THE WORLD BANK
IBRD • IDA

TECHNICAL REPORT:
Evaluation of flood forecasting predictability

Development of a Flood Forecasting System for the Ganges and Brahmaputra Basins using Ensemble Weather Forecasts and Satellite-based Precipitation, River Widths and Heights

SAWI SOUTH ASIA WATER INITIATIVE

Contact: Thomas Hopson
Research scientist
National Center for Atmospheric Research
hopson@ucar.edu

NCAR
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

For the attention of:

Dr. Satya Priya, Program Leader,
South Asia Water Initiative
The World Bank
The Hindustan Times House (Press Block) 18-20
Kasturba Gandhi Marg New Delhi - 110001
Email: spriya1@worldbank.org

Company Information:

Name	National Center for Atmospheric Research
Registered Office Address	1850 Table Mesa Dr Boulder, CO 80305 USA
Corporate Office Address	University Corporation for Atmospheric Research P.O. Box 3000 Boulder, CO 80307-3000 USA

Acknowledgements

We are most thankful for the support and constructive suggestions provided throughout this project by Dr. Satya Priya, Senior Water Resource Management Specialist, World Bank, New Delhi, India.

NCAR is also thankful to Dr. Anju Gaur, Senior Water Resource Management Specialist, World Bank, New Delhi, India, and Dr. William Young, Program Director, World Bank, Washington DC, USA, for their support provided throughout this project.

This work was made possible through funding support provided by the **World Bank Group** under the **South Asia Water Initiative**, through the award:
Title: Development of Flood Forecasting for the Ganges and Brahmaputra Basins
Award No.: 7175247
Contract ID: 7175247

Table of Contents

List of Figures

List of Tables

Abbreviations Used

Abbreviation/Acronym	Expanded Form
App	Self-contained software application
BUET	Bangladesh University of Engineering and Technology
CFAB	Climate Forecast Applications for Bangladesh
CMA	China Meteorological Administration
CMC	Canada Meteorological Centre
CMORPH	CPC MORPHing technique
CPC	NOAA Climate Predition Center
CNES	Centre National d'Etudes Spatiales
CPTEC	Centro de Previsao de Tempo e Estudos Climaticos
CSV	Comma separated values
CWC	Indian Central Water Commission
DEM	Digital Elevation Map
DFO	Dartmouth Flood Observatory
ET	Evapo-transporation
ECMWF	European Centre for Medium-Range Weather Forecasts
Emissivity	Emission of electromagnetic radiation in the microwave-frequency
GBM	Ganges-Brahmaputra-Meghna basin
GDR-D	Geophysical Data Records, version D
GIS	Geographical Information System
GSMaP	Global Satellite Mapping of Precipitation
GTS	Global Telecommunication System
JASON II	Joint Altimetry Satellite Oceanography Network II
JAXA	Japan Aerospace Exploration Agency
JRC	European Union Joint Research Council
Leaflet	Open-source JavaScript library for interactive maps
MeteoFrance	French Meteorological Service
NASA	U.S. National Aeronautics and Space Administration
NSE	Nash-Sutcliffe Efficiency
NCEP	U.S. National Center for Environmental Prediction
NETCDF	Network common data format

NOAA	U.S. National Oceanic and Atmospheric Administration
NWP	Numerical weather prediction
OSTA	Office of Space and Terrestrial Applications
P	Precipitation
Plotly	Online analytics and data visualization tool
Q	River discharge
QC	Quality control
QR	Quantile regression
Q2Q	Quantile-to-quantile mapping
R	R statistical computing language
RMSE	Root mean square error
S	River stage height
SS	Skill score
SARAL	Satellite with ARgos and ALtika
Shiny	Web application framework for R
Stg	River stage height
SWOT	Surface Water and Ocean Topography Mission
THORPEX	The Observing system Research and Predictability Experiment
TIGGE	THORPEX Interactive Grand Global Ensemble
Time of Concentration	time needed for water to flow from the most remote point in a watershed to the watershed outlet
TRMM	Tropical Rainfall Measuring Mission
UKMET	U.K. Meteorological Office
UMD	University of Maryland
WRF	Weather Research and Forecasting model

Overview of Report

This project and report covers the development of a flood forecasting system for the Ganges and Brahmaputra Rivers, along with an evaluation of the sources of forecasting predictability. It should be noted that many of the findings also have implictions for the surrounding South Asia region. The aspects covered here include issues related to input data availability, quality, and utility; issues related to the engineering and technical feasibility of the development of forecasting systems for the region; and general issues related to the planning and optimal resource allocation for the implemention of such systems, and planning for future technical developments. All of these topics are covered under the objectives of this consultancy, which provide the organizational structure for this report. These objectives are:

Objective 1

Assess the skill and estimate the errors in some of the foundational data sets available for use in flood forecasting for India, focusing on remotely-sensed products useful for catchments with limited on-the-ground monitoring

Objective 2

Integrating these data sets into a common hydrologic forecasting framework, showing how this can be achieved from an "engineering perspective", but also provide operational forecasts with potential for societal benefit

Objective 3

Provide effective displays for these products, including maps of areas of inundation corresponding to forecasted discharge that have potential benefits for "on the ground" operations and decision-making

Objective 4

Utilizing estimates of the errors in the hydrologic forecasting framework and input data sets, determine the overall predictability of the system to forecast flooding events of differing levels of severity in the Ganges and Brahmaputra watersheds at different locations and forecast lead-times; from this assessment, recommendations can be made on where investments should be focused to increase flood forecasting capacity throughout the basins to meet World Bank development goals

In the context of these objectives, the structure of this report is as follows: we first present an Executive Summary covering the key findings of this project, covering data and forecast products and displays delivered that are useful for the implementation of effective flood forecasting warning systems over the Indian subcontinent, along with recommendations for future efforts. Within the main body

of this report, we first provide a forecasting system overview to build context for materials presented in the rest of the report. This is followed by a review and evaluation of our data sources useful for flood forecasting. These products include satellite-derived rainfall estimates, ensemble weather forecasts from a range of global forecasting centers, in situ river stage measurements and their quality control, satellite emissivity signals correlated to changes in river width (or simply, "river width measurements"), and satellite altimetry measurements detecting relative changes in river height. We also provide visualizations and links to access many of these data in real-time, with many of these quality controlled for user ease. We then discuss forecasting methodologies we've utilized and introduced in this project, along with a discussion on the validation protocol we utilize for establishing operational reliability and limiting conditions on the forecast utility and skill. This is followed by a discussion on the skill of the forecasting products themselves, both catchment-integrated precipitation and river discharge forecasts. We then conclude the main body of this report by presenting the visualization and information sharing tools we've developed to communicate our operational forecasts. In the appendices, we cover many of these topics in more technical detail, along with an "app" we're developing to provide guidance in the placement of rain gauges (developed as separate from the main direction of this project).

Executive Summary

Global impacts of river floods are substantial and rising, and the Indian subcontinent is no exception due to population pressures and projected increases in the South Asian monsoonal average strength, variance, 5-day seasonal maximum, and duration (IPCC 2013). However, studies have shown that future increases in flood risk can be largely contained using effective disaster risk reduction strategies, such as through the implementation of advanced lead-time warning systems (Hopson and Webster 2010, Webster et al. 2010, Jongman et al. 2015). In the context of the particular advanced lead-time warning system we developed for locations within the Ganges and Brahmaputra River Basins as part of this project, in this Executive Summary we highlight the key findings of our investigation into the data, forecasting approaches and products, and displays that can communicate forecast information to users. Although these findings are useful in our particular context, we also feel many of these findings have general applicabitlity to the Indian Subcontinent and greater region as well. In addition to discussing the data and forecast products in this report, we also provide visualizations and links to access many of these data in near-real-time, along with recommendations for future efforts.

Key Findings

Forecasting System Overview

At its original core, the forecasting system implemented for this project is based on the CFAB forecasting system developed for Bangladesh (Hopson and Webster, 2010; Webster et al 2010). However, because a number of new data sets were introduced into this project in many cases requiring the introduction of new forecasting technologies, and because many of these data sets are intermittent, and because one of the objectives of this project was to assess where forecasting skill was derived from (or limited by), the CFAB model system was restructured such that multiple forecasting methods could be operated in parallel (and combined into a "multi-model" at later stages of the system). A summary of the data sources used in the forecasting system is as follows: in situ river stage measurements; gridded satellite and raingauge estimates from 4 sources, ensemble weather forecasts from 8 weather centers; upstream river stage measurements, upstream river "width" esimates, and upstream altimetry measurements.

Satellite altimetry measurements

Satellite altimetry measurements have been shown to detect changes in ocean height with 2cm precision; but at what precision can altimetry measure river heights, especially over progressively smaller river channels (to the obvious benefit of detection of flood waves emanating higher up in the river basin, providing forecasts of downstream river flows at longer lead-times)? Our analysis showed measurement precision of the JASON II instrument on the order of 10-20cm, even

over channels only 100s of meters in width. (Data from the SARAL mission has shown significantly less skill, to the point of limited usefulness over river channels, possibly due to drift about its preset orbital path). However, in practical terms, the temporal repeat period over the same location of 10-days (exclusive to the JASON II orbit), and gaps between river sampling locations means there is a significant chance many potential flood peaks would be inadequately captured, arguing that the instrument alone would not be sufficient in and of itself to provide complete flood wave monitoring system; but showing significant utility when used in conjunction with other data sources. However, the new Surface Water and Ocean Topography mission (SWOT) high-denisty sampling mission (launched around 2020) is to provide complete global spatial-coverage at least twice every 21 days, with precision expected to be two times better than current instruments in orbit. (Further information and additional time-series and error assessments found at http://ral.ucar.edu/~hopson/WorldBank/satellite_error/Altimetry/.)

Satellite-based microwave emissivity measurements

Satellite-based microwave emissivity measurements can detect changes in river channel widths that have utility for flood forecasting. These measurement provide global coverage roughly once every two days, with less than a one day lag time between satellite pass and data availability. However, useful measurements are not consistently found from all sites we examined through comparison with nearest river gage locations (ranging from .9 correlations, to 0). Quality control of the data is also important, with the signal significantly "noiser" than the (more infrequent) satellite altimetry. In general, river "width" data appears to have larger errors than satellite altimetry. However, spatial and temporal sampling frequency is far superior (once per day, with an along-river measurement roughly every 50km), but care must be taken to ensure usabilty. (Further information below and additional time-series and error assessments found at http://ral.ucar.edu/~hopson/WorldBank/satellite_error/Width/.)

Precipitation Estimates and Forecasts Analysis

To derive our best estimate of "observed precipitation", in this project we relied on a near-realtime straight average of rain gauge and satellite precipitation estimates from JAXA, NASA, and NOAA. We found the following in assessing their relative skill amongst themselves. The rain gauge and satellite precipitation products are mostly similar in terms of their seasonal cycles. The largest discrepancies (in the Meghna and Middle Brahmaputra) occur in regions that have limited gauge coverage. Further analysis of the Meghna suggests that the gauge data has approximately 3x as many days with zero or very low precipitation, possibly suggesting that rainfall events are being missed to lack of coverage.

The skill of river catchment-integrated rainfall forecasts over the Ganges and Brahmaputra basins varies appreciably by weather forecasting center and by

region. Past results of older versions of these models have shown similar results (during 2008-2011; Ranade et al. 2014). We examined forecasts from what are typically considered four of the best global weather forecasting centers (ECMWF, NCEP, UKMET, and Canada Met Centre [CMC]) during the 2014-2015 monsoon seasons. In general, each center captures the general spatial and temporal distributions of the Indian monsoon rainfall. However, significant biases exist in differing regions of the basins, stressing the importance of calibrating these products before their use in flood forecasting models. The variation in the models' spatial representation of climatological rainfall indicating differing interactions between rainfall and topographic features. Models with the smallest biases are not necessarily the ones with the best skill. The relationship between forecast skill and basin size is apparent in these data, though other geographic factors are clearly important as well. These varying biases and skill (varying by center and by region) also highlight the potential benefits of combining the forecasts from these centers to offset their individual limitations (Mitra et al. 2011, Hamill 2012). For operational considerations, ECMWF's model spatial resolution is roughly 25km (deterministic member), with NCEP, CMC, and UKMET roughly 50km. Likewise, ECMWF, UKMET, and CMC provide new initialized forecasts twice daily (00Z and 12Z), while NCEP provide new forecasts four-times daily, with all four centers providing forecast outputs at 6hr intervals going out to around 15day lead-times. At these spatial and temporal resolutions, it is important to point out that rainfall events of much smaller spatial or temporal scale than these resolutions will not be captured. In terms of ensemble members, ECMWF provides 50, while the other centers produce roughly ½ that number each.

Discharge Forecast Findings

Discharge forecasts derived solely from altimetry and upstream stage both appear to have some skill in forecasting downstream, though the skill is variable. This could be due to their relative locations, the size of the river between the two locations, or a upper predictability bound we can't currently cross with our single fit rating curves. Data is limited for the stage so the relationships may not be robust with 1 year (or less) of data. Forecasts derived similar to the CFAB system are found to be more robust; in part, because the depend on a variety of data inputs whose errors can mitigate each other. Forecast skills for the CFAB system vary from maintaining utility out to 3 days only in some cases of smaller river catchments, to out to 16days for the larger river catchments, this analysis based on forecasts of discharge in the upper 75%.

Displays

The interactive web display that was developed expands Geographic Information Systems (GIS) mapping capabilities in a simple but effective way to communicate forecast precipitation accumulation averages for catchment and sub-catchment basins. This application uses state-of-the-art technologies to visualize the spatial and temporal aspects to weather forecast. Data Ingest and Web Services have been

developed to support the current web display of flood risk showing river discharge observations and forecasts. Improved data standards and services could be foundational to better forecasting and sharing. Complex spatial and temporal data query tools will benefit modeling and analysis. Decision support systems can greatly improve effectiveness of operators. Mobile access is necessary for flexibility and convenience. Discharge forecasts can be effectively coupled to imagery data bases to show probability of regions being inundated at differing return levels.

Key Forecasts, Data Products, and Visualizations Delivered

Forecasting Sites

Below in the next three figures, we provide maps of subcatchments we are categorizing our data under, along with the discharge forecasting sites we are currently focusing on in this report for reference. The selection of these points is discussed in more detail further below in this report. All of our gage locations (in situ river stage, satellite river "width" signal sites, and satellite altimetry sites used in this consultancy are shown in Figures 2 and 3 below. In addition, higher resolution versions of both of these two figures, along with the upstream area above each of these sites, can be downloaded at http://ral.ucar.edu/~hopson/WorldBank/.

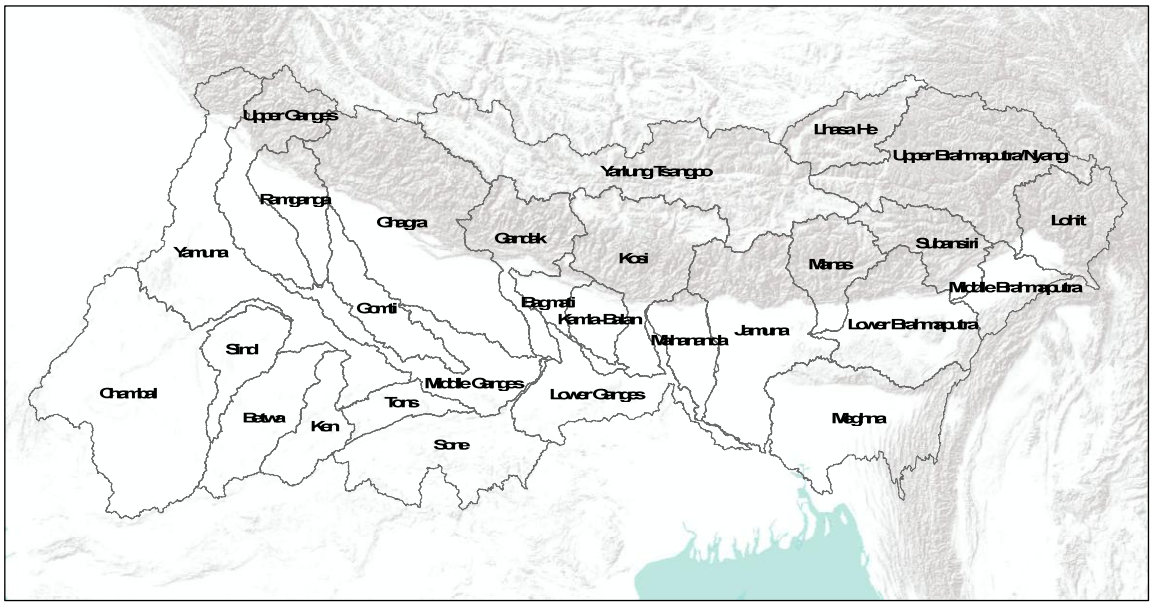


Figure 1: the subcatchment names that are used in our reporting to segregate our results.

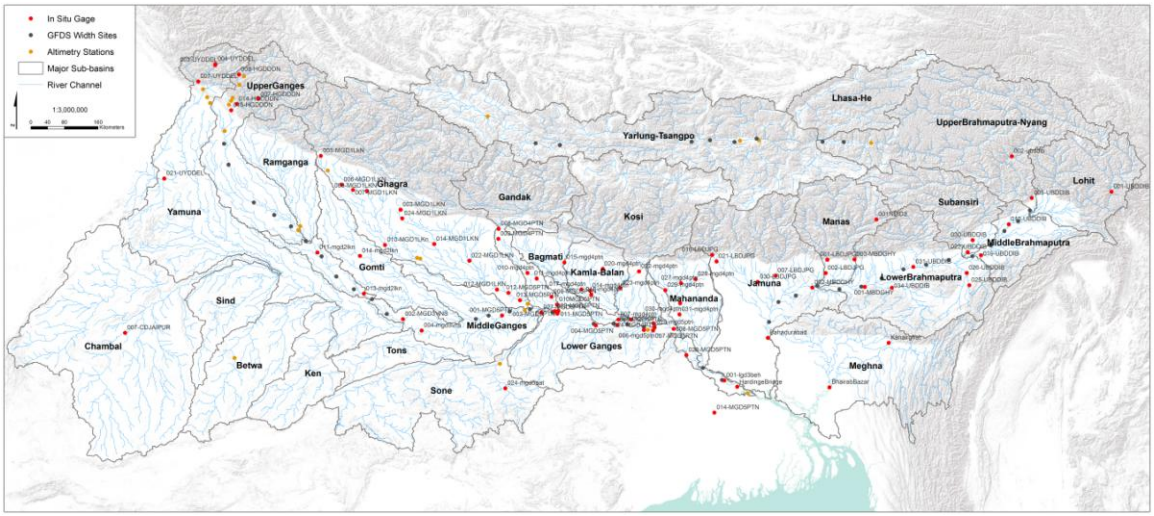


Figure 2: all gage locations (CWC in-situ, altimetry, and river "width" locations) are also shown (with only CWC in-situ gage location names provided).

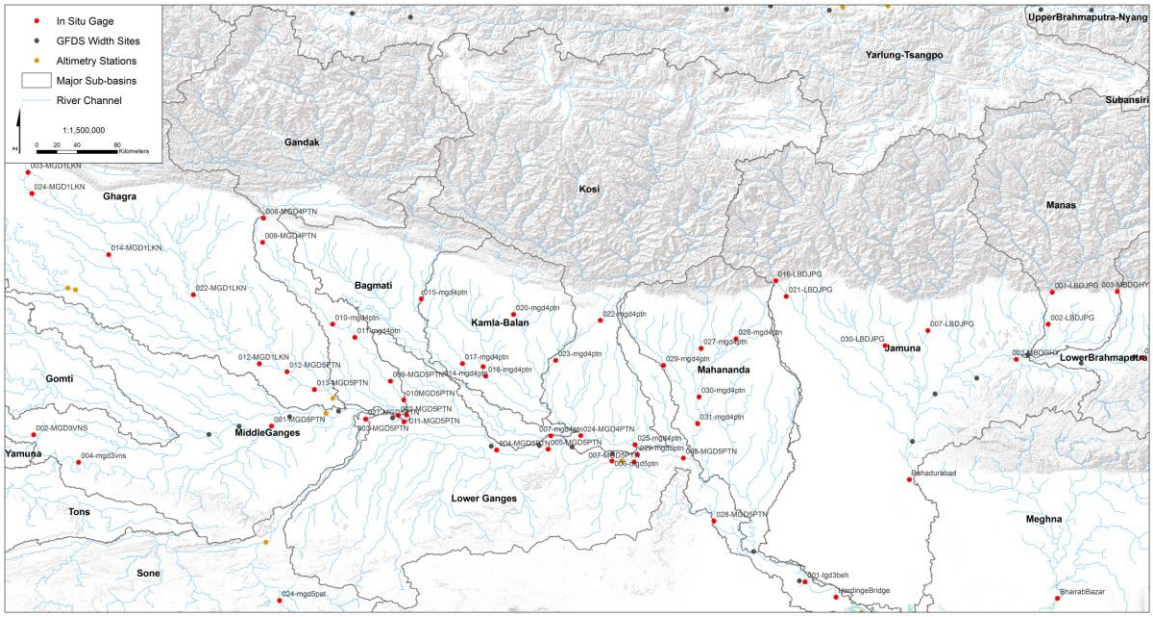


Figure 3: a "zoomed-in" version of Figure 1, with easier-to-read in-situ CWC gage names visible.

A foundational driver for flooding is precipitation. In fact, for many smaller spatial- and quicker time-scale flood events, the severity of the event is often strongly controlled by the severity of the local rainfall. As such, guidance can be gained by visualizing the rainfall itself over the monitored flood forecasting region. In this report, we provide an example of an operational web display of weather forecast and satellite precipitation, where current and forecasted rainfall conditions can be viewed over both large sub-catchments and zoomed in smaller catchments of the

Ganges and Brahmaputra. These displays show both 24hr and 5day averaged rainfall amounts from satellites (NOAA, JAXA, and NASA) and forecasts from 1- to 15-day forecast lead-times for 4 weather forecast centers (ECMWF, NCEP, UKMET, and Canada Met Centre). (See http://indiawbg.rap.ucar.edu/precip/.)

This report also discusses and presents the results of the derivation of approximate rating curves curves (estimates of river flow from stage measurements) where there are no discharge measurement, using discharge measurements from other locations in the river network (upstream or downstream). We applied our approach to 253 in situ Indian stage gaging sites that we are developing operational river discharge forecasts for as part of this initiative. From a flood forecasting standpoint, the generation of rating curves to provide appproximate values where there are none serves a number of purposes. For one, these curves estimate the carrying capacity of the river channels and sensitivity of inundation of a given location to a unit change in river flow. Secondly, they estimate the rivers' potential degree of impact on flows in lower river reaches downstream, and help to determine the monitoring prioritization for these upstream rivers. Third, even simple traditional flood forecasting techniques such as gage-to-gage correlation are improved if the forecasting of the advected state variable is river flow as opposed to the (non-state) variable of river height (given the nonlinear relationship between the two). Fourth, we present a twist on the traditional rating curve relationship to also derive a "rating curve" for the relationship of upstream stage to downstream discharge (as an additional function of time lag). These relationships can be directly used for the forecasting problem, with the optimal lags indicating flood wave travel times, and the "scatter" (error) about the rating curve fit also providing a metric for the degree of predictability the relationship holds, as well as a weighting metric for combining with other upstream gages. (Results shared at http://indiawbg.rap.ucar.edu/Ratefit/).

This report also presents an automated quality control method implemented for river stage, discharge, and satellite-derived "river width" time series data. The algorithm is provided so that this approach could also be implemented by others for similar (or different) time-series data. Real-time updated quality control and flagging of these data can be viewed and downloaded from web sites (http://ral.ucar.edu/~hopson/WorldBank/india_stageQC/, http://ral.ucar.edu/~hopson/WorldBank/india_riverQ/, http://ral.ucar.edu/~hopson/WorldBank/india_widthQC/), along with FTP sites. (provided further below).

For viewing and accessing many of the products discussed above, we have compiled a project website for ease of access: http://gis.ucar.edu/india-world-bank-flood-forecasting. This site provides a overview of the displays and data provided through this project, including three of our project outputs: a) daily-updated river stage

readings and automated quality control of the same, including displays; b) displays of current and forecasted catchment-integrated rainfall conditions over subcatchments of the Ganges-Brahmaputra catchments, including data fields; c) displays of operational river discharge forecasts, d) rating curve fitting parameters for all of our stage gaging sites, derived discharge values, and their display via a separate interactive online application; and e) quality-controlled data and an evaluation of river "width" measurement along the mainstems of the Ganges and the Brahmaputra rivers.

We also refer to our FTP sites below. These data are archived under folders of subcatchments of the Ganges and Brahmaputra, shown above in Figure 1 for reference:
a) at ftp://ftp.rap.ucar.edu/incoming/irap/india_stageQC/ we are providing hourly-updated quality-controlled river stage data available within subcatchments of the the Ganges and Brahmaputra catchments;
b) at ftp://ftp.rap.ucar.edu/incoming/irap/india_riverQ/ we are providing hourly-updated quality-controlled river discharge measurements using the quality-controled stage measusmernts and our derived rating curves;
c) at ftp://ftp.rap.ucar.edu/incoming/irap/tigge/ we are providing operational daily ensemble gridded precipitation (and other met variables) forecasts (going out to 16 day lead-times for eight weather centers: CMA [China], CMC [Canada], CPTEC [Brazil], ECMWF [European Union], MeteoFrance, NCEP [USA], and UKMO). These products will be discussed further in the sections below.

Finally, we provide an overview of an approach we developed for assisting in the optimal siting of rain gauges for under-sampled regions of India. The approach itself requires as inputs the number of gages to be sited, locations of the current in situ gage network, an historical record from satellite estimates, along with geospatial data of roads and other features near which the gages would need to be accessible, to estimate both the optimal locations, along with estimates of the error in using this network to estimate spatially-integrated rainfall.

Recommendations for Future Efforts

Presentations

Presentations discussing many of these findings can also be downloaded from http://ral.ucar.edu/~hopson/WorldBank/FinalReport/.

Rain Gage Siting App

As additional tasking added to this consultancy, we are developing a GIS-based rain gage siting tool to assist in the optimal location of rain gages for use by India's states in their expansion of their rain gage network under the National Hydrology Project.

To adequately capture the local spatial and temporal weather phenomena, an observation network has to be well designed. The design of the observation network requires evaluation of existing observations to quantify the variability of weather conditions, which inform the spatial and temporal monitoring requirements. The network design also depends on available resources (number of available instruments), land-surface/land-use conditions (rough/flat terrain, woodland/farmland), available infrastructure (roads, secure locations, communication networks), and location of population centers, rivers, or other critical monitoring regions.

A GIS framework is well suited to optimally design the network. A GIS network design tool will ingest all relevant and available geographical information, and objectively locate possible weather sites based on spatial design requirements. See Appendix C for further information on the ongoing development of this tool.

Forecasting System Overview

Key Message

At its original core, the forecasting system implemented for this project is based on the CFAB forecasting system developed for Bangladesh. However, because a number of new data sets were introduced into this project in many cases requiring the introduction of new forecasting technologies, and because many of these data sets are intermittent, and because one of the objectives of this project was to assess where forecasting skill was derived from (or limited by), the CFAB model system was restructured such that multiple forecasting methods could be operated in parallel (and combined into a "multi-model" at later stages of the system). A summary of the data sources used in the forecasting system is as follows: in situ river stage measurements; gridded satellite and raingauge estimates from 4 sources, ensemble weather forecasts from 8 weather centers; upstream river stage measurements, upstream river "width" esimates, and upstream altimetry measurements.

Introduction

In this subsection, and shown in the figure below we discuss the overall forecasting structure used in this project and the subsequent forecasting skill analysis. Shown in this figure below is the flow chart of the components used in this project. This flow chart is similar to that used in the CFAB project (Hopson and Webster 2010; Webster et al 2010), with ensemble weather forecasts, satellite precipitation estimates, and river flow measurements used as "I. Initial Data Input" at the top. However, in this application, we introduce forecasts from 8 centers (TIGGE forecasts, now going out to 16-day lead-times -- previously 10days for the CFAB system), an additional satellite precipitation estimate (JAXA), and new upstream river flow information provided by Central Water Commision upstream gages, DFO-JRC upstream river widths, and altimetry measurements. These new upstream data sets required development of new algorithms (shown schematically in the "Stage-Discharge Relationship" box under "III. Hydrologic Modeling", and discussed below in a lower section). As a result, we are producing independent ensemble river discharge forecasts based on these different approaches. One of our primary goals in this project was to assess the contributions to river flow predictability of these different input data sets (and resultant modeling approaches). We then evaluate and compare our different forecasting inputs and algorithms in what follows later in this report.

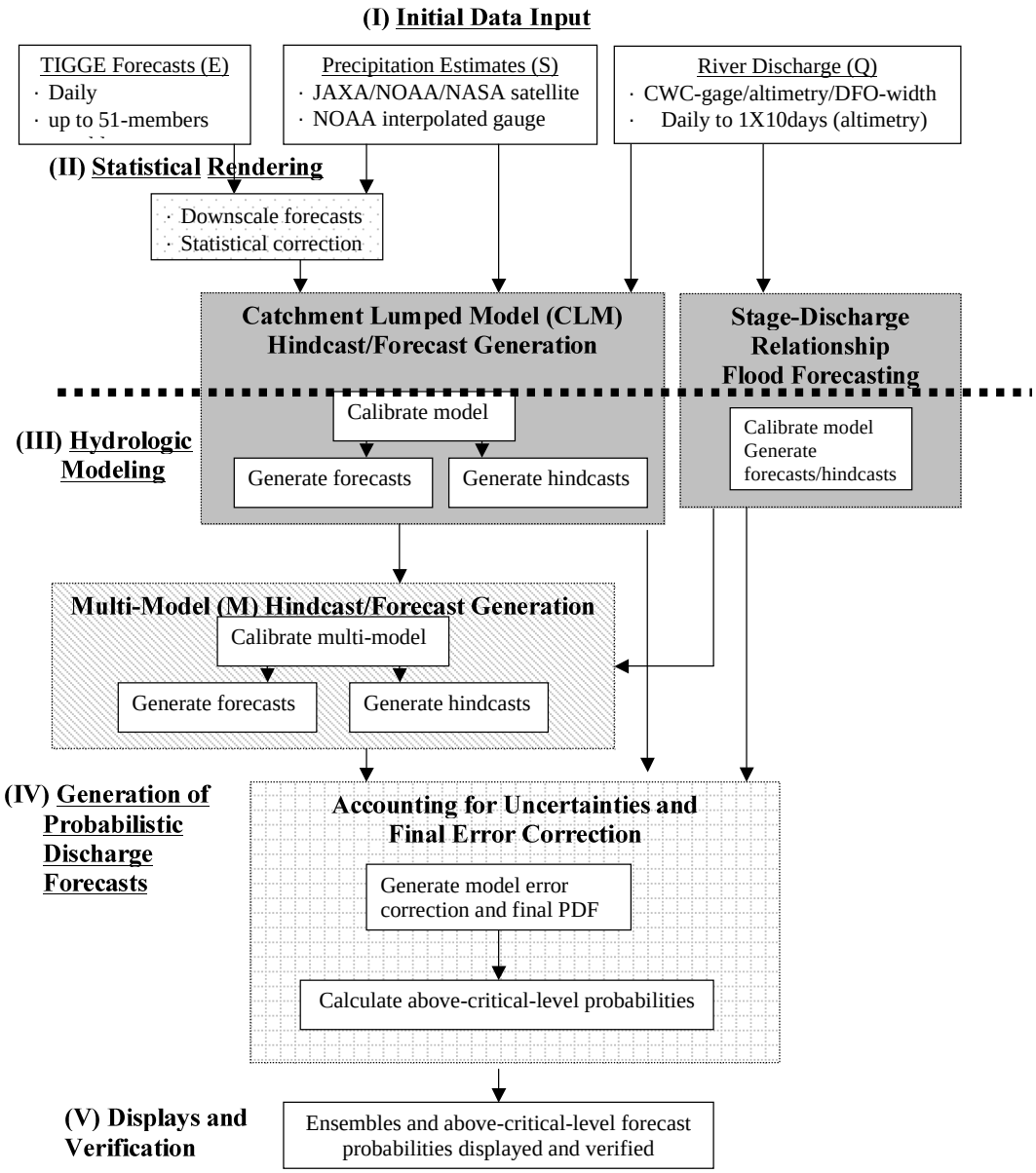


Figure 4: Flowchart of the modified CFAB short-term forecasting scheme. Five steps (I-V) are outlined schematically and within each is a brief description of the data used, tasks performed and etc. The arrows show the path of the data through the system. The components of Step I (clear boxes) represent daily inputs into the scheme. At Step II (dotted box), the TIGGE ensemble forecasts are corrected statistically to reduce systematic error. Step III (hatched boxes) represents the hydrologic modeling

process, where all inputs are integrated to produce river flow forecasts at our 87 forecasting points. Step IV (checkered box) represents the process by which all hydrologic uncertainties are accounted for in the probabilistic forecasting process. In Step V, the ensemble forecasts are tailored to produce probabilistic information necessary for displays and user needs. Note that all processes below the horizontal dashed line are done independently for each day and each 1- to 16-day forecast lead-time.

