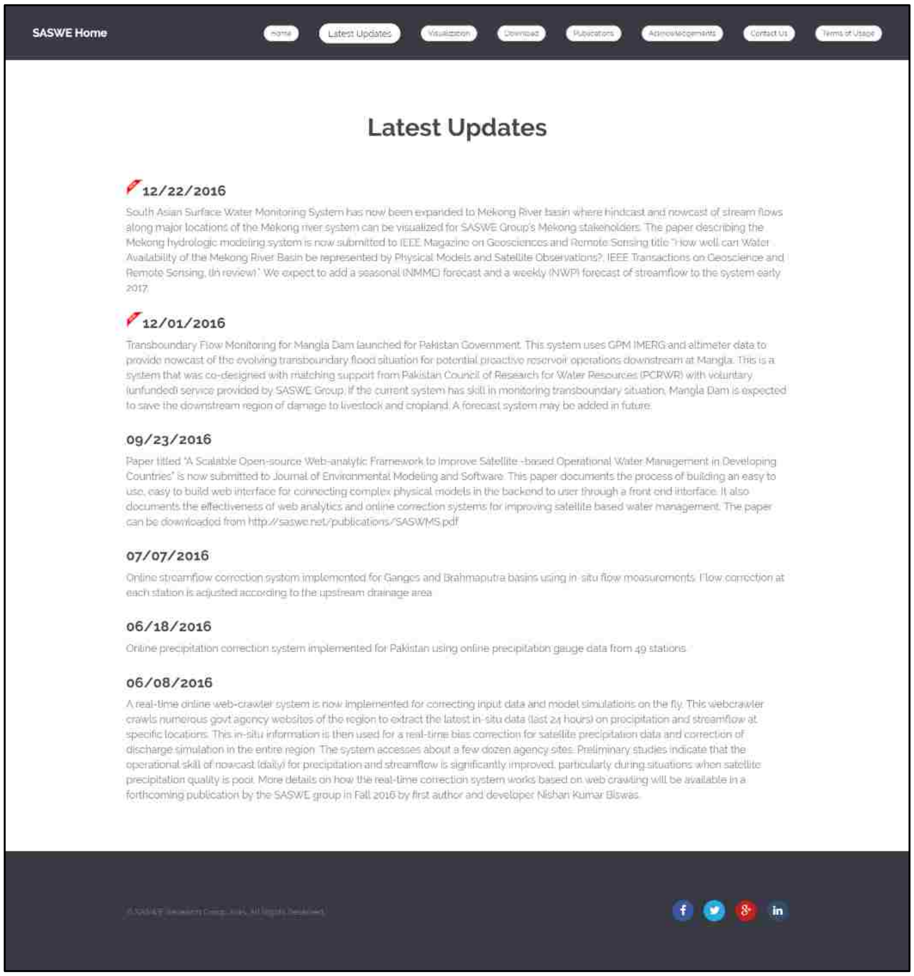


Figure A.1. Latest update page of SASWMS Portal

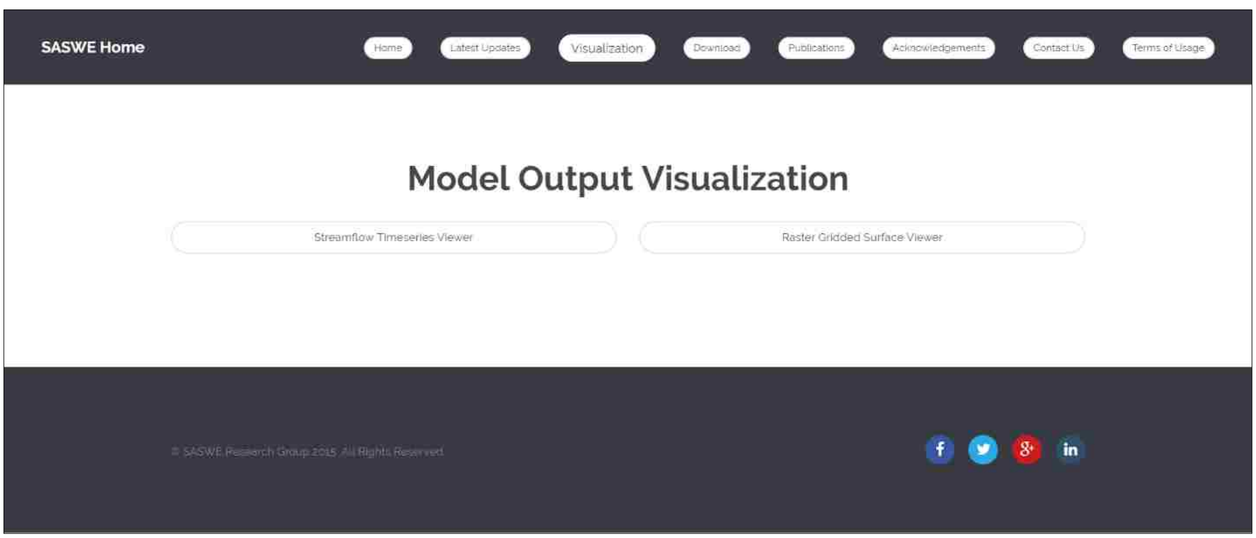


Figure A.2. Visualization parent menu of SASWMS Portal

Figure A.3. Initialization of Raster Gridded Visualization page of SASWMS Portal

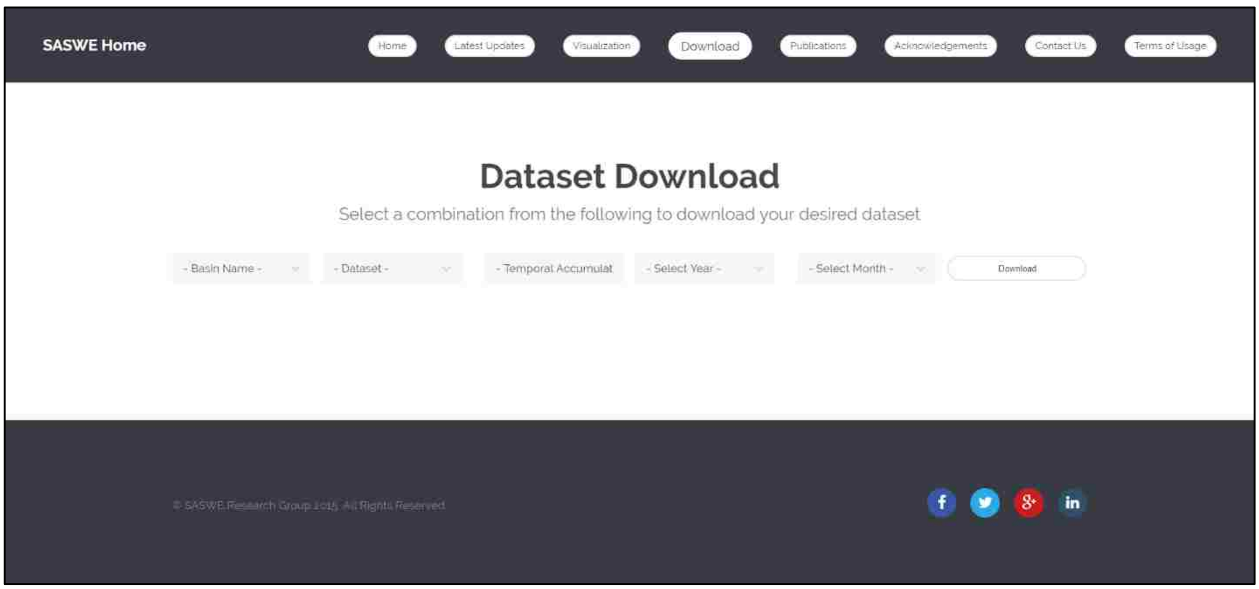


Figure A.4. Data Download parent menu of SASWMS Portal

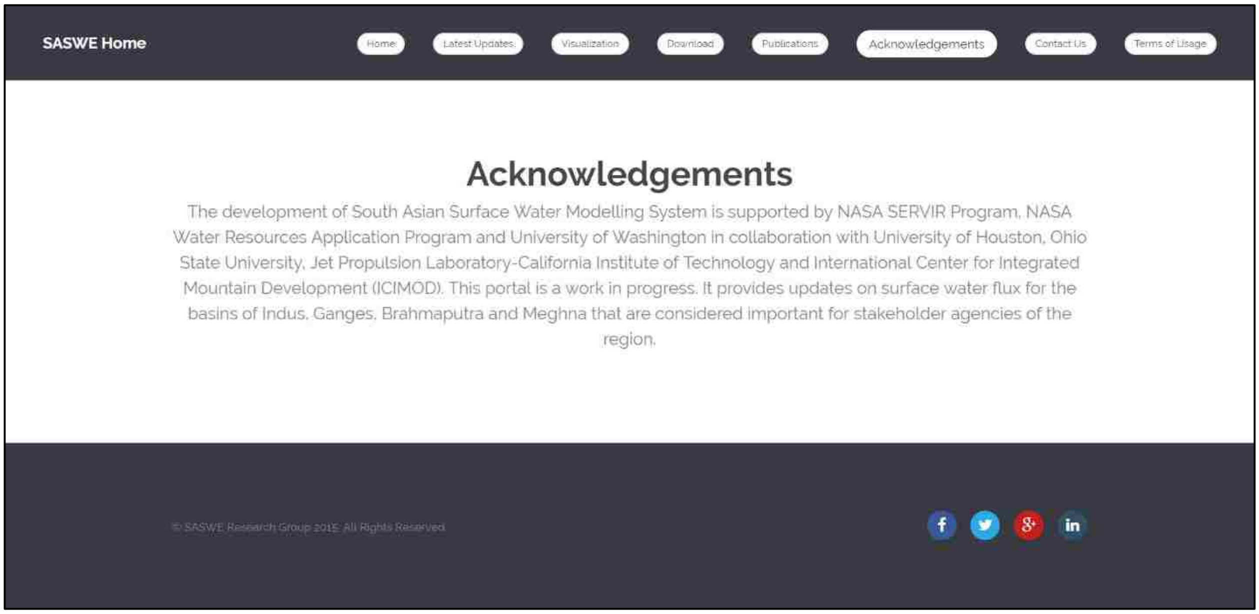


Figure A.5. Acknowledgements page of SASWMS Portal

Figure A.6. Acknowledgements page of SASWMS Portal

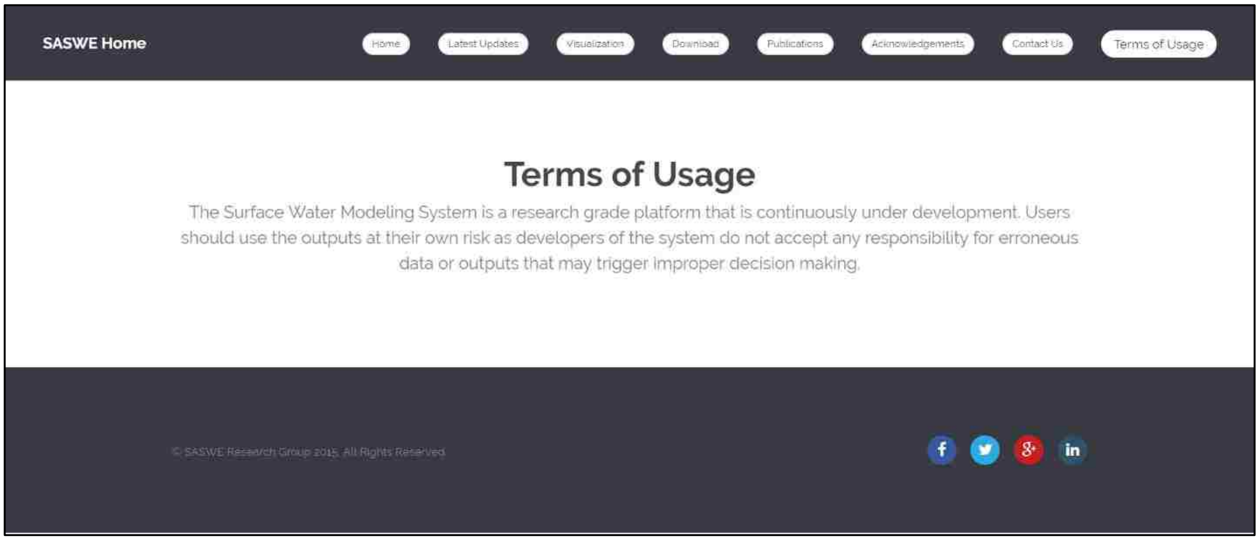