Data Components and Evaluation

In this section we discuss the data inputs used in the river discharge forecasting process. We examine the skill of the merged satellite precipitation product combining NASA TRMM, NOAA CMORPH, and JAXA GSMaP through comparisons to rain gauge values, and discuss the implications for our earlier findings on the TIGGE ensemble NWP skill reported in our previous report. We also discuss briefly an additional improvement required in our river stage quality control procedure, and the set of river gauges we are using in our final comparisons, based on findings in our rating curve fitting procedure.

River Stage Data: current data management

Real-time river and reservoir stage data are currently collected from the Flood Forecast Monitoring (FFM) Directorate of the Central Water Commission (CWC). Data for each station of interest are downloaded every hour from the web services using the site's custom protocol. The stage level, precipitation, and trend are stored in a geospatial relational database for easy access by scientific and display processes. Those records are retrieved for use in initializing the forecast model, generating forecast plots, and populating the web display. While data are currently only retrieved for stations in the Ganges and Brahmaputra basins, the process could easily be scaled to include all stations around the country. Current records extend from June 2015 to the present.

Several services have been created to retrieve subsets of the data in the database and return that data in a specific format. The web display's service is of particular utility because it returns the data prepared in a standard GeoJSON format. Web services such as this are the fundamental building blocks of data interoperability. While this particular service does not support the full Web Feature Service (WFS) specification, adding such capability would immediately allow it to be used with a wide array of WFS clients: model initialization, data visualization, and system monitors.

Delineated catchment-scale distribution

One of the primary benefits of the stage data towards meeting the goals of this consultancy is that each location can provide a test of the utility of different inputs (e.g. upstream gages, satellite precipitation, weather forecasts, etc.) to improving forecast predictability. And we expect this predictability to be a function of upstream catchment area. As such, we have delineated the catchment area above each of our gages (a necessary step for the application of the CFAB model to be independently employed at each location), and it is informative to see what is the

distribution of these areas, as shown in Figure 5 below (in log scale), with maps of these upstream areas also shown in Figure 2 above. Note the dearth of areas falling between 1-30km^2. We expect that one reason for this gap in available catchment areas is potential mis-siting of the gages onto our delineated stream network. Although much time was spent to produce a network as accurate as possible, we recognize there are still deficiencies in the co-locating of this network, and as a result, many of these smaller catchment gage locations (i.e. below 1km^2) will be excluded from our forecast predictability results.

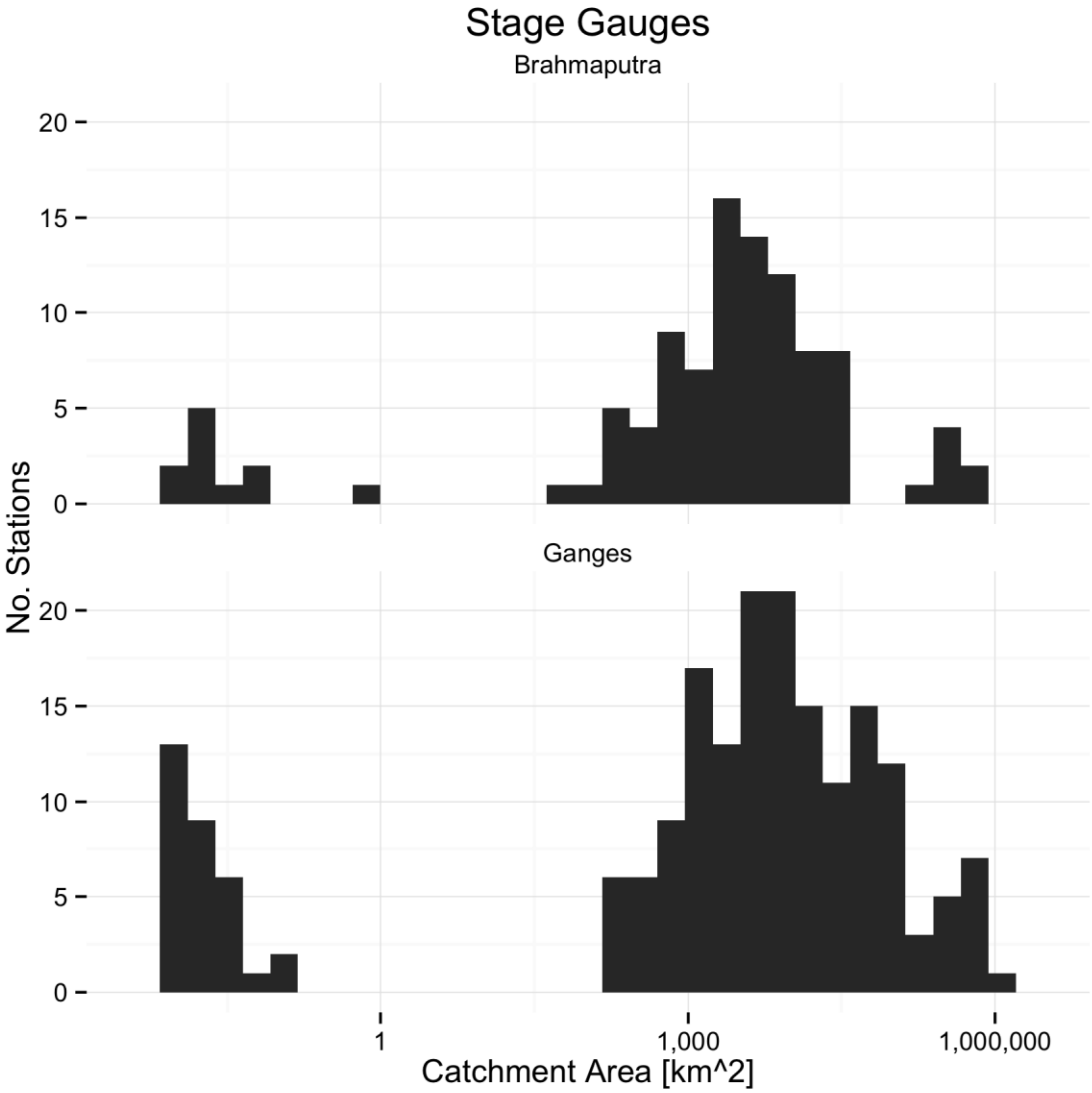


Figure 5: histograms of the catchment areas above each of the stage gaging sites used in this consultancy. Brahmaputra gaged subcatchment counts at top, Ganges below.

Stream Gage Data: Integrating into the forecasting process

In many parts of the globe, one may have access or approximate river height (stage) measurements, but without access to the corresponding river flow. In this section we discuss and present the results of the derivation of approximate rating curves curves (estimates of river flow from stage measurements) where there are no discharge measurement, using discharge measurements from other locations in the river network (upstream or downstream). We applied our approach to the 253 in situ Indian stage gaging sites (that we are providing operational river discharge forecasts for) as part of this project. Later in the section, we also present a twist on the traditional rating curve relationship to also derive a "rating curve" for the relationship of upstream stage to downstream discharge (as an additional function of time lag). These relationships can be directly used for the forecasting problem, with the optimal lag anticipating flood wave travel times, and the "scatter" (error) about the rating curve fit also provides a metric for the degree of predictability the relationship holds, as well as a weighting metric for combining with other upstream gages.

This section begins by presenting our automated quality control method implemented for river stage, discharge, and satellite-derived "river width" time series data. The algorithm is provided so that this approach could also be implemented by others for similar (or different) time-series data.

Quality Control

For operational flood forecasting over a large domain with multiple stage reading sites, automated quality control (QC) of measurements is essential to providing accurate forecasts, without spurious values being ingested and corrupting the forecasting algorithms. Beginning early in this consultancy, we have been collecting stage measurements from www.india-water.gov.in/ffs, and archiving them into a MYSQL database (our filtered set useful for operational forecasting is comprised of slightly fewer than 300 sites). However, there are appreciable errors found in almost every one of these sites. As but one example, going to the "River Gage" link found under our project website, clicking on the "CWC Stage Height" layer, clicking on the "back" time button to a date in October, and clicking on the station "011-MGD3VNS" (in the lower center of the screen), shows the original time series of record for this station in the lower part of the screen, with the quality-controlled time series shown at right. Notice that in the QC figure that there is a red point early in the season flagged as "bad", a turquiose point flagged as a "duplicate", and a green point flagged as "questionable" (which is given a QC number varying between 0-100). Please see Appendix A for more details on our quality-control procedures.

Please find also at our ftp site ftp://ftp.rap.ucar.edu/incoming/irap/india_stageQC/ where we are providing hourly-updated quality-controlled CSV (directory /ascii) and NETCDF (directory /netcdf) data files of all of the relevant river stage gages (noting that the last column of the CSV files provides the QC metric, as described in Appendix A). Plots of these most current QC'd data also are provided at that site with highlighted data points that are "QC'd" (under directory /plots), and can also be viewed directly at the website as just discussed above.

River Stage Quality Control Updates

In previous reporting we have provided a discussion of the quality control (QC) algorithms that we have implemented to filter our CWC in-situ river stage gaging measurements and our DFO-JRC river "width" measurements. However, in application of our river discharge forecasting algorithms, calibration of our models was failing due to excessive (unphysical) repeated measurements of the CWC in-situ stage values. As such, we implemented an additional feature to the QC algorithms to identify river gauge levels that have been constant for an unusual length of time, indicative of a possible stuck gauge. The details of this and the whole QC procedure are provided in the appendix below. Immediately below, we list the questionable data features we identify and how they are flagged.

QC score	meaning
100	good
0-99	questionable rate of change or possible stuck sensor
-1	missing
-2	isolated
-3	too high/low
-4	extreme rate of change
-5	stuck sensor
-11	duplicate

Table 1: QC score values, and their meaning.

In Situ and Forecasting "Rating Curves"

We have designed an approach for estimating rating curves for each of the relevant river gaging sites, converting the stage ([m]) into discharge ([m^3/s]), where there is no (or at least, we have no access to) in situ river flow observations. Similarly, the approach has also derived a "rating curve" of downstream river flow (at a future lagged time) from upstream stage measurements, which is something very useful for hydrologic forecasting purposes. For our data, we used available stage measurements for locations along the Ganges and Brahmaputra, along with

downstream discharge measurements at Hardinge Bridge (Ganges) and Bahadurabad (Brahmaputra) in Bangladesh, respectively, to estimate hydrologic rating curves at each gaging location upstream. Before we discuss how these "rating curves" were derived further below, we point out that an interactive display of this information can be found at http://indiawbg.rap.ucar.edu/Ratefit/ (link also found under our main project website). This app displays both the raw stage and discharge observations and the fitted curve. When a station is selected, it defaults to optimal lag, but allows the user to vary the lag and observe how the raw data and fitted curve change. (This app was built using Shiny, a web application framework for R. The interactive mapping is provided by Leaflet, and the interactive plotting by Plotly).

The underlying process to derive the "rating curves" at the same location where the stage measurements were collected is discussed first below:

Stage measurements of the location we want to derive discharge values at ($Stg_{upstream}$) were first quality controlled as described in the previous section. For each location, this cleaned stage data were matched with corresponding discharge values at the respective downstream location ($Q_{downstream}$) where we have discharge data available. Stage and discharge data were matched by date; stage measurements +/-12 hours were considered, and if there were multiple stage observations within the window, the closest to 00:00GMT was selected, as this is nominally the time of the discharge observations.

Once the pairing was done, the downstream discharge values were "normalized" by the ratio

$$\frac{\overline{(P-ET)}_{upstream}}{\overline{(P-ET)}_{downstream}}$$

which is the spatial and temporal average of the precipitation minus evapotranspiration of the upstream catchment area above $Stg_{upstream}$, divided by the same quantity but averaged over the whole catchment area above the downstream discharge gage location ($Q_{downstream}$). Precipitation "P" was derived from our merged satellite precipitation estimates, and evapotranspiration "ET" was derived from the average over all the "analysis" fields of our 8 THORPEX-TIGGE weather forecast products. The temporal averaging period was the entire 2015 monsoon season.

To reflect the travel time between the upstream stage measurement location and the downstream discharge gauge, data were matched at an optimal lag. Two methods of determining this optimal lag were employed: in the first, stage measurements were correlated with discharge at lags of 0-30 days, and the lag that produced the maximum correlation value was selected. In the second, rating curves were fit on stage and discharge at multiple lags, and the lag producing the best fit, as determined by Nash-Sutcliffe efficiency (NSE) was selected. These lags were

compared to a calculated travel time determined by distance between the stage measurement location and the discharge gauge, as well as channel slope. The traditional rating curve form was used for stage S_i converted to discharge Q_i,

$$Q_i = A (S_i - S_0)^n$$

with the parameters A, S_0, and n determined by nonlinear least squares optimization. (Multiple-part rating curves were also considered; however, visual inspection of the one-part fitting for a variety of sites didn't show clear enough distinction between the fit to low and high flows to justify introducing the added complexity.)

Starting at the most downstream gages, the process was then extended gage-by-gage upstream: using newly-estimated rating curves and thus discharge values at sites just upstream of Hardinge Bridge (Ganges) and Bahadurabad (Brahmaputra), these sites and their discharge values were then used to estimate rating curves at sites further upstream, and so on. At each location, the optimal-fit rating curve (as determined by the largest NSE) was selected by comparing with previously derived rating curves that used discharge from locations further downstream.

The underlying process to derive the "rating curves" at downstream (i.e. "forecasting") locations was done identically to what was just discussed, except the downstream discharge values ($Q_{downstream}$) were not first "normalized" before the optimal fitting was carried out.

As a point of comparison, we provide "forecast" rating curve fits derived for upstream gages using only the Hardinge Bridge and Bahadurabad discharge values. Figure 3 shows the optimal lags (relative to Hardinge Bridge and Bahadurabad gaging locations) derived, showing what we generally expected (hoped) to see: general increase in lag further upstream. Similarly, Figure 4 shows the corresponding NSE for these optimally-lagged rating curves fits, showing again what we expected to see: generally decreasing NSE with further distance upstream.

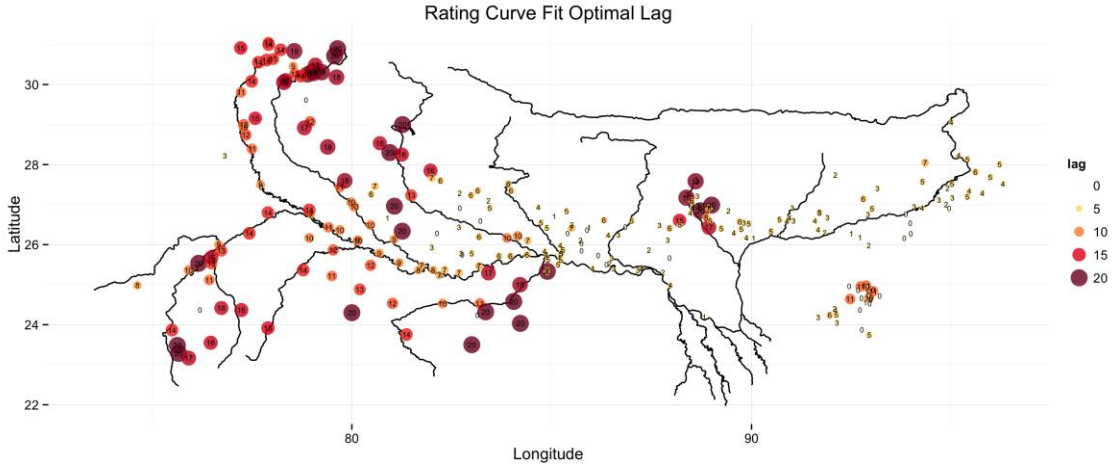


Figure 6: optimal lags of the rating curve fits for gages upstream and utilizing the discharge of Hardinge Bridge (Ganges) and Bahadurabad (Brahmaputra), showing the general increase of lag further upstream.

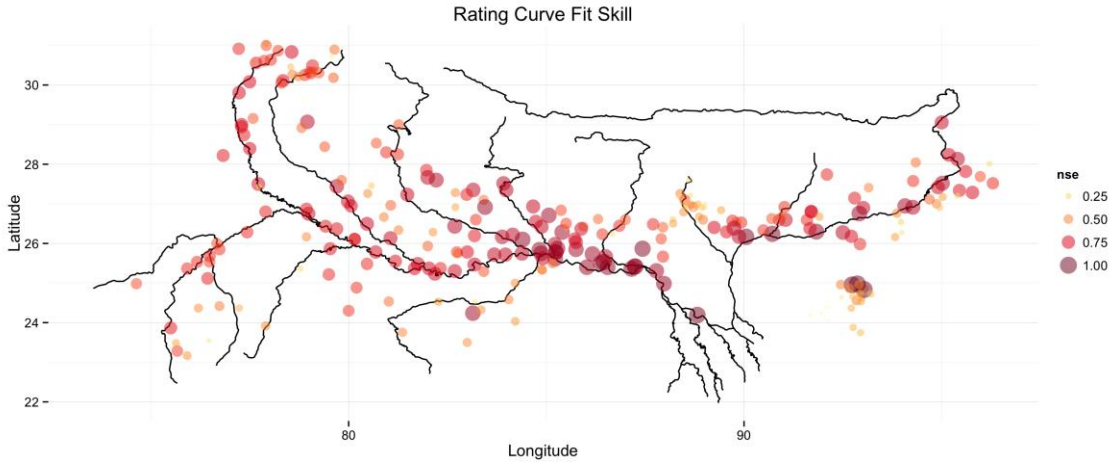


Figure 7: corresponding Nash-Sutcliffe efficiencies (NSE) for the optimal rating curve fits utilizing the discharges measured at Hardinge Bridge (Ganges) and Bahadurabad (Brahmaputra) only, generally showing high NSE throughout the basins, but with an (expected) decrease of NSE with upstream distance.

To provide insights into the strength of the rating curve fits derived in this procedure, we suggest viewing the results found under http://indiawbg.rap.ucar.edu/Ratefit/, which shows the results of the fitting process, along with our best-estimated discharge at the corresponding stage gaging

site. In Appendix B we discuss this App in more detail and provide screen shots for reference. (We anticipate using similar technology for presentation of forecasting results of this consultancy.)

Quality of Rating Curve Fit Filtering

In our previous reporting, we provided a discussion and displays of our rating curve estimate fitting for the majority of our approximately 300 CWC stage gaging locations (results shared at http://indiawbg.rap.ucar.edu/Ratefit/). However, this set needed to be further reduced down for forecasting purposes, given that the (unconstrained) fitting of our rating curve functions converged to unphysical exponents. The The traditional rating curve form was used for stage S_i converted to discharge Q_i,

$$Q_i = A\ (\ S_i - S_0\ )^n$$

with the parameters A, S_0, and n determined by nonlinear least squares optimization. However, the exponent "n" found for many of our gauging locations was either significantly less than a value of one, or significantly greater than a value of 5. As such, these gauging locations were (temporarily) removed from our list of forecasting sites (but can very easily be added back in if additional rated stage-Q valued could be provided). The final gauges can be seen in Figures 1 and 2 above, and also seen in the table found here: http://ral.ucar.edu/~hopson/WorldBank/filtered_gauges.txt.

Satellite Precipitation Accuracy and Thorpex-Tigge Numerical Weather Prediction Evaluation

In this section, we present some of our results in evaluating the biases of ECMWF and three other globally skillful (NCEP, UKMET, and Canada Met Office) ensemble precipitation forecasts, made available through the THORPEX-TIGGE project. These forecasts are compared relative to our merged satellite precipitation estimates (merged product from NASA TRMM, NOAA CMORPH, and JAXA GSMaP).

Utility of Satellite Precipitation

As we argued in the section further above, just observations of rainfall falling over the catchment can provide forecasts of warnings due to the travel times required for the rainfall to migrate through the catchment. Ideally, observed fields of precipitation would be based on densely-spaced ground observations from meteorological stations or radar data, which provide precipitation measurements with the highest available accuracy. Figure 5 below shows an example of the correspondence possible between rain gauge and radar observations (note: this is a "good" example case).

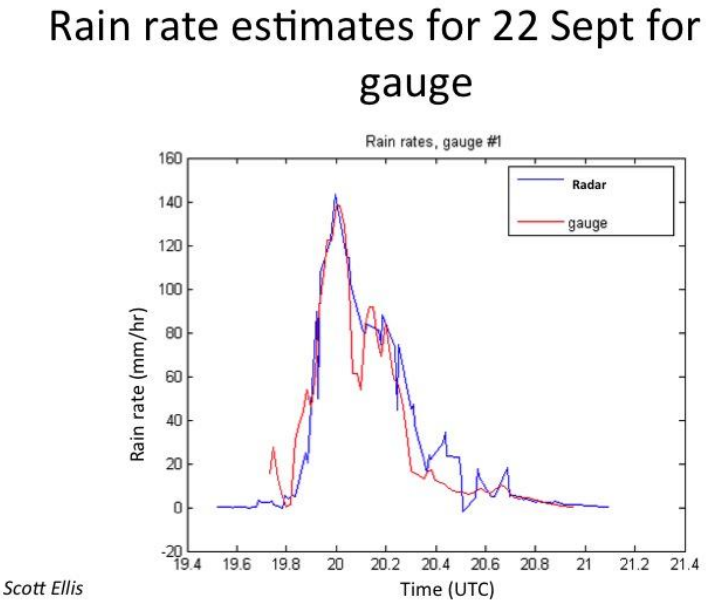


Figure 5: a "good" example of the comparisons one can achieve between rain gauges and radar for one particular experimental site.

However, the spatial coverage of ground networks is very scarce over several regions of the world, such as most parts of Asia. To understand the impact of rain gauge spacing on accuracy, consider Figure 6 below, which shows the impact of spatial separation and temporal averaging time on direct comparisons between two ground-based rain gauges over Florida, and how the correspondence degrades with increasing spatial separation and shorter temporal accumulation times. This shows the importance of adequate rain gauge density when using rain gauges to provide representative measurements of rainfall over larger regions than just their local vicinity.

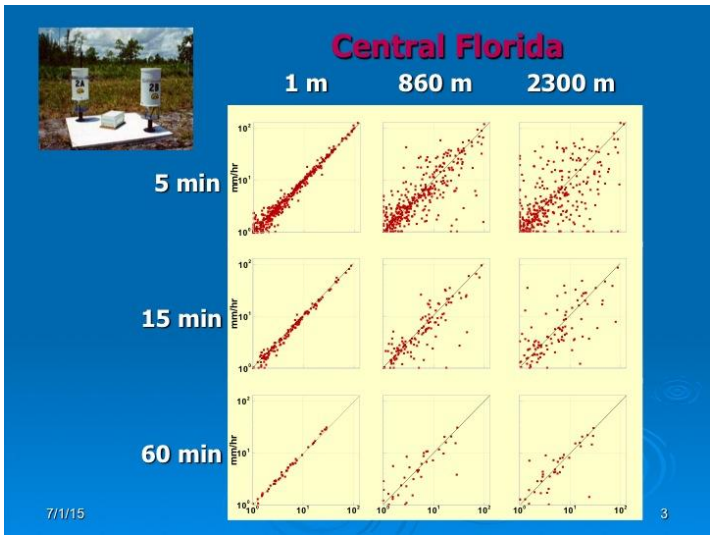


Figure 6: the correspondence of accumulations from two rain gauge when they are separated by differing distance and compared over longer temporal integration times. Shown are scatter plots of the two gauges for greater separation along the x-axis, and longer temporal integration times going down the y-axis. Note how the one-to-one correspondence degrades with larger separation and shorter integration times (lower left best, upper right worst).

Because of this "representativeness error" of sparse rain gauge networks, satellites estimates have been widely used as an alternative or to supplement station observations over the last 30 years. While exhibiting a lower accuracy compared to ground observations especially at short temporal scales, excellent spatial coverage is the main strength of the satellite products. To partly overcome accuracy issues, many precipitation estimates combine data from multiple satellite sensors (typically thermal infrared and passive microwave sensors) [Joyce et al., 2004; Huffman et al., 2007; Kidd and Huffman, 2011; Kidd and Levizzani, 2011] and ground observations [Grimes et al., 1999]. Because of large areas of the Ganges and Brahmaputra catchments (especially those regions outside of India) with sparse rain gauge densities (or radar coverage), satellite precipitation estimates are an important data source for producing skillful flood forecasts throughout the catchments.

However, just as there are "representative errors" with rain gauges representing rainfall rates over larger areas, satellite precipitation estimates also have their own errors, as seen in Figure 7 below, where we show comparisons between two satellite precipitation products over the Ganges catchment.

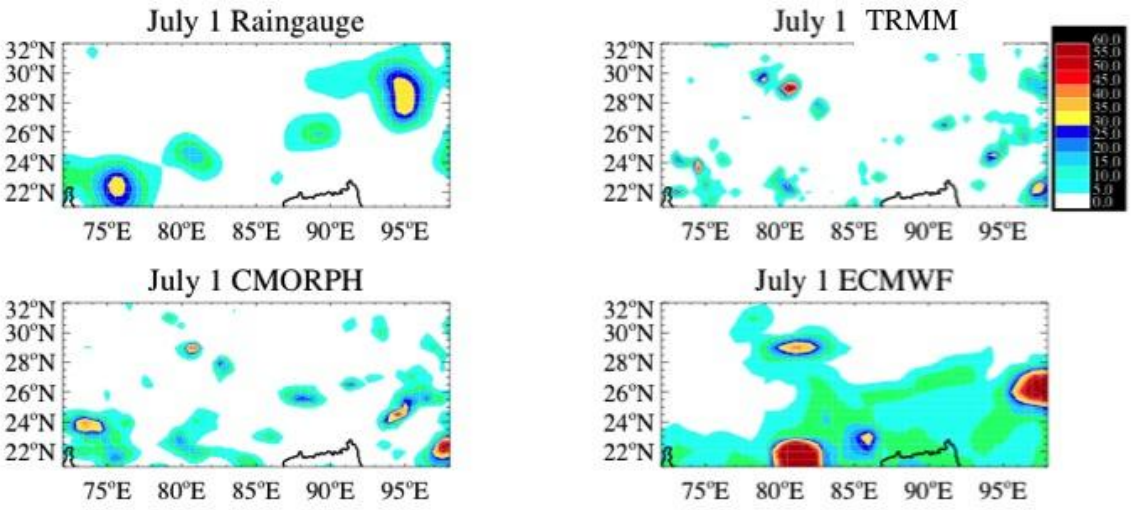


Figure 7: 24hr accumulations of rainfall over India comparing NASA-TRMM, NOAA-CMORPH, GTC rain gauge, and ECMWF 24hr rainfall forecasts for 2004. Comparing the

TRMM and CMORPH products, there are clear differences in locations and intensities of rainfall at finer scales, although the general large-scale features are generally in common.

But contrary to the rain gauge estimates shown in Figure 7, satellite-derived precipitation errors decrease (instead of increase) with spatial-scale. Figure 8 shows Ganges catchment-averaged comparisons between the same two products shown in Figure 7 (NOAA CMORPH and NASA-TRMM) over the 2007 monsoon season, showing how they now correspond well with one another at these larger spatial scales. The implications for this project are that we expect accuracies in flood forecasts to improve for larger river basins, when utilizing these products.

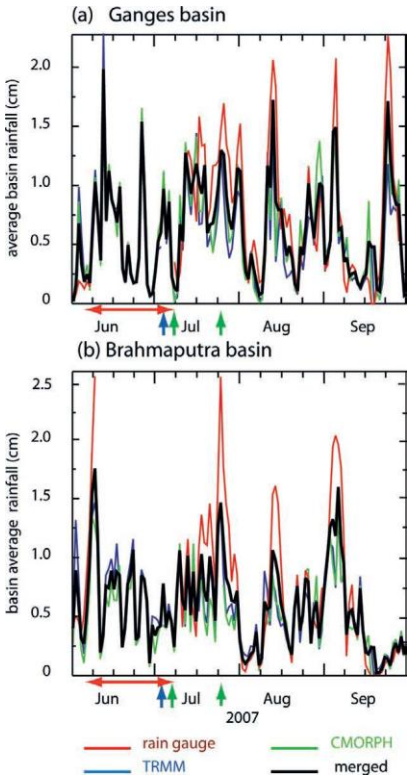


Figure 8: 2007 daily catchment-averaged accumulations of rainfall from GTC rain gauges (red), NOAA-CMORPH (green), NASA-TRMM (blue), and their average (black), showing the general similarities of the satellite products, averaged over these large spatial scales.

Another essential data set for this consultancy are satellite-derived gridded rainfall estimates that are used here to calibrate hydrologic models, given these provide the best estimate of historical precipitation over the watersheds due to the prevalence of raingauge data sparse regions. These data are also used to initialize the states of the soil moisture and in-stream flows before the hydrologic model integrates these states forward in time using weather forecasts.

- Collect archived data of the NOAA-CMORPH, NASA-TRMM, and JAXA/EORC data products used for calibration purposes
  Duration/start date or status: completed

- Automate real-time download of these products and combine (using their straight average) for operational hydrologic model and in-stream flow initialization. These data will be used (initially, before their optimal combination determined) for 1) locations where no rain gauge data are available, but still required for hydrologic model calibration, and for 2) for quicker responding operational flood forecasting, given that satellite data availability times are typically significantly quicker than rain gauge reports, and higher sampling rate.

Duration/start date or status: completed

- Optimal combination of products determined, based on calibrations with available rain gage values were sufficiently dense and available. The calibration will be done for: 1) probability of precipitation, and for 2) the rainfall amount, given a rainfall event occurs. In additional to optimal weights, uncertainty estimates for the optimally-combined product will also be generated, which will be useful for determining the overall predictability of the system via monte carlo simulations.
  Duration/start date: completed Aug-Sept

Utility of Ensemble Precipitation Forecasts: THORPEX-TIGGE

Referring to Figures 1 and 2 above and their associated discussion, past the *time of concentration* for a particular forecasting location within a watershed, long-lead forecasts of river flows will necessarily rely on precipitation forecasts for their skill since all modeled surface flows from observed rainfall will have already traveled past that particular location (neglecting deeper and slower baseflow response in the watershed). For the longer-lead river forecasts generated for is project, we will utilize the THORPEX-TIGGE data set for this use. THORPEX stands for **Th**e **O**bserving system **R**esearch and **P**redictability **Ex**periment, while TIGGE stands for the **T**HORPEX **I**nteractive **G**rand **G**lobal **E**nsemble. THORPEX is an international research programme established in 2003 by the World Meteorological Organization to accelerate improvements in the utility and accuracy of weather forecasts up to two weeks ahead. It is part of the World Weather Research Programme and is a key component of the WMO Natural Disaster Reduction and Mitigation Programme. TIGGE is a key component of THORPEX, and is a World Weather Research Programme designed to accelerate the improvements in the accuracy of 1-day to 2 week high-impact weather forecasts for the benefit of humanity by providing access to global ensemble weather forecasts from eight NWP centers around the globe (see Figure 9 below).

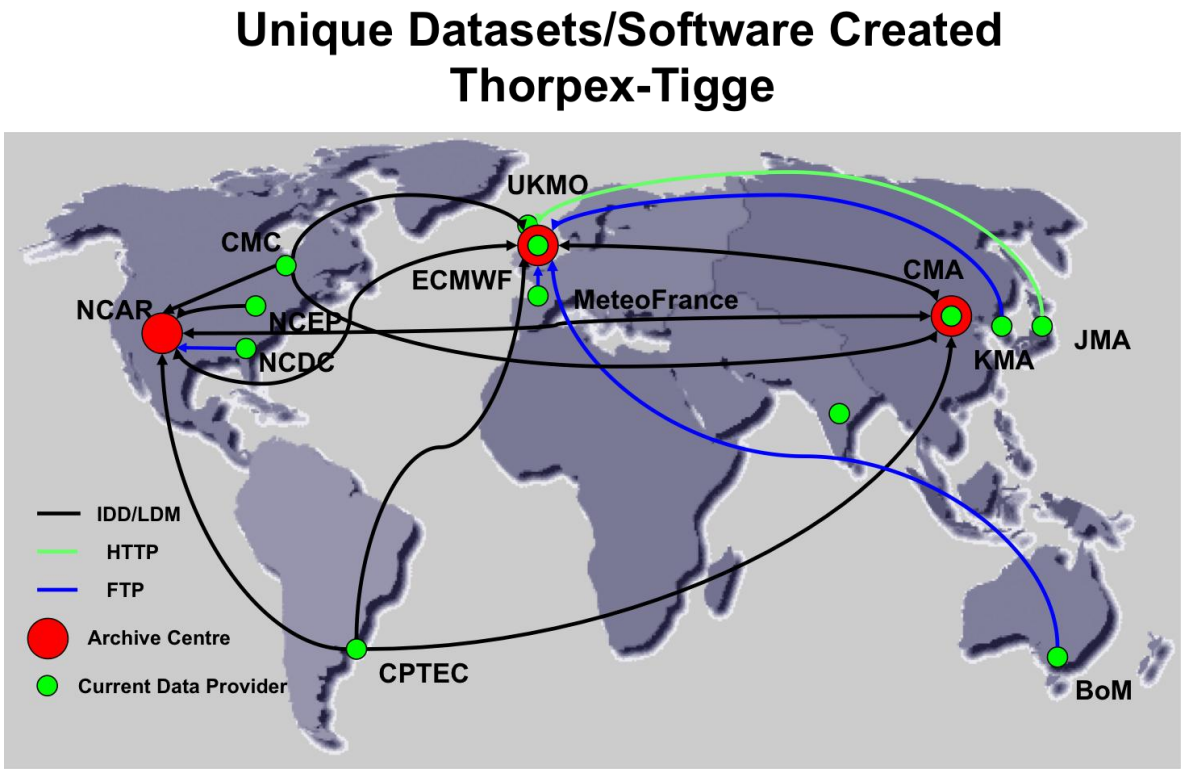


Figure 9: Green dots show the locations of the 12 NWP centers that provide data to the (nominally 3-day delayed) near-real-time THORPEX-TIGGE data archive. This project will primarily utilize forecasts from the following centers: CMA, CMC, ECMWF, UKMO, NCEP, MeteoFrance, JMA, and CPTEC

We will utilize and evaluate data from eight of the contributing NWP data centers for this project, which provide forecasts every day and for every 6hrs ahead out to 15 days in-advance. All told, these data constitute over 300 ensemble members being generated each day.

The skill of these NWP ensemble forecasts degrade as the forecast lead-time increases. However, as with satellite precipitation estimates, the skill of NWP precipitation forecasts also increases with spatial scale. Figure 10 below shows the RMSE skill scores for 1- to 5-day lead-time ECMWF forecasts over the Ganges-Brahmaputra catchments, showing how forecast skill increases (logarithmically) with increasing spatial scale. As with the satellite precipitation estimates, these increases in forecast skill with spatial scale also imply that river discharge forecast accuracies utilizing these NWP ensemble forecasts will likewise improve as the river basin spatial scale increases.

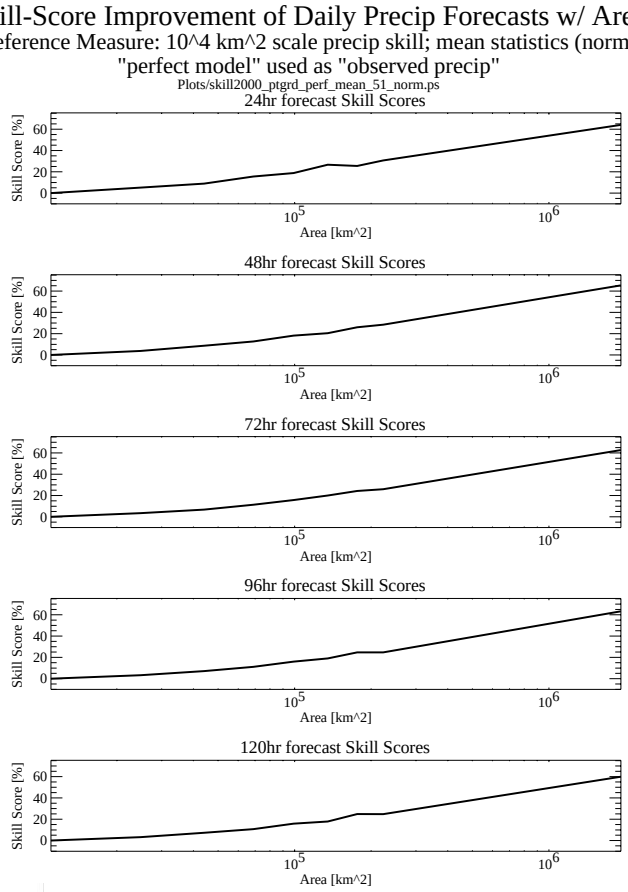


Figure 10: plots of the increase in root-mean-square error skill score (y-axis) versus spatial scale (logarithmic x-axis) for ECMWF ensemble precipitation forecasts over the Ganges and Brahmaputra catchments, showing how skill (logarithmically-) increases with increasing spatial scale, which holds true for 1- (top), 2-, …, 5-day (bottom) forecast lead times.

To get an intuitive idea on the actual variability in these ensemble forecasts, consider the variability in24-hr accumulated precipitation forecasts from five ECMWF ensemble members during a severe flooding even that occurred over SE Africa in 2011, as seen in Figure 11 below. The general large-scale correspondence with the actual precipitation (satellite observations in the middle of the figure) is seen, but also the large degree of variability at the small scale.

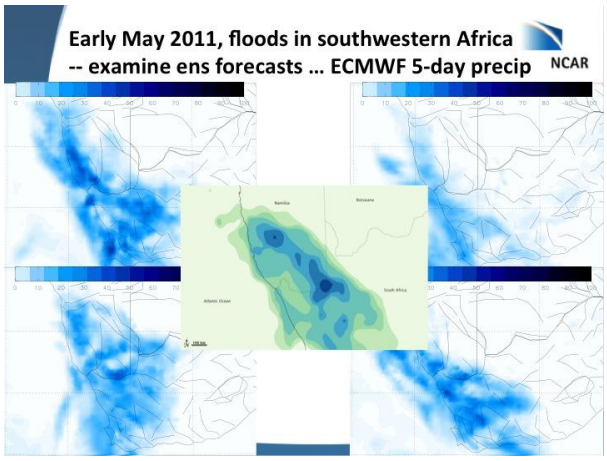


Figure 11: comparisons between five ECMWF ensemble members of 24-hr accumulated rainfall at 5-day lead-times, compared to satellite-derived estimates shown in the center.

In Figure 6 below, we compare ECMWF 5-day (0 to 120hr) accumulated precipitation forecasts (forecasts initialized at 00Z) for the month of August climatology of years 2011 to 2014, compared to our merged precipitation product

(same time period), over our high resolution sub-catchments of the Ganges and Brahmaputra rivers (up to their entry points into Bangladesh), along with their difference (i.e. mean August bias).

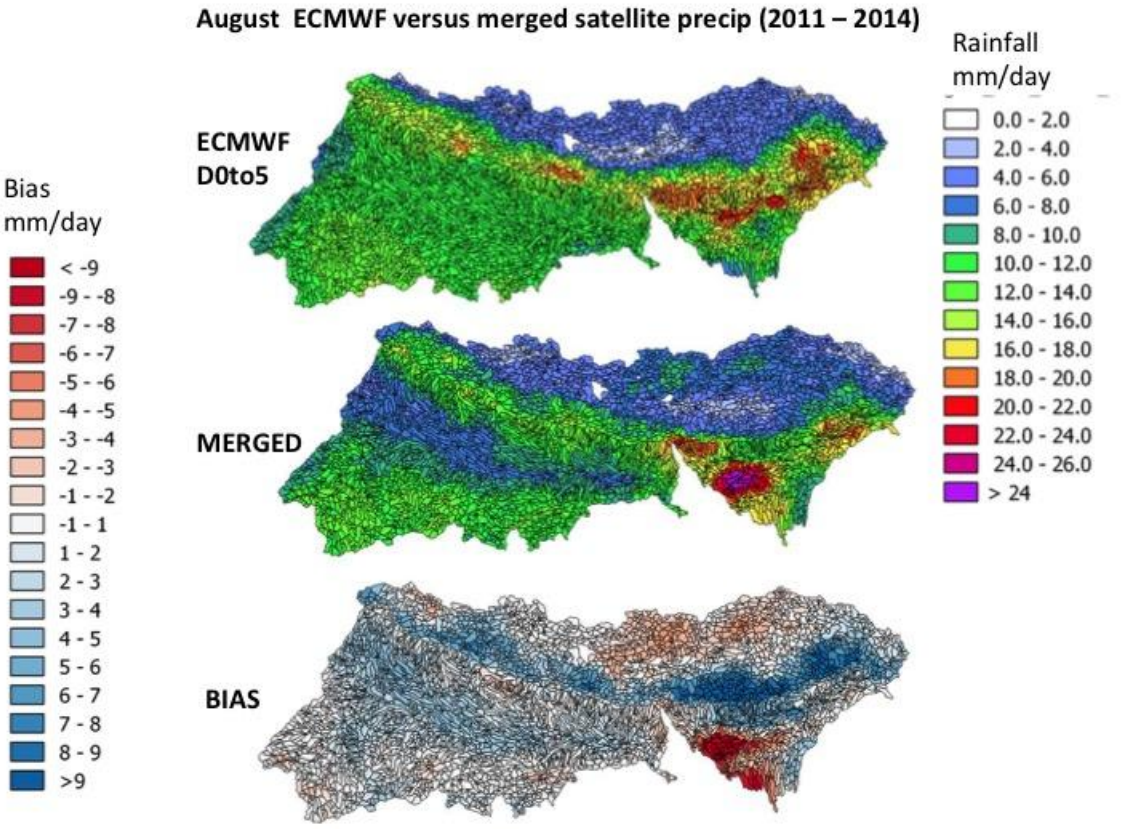


Figure 8: ECMWF 0 to 5day accumulation period rainfall forecasts climatology for Augusts of 2011 to 2015, compared to our merged satellite precipitation product.

What we observe is a rough correspondence in the spatial distribution of precipitation, along with a general over bias in ECMWF. But there are also clear and appreciable (i.e. up to half a cm/day) regional over- (mostly) and under- (less so) biases. In particular, we see insensitivity to topographic variation in the lower-in-elevation central Indo-Gangetic plain, along with a more northerly extent of monsoonal rainfall up the Brahmaputra river valley than is observed (at least, as so far as observed in the merged precipitation).

It is helpful to further compare ECMWF's (as arguably, the world's most skillful global forecast center) biases with those of three other well-respected centers: NCEP, UKMET, and Canada Met Office, as seen below in Figures 7 and 8.

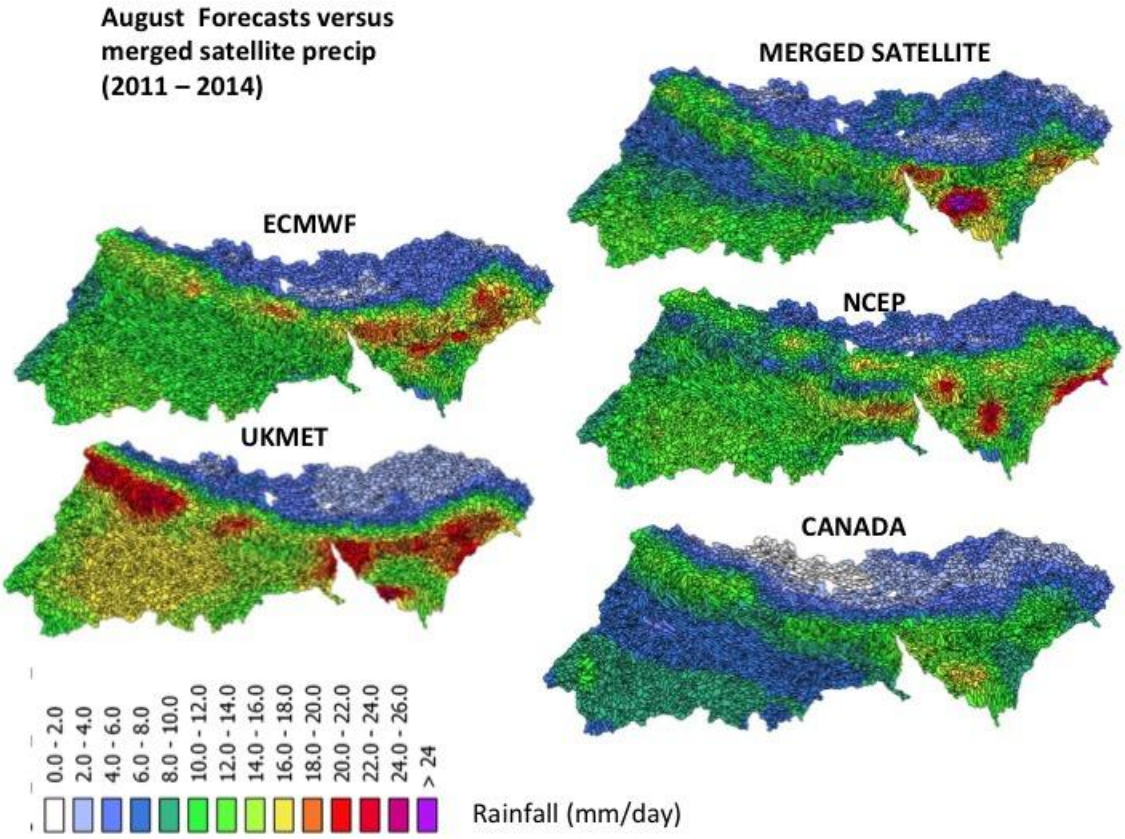


Figure 9: Similar August rainfall climatological distributions to those shown in Figure 4, but now for the additional centers of UKMET, NCEP, and Canada Met Center.

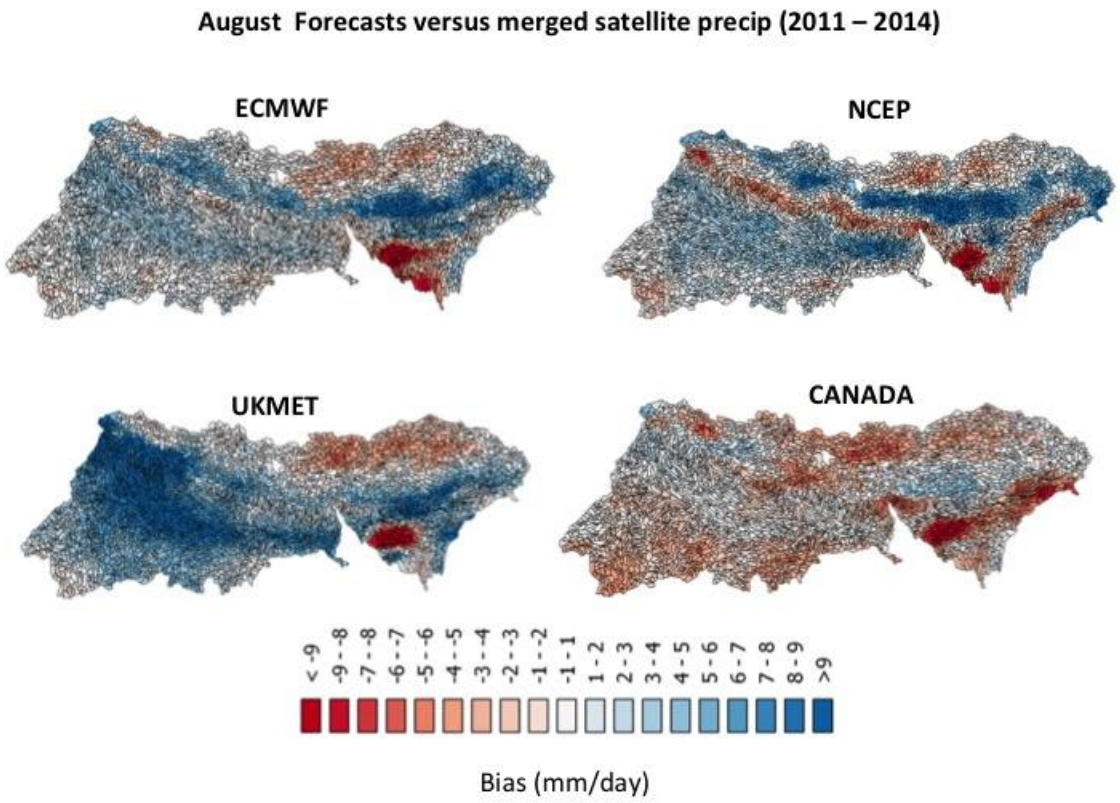


Figure 10: similar to Figure 7, but now showing the biases themselves (relative to the satellite merged product).

In Figures 7 (climatologies) and 8 (biases), we again see general rough correspondence in the spatial distribution of precipitation for all four centers, but also, as with ECMWF, general across-the-board biases -- but in some cases, very different spatial structure as compared to what is observed with ECMWF. For UKMET, we see a similar over-bias (Gangetic Plain)/under-bias (upper Brahmaputra river valley) structure, while that of NCEP seems to have a phase-shift offset in the region of topographically influenced precipitation (Himalaya foothills). Canada appears to have perhaps the best spatial pattern match to the merged satellite product, but with an appreciable general under-bias, though.

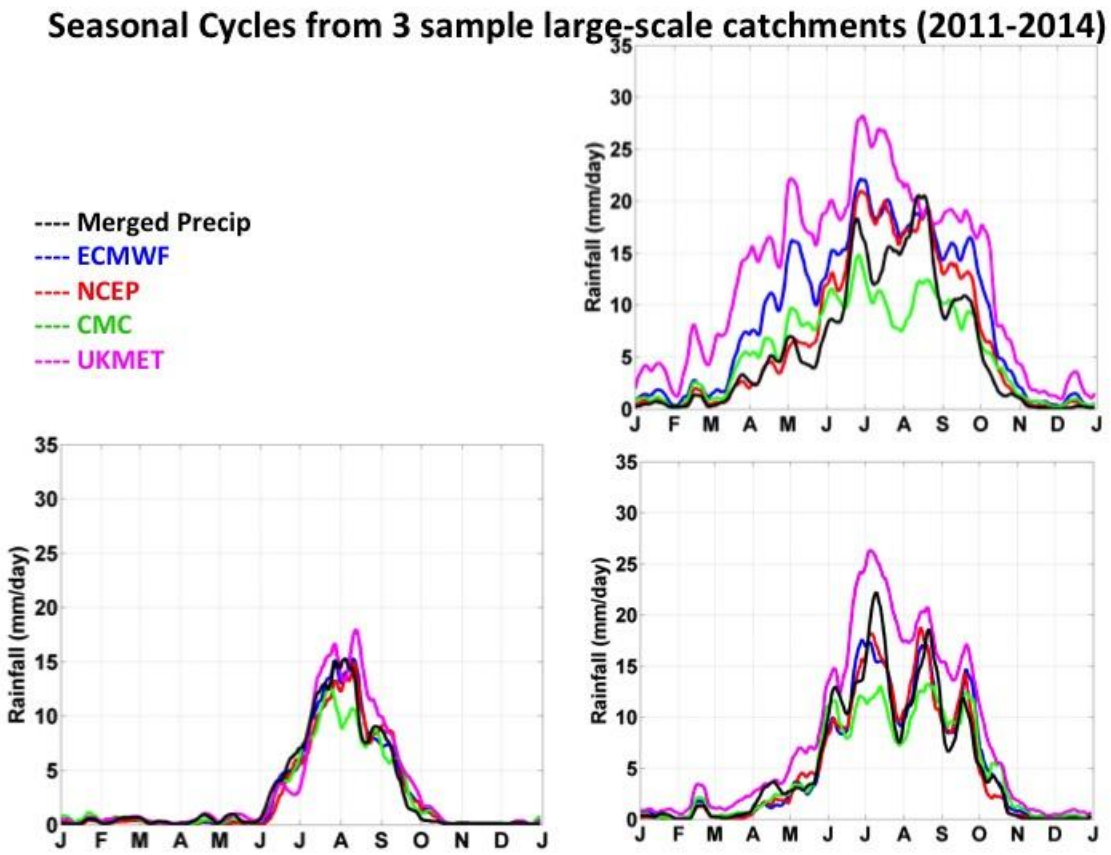


Figure 11: annual catchment spatially-averaged climatological distributions of the same products as shown in Figure 8, for three randomly-chosen watersheds in our study region.

To get a better sense for how these biases may impact hydrological forecasting over the annual cycle, we have chosen (randomly) three larger-scale catchments in our region to spatially-average the forecasts over (again, 0 to 5-day accumulations averaged over 2011-2014) for each day of the year, shown in Figure 9. One can see how the structure of these biases varies not only by region, but also by month-of-year (especially with ECMWF and UKMET which, in the upper right, show appreciable early- and late-monsoonal season biases, but less so [at least in the case of ECMWF] during the peak of the monsoon season).

What these observed biases point out is the need for the precipitation forecasts to be calibrated. This is especially relevant when initializing (i.e. during the start period) the hydrologic forecasts with an observed (i.e. estimated merged satellite) precipitation product, but then driving the hydrologic model forward in time using a ***relatively*** biased precipitation forecast product. The reason for the emphasis on "relatively-biased" is that hydrologic models can often be calibrated themselves to implicitly remove rainfall biases – but the hydrologic model cannot also simultaneously remove relative bias between the observed and forecasted rainfall

products – this causes the derived river forecasts to drift away and incur larger and larger biases in its initial optimized state as the forecast lead-time increases.

It is also important to point out that just because a precipitation forecast is biased, that does not necessarily imply that the forecast itself is necessarily unskillful *after* the forecast is calibrated. The calibration process and evaluation of the skill of these weather forecasting products will be further discussed in our next technical report, in particular, in the context of inputs into hydrologic forecasting.

Conclusions in the skill of our TIGGE ensemble numerical weather prediction precipition forecasts discussed in our previous report were based on analsyses performed with our merged NASA TRMM, NOAA CMORPH, and JAXA GSMaP satellite precipitation product (merged to a common 0.1X0.1 degree resolution, accumulations every 3hrs). Here, we want to confirm the similarity of these satellite products to available (daily) rain gauge estimates over the same domain. For this study we compared sub-basin wide merge satellite climatologies and daily correlations with the same, using daily 0.5X0.5 gridded rain gauge analyses that used rain gauges reporting to the WMO GTS network (Xie et al. 1996). The figure below shows a comparison of seasonal cycles of rain gauge and merged satellite precipitation for eight different catchments in the Ganges Brahmaputra watershed. The seasonal cycles are for a 15 year average (2001-2015) and smoothed with a 17 day running mean. The map shows the merged satellite precipitation.

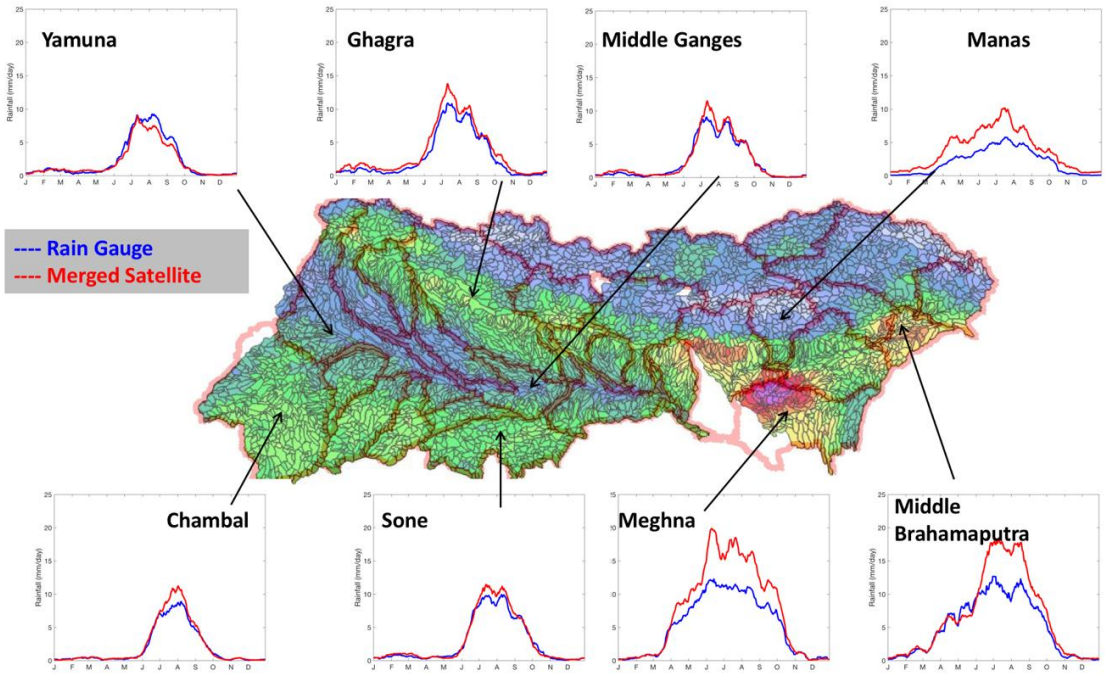


Figure 12: Comparison of seasonal cycles of rain gauge and merged satellite precipitation for eight different catchments in the Ganges Brahmaputra watershed. The seasonal cycles are for a 15 year

average (2001-2015) and smoothed with a 17 day running mean. The map shows the merged satellite precipitation.

Our main conclusions from this analysis are as follows:

- The rain gauge and satellite precipitation products are mostly similar in terms of their seasonal cycles.
- The largest discrepancies (in the Meghna and Middle Brahmaputra) occur in regions that have limited gauge coverage. Further analysis of the Meghna suggests that the gauge data has proximately 3x as many days with zero or very low precipitation, possibly suggesting that rainfall events are being missed to lack of coverage.

Given this reasonable similarity in the gauge-satellite climatologies, we repeat our main findings again from our previous report's analyses of the TIGGE forecasts below (based only on merged precipitation comparisons):

- Precipitation forecasts show large regional biases in rainfall, highlight the importance of calibration and bias correction.
- The models vary in their spatial representation of climatological rainfall, indicating differing interactions between rainfall and topographic features.
- Models with the smallest biases are not necessarily the ones with the best skill.
- The relationship between forecast skill and basin size is apparent in these data, though other geographic factors are clearly important as well.

Similar daily sub-basin time-series correlation analysis on these products done for July was also done and is provided in the appendix below. However our findings are similar to the above, in that we find correlations between these products of between 0.6-0.7; which although not extremely high, we argue they do prove a strong-enough relationship between these products to affirm the use of satellite precipitation for river discharge forecasting over the Indian subcontinent.

However, future work would suggest the need for in-depth analyses of both NWP and satellite precipitation using a higher density network of rain gauges available for localities over India (given that our gridded rain gauge product had a limited density), especially over strong terrain transition zones.

Satellite River "Width" Evaluation

In this section, we discuss the workflow and implementation of remotely-sensed satellite-based data sets estimating river widths and heights into the hydrologic forecasting framework. Note that these components are not essential for generating the ensemble hydrologic forecasts based on the hydrologic forecasting framework of Figure 3 (which itself is primarily based on rainfall products and river stage observations only), but these data sets can contribute to enhanced forecast accuracy, as discussed in the first section of this report, especially as they become directly data assimilated into the hydrologic modeling structure during latter stages of development. As well, because of advective lags in water flow travel times, these components can also be used to generate stand-alone river discharge forecast products based on correlations with lower sections of the river reaches (see the earlier section and Figures 1 and 2 of this report that discusses the limits of predictability of these data sets). In what follows, we have broken down workflow components into those relating to the Dartmouth Flood Observatory ("DFO") river width measurements, and those of the University of Maryland satellite altimetry measurements ("altimetry").

Lack of data availability of near-real-time river stage measurements upstream in the Ganges and Brahmaputra catchments greatly restricts short-term predictability of river flooding. One approach to instead provide a "work around" on this data limitation is to explore remotely-sensed information that do not rely on in situ measurements being taken and being timely provided. The DFO RiverWatch data provide one such source of such data, especially since the imagery this system relies on (microwave) can penetrate cloud cover (which should be potentially plentiful during times of flood risk). Figure 12 shows locations of measuring sites of the DFO River Watch system that will be used in this consultancy (produced in collaboration with the Joint Research Council; see JRC-Ispra, http://www.gdacs.org/floodmerge/ and DFO http://www.dartmouth.edu/~floods/ -- sites along these river channels will be further expanded and filtered as part of this consultancy).