Figure A.7. Terms of Usage page of SASWMS Portal

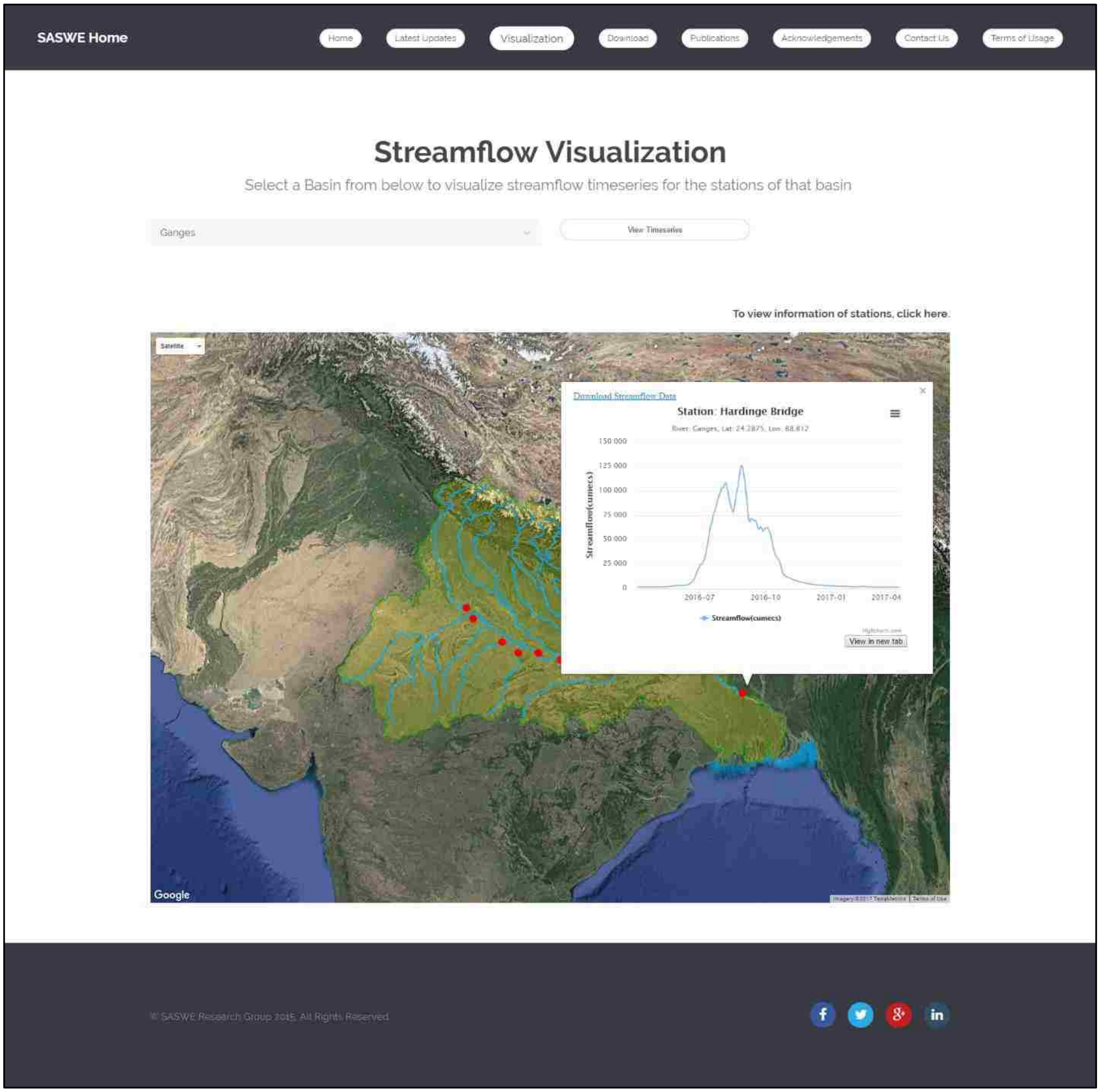


Figure A.8. An example of iframe uses in SASWMS portal (after clicking on the location of Hardinge Bridge, the iframe is shown with the streamflow of the station)

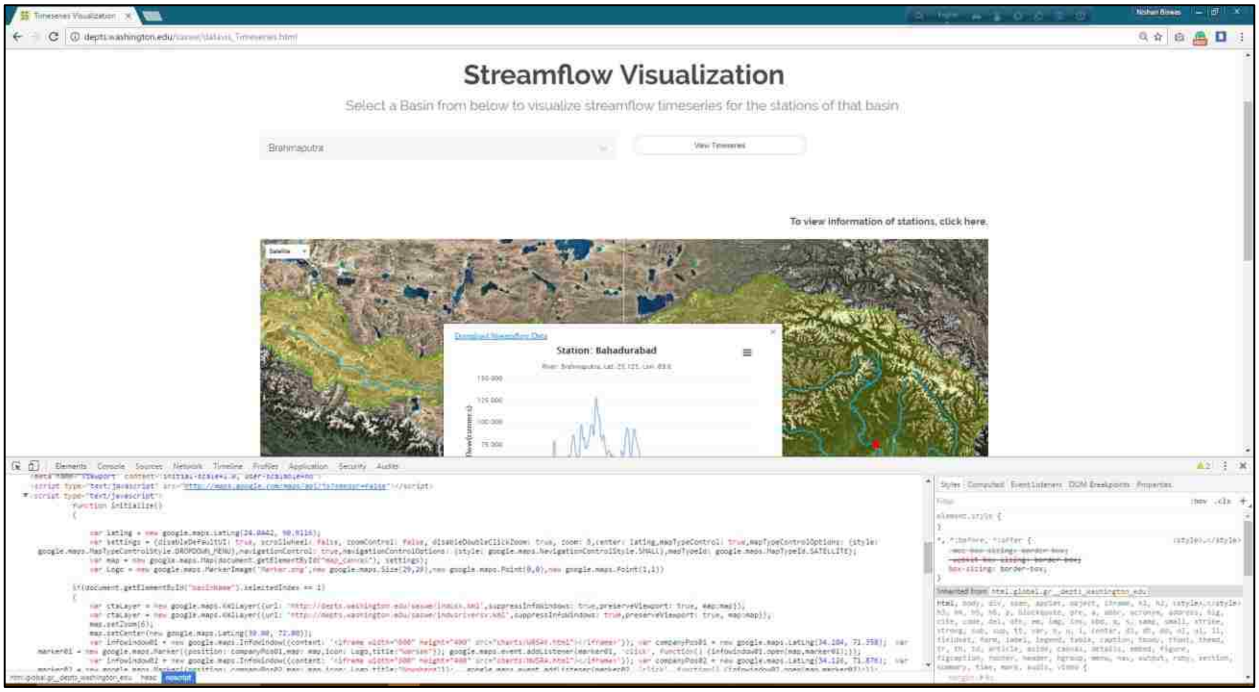


Figure A.9. Sample JavaScript code for background google map connection and timeseries visualization (this code can be opened by going to the Time Series Visualization and then right click in anywhere, inspect, and then initialize function under the head HTML tag)