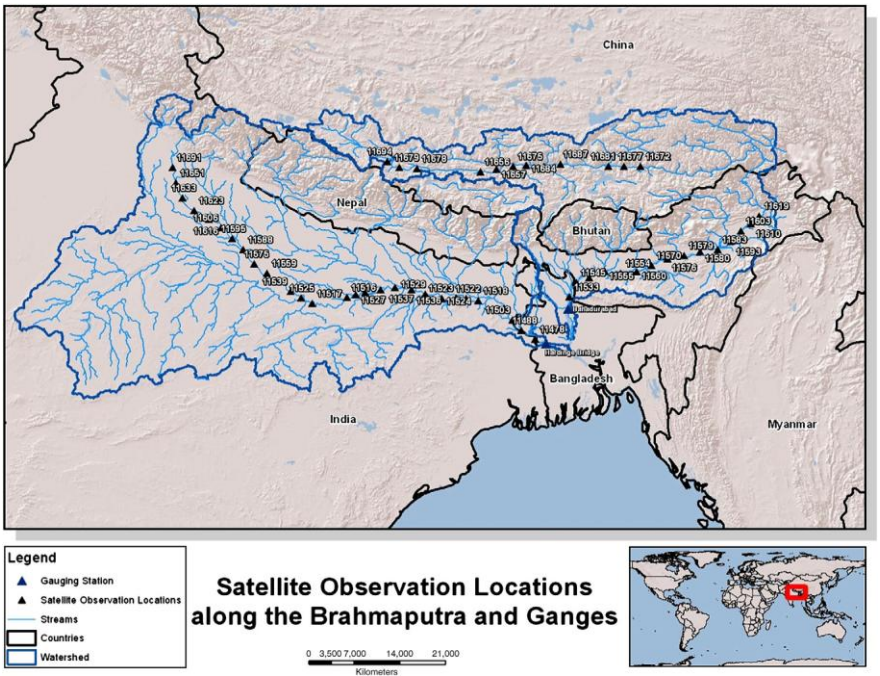
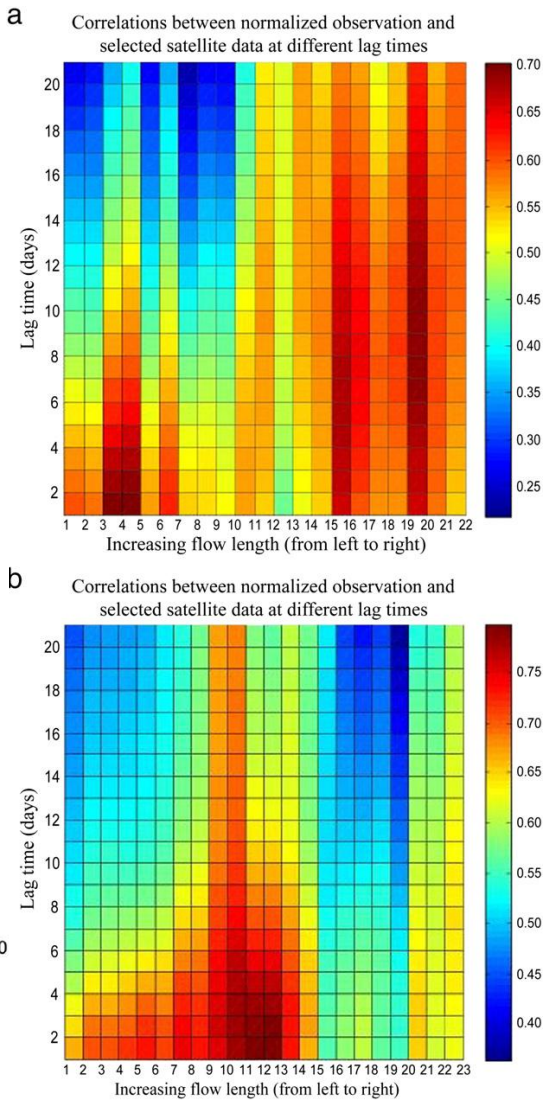


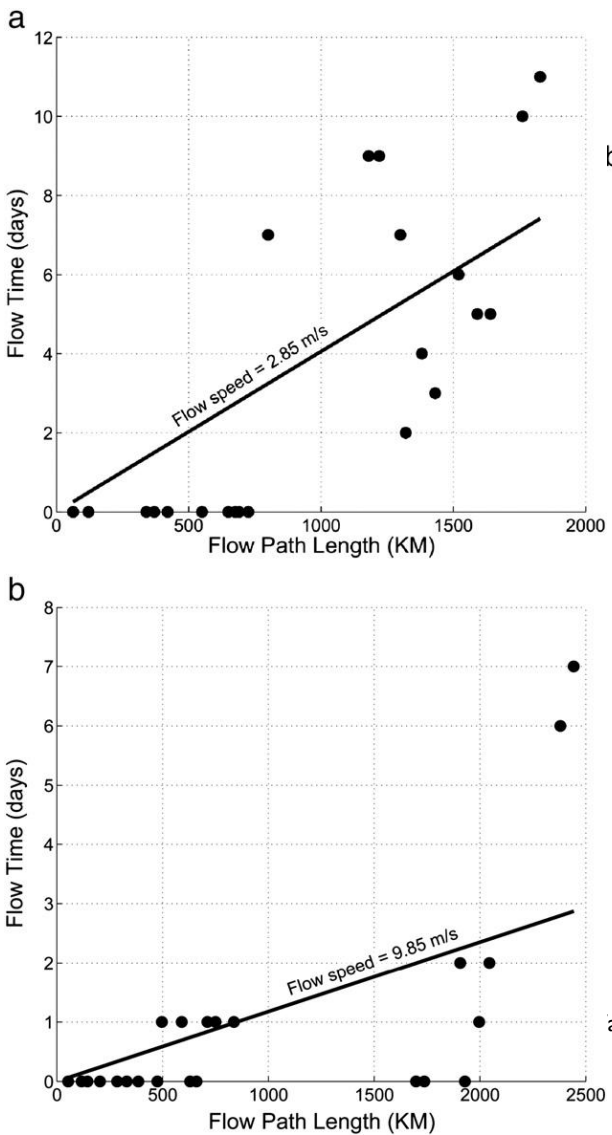
Figure 12: The Brahmaputra and Ganges Rivers in South Asia. The satellite flood signal observations are located on the main streams of the

Brahmaputra (top right) and the Ganges (bottom left) rivers. The observation sites are shown in small dark triangles.

If one were to correlate these upstream signals at different time lags with actual discharge values at the catchment outflow points, one would get the correlation maps shown in Figure 13, which clearly show that further upstream, correlations peak at longer and longer lead times.

Figure 13: a. Lagged correlation map of daily satellite-derived flow signals calculated against the discharge observation at Hardinge Bridge for Ganges River. The horizontal axis shows the satellite flood signal sites (see Fig. 1) arranged in the order of increasing flow path length and the vertical axis shows lag time (days). b. Same as Fig. 4a, but for Brahmaputra.





Taking the peaks in these correlations and plotting them against the upstream distance, we can generate a scatter plot, whose trend provides an estimate of the average travel time in the river reach, as shown in Figure 14.

This lagged correlation with observed discharge downstream shows the potential for these signals to provide forecasts of downstream flows, as we discussed further above in the first section.

Figure 14: a. Plot of flow time (as estimated from the satellite flood signal data) versus distance from the satellite flow detection point to the outlet (Hardinge bridge station) of the Ganges River. The flow time is the lag time at which the peak correlation occurred, as shown in the figure above. The flow speed estimated from the slope of the fitted line is 2.9 m/s. b. Same as Figure 3a, but for Brahmaputra river. The flow speed estimated from the slope of the fitted line is 9.6 m/s.

In this section we provide an evaluation of river "width", as measured from surface emissivity variations, and their potential utility for upstream monitoring of potential flood wave conditions. Although we refer to our sites as "river width" sites, they are not, really. What is being measured, via a spatial ratio approach, is the differential brightness temperature (upwelling microwave radiance, or brightness) of pixels strategically located over river channel and floodplain reaches (square 10 km parcels of land and water). When the river rises, the mixed water/land monitored by the pixel now includes more water; meanwhile, the array of surrounding pixels is also monitored, the driest (brightest) value selected, and a ratio calculated for the measurement pixel. In short, the background is intentionally removed (because microwave radiance responds strongly to other variables, such as land surface temperature). The differential wetting and drying of the measurement pixel is thus sensitively recorded, but sometimes at the limit of detection. That ratio value is indeed sensitive to flow width, averaged over the ~10 km length of the reach, but is more directly responding to reach flow area.

One word of caution and background on the utility of the data used in this consultancy: many of the sites we are using were almost "automatically" located, and as such, are not acceptable for use without oversight. One issue contaminating the signal is that many of the sites are not in fact located over rivers, perhaps with the river just hitting one corner of the pixel. Another major issue is that pixels can be saturated at high water: there is enough flooded area to cause the entire pixel to only observe water, including for the Ganges. That is why it is important to use multiple pixels to monitor one site. However, complete saturation of the pixel at high flow does not happen often. The pixel may look completely flooded, but in fact there are islands of dry land, and as floodwater rises in stage, the signal still responds. And yet another major issue, perhaps the main factor in quality control: the ratio values may range from say, .99 to .77 (a very robust site, reflecting large low water to high water difference in the signal) or, instead, say, .81 to .85 (basically, quite possibly, just reflecting the lack of any strong hydrologic signal; the selected pixel is varying about at same level as the background pixels, and as a result, the ratio plot is just noise).

Analysis of sites compared to farthest downstream outflow locations

Because we have only one year, at best, of in situ river stage values over India (collection began February 2015), we first look at how well the upstream "width" sites correlate with downstream river discharges at two gaging sites in Bangladesh where we have a long historical record: Hardinge Bridge for the Ganges, and Bahadurabad for the Brahmaputra. Shown below in Figures 10 and 11.

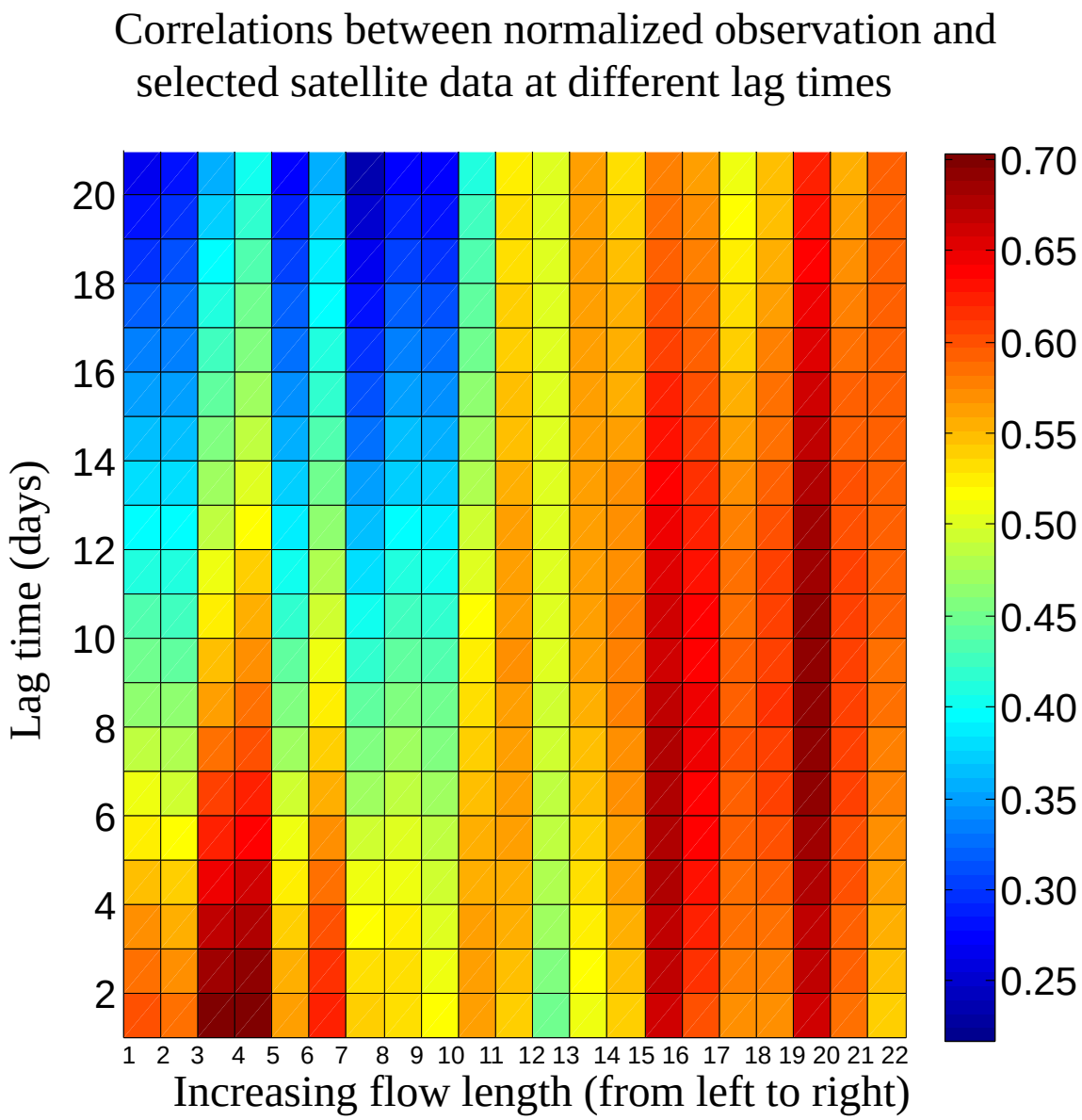


Figure 13: Lagged correlation map of daily satellite-derived flow signals calculated against the discharge observation at Hardinge Bridge for Ganges River. The horizontal axis shows the satellite flood signal sites (see Figure 2) arranged in the order of increasing flow path length and the vertical axis shows lag time (days).

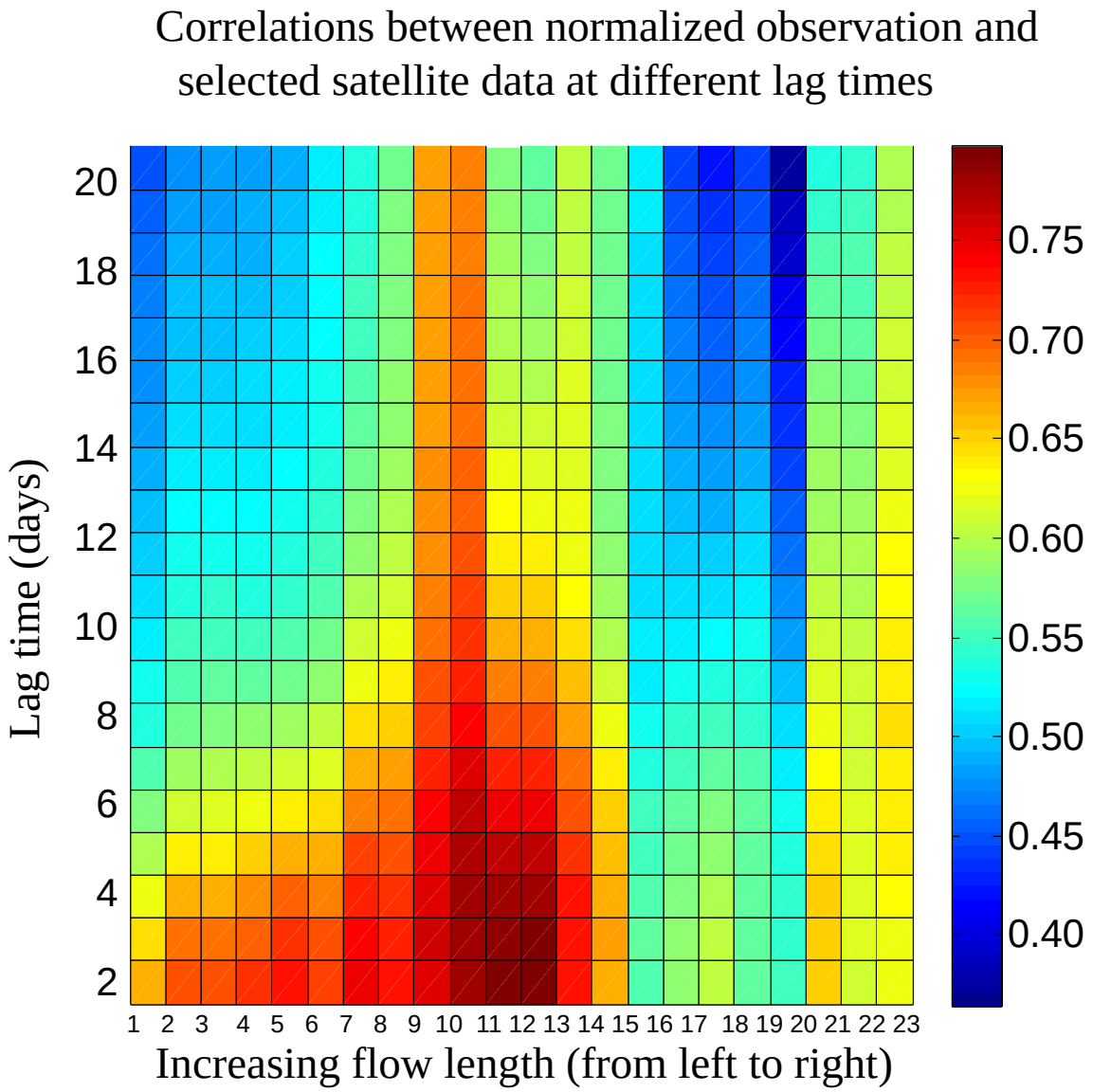


Figure 14: Same as Figure 10, but for the Brahmaputra.

What we observe in Figures 10 and 11 are certain site with high correlation (0.8), and others with poor correlations (below 0.4). Also, the correlation values generally increase with increasing lag as one goes further upstream for both the Ganges and the Brahmaputra. For the Brahmaputra there is a disappointing lack of high correlations higher up in the catchment, which would have been most useful for long lead-time forecasting. Can we see improvements when comparing with in situ gages within India (but with a shorter time series of one season to work with)?

Analysis of sites compared to nearest stage locations

In Figure 12 we show a map of the main channels of the Ganges and Brahmaputra with the locations of the "width" measuring sites shown with red "+" symbols, and the in situ gages we correlated them with shown in blue triangles, ideally choosing one gage above and one below each "width" ("+") site, with dotted lines connecting each "width" site to the in situ gages sites compared with.

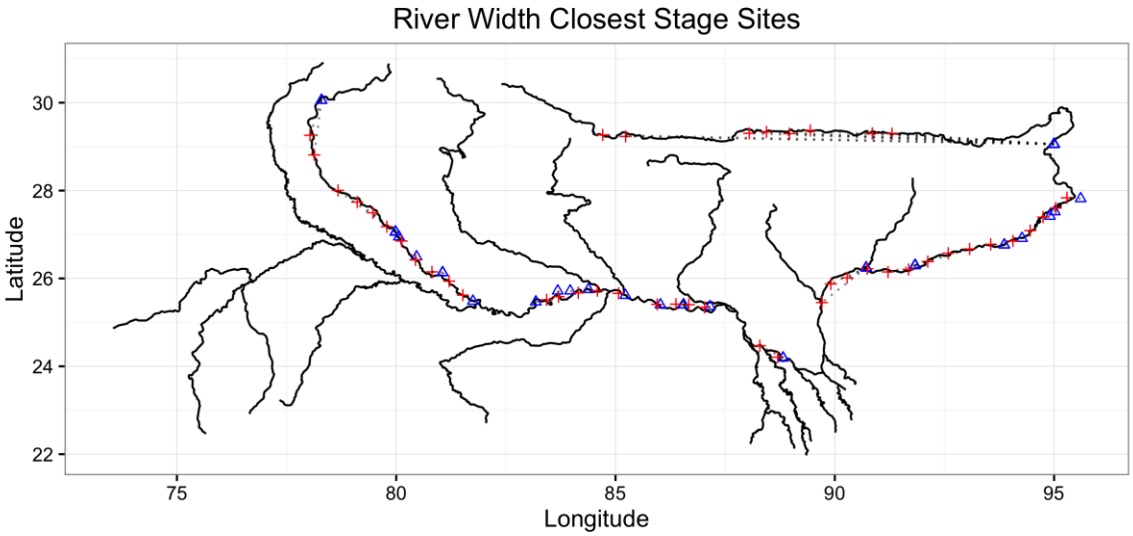


Figure 15: map of the main channels of the Ganges and Brahmaputra with the locations of the "width" measuring sites shown with red "+" symbols, and the in situ gages correlated them with shown in blue triangles. Dotted lines connect each "width" site to the in situ gages sites compared with (ideally one gage "above" and one "below", bounding the "width" site).

Our "width" site correlations to the nearest in situ gages are shown in Figures 13 (Ganges) and 14 (Brahmaputra) below.

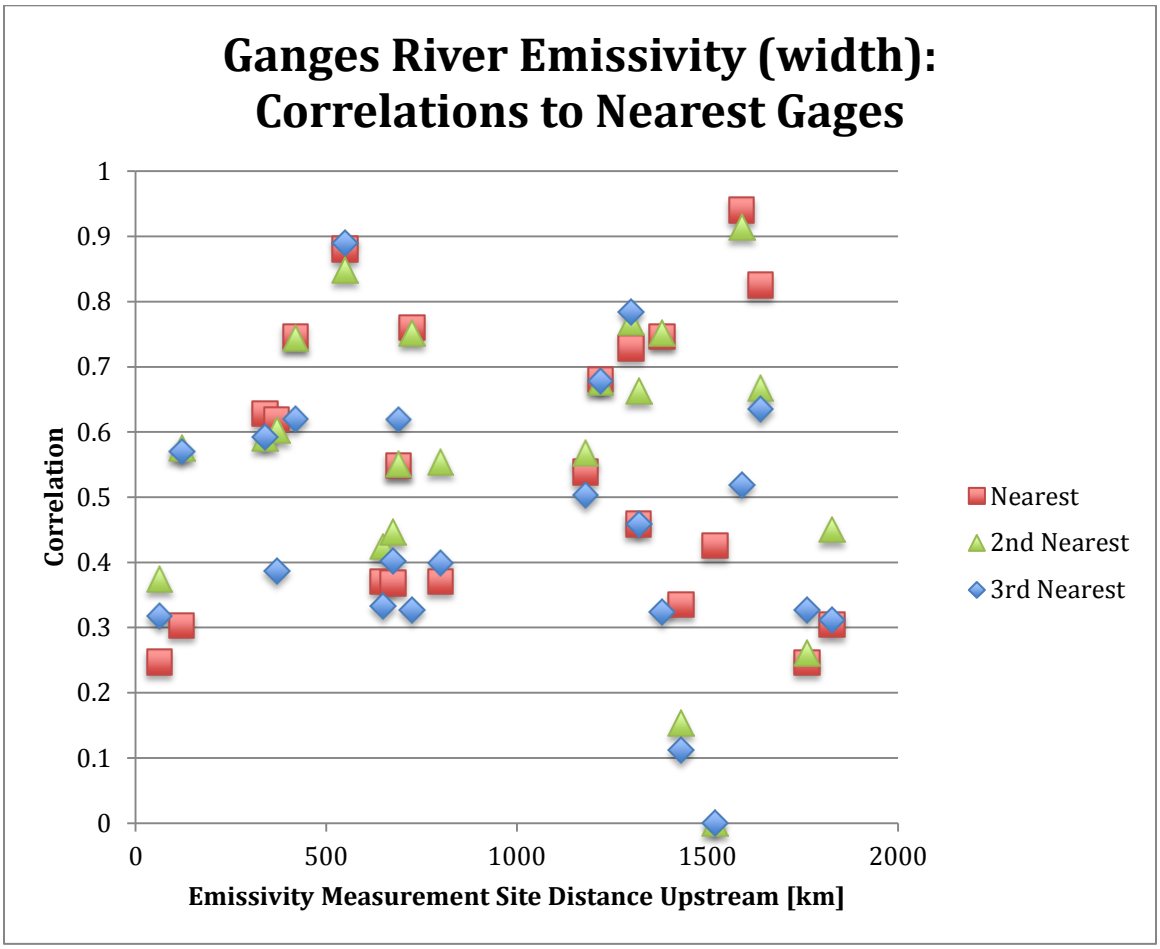


Figure 16: The correlation (y-axis) of the "width" sites with their respective nearest in situ stage gage site for locations along the Ganges main stem. The "width" site locations are labeled by distance upstream from Hardinge Bridge (x-axis).

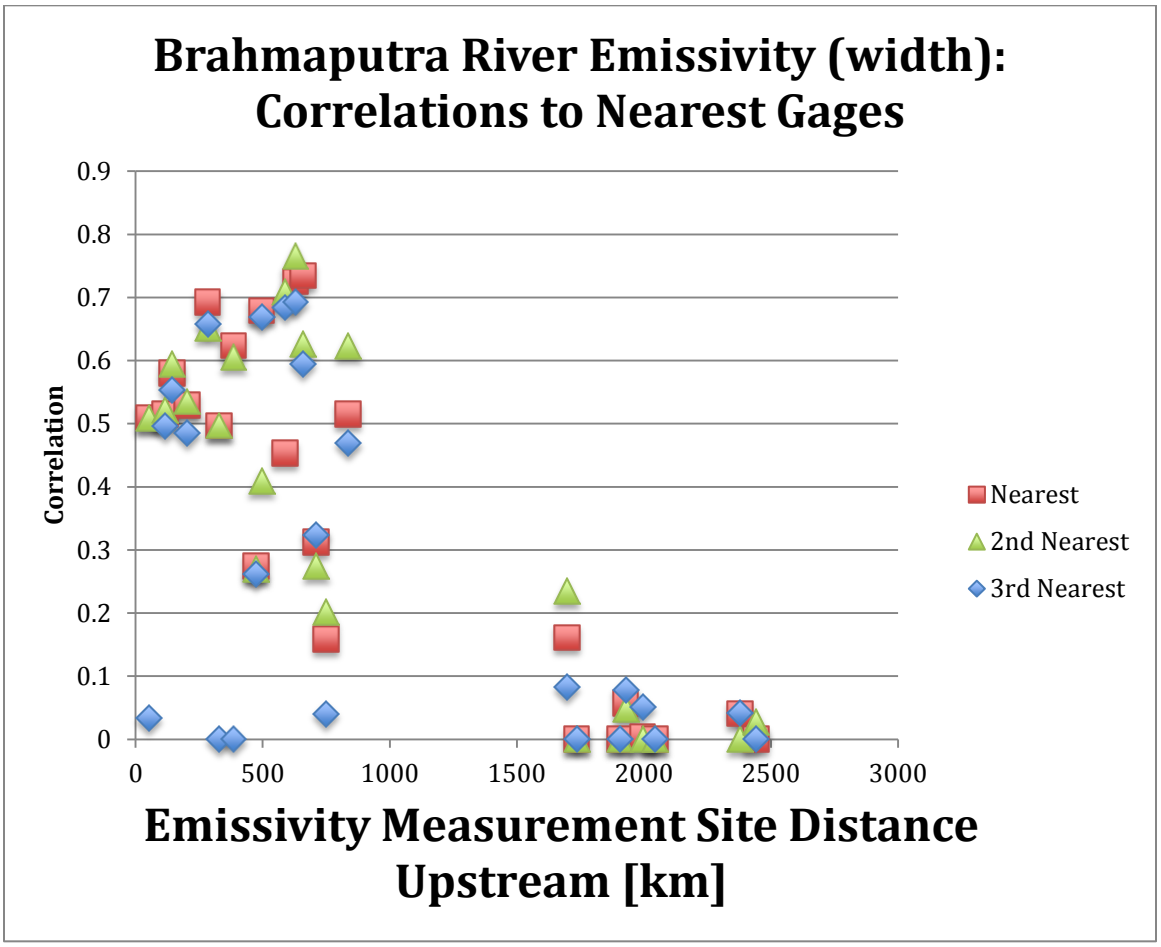


Figure 17: Same as Figure 13, but for the Brahmaputra, and locations upstream referenced to the Bahadurabad gaging site.

What we observe in Figures 13 and 14 are some very high correlations (0.9 in the Ganges, close to 0.8 in the Brahmaputra), higher than those observed with the downstream-only correlations of Figures 11 and 12. But we also see some very poor correlations (some with essentially none), which are likely due to a combination of poor "width" measurement sites, poor in situ river stage values, which potentially are mis-sited off of the main stem. But, as with Figure 12, we also observe a lack of strong correlation with "width" sites higher up in the Brahmaputra river catchment, limiting the usefulness of this data set for long-lead flood wave monitoring.

Our "take home message" is that with care, satellite "width measurements" may also have an even stronger potential to forecast local flood waves given the strengths of some of the signals. But they must be used with care, given the high potential for having no utility whatsoever. In addition, enthusiasm for forecasting must be tempered by the reality of at least a 24hr delay before the data are available for use.

For paired data comparisons used in this analysis, please see http://ral.ucar.edu/~hopson/WorldBank/satellite_error/Altimetry/ (listed by the subcatchment each falls into). Appendix D below also provides figures and more information and time-series of some of the altimetry data and their respective sites, used in this study.

Satellite River Altimetry Evaluation

In part because of the "noise" seen in the upstream-river-width-to-downstream discharge relationship, satellite altimetry has the potential to provide important additional information. Figure 15 shows a cartoon of two earlier such instruments in action (Topex/Poseidon and Jason-1) recording ocean height.

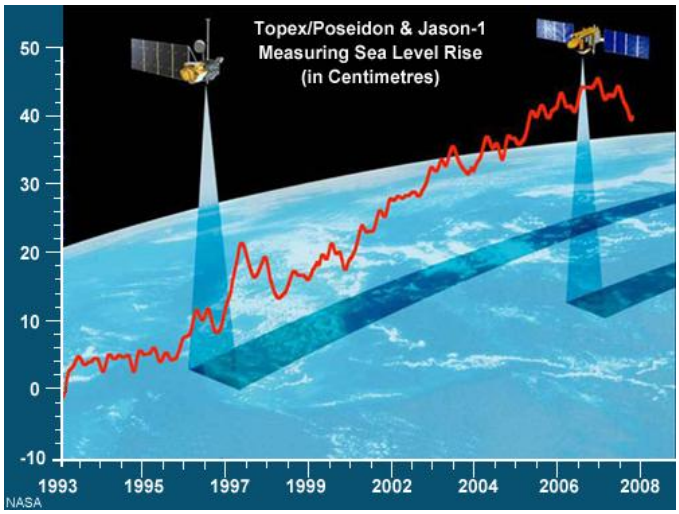


Figure 15: earlier versions of satellite altimetry instruments recording ocean heights. The same principles can also be used for heights of large river channels.

Satellite radar altimeters have a certain capacity for monitoring the variations in surface water level over the world's largest lakes, reservoirs, wetlands and river channels. As such the altimeters have played a role in basic research for the last two decades, and they are also currently contributing to applied programs via the delivery of near real time products to meet operational objectives. Just as with the DFO RiverWatch measurements discussed above, satellite altimetry measurements can also be employed to measure upstream river conditions (where river stage gauges are not available) whose information can then be used (due to advective travel-time delays) to forecast the onset of flood waves downstream. The benefits of satellite altimetry are its greater accuracy as compared to the RiverWatch data (very roughly a factor of 10), but its drawbacks are its lower sampling frequency (roughly every 10-days as compared to daily for RiverWatch at a fixed location). However, the sampling frequency for a given *river* (as opposed to fixed location) is much higher, and in particular, we believe combining both the RiverWatch and altimetry approaches together can have real power to monitor upstream conditions of rivers with no real-time river stage reporting. *An important anticipated finding of this consultancy will be to test how far upstream (as river widths diminish) the altimetry approach can retain some level of accuracy to benefit flood forecasting, and over what types of terrain and at what sampling frequency and spatial intervals.*

In addition, this approach could also test monitoring of reservoir levels, to remotely-assess water release schedules of management agencies.