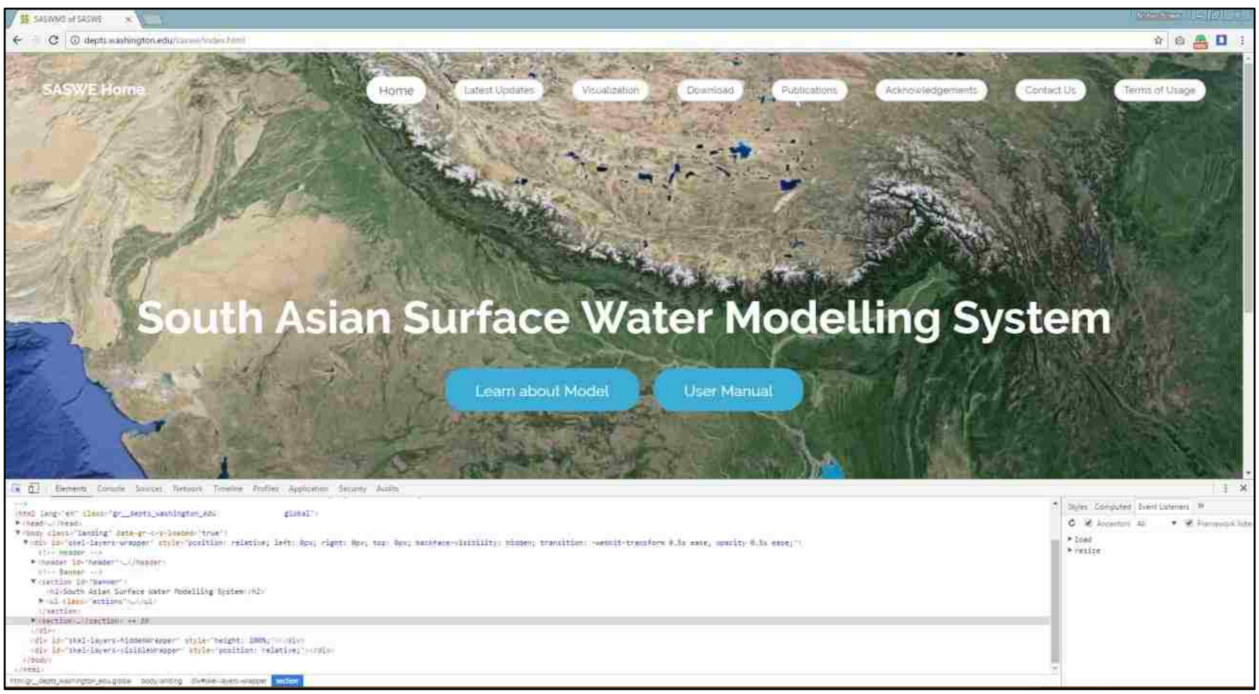


Figure A.10. Sample HTML Code for building the portal

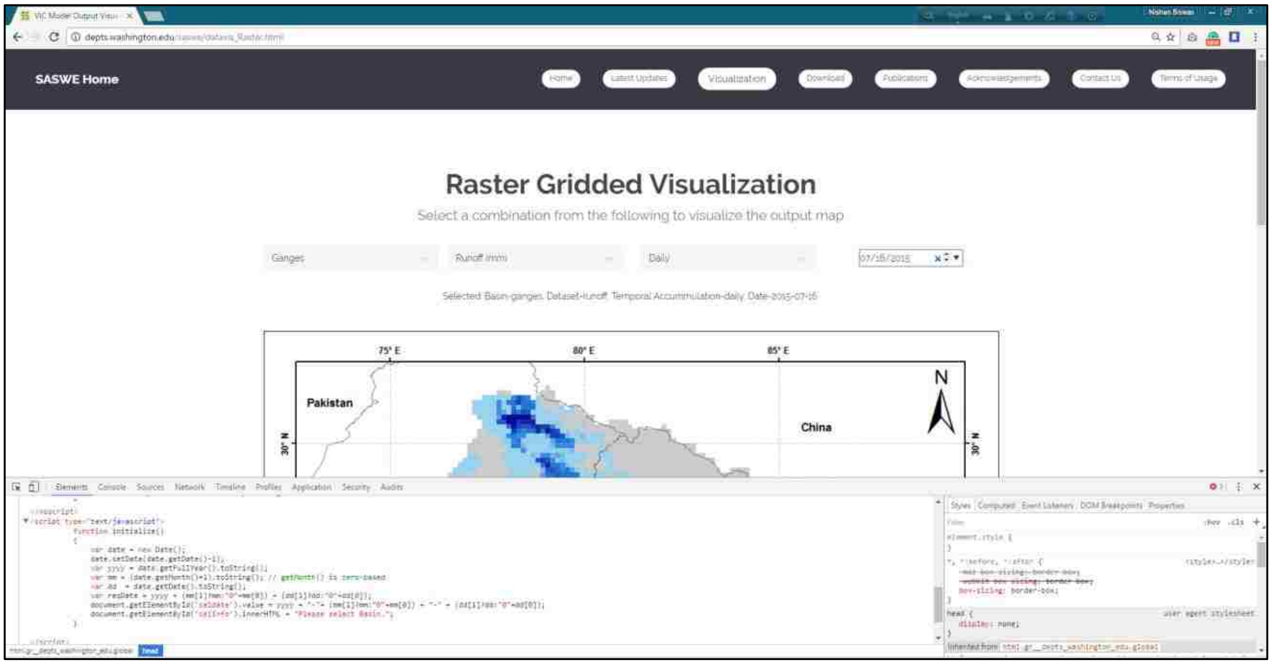


Figure A.11. Date Time Initialization using JavaScript

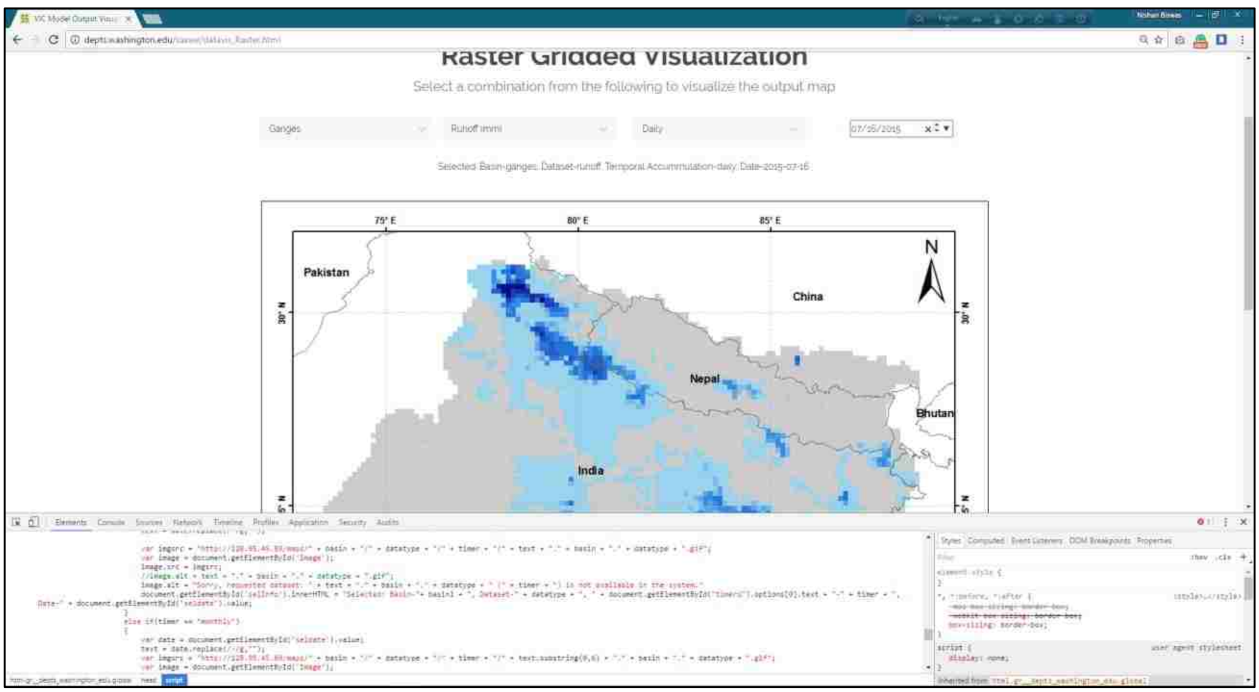


Figure A.12. Sample JavaScript code for Raster Map connection (this code can be opened by going to the Raster Grid Visualization and then right click in anywhere, inspect, generate raster function under the head HTML tag)