For this consultancy, Sub-Consultant Dr. Charon Birkett at the Earth System Science Interdisciplinary Center (ESSIC) of the University of Maryland will be employed. Dr. Birkett is the Principal Investigator of the G-REALM program, the Global Reservoir and Lake Monitor, is a NASA/USDA funded program that ingests raw altimetric data sets and delivers water-level variations products for the world's largest lakes and reservoirs. The products are utilized by USDA/FAS regional analysts to determine i) short-term agricultural drought conditions, and ii) long-term hydrological drought status. They also help assess irrigation potential in basins for which *in situ* data is often sporadic or has restricted access. The technique of deriving lake water-level variations is similar for rivers, floodplains and wetland regions, and the G-REALM system could equally be used to output products for identified river channel crossings.

The G-REALM water level products (in graphical and text format) are created using a suite of altimeter data sets. Current operational products are based on IGDR data from the NASA/CNES Jason-2/OSTM mission, and in the near future, from the ISRO/CNES SARAL mission. The two figures below provide preliminary analysis from the use of this system to meet this particular objective of this consultancy: testing the sensitivity of satellite altimetry to Ganges river height variations.
In this section we provide an evaluation of river height, as measured through satellite altimetry, and its potential usefulness for flood forecasting purposes. Two satellite altimetry missions were the primary focus of this consultancy: Jason-2 and SARAL, whose fixed track locations crossing the Ganges and Brahmaputra basins are shown below in Figure 15. Notice the large differences in the spatial and temporal (10-day repeat for Jason-2, 35-day for SARAL) sampling frequencies.
Disappointingly, SARAL data quality were found to be insufficient in skill to have any real utility (possibly due to inaccuracies in the satellite repeat tracking consistency), so results discussed below are only in terms of Jason-2.

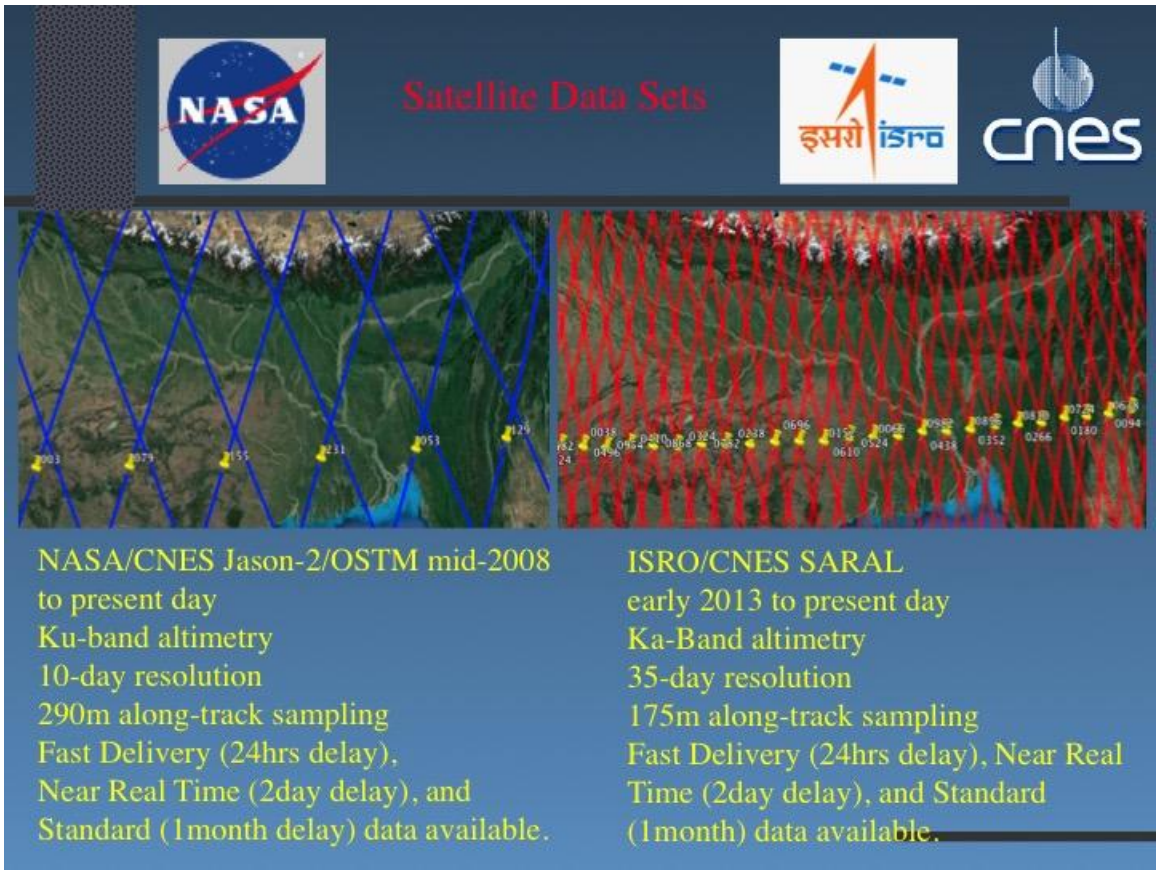


Figure 18: Jason-2 and SARAL fixed track locations crossing the Ganges and Brahmaputra basins.

In Table 1 below, we provide locations and descriptions of the sites monitored in this consultancy. Notice the description of the channel size provided in the last column, with some sites over channels of only a few 100s of meters in width.

Site	Latitude	Longitude	Width [m]	Comments
1	30.520	78.367	700	Tehri Reservoir, 70m seasonality
2	30.701	78.463	100	narrow
3	30.090	78.141	350	narrow, multiple channels, complex
4	30.177	78.186	330	noisy time-series but poss 2nd peaks?
5	30.241	78.221	650	many cycles defaulted
10	30.422	77.585	370	100m channel, 800m island
11	30.130	77.737	260	Narrow
12	30.252	77.674	380	Narrow
13	29.534	78.048	6500	multiple channels, complex, channels ~150m; time-series for floodplain only, possible 2nd peaks?

24	28.686	80.256	2000	Main channel 280m, complex land/streams
25	27.498	79.667	1750	Multiple channels, complex, channels 100-400m
26	27.410	79.624	1700	Multiple channels, complex, channels 200-700m
27	24.682	78.263	<10k	Rajghat Reservoir, 15m variability
28	26.820	82.171	3000	Two channels 300m, 500m, channel+floodplain timeseries
29	26.803	82.240	<12k	channel 1,200m, complex, steep approach angle
30	25.835	84/537	5,700	2 channels, 400m wide, island
31	25.702	84.474	1,300	1 channel 600m wide, island
32	24.554	83.940	<18k	Main channel 500m, multiple channels
50	25.279	87.112	21.7k	Floodplain, lake, 6,600m wide channel (islands)
60	23.925	89.252	4000	Main channel 600m wide
90	29.274	91.893	250-500	Few good data points
91	29.334	89.489	2380	complex, multiple channels, islands, altimeter crosses where 3 channels merge; channel widths 100-700m in total
92	29.324	89.085	750	Crossing is 2 channels, each 100m wide
93	29.845	83.682	1740	Height variability noisy; sigma0 variability clear with 2 peaks per year

Table 2: Description and locations of the altimetry sites monitored as part of this consultancy, for locations along both the Ganges and Brahmaputra. Notice the description of the channel size provided in the last column, some being quite narrow.

Analysis of sites compared to farthest downstream outflow locations

We begin the discussion and evaluation of the satellite altimetry measurements by examining their correlation to the farthest downstream gaging sites for which we have data. Because we have only one year, at best, of in situ river stage values over India (collection began February 2015), we first look at how well the upstream "width" sites correlate with downstream river discharges at two gaging sites in Bangladesh where we have a long historical record: Hardinge Bridge for the Ganges, and Bahadurabad for the Brahmaputra.

Shown below in Figures 16 are the lagged correlations of the altimetry sites with Hardinge Bridge discharge values. What we see at some sites are their highest correlations occurring at lags many days (up to 19 at site 24) preceding the observed river flow at Hardinge Bridge, with fairly high correlations (very close to 1.0 for site 30), clearly pointing out the potential for these data to provide long-lead

forecasts of flood waves many days preceding their arrival at Hardinge Bridge. However, we also see some sites with essentially no such skill (site 4).

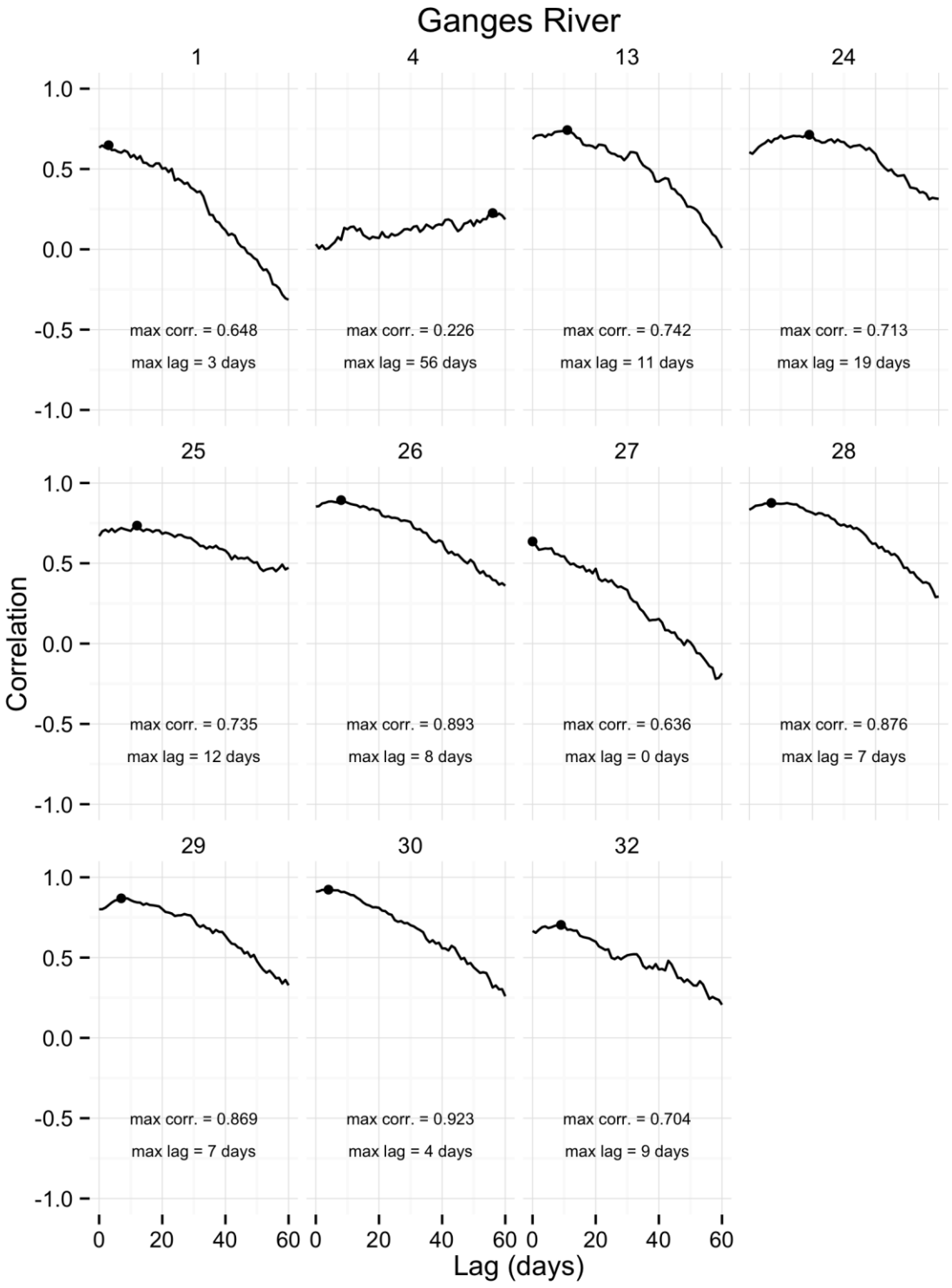


Figure 19: Lagged correlations of 11 satellite altimetry sites in the Ganges basin with river discharge observed at Hardinge Bridge.

Figure 17 below shows scatterplots of the data that went into Figure 16, but only at the optimal lag. What can be seen are some locations with a great deal of strong signal (sites 26, 28, 29, 30), and some with essentially none (site 4). Note that even for perfect measurements of river height, we would not expect perfect correlations due to the fact that the monitored river height at locations for up-catchment can only represent a small amount of the net river flow flowing into the downstream gage at Hardinge Bridge. Notice also the "classic rating curve" behavior in these scatter plots, where there is a non-linear increase in river flow with a unit increase in river stage at high flows, as one would expect for rivers topping over into flood plains.

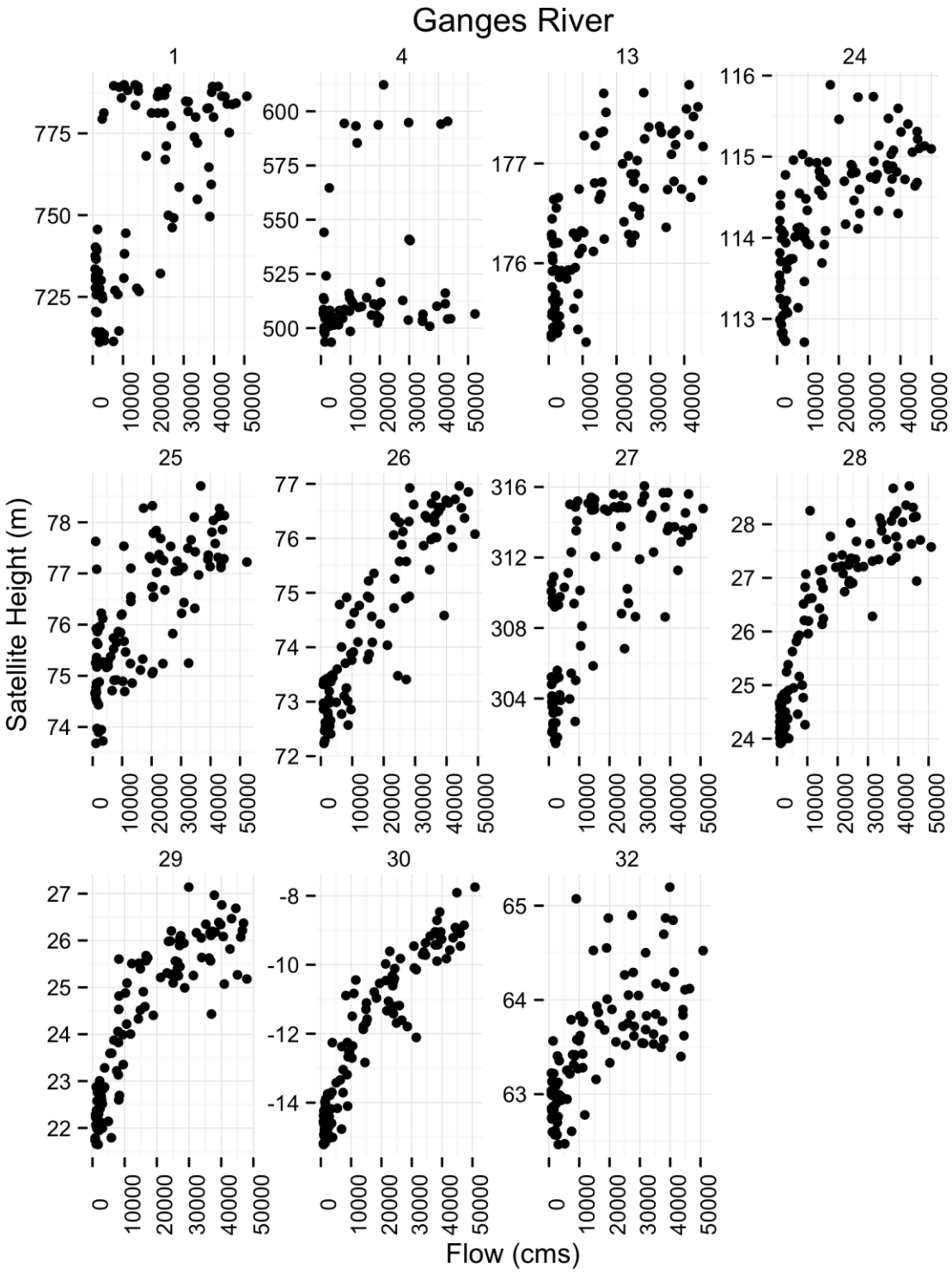


Figure 20: Scatter plots of the data shown in Figure 19, but only for the optimal lag corresponding to the peak in the correlation curve. Discharge is shown on the x-axis, with altimetry-measured river height (relative to different datum) shown on the y-axis.

In Figures 18 and 19 below for the Brahmaputra (with downstream river discharge gaging site at Bahadurabad), we see similar results to those observed for the Ganges, but with overall highest correlations not as high (i.e. peak correlations of 0.64 to 0.78), but consistent for all 4 sites monitored, and still with utility with correlations this high. Note that all four of these sites lie high up in the Brahmaputra basin, implying good potential for long-lead predictability of flood waves before they reach Bahadurabad.

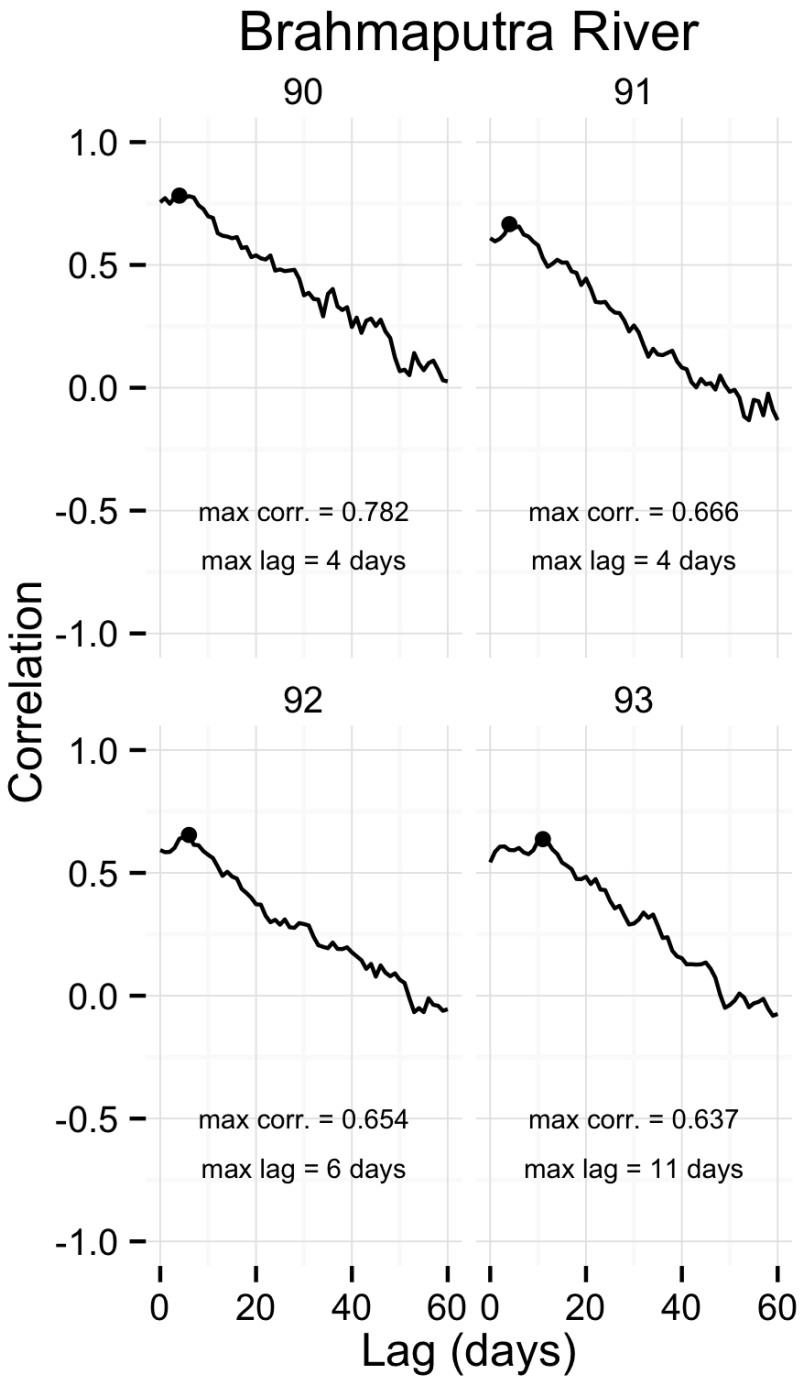


Figure 21: Same as Figure 19, but for altimetry sites in the Brahmaputra basin correlated with river discharge at Bahadurabad.

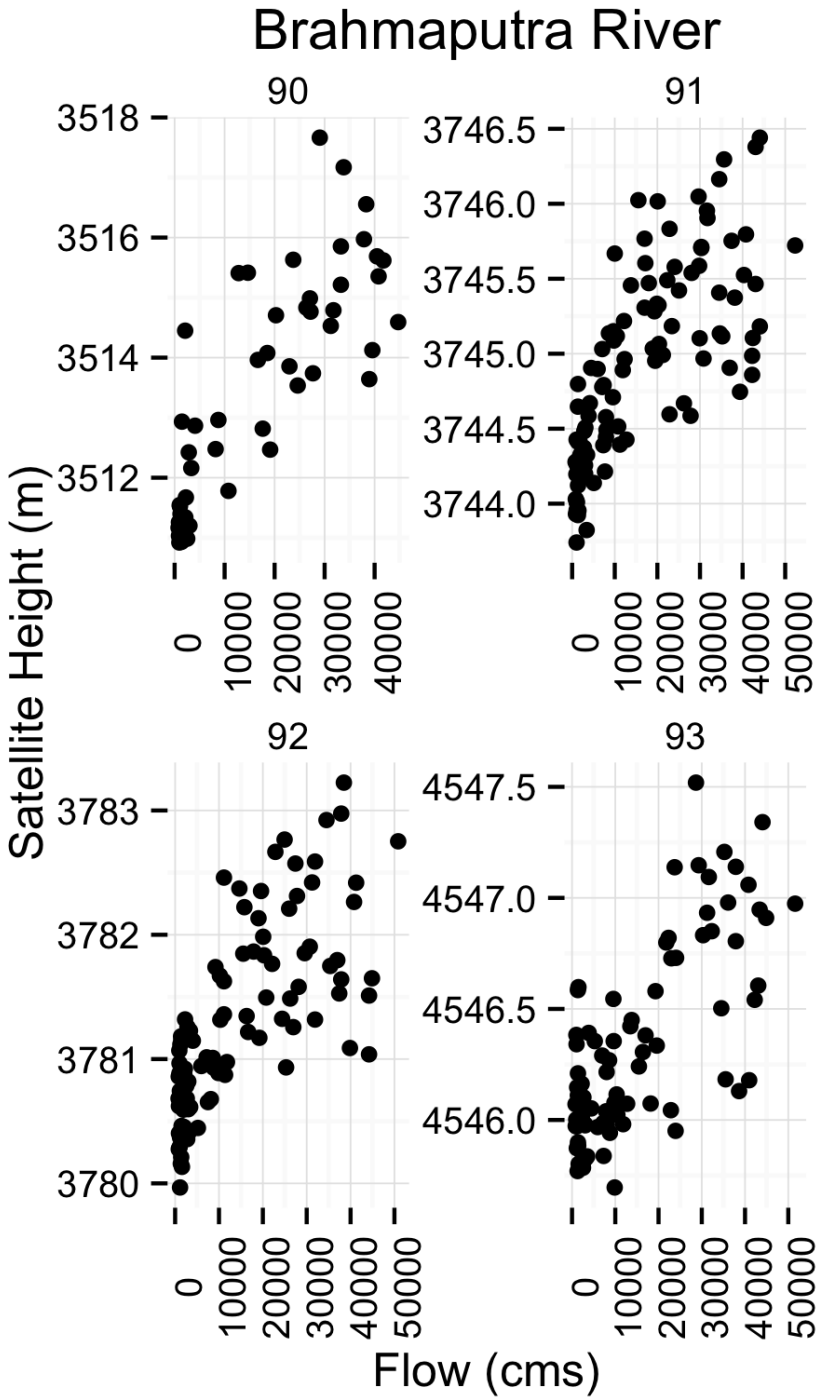


Figure 22: Same as Figure 20, but for the Brahmaputra.

Analysis of sites compared to nearest stage locations

Next, we would like to directly compare altimetry with closest in situ river gage (stage height) sites in locations within India to get an estimate of the measurement error. As shown in Figure 20, not all altimetry sites of nearest gaging sites that are that close, so we expect results to only be indicative of errors in certain locations.

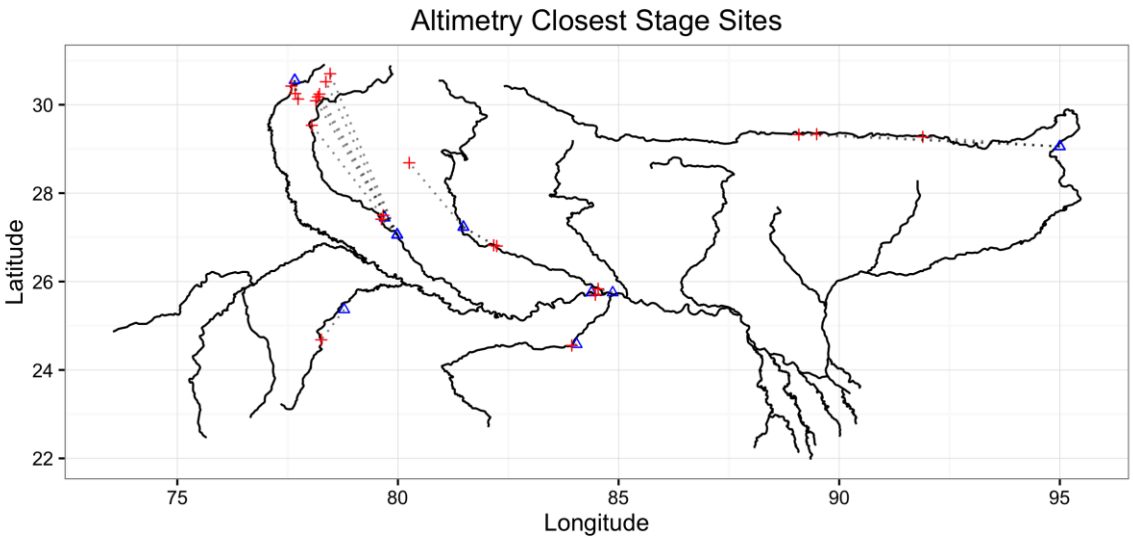


Figure 23: map of the main channels of the Ganges and Brahmaputra with the locations of the altimetry measuring sites shown with red "+" symbols, and the in situ gages correlated them with shown in blue triangles. Dotted lines connect each "width" site to the in situ gages sites compared with (ideally one gage "above" and one "below", bounding the altimetry site).

Comparisons of seven of our altimetry sites with their nearest in situ gaging sites are shown in Figure 21 below. In this figure, we plot sites found in both the Ganges (24, 28, 29, 30) and Brahmaputra (90, 91, 92). Altimetry data were paired (in time) with in situ stage measurements for the same days for the three nearest in situ sites to each altimetry site. Not all data could be paired, so some results are "blank" in Figure 21. The points in this figure are the absolute errors between the pairings after the mean was removed from each data set (i.e. in situ gages and altimetry measurements were referenced to different datum; so the average of each data set was removed before they were paired with each other). Notice that this figure has different colored dots to represent different "filters" on the altimetry signals that were tried by University of Maryland. Taking the best results for each site as being the most accurate, we see almost all sites having "errors" of less than 25cm, which is quite good given the lack of co-location of gage sites (with potentially very different stream channel cross sections).

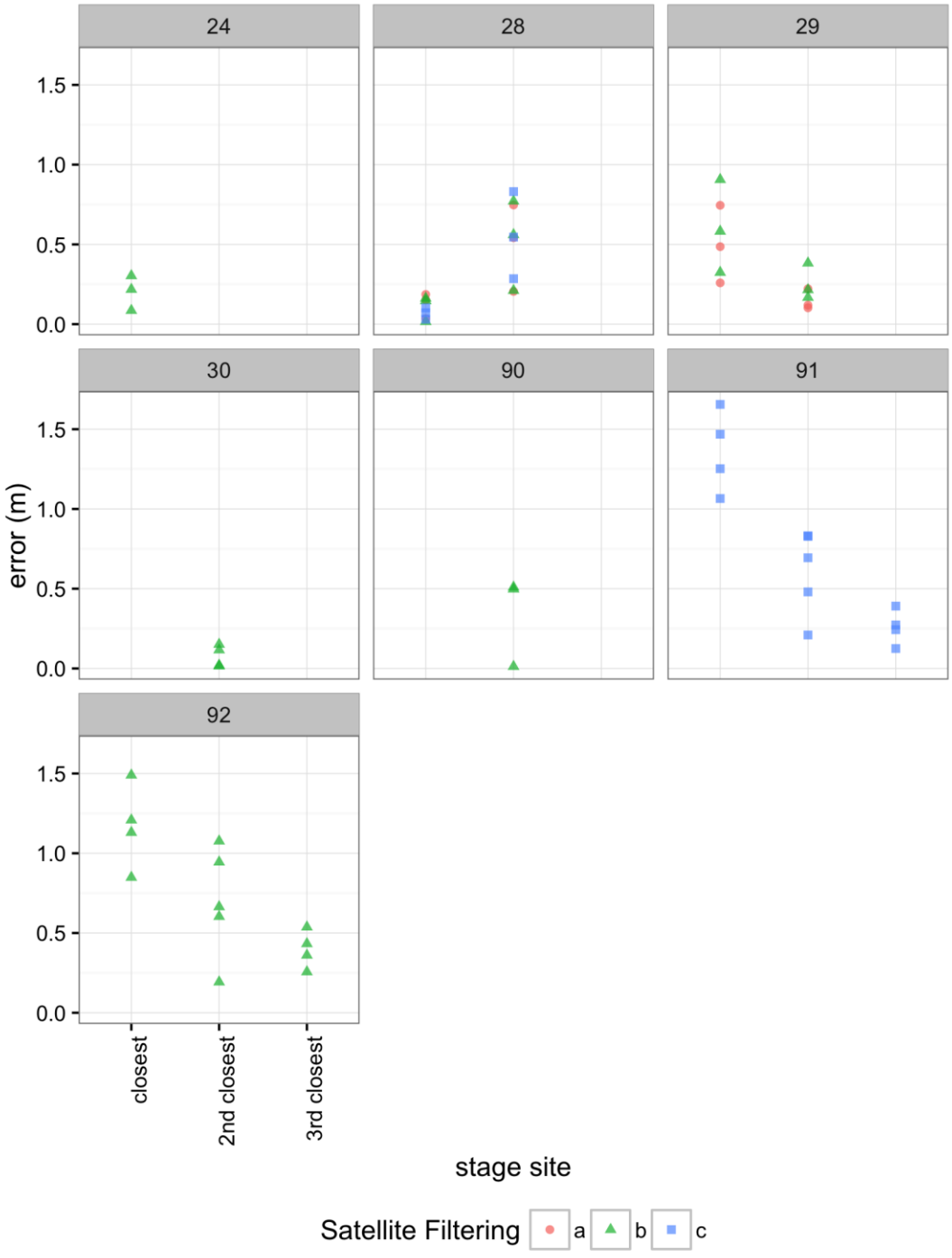


Figure 24 Comparisons of seven satellite altimetry sites (24, 28, 29, 30 in the Ganges; 90, 91, 92 in the Brahmaputra) with their nearest ("closest", "2nd closest", "3rd closest") in situ gaging locations, as shown in Figure 20. The mean of each data set was removed before they were paired (in time), and then the absolute value of the differences between the pairs is shown in this figure (y-axis, with units in meters).

For maps of all time series of the monitored altimetry sites used in this consultancy throughout the Ganges and Brahmaputra basins (listed by the subcatchment each falls into), please see http://ral.ucar.edu/~hopson/WorldBank/satellite_error/Altimetry/. Appendix D below also provides figures and more information and time-series of some of the altimetry data and their respective sites, used in this study.

Forecasting Methods

In the section we review the individual forecasting components that go into our hydrologic forecasting system. Many of the components are similar to the Climate Forecast Applications for Bangladesh (CFAB) system (Hopson and Webster, 2010). However essentially all the codes and procedures developed for this original project are newly-developed, either by providing brand new technologies, or making significant changes to the original CFAB approach. There are a number of reasons for this "rebuilding", dealing with irregular data sets with quality control issues (CWC gages and DFO-JRC river "width"), to intermittently-available data sets (altmetry), to the scale of the processing (e.g. from one weather forecast center to eight, requiring approximately 20K gridded fields were processed daily; and going from 2 river forecasting sites to close to 300). We highlight further below in this section two particular algorithms that are used extensively in this project for time-series forecasting (quantile regression), and for calibrating our TIGGE NWP ensemble precipitation forecasts (quantile-to-quantile mapping) in novel ways than had been applied before. We also introduce the stage-discharge forecasting relationship approach that we also use extensively in this project for both forecasting downstream flows based on upstream stage measurements, whether these are derived from CWC in-situ stage gaged values, altimetry measurements, or DFO-JRC river "width" measurements. We conclude this section on modeling approaches by discussing the application and investigation of the NCAR Weather Research and Forecasting (WRF) model over the Ganges and Brahmaputra domains to capture small scale "flash flood" events that would be at smaller resolution than the medium-range TIGGE ensemble NWP forecasts that are used in this project.

Quantile Regression

In this subsection, we briefly discuss a statistical forecasting and correction approach we introduced into this project that we rely on heavily in this project, called "quantile regression" (QR). In particular, we use this approach in both III. CLM Generation, and IV. Accounting for Uncertainties and Final Error Correction in the figure shown above. We defer more technical details of this approach to Appendix A – Quantile Regression below. Quantile regression has been little used in the atmospheric and hydrologic communities. We introduced this approach for this project, relaying on it heavily to generate ensemble forecasts and forecast error corrections. Similar to simple linear regression which minimizes the square error of a model to-be-fit, QR instead minimizes the absolute error, but done for each "quantile" of the distribution one is interested in. The result is that multiple models are generated, one for each quantile, and each provides a reliable probability of what values the river flow will exceedance on a daily basis (such that "flat

histograms" of the forecasts are produced – described further in an appendix below). As we apply it for our forecasting application, we ensure that the resultant ensemble discharge forecasts are no worse than the climatological average flow after the model fitting process.

Quantile-to-Quantile Mapping

In this subsection, we discuss another somewhat similar approach to model fitting as described by QR above, called "quantile-to-quantile mapping (q-to-q), but with a different twist such that this approach is less stringent on model fitting when the regressand (observed values) may be error-prone themselves. This approach is applied to II. Statistical Rendering shown in the above figure. More technical details of this approach are discussed in Appendix B – Quantile-to-Quantile Mapping below.

Generally, all practical hydrologic models require some form of calibration arising from reasons such as incomplete knowledge of watershed properties and the necessary parameterization of transport at unresolved scales. Through the hydrologic model parameter calibration process, some of the additional errors in the watershed inputs (rainfall, in particular) can also be implicitly reduced (e.g. runoff errors arising from an over-bias in rainfall forcing can be implicitly minimized by increasing the parameterization of evapotranspiration). However, such bias reduction through calibration does not occur if there are relative errors between more than one input of the same physical quantity. In particular, using both weather forecasts and observationally-based estimates of precipitation concurrently to generate unbiased discharge forecasts, requires that these two data sources maintain their statistical similarity such that there are no relative biases between their statistical moments (mean, variance, and skewness, in particular). Note that quantile regression (discussed above) could also provide the required bias reduction on a forecast-day-by-forecast-day bases. However, because we feel that the merged satellite precipitation product (i.e. "observations") inherently have their own error, we hypothesize it is more optimal to only utilize a more "global" statistical adjustment of the TIGGE precipitation forecasts to require them to only be drawn from the same probability distribution function as the observations. Quantile-to-quantile mapping is designed to exactly do this.

Stage-Discharge Relationship Flood Forecasting

This approach builds from the "in situ and forecasting rating curves" we discussed in our last report (found under http://indiawbg.rap.ucar.edu/Ratefit/) that derived a "rating curve" of downstream river flow (at a future "optimal" lagged time) from upstream stage measurements; which is something very useful for hydrologic forecasting purposes. Using these rating curves with observed stage (whether from

CWC stage gage, altimetry measurements, or river "width" measurements), forecasts of downstream discharge can be produced at 0 day (current condition) to 16 day forecasts (or beyond).

The figures and figure captions immediately below shows conceptually the idea behind these forecasts and how they are combined, where we are capitalizing on travel-time delays from upstream to downstream locations, with the forecasts themselves provided at: http://indiawbg.rap.ucar.edu:8080/QForecast/)

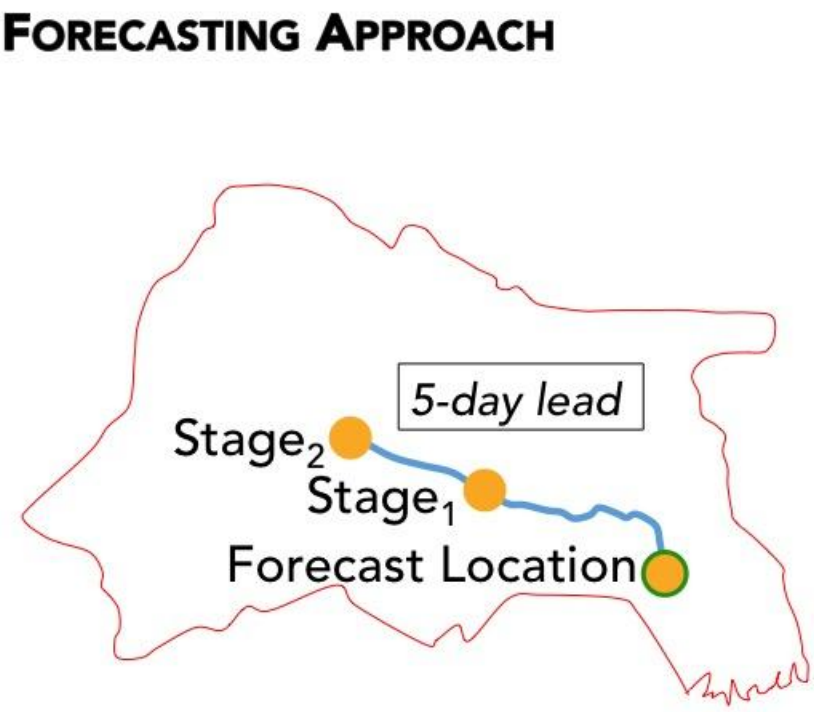


Figure 25: conceptual idea of using upstream stages (stage 1 and stage 2) to forecast a downstream location based on a 5-day flood wave travel time from the most upstream to downstream locations.

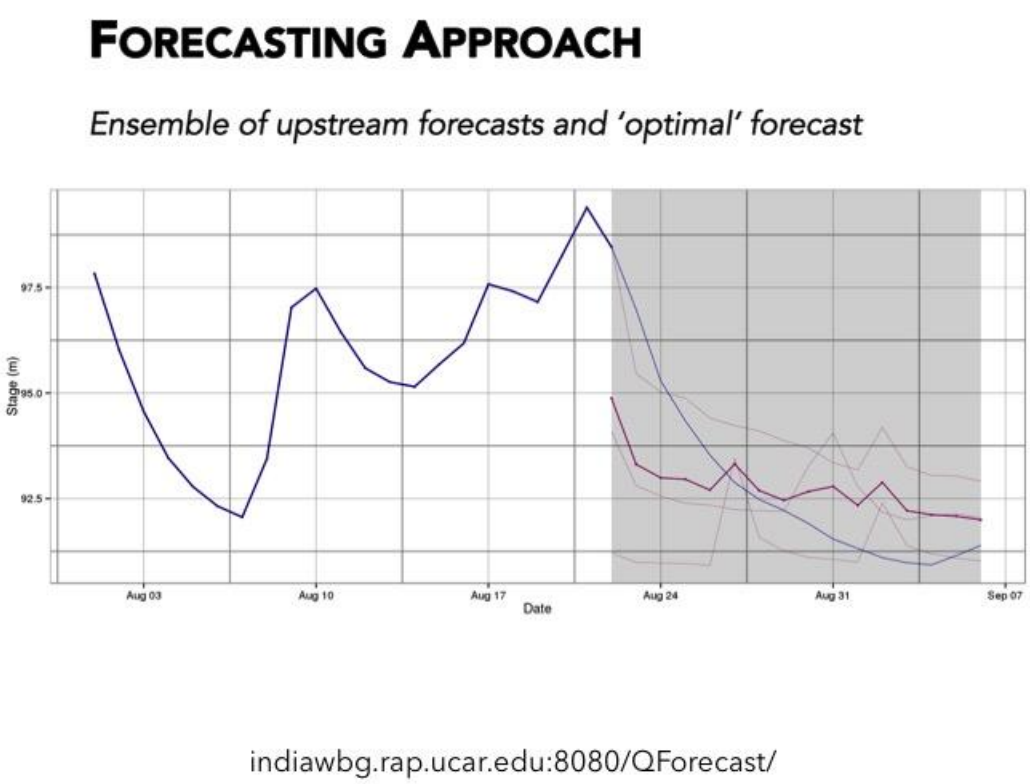
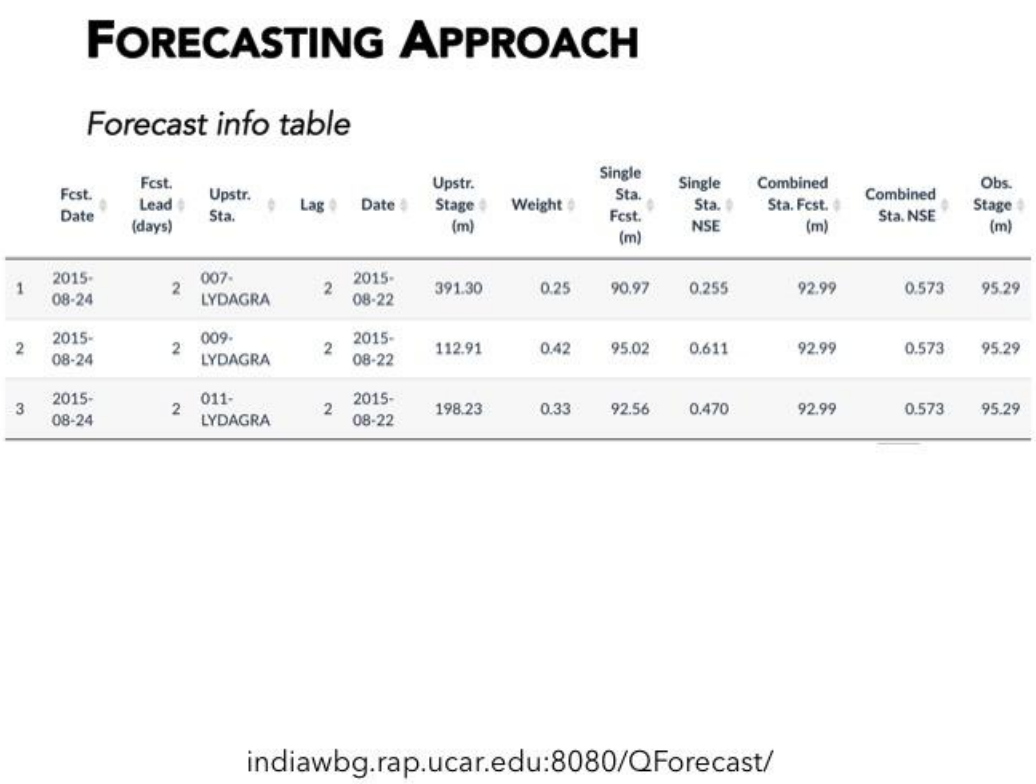


Figure 26: depending on the number of upstream gauges, multiple forecasts can be estimated, one for each gage.



FORECASTING APPROACH

Forecast info table

	Fcst. Date	Fcst. Lead (days)	Upstr. Sta.	Lag	Date	Upstr. Stage (m)	Weight	Single Sta. Fcst. (m)	Single Sta. NSE	Combined Sta. Fcst. (m)	Combined Sta. NSE	Obs. Stage (m)
1	2015-08-24	2	007-LYDAGRA	2	2015-08-22	391.30	0.25	90.97	0.255	92.99	0.573	95.29
2	2015-08-24	2	009-LYDAGRA	2	2015-08-22	112.91	0.42	95.02	0.611	92.99	0.573	95.29
3	2015-08-24	2	011-LYDAGRA	2	2015-08-22	198.23	0.33	92.56	0.470	92.99	0.573	95.29

indiawbg.rap.ucar.edu:8080/QForecast/

Figure 27: the error of each upstream gauge (an example of which is shown here from our website provided) forecast is available directly from the derived "rating curves".

FORECASTING APPROACH

$$Q_{fcst} = \frac{\sum_1^n C_i (Stage_i + a_i)^{n_i} \cdot \frac{1}{\varepsilon_i}}{\sum_1^n \frac{1}{\varepsilon_i}}$$

Figure 28: the error from the derived rating curves of upstream gauges is then combined in the manner shown in this figure, to generate an "optimal" combined forecast.

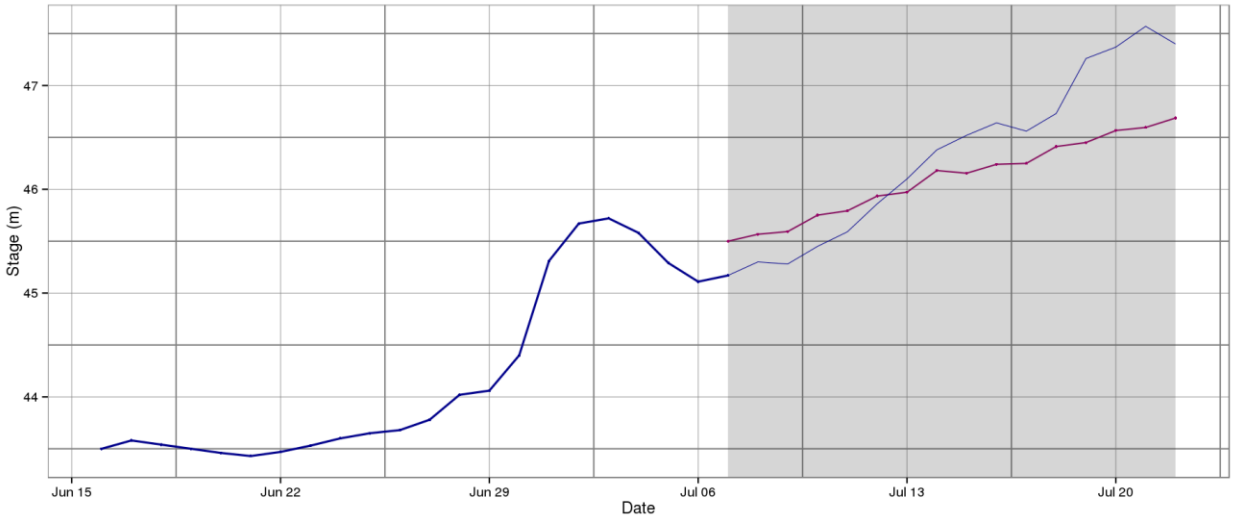


Figure 29: results of our combined upstream stage forecast for one forecasting site. The blue is the observer, red is the forecast, and the grey represents to the forecast period.

Future Improvements

- Combine with additional upstream stage gauge measurements –
- addition of satellite altimetry and satellite river width being tested to improve skill

- Satellite precipitation could improve forecasting by accounting for additional streamflow entering between the stations

- By using additional proper weighting, more than one measurement from upstream gages can be used simultaneously for each forecast

- Employing an error correction algorithm (forecasting from the most recent time series of observed errors) can lead to further improvements

Improving Flood Forecasting using WRF Modeling

As part of this project, we wanted to explore the skill of employing the NCAR Weather Research and Forecasting (WRF) model over the Ganges and Brahmaputra domains to capture small scale "flash flood" events that would be at smaller resolution than the medium-range TIGGE ensemble NWP forecasts that are used in this project. Although in this study one of the WRF modeling domains is large enough to cover large portions of the Ganges and Brahmaputra, the primary focus of this ongoing study is on the northeast Haor regions of Bangladesh. The northeast Haor regions are typical lowlands which experience flash floods every year. If flash floods occur before the harvesting of the only major crop in the region (Boro rice), it becomes critical. According to the officials of the Department of Agricultural Extension, Boro rice crop on over 22, 000 hectares out of 4,00,000 hectares was damaged by flash floods this year. Providing accurate early warning of the flash with sufficient lead time is essential for that region. At present, the early warning and flash flood mechanism of that region is neither efficient nor timely. For more context into this issue and this area, see http://bangladeshchronicle.net/2016/04/rice-crop-damaged-by-flash-floods-in-haor-belt/.

This effort was in collaboration with our colleague Prof. Saiful Islam at the Bangladesh University of Engineering and Technology (BUET). This project funded Prof. Islam's visit to NCAR on Jan 23 – Feb 3, 2016 to attend the Weather Research Forecasting Tutorial 2016, to meet with potential collaborators for improving flood forecasting in the GBM region (with special focus on the northeast Haor regions of

Bangladesh), and to gain access to the yellowstone super-computer for the rainfall forecasting effort.

Preliminary conclusions from this ongoing study of one recent rainfall event are as follows:

- This study is investigating the ability of a cloud-resolving WRF to reproduce the convective cells associated with a heavy rainfall event over the Haor region of Bangladesh;
- The ability of WRF model with a highest resolution of 18 and 6 km horizontal grid spacing to predict heavy rainfall near Sylhet in the northeastern part of Bangladesh on 17 and 18 April 2010 was evaluated;
- The model underestimated the strength of the storm in general in terms of the rainfall. The simulated rainfall was 94 mm day^{-1} for the outer domain and 98 mm day^{-1} for the inner domain, but the observed (rain gauge) amount was 161 mm day^{-1}. TRMM retrieved 24 hours rain amount was 100 mm, which compares favorably with the WRF simulations, but was also less compared to rain gauge observation.

Future work for this study is to continue to compare WRF simulations to ongoing small scale intense rainfall to generate more statistics. As well, we are planning to validate the WRF simulations with radar driven rain rates provided by the Bangladesh Meteorological Department's S-band weather radar.

More information on this effort is provided below in the appendix below.

Verification and Extreme Event Predictability

In this section we evaluate

Verification and Evaluation Metrics

In this section we briefly discuss the three evaluation metrics we apply to our discharge forecasting systems to assess their skill and calibration: the Brier Score, the Brier Skill Score, and the Rank Histogram.

The Brier Score

The Brier score is a function that measures the accuracy of probabilistic predictions. It is applicable to tasks in which predictions must assign probabilities to a set of mutually exclusive discrete outcomes. The set of possible outcomes can be either binary or categorical in nature, and the probabilities assigned to this set of outcomes must sum to one (where each individual probability is in the range of 0 to 1).

The Brier score can be thought of as a measure of the "skill" of a set of probabilistic predictions. More precisely, across all items $i \in 1...N$ in a set *N* predictions, the Brier score measures the mean squared difference between:

- The predicted probability assigned to the possible outcomes for item *i*
- The actual outcome O_i

Therefore, the lower the Brier score is for a set of predictions, the more skillful the predictions are. Note that the Brier score, in its most common formulation, takes on a value between zero and one, since this is the largest possible difference between a predicted probability (which must be between zero and one) and the actual outcome (which can take on values of only 0 and 1).

The Brier score is appropriate for binary and categorical outcomes that can be structured as true or false.

The most common formulation of the Brier score is

$$BS = \frac{1}{N}\sum_{t=1}^{N}(f_t - o_t)^2$$

In which f_t is the probability that was forecast, o_t the actual outcome of the event at instance t (0 if it does not happen and 1 if it does happen) and N is the number of forecasting instances. In effect, it is the mean squared error of the forecast. This formulation is mostly used for binary events (for example "rain" or "no rain").

For our application, the binary forecasting event we use is whether the discharge will be above or below the upper 75th percentile of flow values (i.e. a moderately extreme flow event), so we test for how well the forecast system can anticipate high flow values.

The Brier Skill Score

Forecast Skill (or skill score) is a generic term referring to the accuracy and/or degree of association of prediction to an observation or estimate of the actual value of the predictand (i.e., what is being predicted). The term 'forecast skill' can be used both quantitatively and qualitatively. In the former case, skill could be equal to a statistic describing forecast performance, such as the correlation of the forecast with observations. In the latter case, it could either refer to forecast performance according to a single metric, such as our case of the Brier Score. Skill is often, but not exclusively, expressed as the relative representation that compares the forecast performance of a particular forecast prediction to that of a reference, benchmark prediction—a formulation called a 'Skill Score':

$$SS = \frac{A_{forc} - A_{ref}}{A_{perf} - A_{ref}}$$

For our case, our skill metric "A" is the Brier Score (BS, described above), and the "reference forecast" is the climatological probability of the event occurring (for the upper 75% events, this "climatological probability" is 0.25). In this case, a perfect forecast results in a forecast skill metric of zero, and skill score value of 1.0. A forecast with equal skill to the reference forecast would have a skill score of 0.0, and a forecast which is less skillful than the reference forecast would have unbounded negative skill score values.

The Rank Histogram

The rank histogram is not a verification method per se, but rather a diagnostic tool to evaluate the *spread* of an ensemble. The underlying assumption is that the ensemble member forecasts are distributed so as to delineate ranges or "*bins*" of the predicted variable such that the probability of occurrence of the observation within each bin is equal. For each specific forecast, the bins are determined by ranking the ensemble member forecasts from lowest to highest. The interval between each pair of ranked values forms a bin. If there are N ensemble members, then there will be N+1 bins. The outer bins, lowest and highest – valued, are open-ended.

Rank histograms are prepared by determining which of the ranked bins the observation falls into for each case, and plotting a histogram of the total occurrences in each bin, for the full verification sample. It is desirable to use a large sample of cases so that there is likely to be some occurrences in each of the bins. The examples

shown in the figure below highlight ensemble forecasts that are too wide (left panel) and too narrow (right panel).

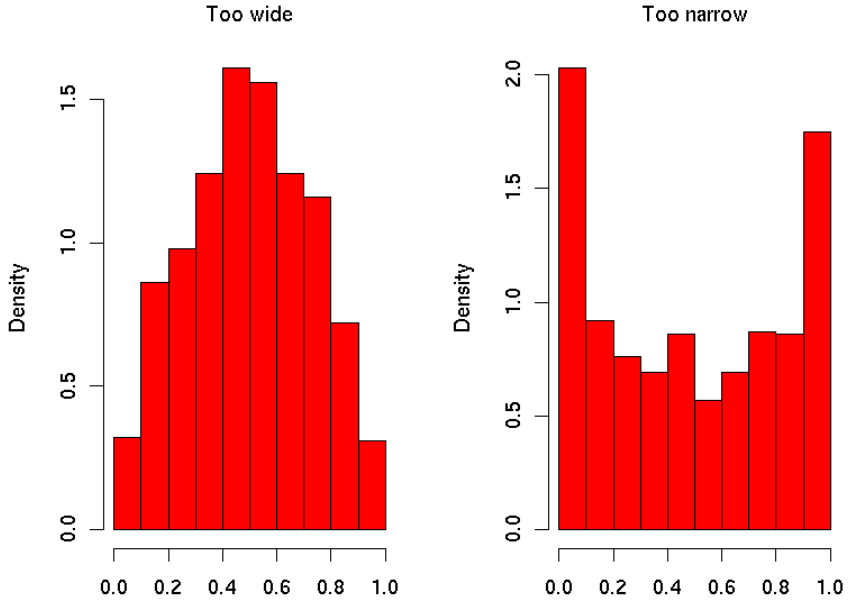


Figure 30: Troubled rank histograms: panel on the left shows an ensemble forecast system whose ensembles are too dispersed (the observation falls too often in the middle of the distribution), while the panel on the right would be for a too narrow ensemble forecasting system. For a perfect ensemble forecasting system, rank histograms would be "flat" (i.e. the observation falling into each bin the same number of times).

Assessing Forecasts of Extremes: sensitivity versus false discovery rate

In this section, we present another metric for assessing the utility of our probabilistic rainfall or river flow forecasts -- focusing on how well a forecast system can forecast the occurrence or "hits" of extreme events, while minimizing the forecasting of "false alarms" – can we forecast many extreme events that actually occurred without forecasting many false alarms? Here, we're interested in events falling above the upper 75th percentile (whether it be rainfall or river flow) at each of our forecasting sites.

Our forecasting systems produce ensemble forecasts (say, with 9 members, 20 members, etc.) of continuous variables (i.e. *mm/day* of rainfall or m^3/s of river flow). We can translate these ensembles into probability forecasts if we prescribe a physical threshold (e.g. say, *50mm/day* corresponding to the upper 25th percentile of events) that we're interested in, and ask how many of these ensemble members (for a given forecast lead-time: 1-day in advance, 2-days in advance, etc.) exceed this threshold. As an example, if 2 out of 9 (total) ensembles exceed this threshold, we have a probability of exceedance of 2/9 = 22%.

For our use here, we look at the upper 75th or greater percentile extreme events, and look at cases where our ensemble forecasts provide a probability of exceeding

this threshold of 50% or greater. Based on these two thresholds, we can then examine two statistics. First, we can ask how many times "hits" occur (e.g. rainfall above the 75th percentile) and how many "false alarms" – because we've set the probability threshold at 50% or greater, we would except at least as many hits as false alarms – do we see this is indeed the case? Secondly, looking at all of the 75th percentile or greater events that do occur, how many of these do occur when our forecast system is forecasting 50% or greater probability, and many corresponding "misses" occur – do we at least capture a significant fraction of all of these extreme events? One way to view the forecast system's skill is in terms of a "contingency table" as shown below, where we say the event was forecasted if the system forecasted 50% or greater, and not forecasted if the probability was less than this.

Event forecast	Event Observed	
	Yes	No
Yes	*Hit (h); C*	*False alarm (f); C*
No	*Miss (m); L*	*Correct non-event (c)*

Table 3: Contingency table, along with Cost-Loss (C, L) value assessment for each category.

Forecast Assessment in Terms of Economic Benefits

The purpose of this validation protocal is to examine the economic benefits of providing reliable, ensemble-based flood forecasts of rainfall and discharge forecasts in the Ganges and Brahmaputra basins within the cost-loss model reported elsewhere in the literature (Mylne 1999; Richardson 2000; Wilks 2001; Zhu et al. 2002). Of course the primary utility of operational flood forecasts has been to save lives: advanced lead-time warnings providing time for adequate evacuation preparedness in the case of predicted major flooding. But beyond this focus, the availability of such forecasts can have significant economic impact, e.g. agricultural decision making, such as when to plant, apply pesticides, and to harvest to optimize yields (Webster et al. 2006).

We give a brief explanation of the cost-loss analysis applied here; however for a more detailed explanation of similar analyses, the reader is referred elsewhere (Mylne 1999; Richardson 2000; Wilks 2001; Zhu et al. 2002).

Consider the contingency table of the relative frequency of outcomes (h, m, f, c) of an imperfect forecast given in Table 2 above, and also the associated economic expense of taking action based on the forecast information (C, L). If the flood does or does not occur, but the forecast system says there will be such an event and the user (farmer, shopkeeper, etc.) acts on this information, his/her mitigating actions have an associated cost C; if the flood occurs but the forecast was not successful, the user will experience the loss L. See Table 1 below.

Based on these relative frequencies and associated costs, consider the expected expense to the user of long-term use of this (imperfect) forecast information. If the user has access to the forecast information, and takes mitigating steps based on this information, his/her expected long-term expense E_f will be

$$E_f = hC + fC + mL. \quad 1)$$

If, however, the user either does not have access to the forecast information, or does not choose to act upon it, his/her expected (climatological) expense E_c will be

$$E_c = Min[\mathrm{o\,L}, \mathrm{C}], \quad 2)$$

where o is the climatological frequency of the event. If, however, the forecast system were perfect and the user acted on such a hypothetical system's information, the expected expense E_p to the user would be

$$E_p = oC. \quad 3)$$

Based on these expected costs, a skill-score of expected value V to the user could be

$$V = (E_f - E_p)/(E_c - E_p). \quad 4)$$

Note that a perfect forecast system would produce an associated value V to the user of 1, while a forecast system with skill no better than a climatological forecast would produce a value of 0.

Substituting into Eq (4) using Eqs (1) - (3), noting $h + m = o$, and defining the cost-loss ratio $r = C / L$, the expression for V can be written as

$$V = (Min[o, r] - r(h + f) - m)/(Min[o, r] - or). \quad 5)$$

Note that each user is expected to have his or her own unique cost-loss ratio r, that itself may vary for different time periods.

The above analysis is based on a binary forecasting system: the forecast system says the event either will or will not occur. However, how should a user act if instead given (well-calibrated) probabilistic forecasting information giving likelihoods of severe flooding events? In this case, it can be shown that the optimal use of the forecast information is for the user to act when the forecast probability of the event

to occur equal to or greater than the user's unique cost-loss ratio *r* (Richardson 2000).

We can then apply this economic cost-loss framework to the probabilistic flood forecast information operationally provided by this project. Prescribing the flooding events (in the form of rainfall and river discharge thresholds) as those falling above the 75% climatological value for each location, what would be the economic benefits to a variety of users (i.e. covering the range $0 £ r £ 1$) who would have acted on this information to mitigate their economic losses? The figure below shows this relative value *V* versus cost-loss ratio ($r = C / L$) of providing 1-day ahead ensemble catchment-integrated rainfall forecasts (top panel) compared to the relative value of issuing just the ensemble mean forecast (bottom panel), for the catchment above the gauge location 001-LBDJPG. All value assessed relative to "climate conditions". The values falling above zero show that the probabilistic forecast has clear economic benefit over the a "climatological" forecast over a large portion of the cost-loss range.