Data Download started for 2016-09-07, started at 2016-09-07 00:01:01
1st Website, IMD front page, Page link: http://14.139.247.11/citywx/citywxnew.php, Downloaded Station: 206
2nd Website, CityWX Earlier Page: http://202.54.31.7/citywx/city_weather.php, Downloaded station: = 70
3rd Website, CityWX updated Page: http://202.54.31.7/citywx/city_weather1.php, Downloaded station: = 16
4th Website, W-Underground Page: http://www.wunderground.com, Downloaded Station= 66
5th Website, Meteorological Forecasting Division of Nepal: http://www.mfd.gov.np/, Downloaded Station= 17
6th Website, Flood Forecasting and Warning Center, Banglaseh: http://www.ffwc.gov.bd/, Downloaded Station= 61
7th Website, Department of Hydrology and Meteorology, Nepal: http://hydrology.gov.np/, Downloaded Station= 68
8th Website, Regional Meteorological Center, New Delhi: http://amssdelhi.gov.in/, Downloaded Station= 30
9th Website, Regional Meteorological Center, New Delhi,: http://121.241.116.157/dynamic/weather/delhiregion.html, Downloaded Station: 18
10th Website, Regional Meteorological Centre, Gangtok-Sikkim: http://www.imdsikkim.gov.in/daily_Forecast.pdf, Downloaded Station: 11
11th Website, Weather Report for NCR Delhi: http://121.241.116.157/dynamic/weather/Delhi.pdf, Downloaded Station: 25
12th Website, Guwahati PDF: http://www.imdguwahati.gov.in/dwr.pdf, Downloaded Station: 0
13th Website, Bangladesh Meterological Department: http://www.bmd.gov.bd/, Downloaded Station: 39
14th Website, Central Water Commission, India: http://www.cwc.gov.in/, Downloaded Station: 285
All station data downloaded and saved successfully.

Figure A.13. Log Content of Web Crawler of SASWMS of 2016-09-07

Precipitation Correction started for Date: 2016-09-03, started at 2016-09-03 08:14:21
IMERG Satellite estimated rainfall Data correcting started...
162 value replaced with exact value from collected rainfall data.
57 value removed from collected rainfall data.
Today rainfall over GBM Basin's CSV file is creating ....
Rainfall CSV file successfully created. No. of stations found: 584
Rainfall data is dividing according to the Basins ...
Rainfall file for Ganges Basin successfully created. No. of stations found: 291
Maximum, Average and Minimum Observed Rainfall in mm over Ganges Basin: 137.00, 6.84 and 0.00
Rainfall file for Brahmaputra Basin successfully created. No. of stations found: 202
Maximum, Average and Minimum Observed Rainfall in mm over Brahmaputra Basin: 158.00, 14.70 and 0.00
Rainfall file for Indus Basin successfully created. No. of stations found: 223
Maximum, Average and Minimum observed Rainfall in mm over Indus Basin: 38.20, 0.65 and 0.00
Rainfall file for Pakistan successfully created. No. of stations found: 137
Maximum, Average and Minimum Observed Rainfall in mm over Pakistan: 76.20, 1.60 and 0.00
Windows PC is now going to copy IMERG precipitation files from SASWMS Server ....
IMERG precipitation files successfully downloaded from SASWMS Server ....
Maximum, Average and Minimum IMERGRT estimated Rainfall in mm over Ganges Basin: 523.03, 34.61 and 0.00
Maximum, Average and Minimum IMERGRT estimated Rainfall in mm over Brahmaputra Basin: 464.79, 50.53 and 0.00
Maximum, Average and Minimum IMERGRT estimated Rainfall in mm over Indus Basin: 408.06, 5.06 and 0.00
Maximum, Average and Minimum IMERGRT estimated Rainfall in mm over Pakistan: 172.46, 0.82 and 0.00
Applying Bias Correction to the IMERG Precipitation using python and ArcGIS....
Python successfully initialized and hopefully adjusted bias. To check the result please see the GIS folder.
Windows PC is trying to send corrected precipitation to SASWMS Server....
Bias corrected IMERG precipitation files successfully sent to SASWMS Server.
Maximum, Average and Minimum of Bias Corrected IMERGRT estimated Rainfall in mm over Ganges Basin: 435.92, 8.21 and 0.00
Maximum, Average and Minimum of Bias Corrected IMERGRT estimated Rainfall in mm over Brahmaputra Basin: 395.75, 15.17 and 0.00
Maximum, Average and Minimum of Bias Corrected IMERGRT estimated Rainfall in mm over Indus Basin: 407.78, 3.12 and 0.00
Maximum, Average and Minimum of Bias Corrected IMERGRT estimated Rainfall in mm over Pakistan: 170.37, 0.97 and 0.00
All scheduled tasks completed at 2016-09-03 08:15:02.

Figure A.14. Log Content of Precipitation Correction System of SASWMS of 2016-09-03

```
Dataset and Map Uploader program started for 2017-02-28, started at 2017-02-28 00:40:17
Weekly corrected precipitation calculated and all files copied successfully for ganges.
Weekly corrected precipitation calculated and all files copied successfully for brahmaputra.
Weekly corrected precipitation calculated and all files copied successfully for indus1.
Weekly corrected precipitation calculated and all files copied successfully for pakistan.
Python is intializing to prepare daily and weekly precipitation maps....
Python successfully initialized and hopefully prepared maps. To check the result please see the Maps folder.
Maps and datasets of daily and weekly precipitation are sending to SASWE Server....
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\datasets\ganges\precipitation\daily\20170227.ganges.precipitation.txt
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\maps\ganges\precipitation\daily\20170227.ganges.precipitation.gif
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\datasets\ganges\precipitation\weekly\20170227.ganges.precipitation.weekly.txt
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\maps\ganges\precipitation\weekly\20170227.ganges.precipitation.weekly.gif
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\datasets\brahmaputra\precipitation\daily\20170227.brahmaputra.precipitation.txt
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\maps\brahmaputra\precipitation\daily\20170227.brahmaputra.precipitation.gif
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\datasets\brahmaputra\precipitation\weekly\20170227.brahmaputra.precipitation.weekly.txt
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\maps\brahmaputra\precipitation\weekly\20170227.brahmaputra.precipitation.weekly.gif
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\datasets\indus\precipitation\daily\20170227.indus.precipitation.txt
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\maps\indus\precipitation\daily\20170227.indus.precipitation.gif
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\datasets\indus\precipitation\weekly\20170227.indus.precipitation.weekly.txt
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\maps\indus\precipitation\weekly\20170227.indus.precipitation.weekly.gif
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\datasets\meghna\precipitation\daily\20170227.meghna.precipitation.txt
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\maps\meghna\precipitation\daily\20170227.meghna.precipitation.gif
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\datasets\meghna\precipitation\weekly\20170227.meghna.precipitation.weekly.txt
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\maps\meghna\precipitation\weekly\20170227.meghna.precipitation.weekly.gif
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\datasets\pakistan\precipitation\daily\20170227.pakistan.precipitation.txt
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\maps\pakistan\precipitation\daily\20170227.pakistan.precipitation.gif
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\datasets\pakistan\precipitation\weekly\20170227.pakistan.precipitation.weekly.txt
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\maps\pakistan\precipitation\weekly\20170227.pakistan.precipitation.weekly.gif
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\datasets\pakistan\precipitation_corr\daily\20170227.pakistan.precipitation_corr.txt
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\datasets\pakistan\precipitation_corr\weekly\20170227.pakistan.precipitation_corr.weekly.txt
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\maps\pakistan\precipitation_corr\daily\20170227.pakistan.precipitation_corr.gif
C:\Users\nbiswas\Desktop\Nishan\SASWE\VIC_HKH\WebContent\maps\pakistan\precipitation_corr\weekly\20170227.pakistan.precipitation_corr.weekly.gif
All steps completed successfully.
```

Figure A.15. Log content of dataset and map upload records from the operational machine to SASWMS Server

```
///--------------------------------------------Data from AMSS-Delhi Webpage---------------------------------------------------
try
{
    Console.ForegroundColor = ConsoleColor.Cyan;
    Console.WriteLine("8th Website, Regional Meteorological Center, New Delhi: http://amssdelhi.gov.in/");
    Directory.CreateDirectory(di + @"\WebPage8");
    Console.ResetColor();

    WebClient client = new WebClient();
    string htmlCode = client.DownloadString("http://amssdelhi.gov.in/dynamic/weather/wxtable.html");
    File.WriteAllText(di + @"\WebPage8\amssdelhi.gov.in.html", htmlCode);
    HtmlDocument doc = new HtmlDocument();
    doc.LoadHtml(htmlCode);

    HtmlNodeCollection date = doc.DocumentNode.SelectNodes(".//b");
    string titleText = date[1].InnerText;
    var tableTitle = titleText.Split(' ');
    DateTime webDate = DateTime.Parse(tableTitle[8] + " " + tableTitle[9]).AddHours(9);
    HtmlNodeCollection tables = doc.DocumentNode.SelectNodes("//table");
    HtmlNodeCollection rows = tables[0].SelectNodes(".//tr");
    for (int i = 4; i < rows.Count - 4; i++)
    {
        HtmlNodeCollection col = rows[i].SelectNodes(".//td");

        rainfallData.AppendLine(webDate.ToString("yyyy-MM-dd HH:mm:ss") + "," + col[0].InnerText.Trim() + "," + col[7].InnerText.Trim());
        Console.WriteLine(webDate.ToString("yyyy-MM-dd HH:mm:ss") + "," + col[0].InnerText.Trim() + "," + col[7].InnerText.Trim());
    }
    logText.AppendLine("8th Website, Regional Meteorological Center, New Delhi: http://amssdelhi.gov.in/, Downloaded Station= " + (rows.Count - 6));

}
catch (Exception error)
{
    Console.ForegroundColor = ConsoleColor.Red;
    Console.WriteLine("Problem in 8th Website, Regional Meteorological Center, New Delhi: http://amssdelhi.gov.in/ cannot be accessed, Error: " + error.Message);
    Console.ResetColor();
    logText.AppendLine("Problem in 8th Website, Regional Meteorological Center, New Delhi: http://amssdelhi.gov.in/ cannot be accessed, Error: " + error.Message);
}
```

Figure A.16. Sample code for parsing html table

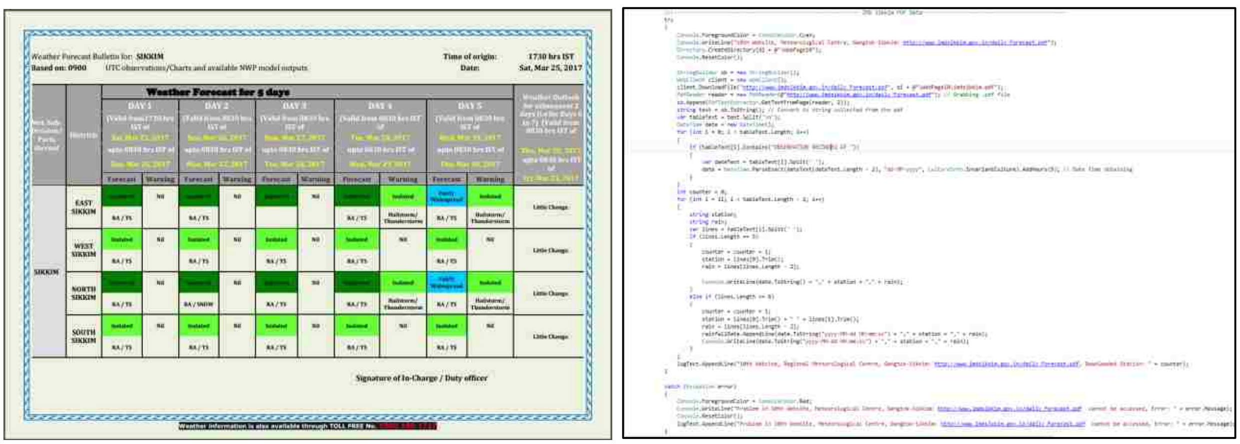
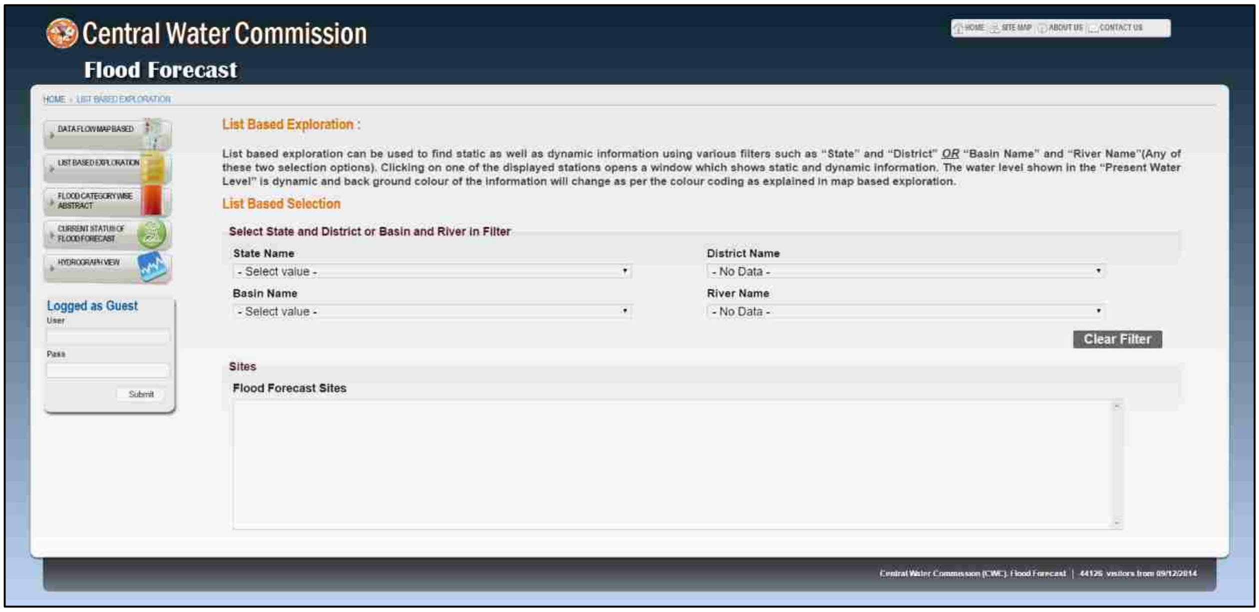


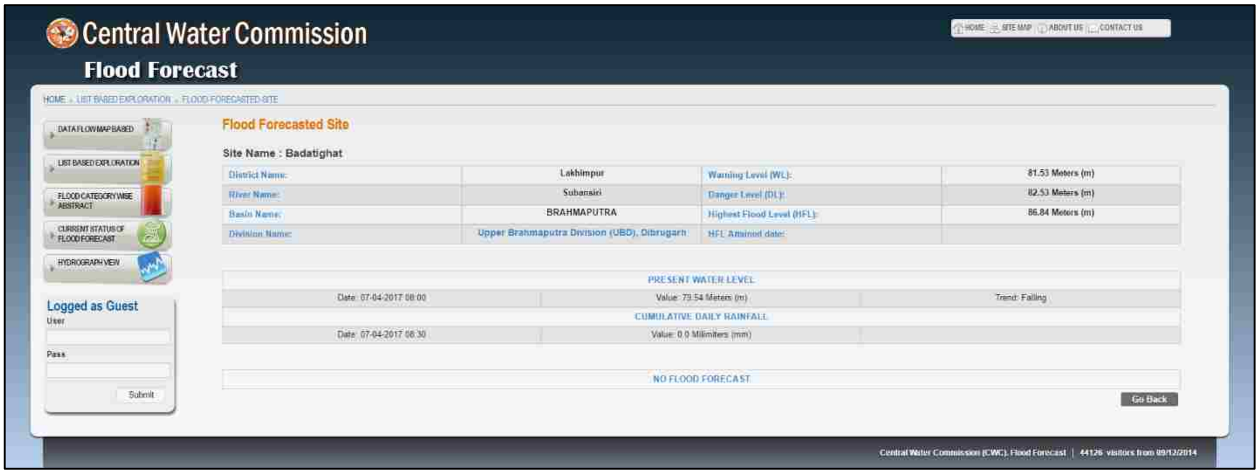
Figure A.17. Left: PDF document with in situ rainfall information shared in the Sikkim Meteorological Department (http://www.imdsikkim.gov.in/daily_Forecast.pdf) and Right: Sample C# code for parsing PDF Document to get rainfall from that website



(a)

```
///---------------------------------------------Central Water Commission(CWC) Data Download------------------------------------------/////
try
{
    Console.ForegroundColor = ConsoleColor.Cyan;
    Console.WriteLine("14th Website, Central Water Commission, India: http://www.cwc.gov.in/");
    Directory.CreateDirectory(di + @"\WebPage14");
    Console.ResetColor();
    string[] stationCode = File.ReadAllLines(iniDi + @"\NecessaryFiles\CWCStationID.txt");
    foreach (string element in stationCode)
    {
        try
        {
            WebRequest req = WebRequest.Create("http://www.india-water.gov.in/ffs/data-flow-list-based/flood-forecasted-site/");
            string postData = "lstStation=" + element;

            byte[] send = Encoding.Default.GetBytes(postData);
            req.Method = "POST";
            req.ContentType = "application/x-www-form-urlencoded";
            req.ContentLength = send.Length;

            Stream sout = req.GetRequestStream();
            sout.Write(send, 0, send.Length);
            sout.Flush();
            sout.Close();

            WebResponse res = req.GetResponse();
            StreamReader sr = new StreamReader(res.GetResponseStream());
            string returnvalue = sr.ReadToEnd();
            File.WriteAllText(di + @"\WebPage14\" + element + ".html", returnvalue);
            HtmlDocument doc = new HtmlDocument();
            doc.LoadHtml(returnvalue);
            HtmlNode header = doc.DocumentNode.SelectSingleNode("//h4");
            string stationName = header.InnerText.Replace("Site Name : ", "").Trim() + "-CWC";

            HtmlNodeCollection tables = doc.DocumentNode.SelectNodes("//table");
            HtmlNodeCollection rows = tables[1].SelectNodes(".//tr");
            HtmlNodeCollection colWL = rows[1].SelectNodes(".//td");
            HtmlNodeCollection colRF = rows[3].SelectNodes(".//td");
            if (colWL[0].InnerText.Trim() != "")
            {
                DateTime dateWL = DateTime.ParseExact(colWL[0].InnerText.Replace("Date: ", "").Trim(), "dd-MM-yyyy HH:mm", CultureInfo.InvariantCulture).AddHours(0.5);
                float valueWL = float.Parse(colWL[1].InnerText.Replace("Value: ", "").Replace(" Meters (m)", ""));
                //sb.AppendLine("WL-" + stationName + "," + dateWL.ToString("yyyy-MM-dd HH:mm:ss") + "," + valueWL);
                Console.WriteLine(dateWL.ToString("yyyy-MM-dd HH:mm:ss") + "," + stationName + "," + "WL-" + valueWL);
                waterLevelData.AppendLine(dateWL.ToString("yyyy-MM-dd HH:mm:ss") + "," + stationName + "," + valueWL);


            }
            if (colRF[0].InnerText.Trim() != "NOT AVAILABLE")
            {
                DateTime dateRF = DateTime.ParseExact(colRF[0].InnerText.Replace("Date: ", ""), "dd-MM-yyyy HH:mm", CultureInfo.InvariantCulture).AddHours(0.5);
                float valueRF = float.Parse(colRF[1].InnerText.Replace("Value: ", "").Replace(" Milimiters (mm)", ""));
                Console.WriteLine(dateRF.ToString("yyyy-MM-dd HH:mm:ss") + "," + stationName + "," + valueRF);
                rainfallData.AppendLine(dateRF.ToString("yyyy-MM-dd HH:mm:ss") + "," + stationName + "," + valueRF);
            }
        }
        catch (Exception error)
        {
            Console.ForegroundColor = ConsoleColor.Red;
            Console.WriteLine("Problem in one of the stations of 14th Website, Central Water Commission, India, Station: " + element + " data not available. Error: " + error.Message);
            Console.ResetColor();
        }
    }
    logText.AppendLine("14th Website, Central Water Commission, India: http://www.cwc.gov.in/, Downloaded Station: " + stationCode.Length);
}
catch (Exception error)
{
    Console.ForegroundColor = ConsoleColor.Red;
    Console.WriteLine("Problem in 14th Website, Central Water Commission, India: http://www.cwc.gov.in/ cannot be accessed, Error: " + error.Message);
    Console.ResetColor();
    logText.AppendLine("Problem in 14th Website, Central Water Commission, India: http://www.cwc.gov.in/ cannot be accessed, Error: " + error.Message);
}
```