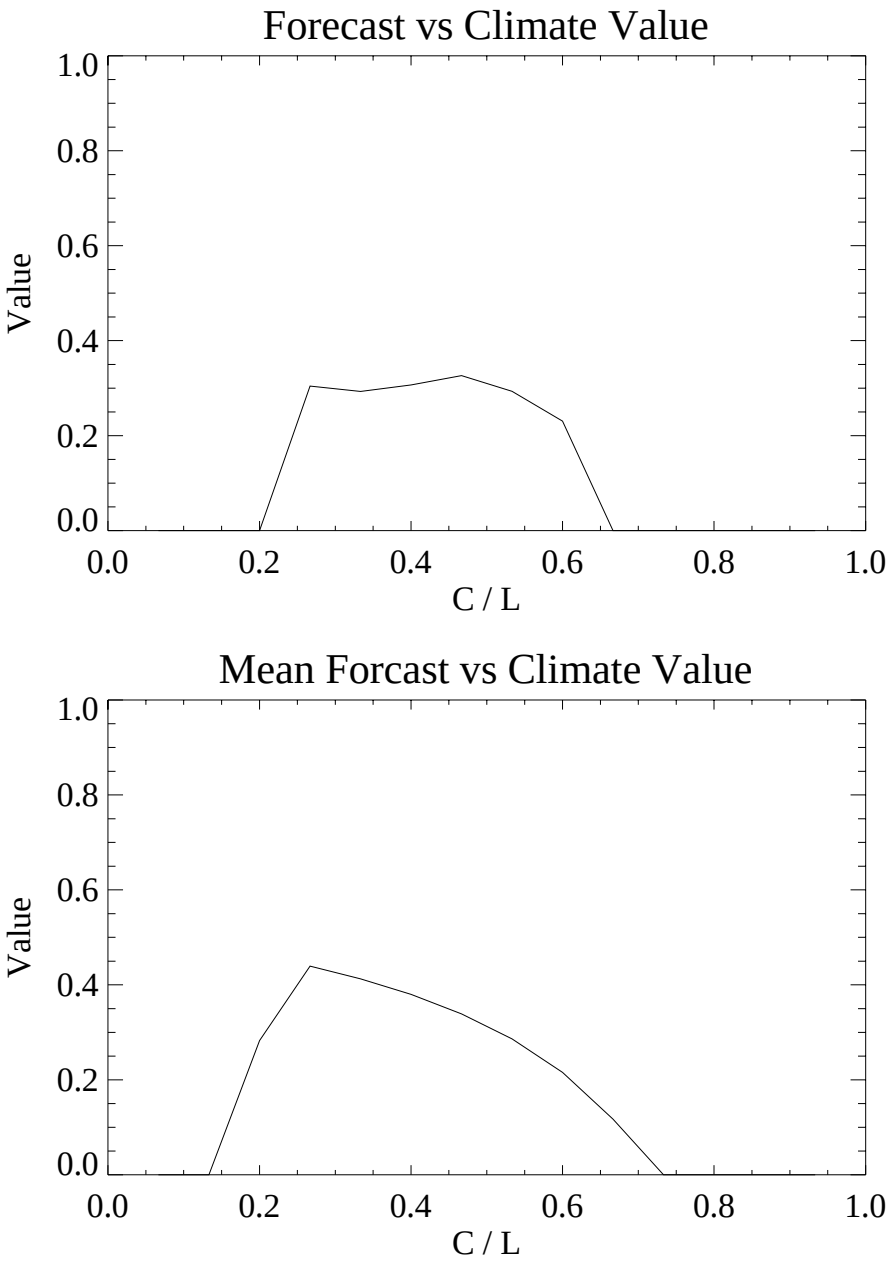


Figure 31: The relative value versus cost-loss ratio (C/L) of issuing 1-day ahead ensemble catchment-integrated rainfall forecasts (top panel) compared to the relative value of issuing just the ensemble mean forecast (bottom panel), for the catchment above the gauge location 001-LBDJPG. All value assessed relative to "climate conditions".

Stage-Discharge Relationship Flood Forecasting

In this subsection, we discuss results and analyses of the upstream "rating curve"-derived forecasting approach (discussed in the earlier section above). Results are provided under http://ral.ucar.edu/~hopson/WorldBank/StageDischargeRelationshipFloodForecasting/.

Based on Brier skill score analyses, our results show that forecasting skill based strictly on this technology can vary from 6 days (one example: http://ral.ucar.edu/~hopson/WorldBank/StageDischargeRelationshipFloodForecasting/001-MGD5PTN_stage_fcst_brier.png) to essentially no forecast lead-time at all, likely due to limited upstream stage gauging above the site of intereset, and of poor rating curve analyses from limited (in size and quality) upstream stage gage information.

CFAB-based Forecasting

In this subsection, we discuss results and analyses of the upstream "rating curve"-derived from the CFAB-based forecasting approach (discussed in the earlier section above). Results are provided under http://ral.ucar.edu/~hopson/WorldBank/CFABbasedForecasting/. We find that this forecasting system is much more robust than the stage-discharge relationship flood forecasting just discussed. Here we find, again, poor skill for certain forecasting gauging sites, but on average significantly more skillful forecasts out to 16-day lead-times in some cases (e.g. http://ral.ucar.edu/~hopson/WorldBank/Verification/AllGages/007-mgd4ptn/tfq_007-mgd4ptn_SSbrier_fc.ps), with typically the shortest in-length forecasts still being skillful out to 3-day lead-times. Additionally, we consistently see ensemble forecasts that are reliable by virtue of producing flat rank histograms (see, as one example, http://ral.ucar.edu/~hopson/WorldBank/Verification/AllGages/007-mgd4ptn/tfq_007-mgd4ptn_f7d_rankhist.ps).

Altimetry-based Forecasting

In this subsection, we discuss results and analyses of the upstream "rating curve"-derived forecasting approach (discussed in the earlier section above). Results are provided under

http://ral.ucar.edu/~hopson/WorldBank/AltimetryBasedForecasting/. Viewing the brier skill scores found under http://ral.ucar.edu/~hopson/WorldBank/AltimetryBasedForecasting/alt_fcst_brier.png, sadly we see essentially no skill at all lead-times for the strictly altimetry-based forecasts for the Brahmaputra at Bahadurabad. However, the forecasts for the Ganges River at Hardinge Bridge are skillful (although only slightly) out past 11 day lead-times.
ddd

Displays and Forecast Translation to User Communities

In this section we provide an evaluation of river "width", as measured from surface some of the altimetry data and their respective sites, used in this study.

Displays of Ensemble Numerical Weather Prediction and Satellite Precipitation

Because of the essential importance rainfall plays in generating runoff, many flooding events can be anticipated just by locating areas of severe rainfall, especially relevant for flash flood events occurring over smaller catchment spatial scales. With this motivation, we are generating an hourly-updated display of rainfall occurring over sub-catchments of the Ganges and the Brahmaputra (follow the link in our project's web site under "Rainfall Accumulation"). This interactive map displays 24-hour accumulated precipitation forecasts, 5-day precipitation forecasts, and satellite-based rainfall estimates for each watershed. Zoom in to see accumulated precipitation for more detailed sub-catchments. The forecasts are based on ECMWF, UKMET, and Canada Met Centre control run predictions (the display also shows NCEP forecasts – however, the dissemination of these forecasts by NCEP is currently suspended, so the forecasts are not visible). The forecast data are provided on a 2-day delay such that the displayed forecast over the next 24-hrs (0Z to 0Z) is based on forecast model runs initialized two days ago (however, note that for operational purposes, ECMWF, NCEP, and Canada Met Service forecasts are available "real-time" without delay).

The satellite precipitation is a merged product of NASA TRMM, NOAA CMORPH, and JAXA GSMaP precipitation and is provided immediately after the 24-hr accumulation is complete.

Click on a watershed to get the total precipitation value in *mm/day*. Click the *Layer List* icon to display forecasts for other lead times or the most recent satellite precipitation.

In this section, we provide an overview of displays and recent improvements to our THORPEX TIGGE ensemble weather prediction and satellite-based precipitation products. This site was also reported on in our last report, however we have made some significant modifications to these displays, and provide an overview of the process to access this information below, with the site itself found at: http://indiawbg.rap.ucar.edu/precip/index.html.

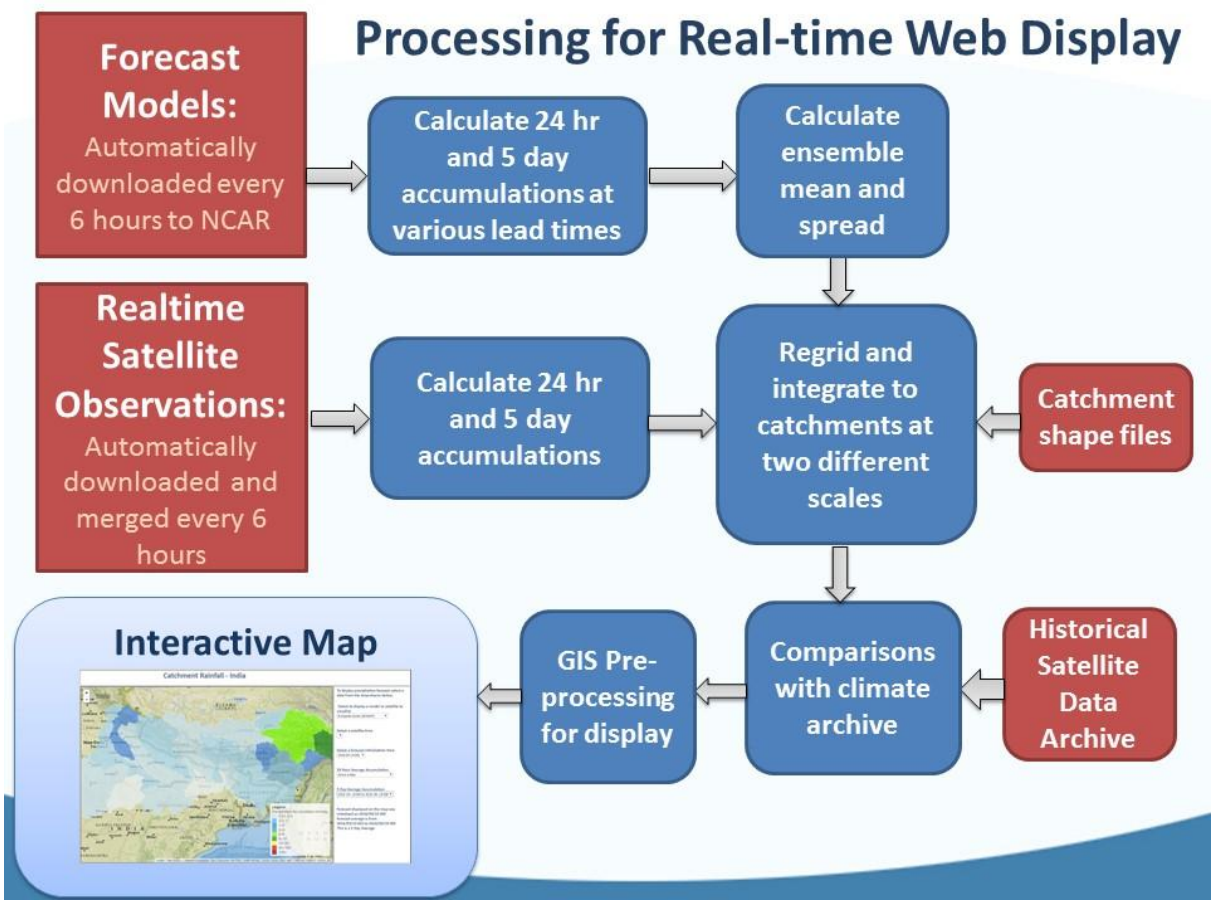


Figure 32: Overview of the processing and display of catchment-integrated real-time satellite precipitation and TIGGE ensemble precipitation forecasts over the Ganges and Brahmaputra catchments.

This interactive web display, shown in outline form in the figure above, gives users an easy way to visualize and compare precipitation accumulation for different weather models and satellite products. The near real time web display (currently) incorporates ensemble precipitation forecasts from four of our eight THORPEX-TIGGE forecast centers and our merged NASA TRMM, NOAA CMORPH, and JAXA GSMaP satellite precipitation product. This portal is an interactive web display that expands Geographic Information Systems (GIS) mapping capabilities to visualize forecast precipitation accumulations averaged over watershed basins. Models and satellite estimates are calculated to 24 hour and 5 day average accumulations. The models that are displayed are the European ECMWF, the UK UKMET, the Canadian CMC, and the USA NCEP models. Each model has a different number of ensemble members. The ensemble average and other statistics are aggregated to a catchment and sub-catchment basin. By default the larger catchments are displayed. When zoomed in the sub-catchment become visible for a more detailed view. Ensemble

average per watershed is being displayed on the map and ensemble statistics for the watershed are available through a popup window.

This portal is built using a number of technologies. All forecast and satellite products are exported to text comma separated values file (.csv). This file format allows data to be saved in a table structured format but as a text file. The text file forecasts are produced every 24 hours. Once the forecast and satellite precipitation products are created, a GIS python script is automatically run to convert these data into a spatial GIS-friendly format. The web application is built using a free open-source JavaScript API called Leaflet. This library is ideal for powerful data driven visualizations. This portal takes advantage of the most appropriate tools and technologies to easily display the information from the weather forecast and satellite products. In what follows, we step through the process of displaying the different products found on this site.

Step 1 – Select a model or Satellite product

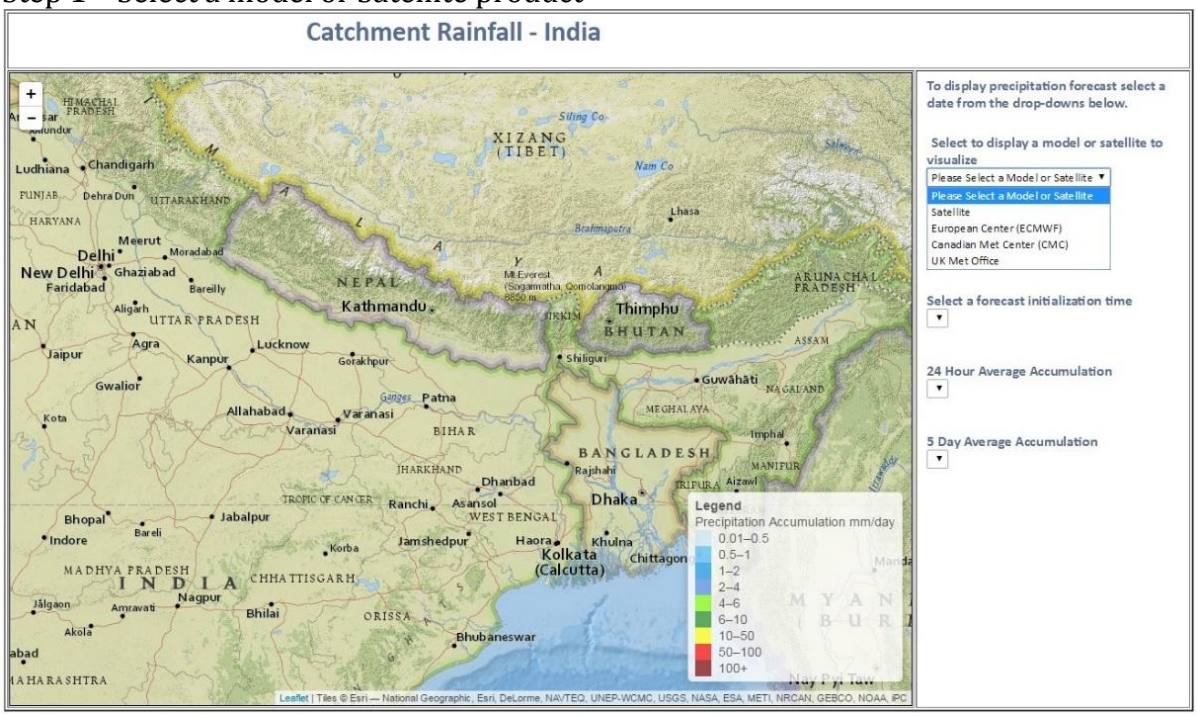


Figure 33: Step 1 – Select a model or Satellite product

Step 2 – Select the forecast initialization time.

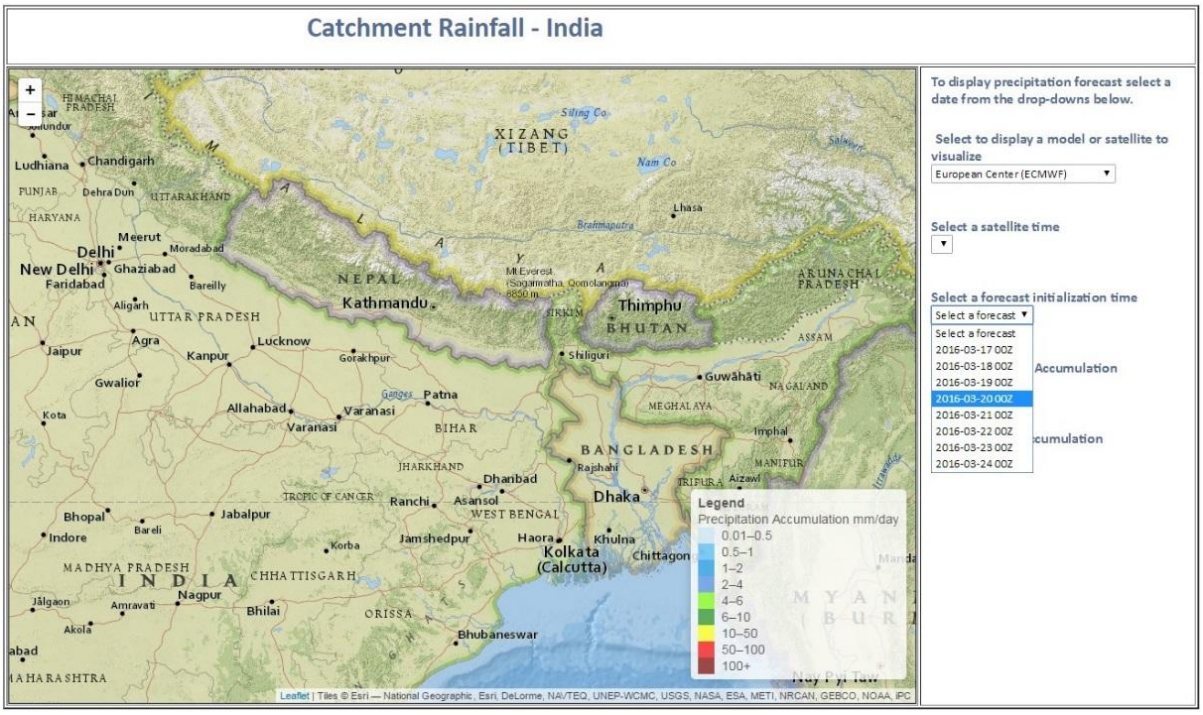


Figure 34: Step 2 – Select the forecast initialization time.

Step 3 – Select the 24 hour or 5 day accumulation date.

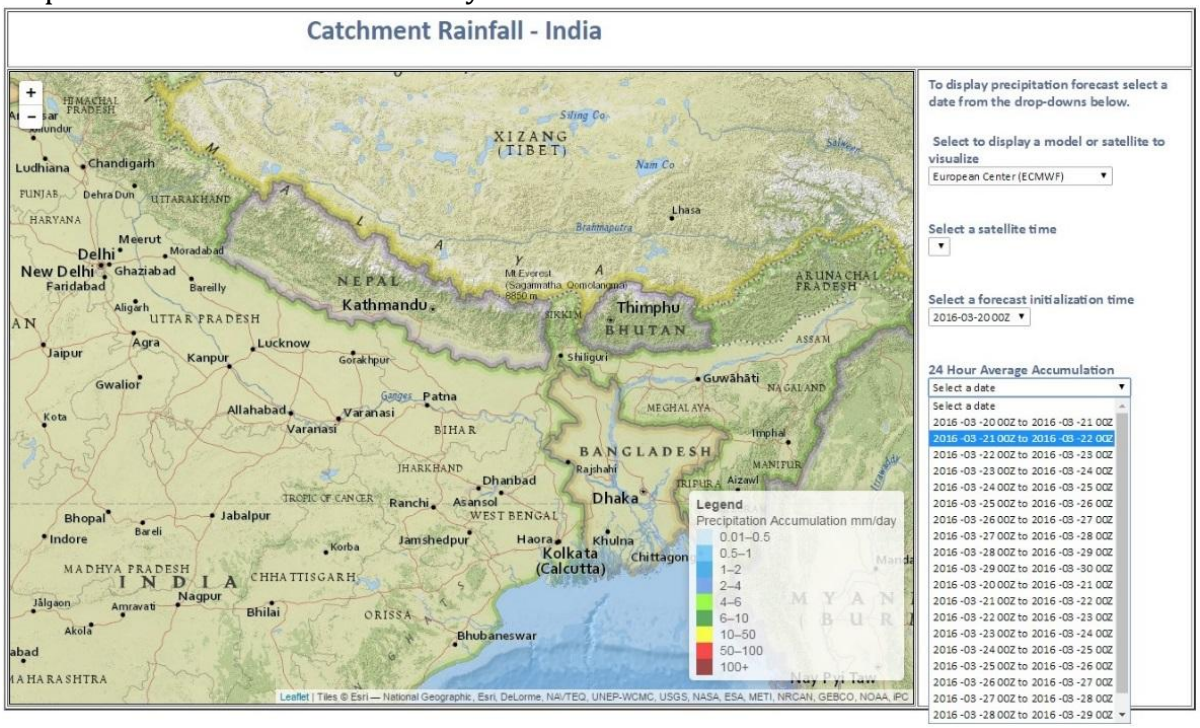


Figure 35: Step 3 – Select the 24 hour or 5 day accumulation date.

Result - An interactive web map is displayed with zoom and pan functionality.

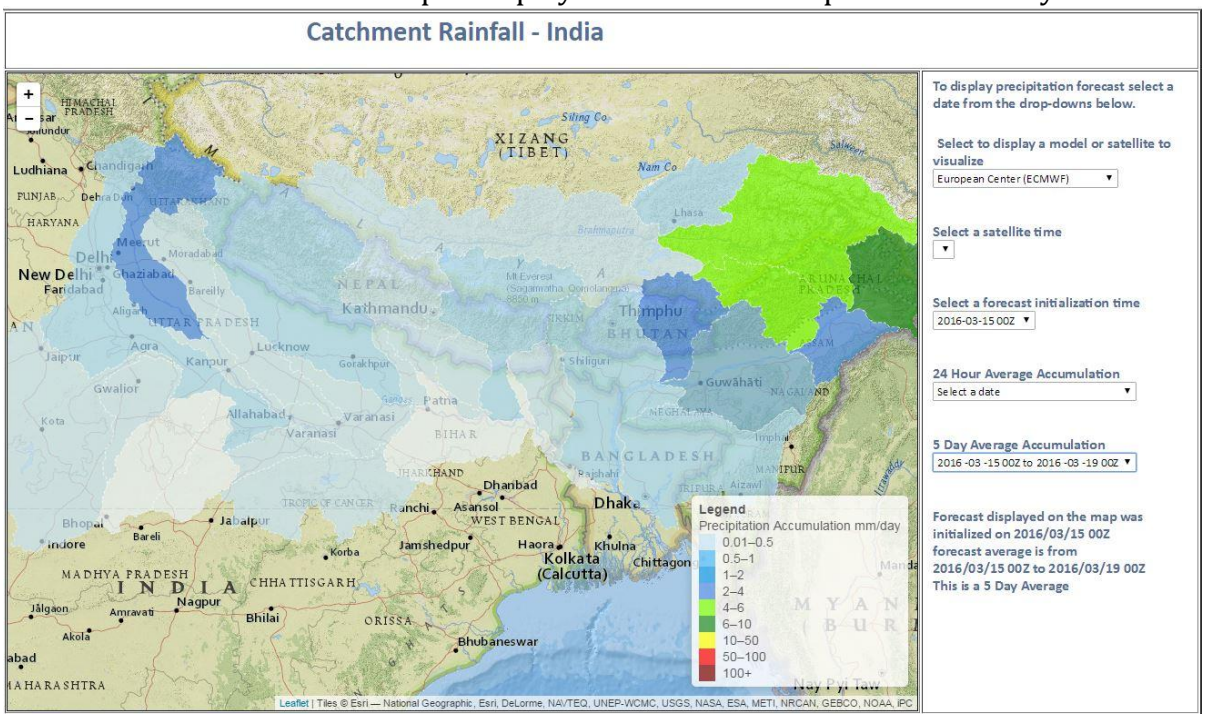


Figure 36: Result - An interactive web map is displayed with zoom and pan functionality.

By clicking on a watershed a popup window appears with more detailed information about the model ensemble statistics.

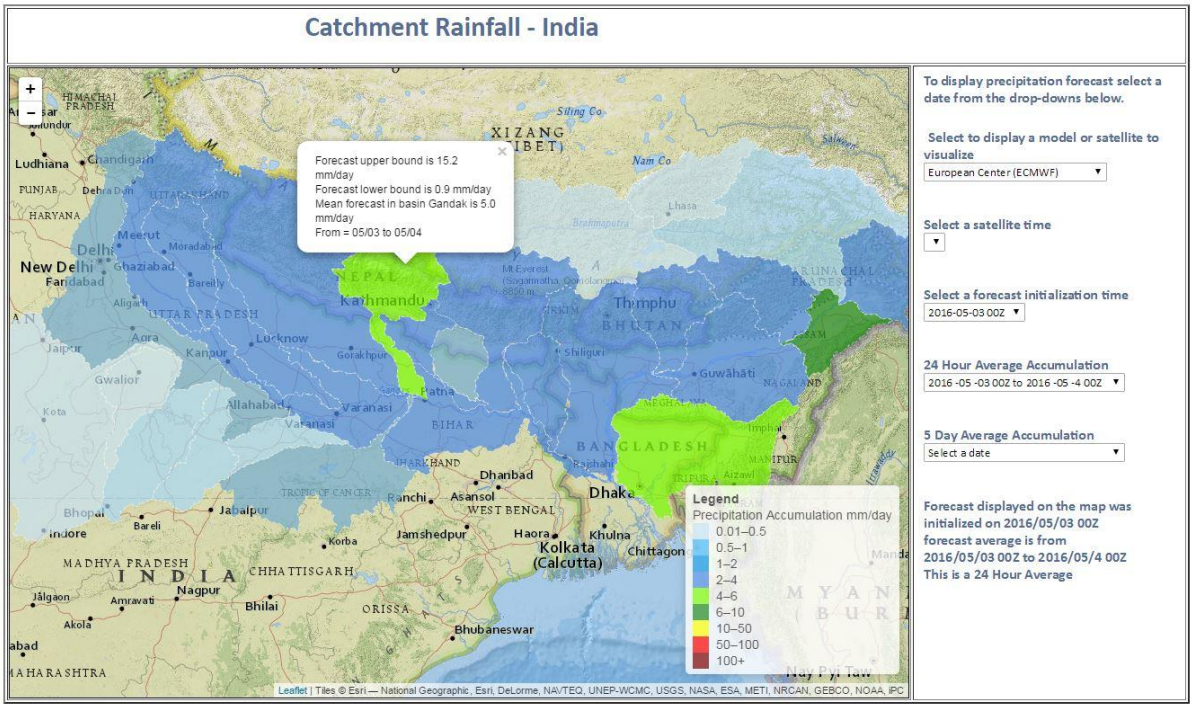


Figure 37: By clicking on a watershed a popup window appears with more detailed information about the model ensemble statistics.

When zoomed in the sub-catchments are displayed for a more detailed view.

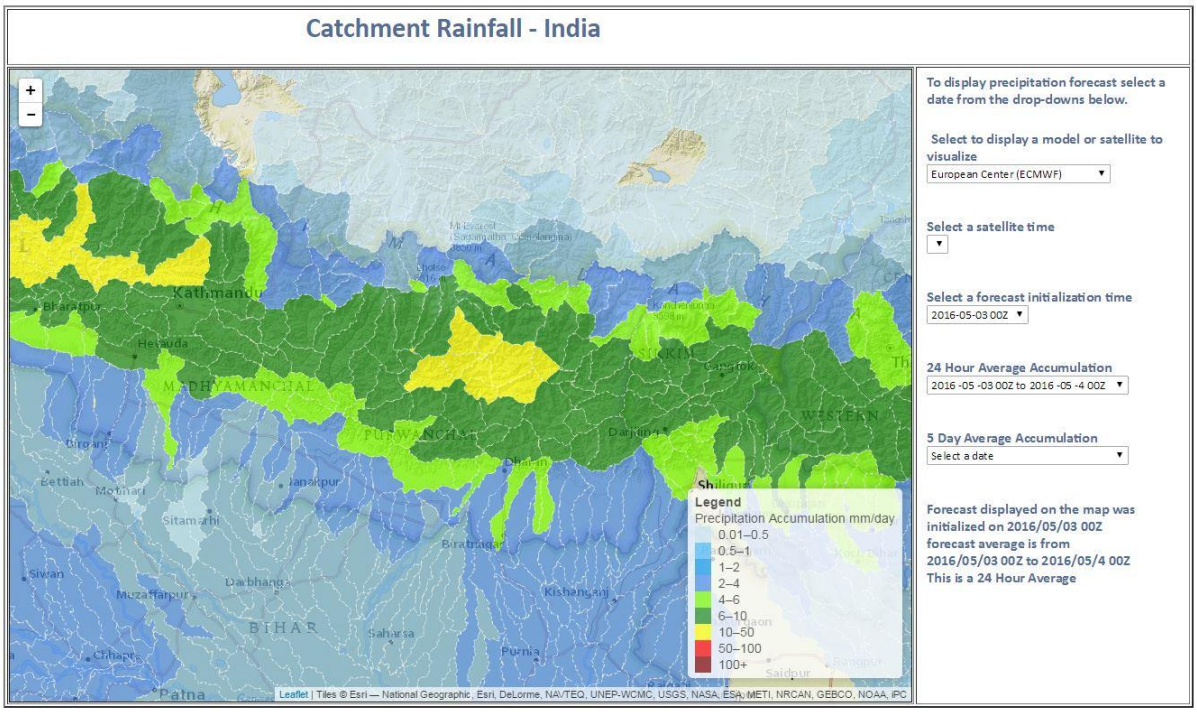


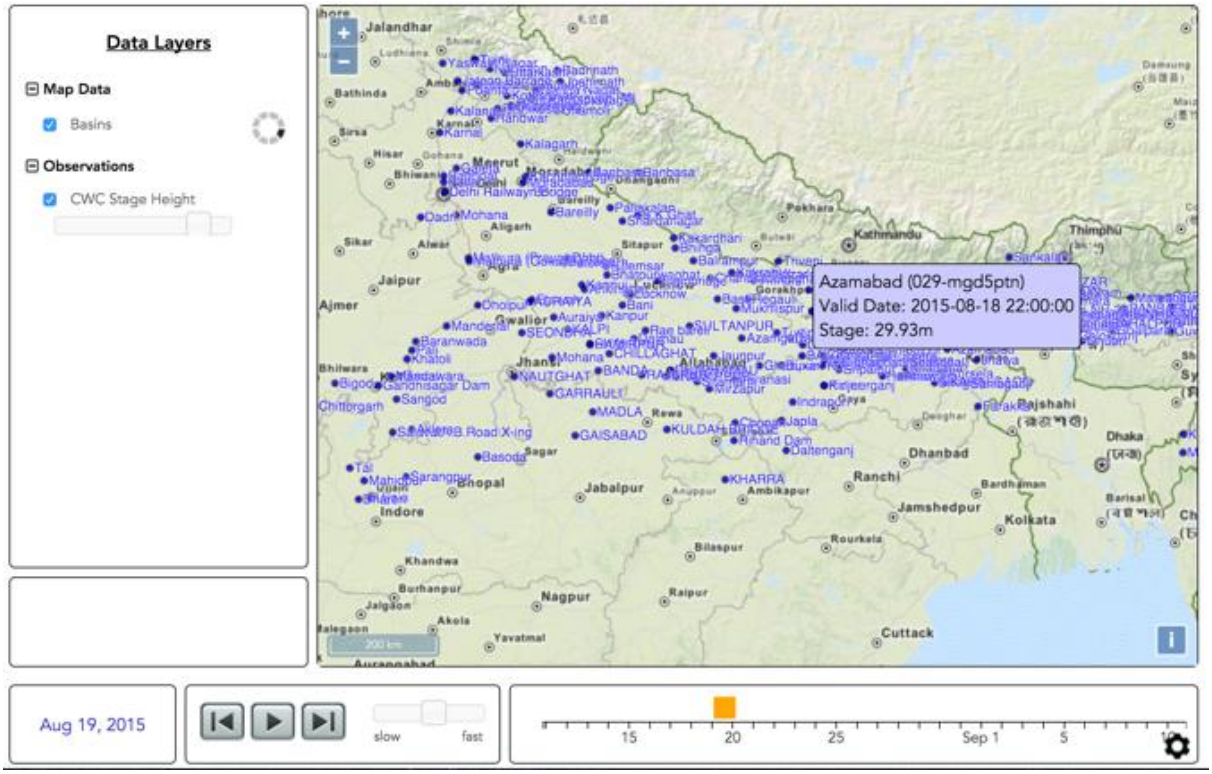
Figure 38: When zoomed in the sub-catchments are displayed for a more detailed view.