(b)



(c)

```
<html class="gr__india-water_gov_in">
▶<head>…</head>
▼<body data-gr-c-s-loaded="true">
  ▶<script type="application/javascript">…</script>
  ▼<div id="cuerpo">
    ▼<div id="fondo">
      ▶<div id="cabecera">…</div>
      ▼<div id="contenedor">
        ▶<div id="migapan">…</div>
        ▶<div id="lateral">…</div>
        ▼<div id="contenido">
            <!-- include the boxes on the center side -->
          ▼<div id="centercontainer">
            ▼<div class="fondo-top">
                <h3>Flood Forecasted Site</h3>
              ▼<div class="editoruys">
                  <h4>Site Name : Badatighat</h4>
                ▶<table style="width: 100%">…</table>
                  <p> </p>
                ▼<table>
                  ▼<tbody>
                    ▶<tr>…</tr>
                    ▼<tr>
                        <td class style="width:33%; height:18px;">Date: 08-04-2017 08:00</td>
                        <td class style="width:33%;">Value: 79.34 Meters (m)</td>
                        <td class style="width:33%;">Trend: Falling</td> == $0
                      </tr>
                    ▶<tr>…</tr>
                    ▶<tr>…</tr>
                    </tbody>
                  </table>
                  <p> </p>
                ▶<table style="width: 100%">…</table>
                ▶<div class="botonera">…</div>
                </div>
              </div>
            </div>
            <!-- InstanceBeginEditable name="contenido" -->
          </div>
        </div>
      </div>
    ▶<div id="pie">…</div>
    </div>
  </body>
</html>
```

(d)

Figure A.18. Dynamic Webpage Posting Method example to get www.cwc.gov.in data; (a) Initial appearance of the webpage, (b) Sample code to process the webpage where the first part of code is used to post the ID of any station and the last part of code is used to process the captured output which is shown in (c). The background *HTML* code of (c) which is captured by *HtmlAgilityPack* is shown in (d)

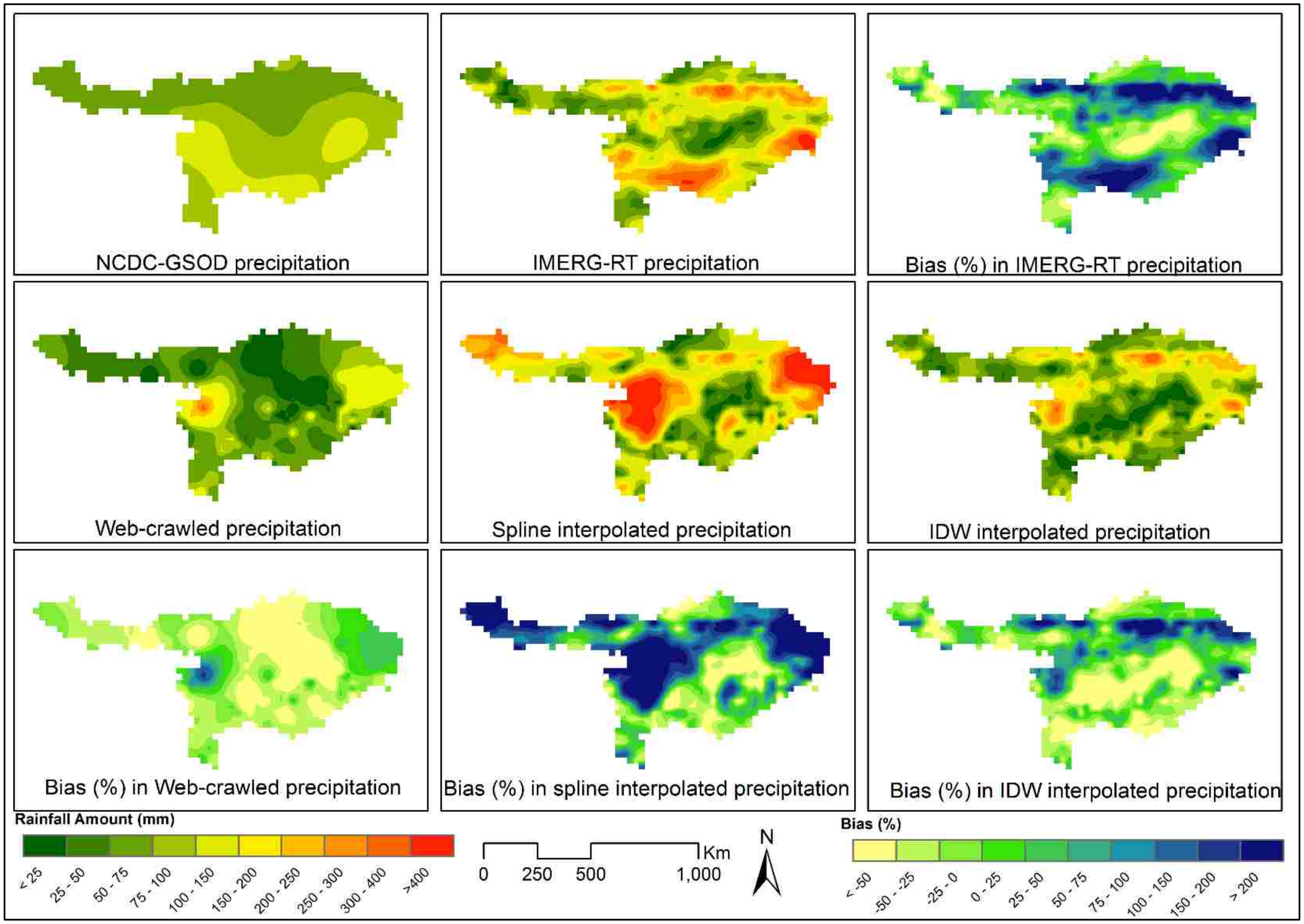


Figure A.19. Comparison of precipitation from IMERG, Web crawled, IDW bias interpolated and spline bias interpolated precipitation along with bias amount (in percentage with respect to NCDC-GSOD precipitation) in every product over Brahmaputra Basin for dry (January-March) season.

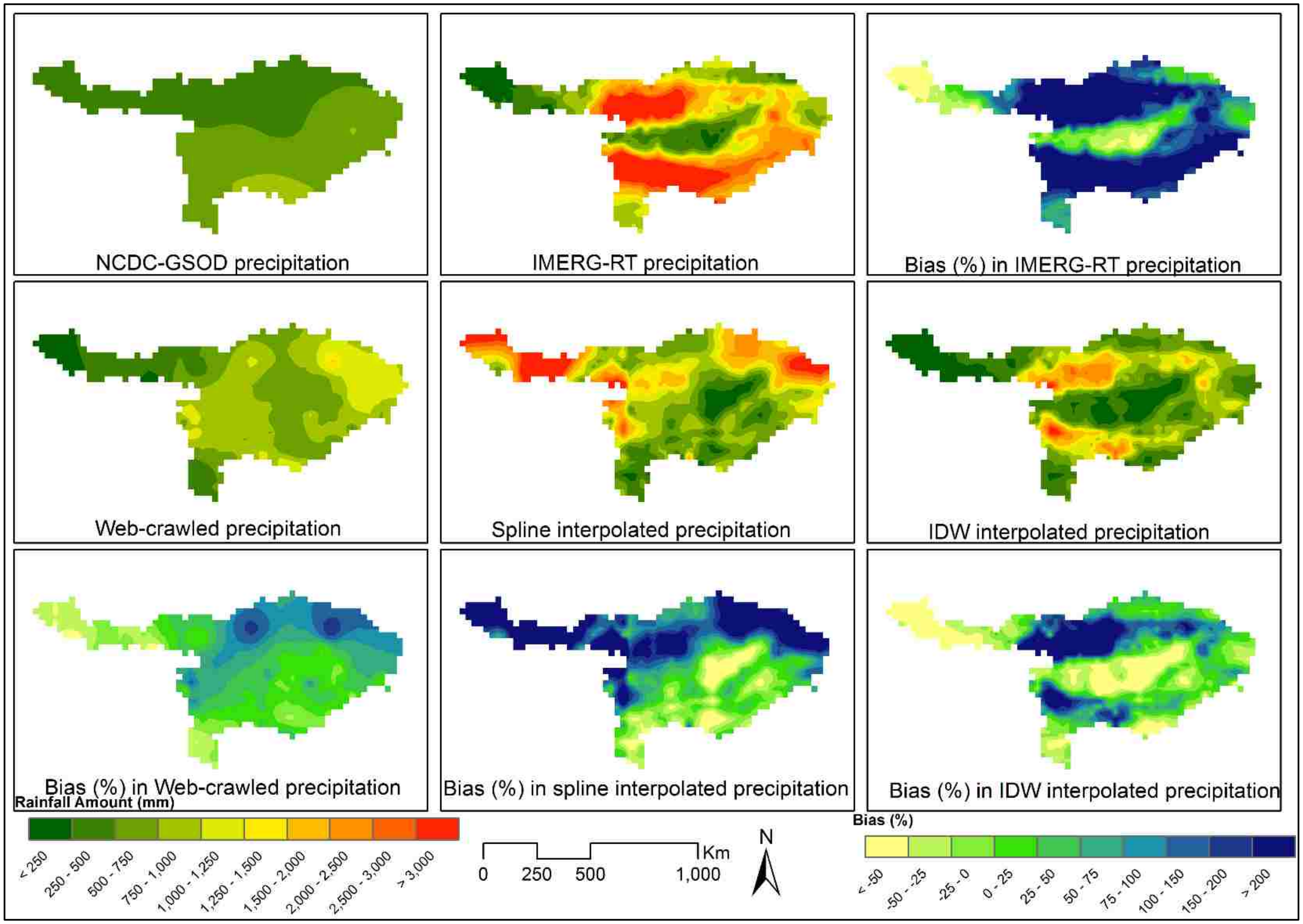


Figure A.20. Comparison of precipitation from IMERG, Web crawled, IDW bias interpolated and spline bias interpolated precipitation along with bias amount (in percentage with respect to NCDC-GSOD precipitation) in every product over Brahmaputra Basin for Pre-monsoon (April-June) season.

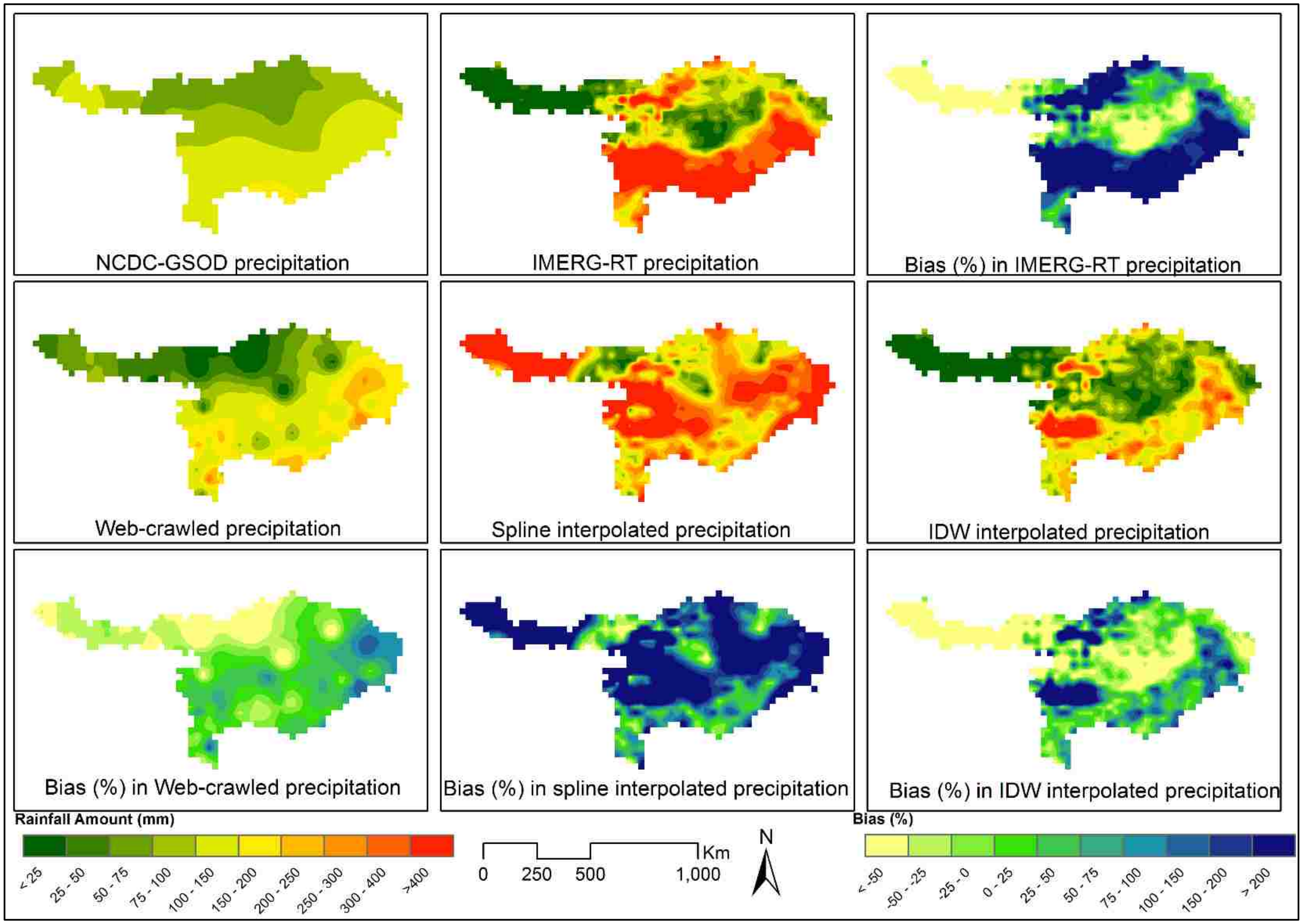


Figure A.21. Comparison of precipitation from IMERG, Web crawled, IDW bias interpolated and spline bias interpolated precipitation along with bias amount (in percentage with respect to NCDC-GSOD precipitation) in every product over Brahmaputra Basin for Post-monsoon (October - December) season.

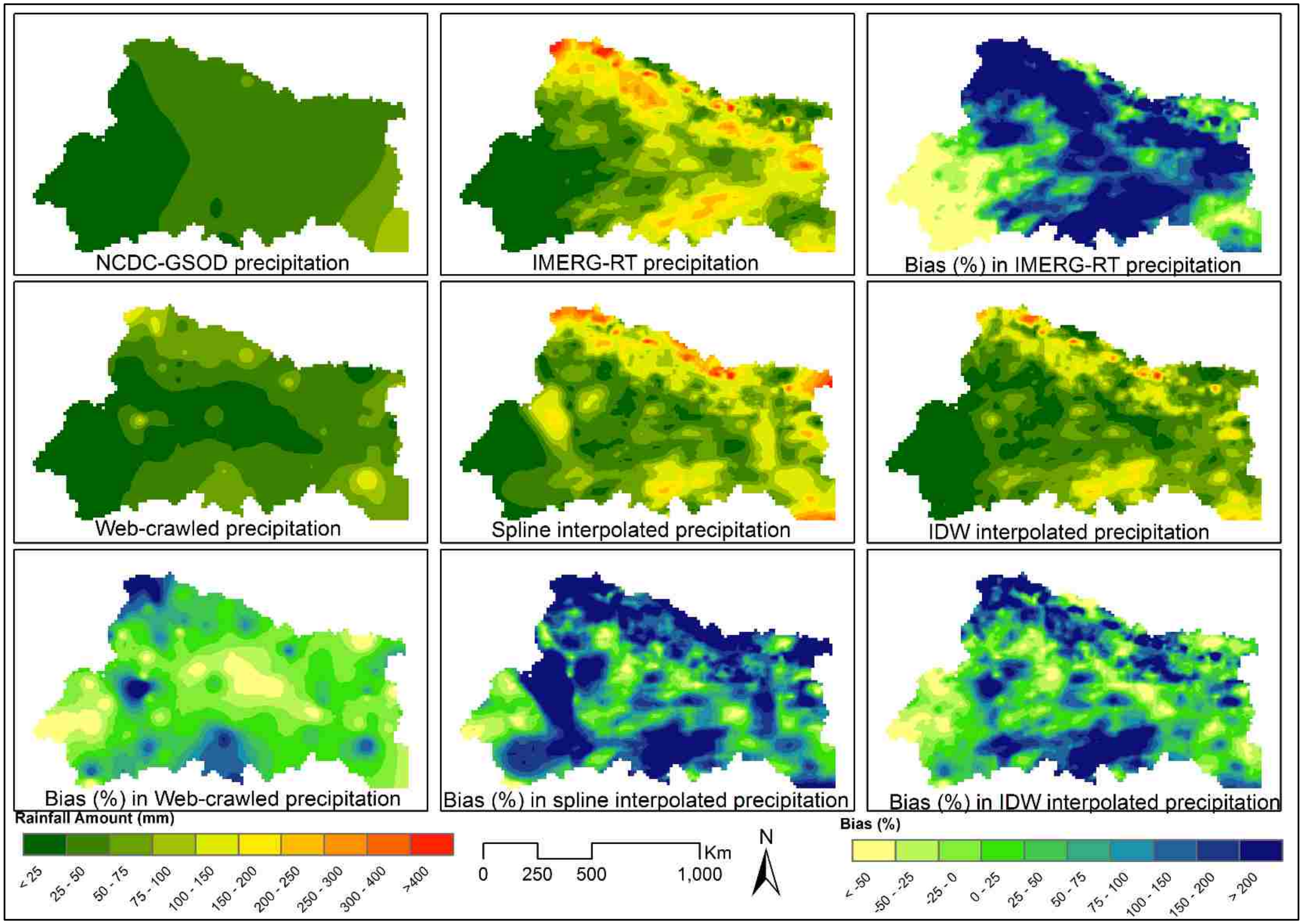


Figure A.22. Comparison of precipitation from IMERG, Web crawled, IDW bias interpolated and spline bias interpolated precipitation along with bias amount (in percentage with respect to NCDC-GSOD precipitation) in every product over Ganges Basin for dry (January-March) season.

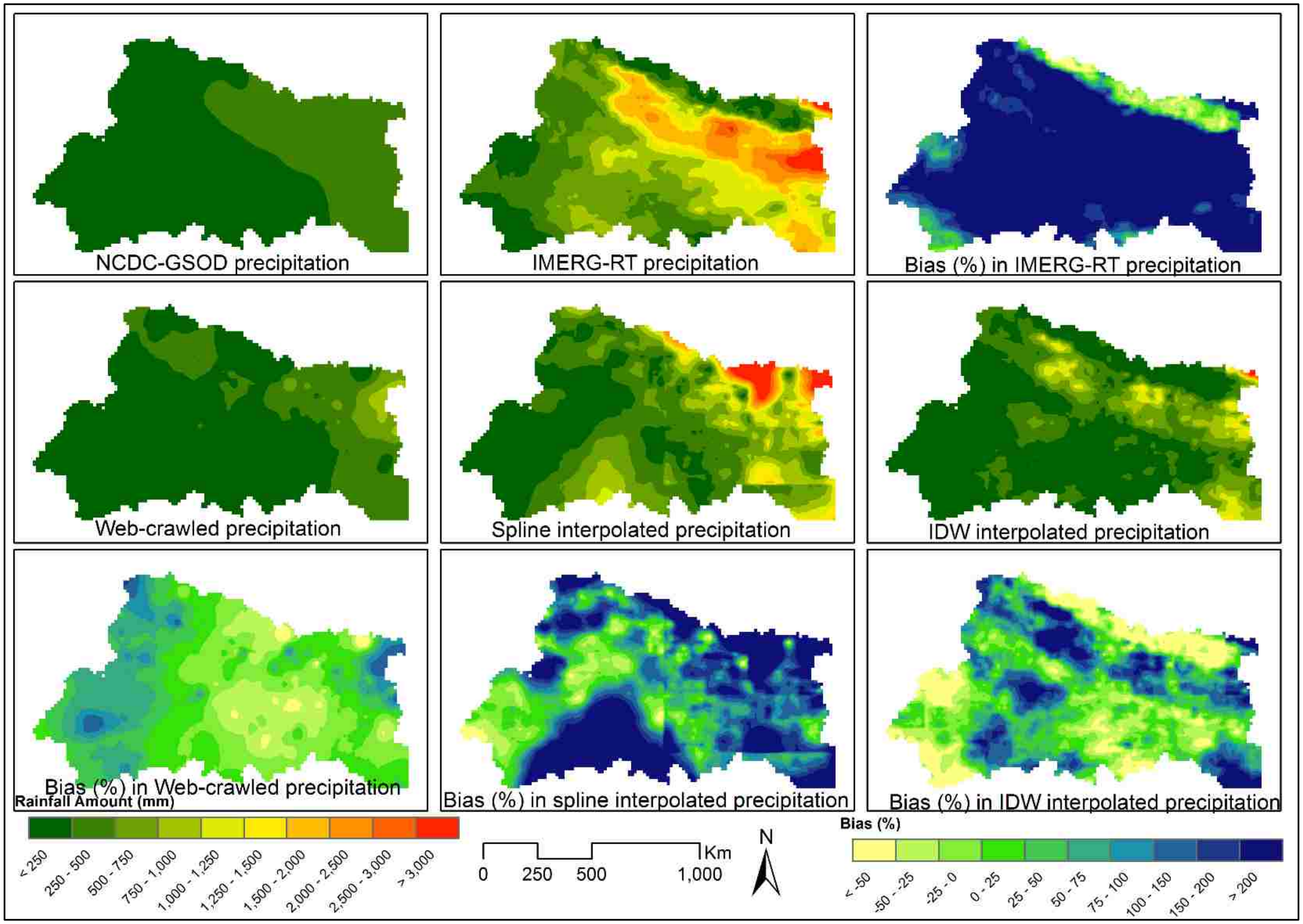


Figure A.23. Comparison of precipitation from IMERG, Web crawled, IDW bias interpolated and spline bias interpolated precipitation along with bias amount (in percentage with respect to NCDC-GSOD precipitation) in every product over Ganges Basin for Pre-monsoon (April - June) season.

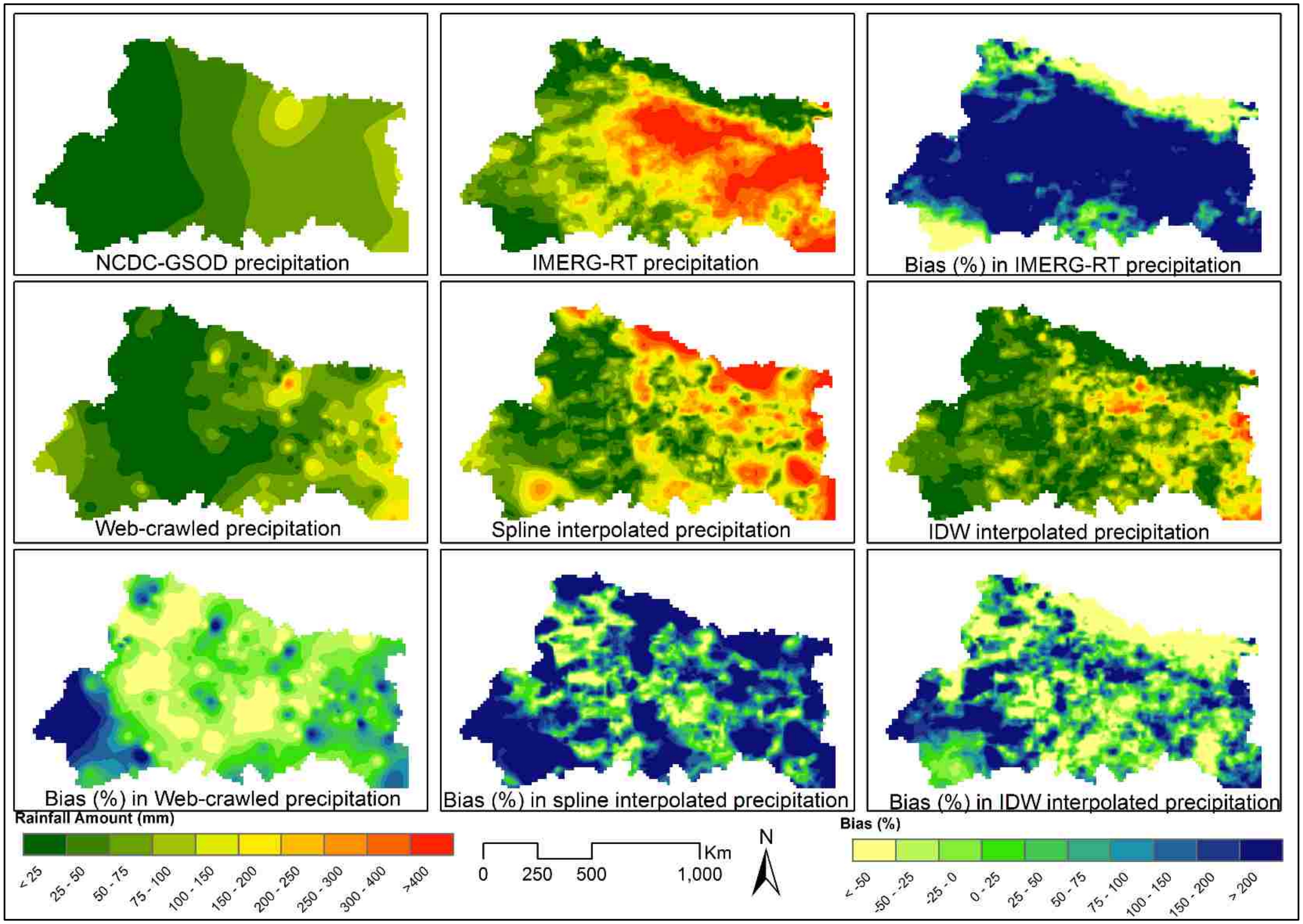


Figure A.24. Comparison of precipitation from IMERG, Web crawled, IDW bias interpolated and spline bias interpolated precipitation along with bias amount (in percentage with respect to NCDC-GSOD precipitation) in every product over Ganges Basin for Post-monsoon (October - December) season.