Stage and Discharge Data – Current Displays

In this subsection, we discuss the processes that led to the display of CWC in situ river stage gauging data found at http://indiawbg.rap.ucar.edu/display/. This site also provides displays of our operational ensemble river forecasts under a pilot testing link. This simple application could have significant improvements to enhance the displays, depending on user needs, some of which is discussed in the following section, "Future Display Enhancements".

Current Data Display

The current web display is map-based in order to visualize the geographic extent of the river basins and to allow gridded model forecasts to be easily displayed. Stations in the display may be selected by clicking them with the mouse. When selected, a station displays its detail information in a pop-up window. Selected stations are also automatically queried to generate plots of observations, ensemble forecasts, and quality control analyses at that location.



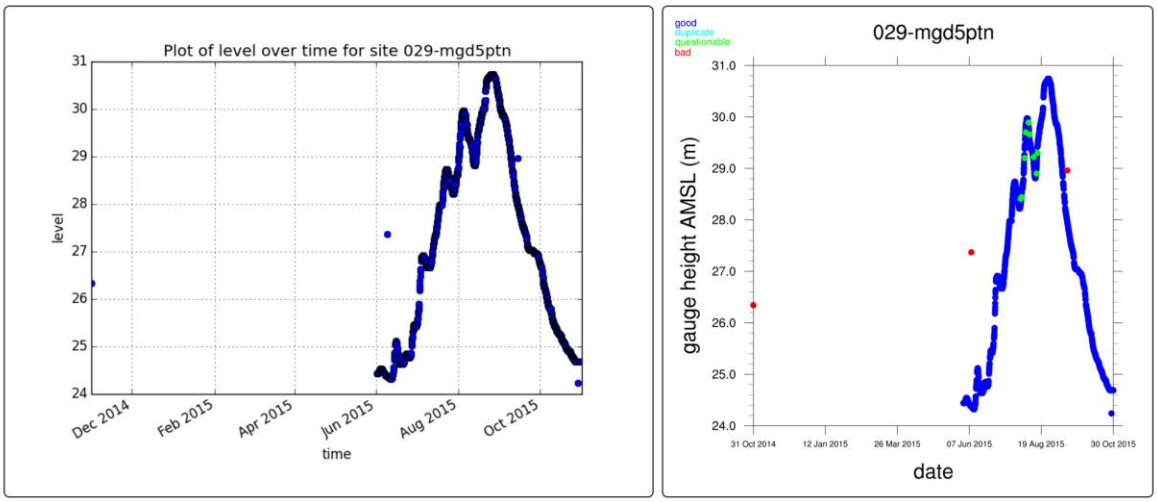


Figure 39: displays provide visualize of the geographic extent of the river basins, observations, ensemble forecasts, and quality control analyses at that location.

The display is written in JavaScript, so that it may be viewed in a web browser on a desktop computer or a mobile smartphone.

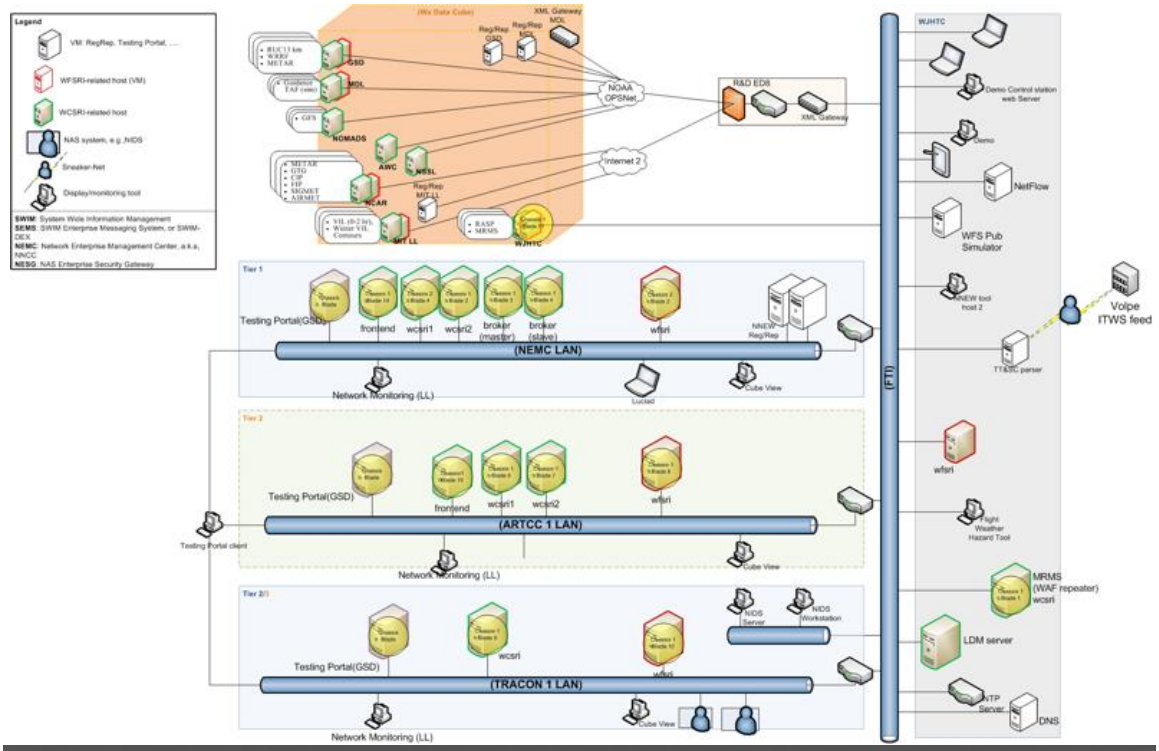
Future Display Enhancements

NCAR's expertise in environmental modeling, data analysis, and visualization can be brought to bear on many areas of India's water resources management capacity building. Several areas in which we see potential for future enhancements are listed below.

Data Interoperability Standards

Sharing of data between multiple processes and users can most easily be achieved through the implementation of standard web data services. The Open Geospatial Consortium (OGC) is an international open standard body which develops specifications for such services. Their most widely adopted services are Web Map Service (WMS), which serves geo-located imagery, and Web Feature Service (WFS), which serves location-specific data at a point, along a line, or for a polygonal region. Implementation of OGC standards should be promoted as the primary mechanism for making data available to users within the government and, if desired, to the public at large.

For the past seven years, NCAR has been the leader in developing weather data exchange standards for the Federal Aviation Administration's (FAA) NexGen modernization effort. Those standards extend OGC, ICAO, and WMO standards to provide the spatial, temporal, and data fidelity detail necessary to support the FAA's challenge. NCAR prototyped and tested the system over several years and is now the FAA's technical expert, overseeing the operational implementation which was recently awarded to a contractor.



Status

Last Updated: 44 minutes ago

MIT-LL-Primary Registry/Repository

http://nnew-regrep1.wx.ll.mit.edu/omar-server

Services

Search:

Service	Type	Endpoint	Status
MIT_WebCoverageService-01	WCS	http://nnew-wcs1.wx.ll.mit.edu/nnew/fy10/wcs/soap	Up
MIT_WebCoverageService-02	WCS	http://ngentst.wx.ll.mit.edu:8080/wcs-spring	Down
MIT_WebCoverageService-03	WCS	http://ngentst.wx.ll.mit.edu:8080/wcs-spring	Down
MIT_WebFeatureService-01	WFS	http://nnew-wfs1.wx.ll.mit.edu/wfsri-2.0/wfs	Down
NCAR_WebCoverageService-01	WCS	http://weather.aero/nnew/fy10/wcs/soap	Up
NCAR_WebFeatureService-01	WFS	http://weather.aero/soap/wfs	Unknown

Showing 1 to 6 of 6 entries

Datasets

Search:

Dataset	Source	Endpoint	Status
Ceiling	WCS	http://weather.aero/nnew/fy10/wcs/soap	Up
CIP SLD - Super Cooled Liquid Droplets	WCS	http://weather.aero/nnew/fy10/wcs/soap	Up
CIWS Vertically Integrated Liquid (CIWS Netcdf4)	WCS	http://nnew-wcs1.wx.ll.mit.edu/nnew/fy10/wcs/soap	Up
CIWS Vertically Integrated Liquid (CIWS Netcdf4)	WCS	http://nnew-wcs1.wx.ll.mit.edu/nnew/fy10/wcs/soap	Up
CIWS Weather Avoidance Field for 29 ft(CIWS Netcdf4)	WCS	http://nnew-wcs1.wx.ll.mit.edu/nnew/fy10/wcs/soap	Up
CIWS Weather Avoidance Field for 35 ft(CIWS Netcdf4)	WCS	http://nnew-wcs1.wx.ll.mit.edu/nnew/fy10/wcs/soap	Up
CIWS Weather Avoidance Field for 39 ft(CIWS Netcdf4)	WCS	http://nnew-wcs1.wx.ll.mit.edu/nnew/fy10/wcs/soap	Up
Echo Tops (CIWS Netcdf 4)	WCS	http://nnew-wcs1.wx.ll.mit.edu/nnew/fy10/wcs/soap	Up
Flight Category	WCS	http://weather.aero/nnew/fy10/wcs/soap	Up

Figure 40: examples of data interoperability.

Advanced Spatio-Temporal Displays

NCAR has extensive experience in developing web services for retrieving data with complex spatial and temporal bounds. Such queries are critical to efficiently supporting modeling and data analysis. In recent years, those types of complex queries have become available in many off-the-shelf relational database management systems and geographic information systems, but working with, or extending, those systems is still difficult. Doing so requires knowledge of the data and its intended uses as well as the software implementations. Providing this kind of fundamental capability for India's water resources data opens the door to advanced scientific analysis and improved forecasting.

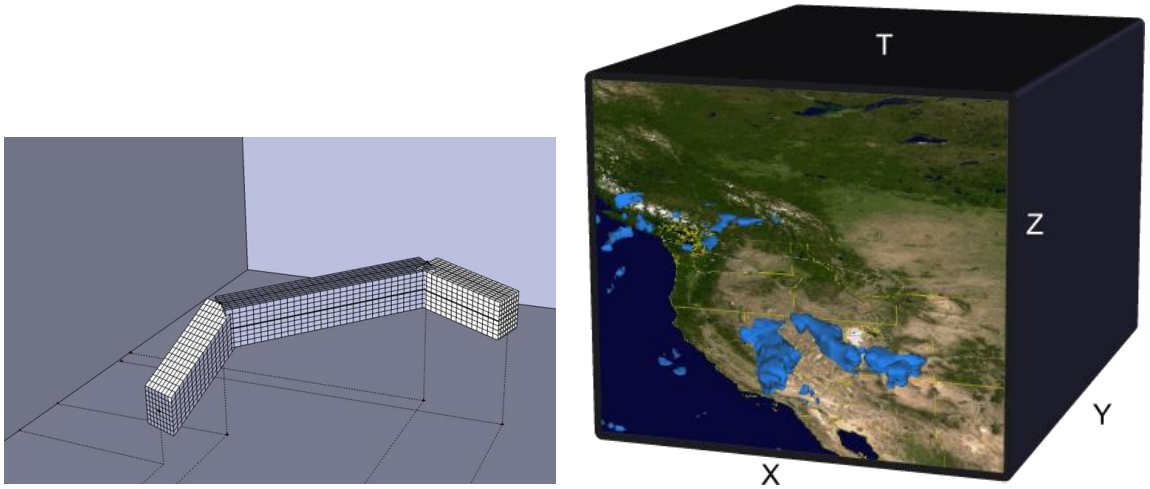


Figure 41: examples of advanced spatio-temporal displays.

Decision Support and Warning Systems

It is increasingly difficult for system operators to make use of the large amounts of data available about their domain. Decision support tools and alerting systems can address that issue by synthesizing data from many sources to provide clear, actionable, information.

NCAR has developed such a system for winter road maintenance. It is in use by a dozen States' Departments of Transportation for keeping highways safe from snow and ice, as well as Denver International Airport for maintaining its runways. The system links a tuned weather forecasting model, a ground temperature model, a chemical dilution model, and a rules engine to provide managers with treatment recommendations and the ability to perform "what-if" scenarios. This type of system would be extremely valuable for managing operations and flood risk on large river systems such as those in India.

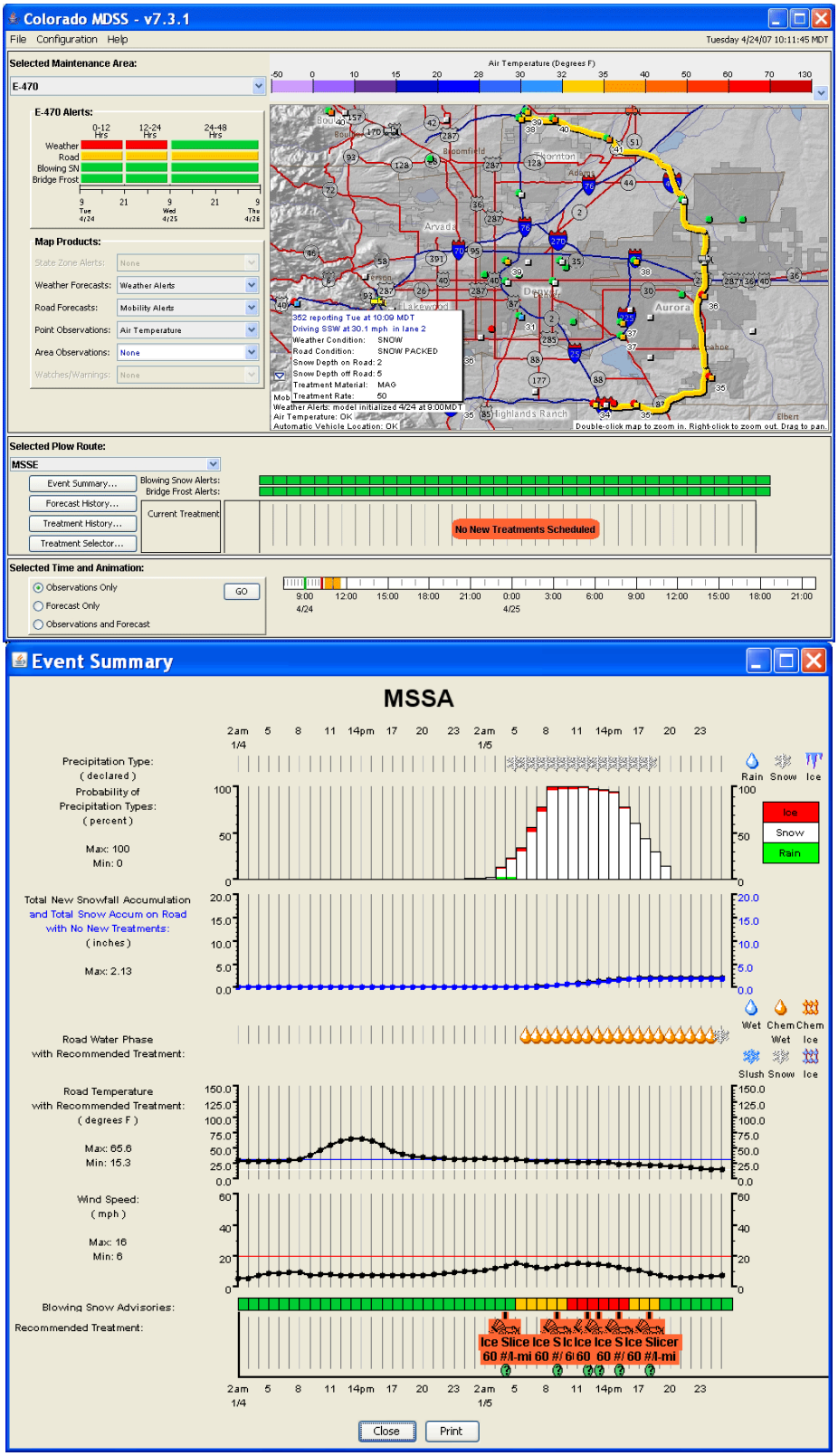


Figure 42: examples of decision support and warning systems.

Mobile Access

Access to information from mobile devices is especially important in developing countries, where operators may be out in the field or network infrastructure may not be mature. Most displays are now designed to be accessed flexibly from a web browser on a desktop computer as well as from a mobile tablet or phone. This should be required of any displays developed for flood forecasting in India.

NCAR has several web-based displays that are used primarily on mobile hardware. The MobileMet display has been our test display for FAA cockpit simulations, aimed at assessing the challenges of showing information on a smaller format while maximizing readability. Our Weather Research and Forecasting (WRF) Hydro display is used to verify riverbasin forecasts for an experiment over the western United States.

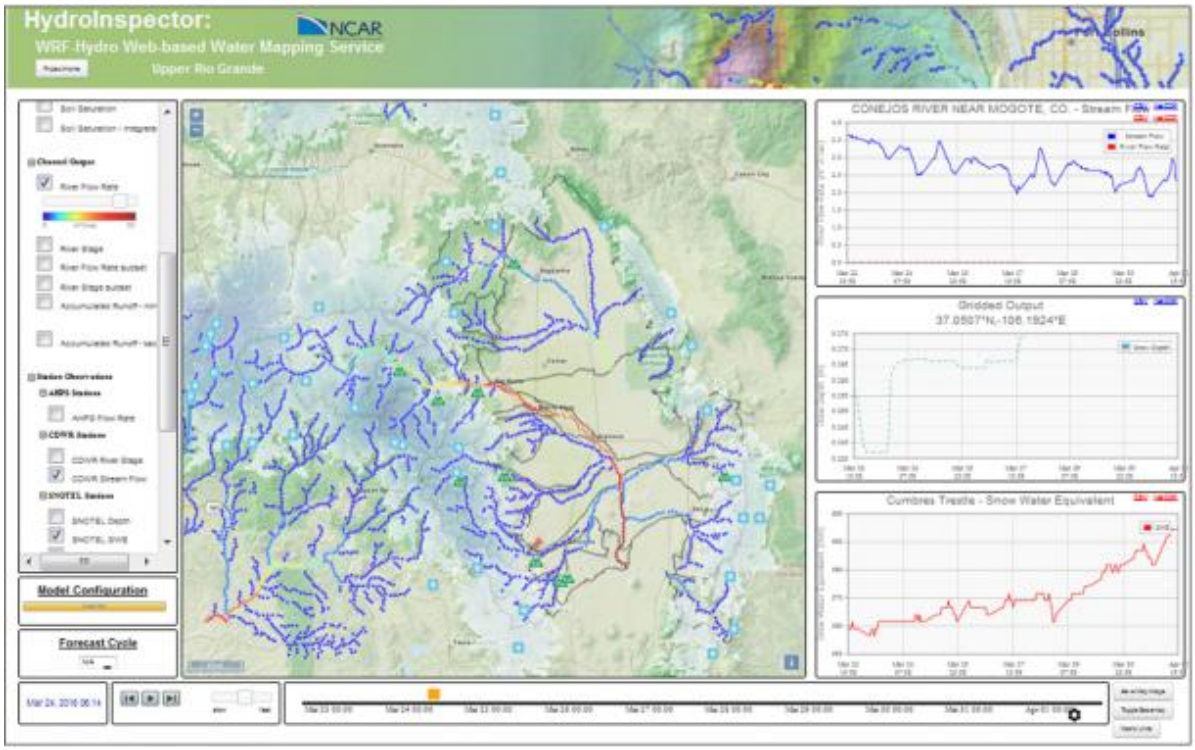


Figure 43: examples of mobile accessible displays.

Dartmouth Flood Observatory Inundation Mapping

Utility of transforming discharge information into inundation extent maps

Providing flood forecasts of discharge magnitude are certainly vital to effectively mitigate flood hazards for vulnerable citizens. If extensive flood mapping surveys have been carried out, then the spatial extent of oncoming flood waters can be *a priori* determined. However, for most regions of the world, such hydraulic modeling and spatial mapping have not been carried out in sufficient detail.

To get around this limitation, historical context could be used: if the discharge magnitudes can be placed into the context of historical flows, a sense of the flood return period can be estimated (10yr flood, 25yr flood, etc.), and citizens with intimate knowledge of their region could have context for the spatial extent the flood waters will occupy. However, preferable to this latter approach would be to relate forecasted flood discharges to archived imager over the region of interest that occurred during periods of similar river flows, thus giving citizens and relief workers more specific a priori knowledge of the extent oncoming flood waters could occupy, without reliance on detailed surveys. Note, however, one of the limitations of this approach is the reliance on the assumption of geomorphological consistency – past and future river channel characteristics (e.g. in cross-section, conveyance, and river channel location) need to remain similar for this approach to be accurate. Thus, this method (or any of the three methods discussed, actually) will be less accurate on rivers that are highly geomorphologically active.

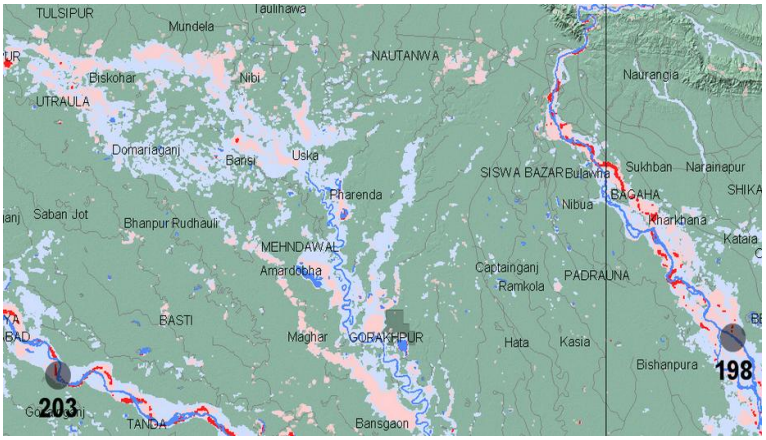


Figure 18. The DFO record of flooding in this portion of the Ganges Basin is shown as light blue (2000 to 20111), light red was flooding in the 10 days prior to map update date, and red was current flooding. The numbers indicate River Watch discharge measurement sites. This is a small subscene of the complete Surface Water Record display for this region.

The method chosen in this consultancy is to utilize the DFO archive of catalogued imagery. NASA's orbital technology has been used at DFO extensively, since the launch of the twin MODIS sensors in early 2000 and 2002, to map flooding in South Asia. Unlike other remote sensing-based organizations active in flood response, DFO maintains a large and growing archive of such

map data, in digital (GIS) format, and for use in making comprehensive regional displays indicating the history of inundation as well as on-going flooding (Figure 18). The archival flood information is exceptionally valuable, providing as it does a view of *flood hazard*.

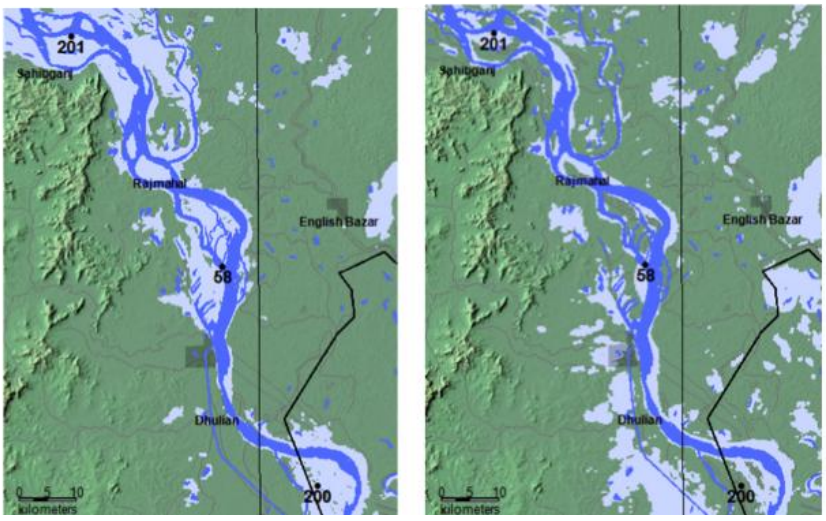


Figure 19. Left: MODIS imaging and mapping of 2003 flooding along the Ganges River between river measurement sites 200 and 201. At site 200 (uncalibrated) peak discharge was 8500 m3/sec. Right: mapping of 2004 flooding. The uncalibrated peak discharge here is only ~3500 m3/sec.

This large archive of such mapped inundation resident at DFO will allow production of this innovative flood prediction product. As illustrated in Figure 19, past inundation extent can be matched to the corresponding remote sensing-derived discharge values (the same approach can be used for any ground station sites for which data output is available publicly). Linkage to the appropriate inundation map can be provided at the individual site displays: when a particular discharge and flood threshold is predicted, the user can call up the inundation that resulted, historically, from the same values. Similarly, mapping inundation maps to the ensemble of river forecasts produced by the flood forecasting framework could then produce a range of possible inundation extent scenarios.

In this section we discuss the collaborative work with the Dartmouth Flood Observatory (DFO; under the direction of Dr. Robert Brakenridge) to integrate the operational river discharge forecasts being produced at our selected sites under this project with historic inundation imagery over the Ganges and Brahmaputra Basins. We are separately calling this effort the Ganges-Brahmaputra Flood Awareness and Prediction System. Note that in addition to the World Bank and South Asia Water Initiative funding provided as part of this project, partial funding to support the DFO efforts by NASA. Below we discuss the efforts and information that is being produced and displayed as part of this project, with further links provided online at:
http://floodobservatory.colorado.edu/GangesBrahmaputra/GangesBrahmaputraIndex.html

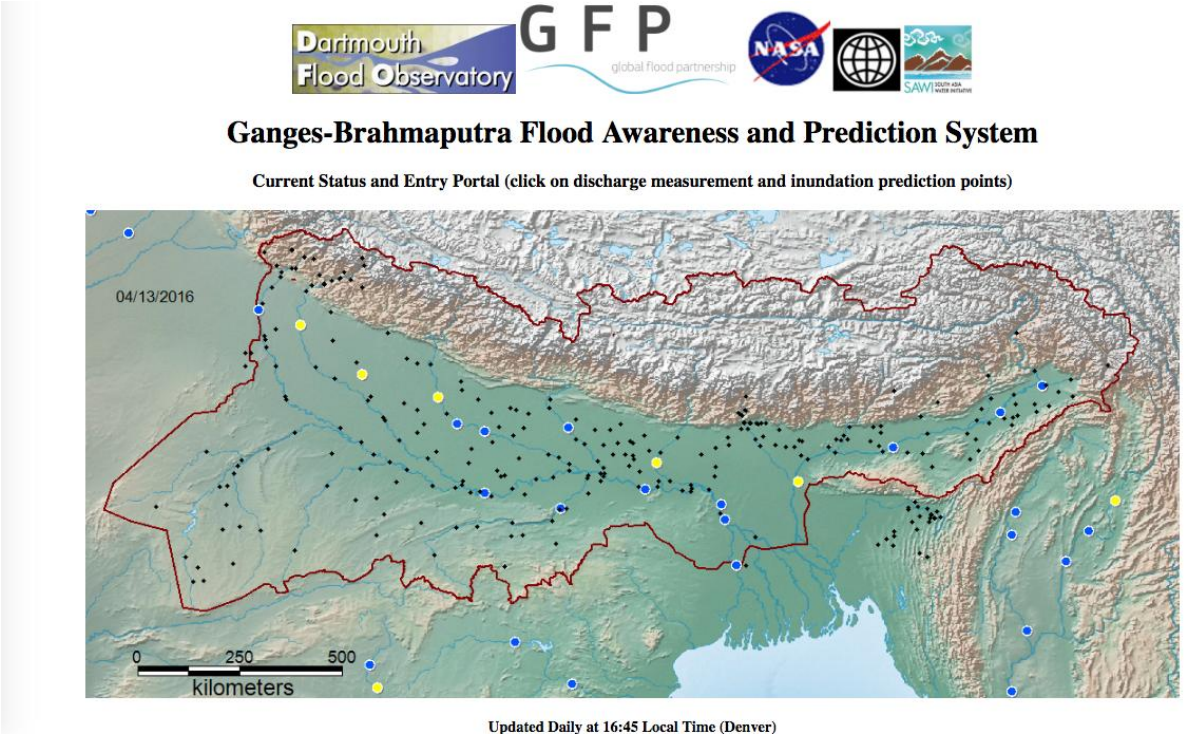


Figure 44: home page imagery and site location of this collaboration. Symbology: Black dots: all forecast points. Colored Dots (clickable): discharge measurement/inundation prediction sites. Yellow dots: low flow (<20th percentile discharge for this day of the year, 2003-2013; Blue: normal flow; Purple: moderate flooding (>1.5 yr recurrence interval); Red, major flooding (> 5 yr recurrence interval).

Co-located river discharge forecasts from the WorldBank-SAWI-NCAR system are being added to each measurement site. A second portal display will show forecast information instead of current status. The following disclaimer is also provided at this site:

Data are experimental, are not official weather or streamflow forecasts, and may be subject to intermittent outages. Viewers requiring official river stage observations and forecasts should consult the Indian Central Water Commission at: india-water.gov.in or other national water ministries.The initial processing of the River Watch microwave data is provided by the Joint Research Centre of the European Commission: *Technical Report and Data Product Specifications, Global Flood Detection System.*

Ganges-Brahmaputra MODIS NRT Current Flood Conditions Display

Here we present examples of the current and (now being integrated) forecasted MODIS imagery online at: http://floodobservatory.colorado.edu/GangesBrahmaputraCenter3.html. These samples from large format, online display for whole basin. Date: May 3, 2016:

Ganges-Brahmaputra MODIS NRT Current Flood Conditions

Updated Daily at 16:45 Local (Denver) Time.

Red: current surface water that is larger in extent than the Blue reference water (obtained in February 2000 by the NASA SRTM mission; spatial resolution of 90 m)

Very light red: all flooded areas imaged by NASA's twin MODIS sensors since January, 2000.

MODIS Water Mapping information at a spatial resolution of 250 m is from NASA's Terra and Aqua satellites via an automated water classifier. The Flood Observatory accumulates daily files from a special processor at the NASA GSFC over 14 days (running forward, updated each day). This provides nearly complete geographic coverage even under cloudy conditions. New artificial reservoirs built since 2000 also appear as red.

River Watch sites (colored dots): Yellow dots: low flow (<20th percentile discharge for this day of the year, 2003-2013; Blue: normal flow; Purple: moderate flooding (>1.5 yr recurrence interval); Red, major flooding (> 5 yr recurrence interval).

WESTERN GANGES-BRAHMAPUTRA BASIN

Obtain Geotif version

Ganges River, no flooding, light pink are previously flooded areas, 1999-present. River Watch measurement sites (large colored dots) and NCAR forecast points (small green dots) are shown.

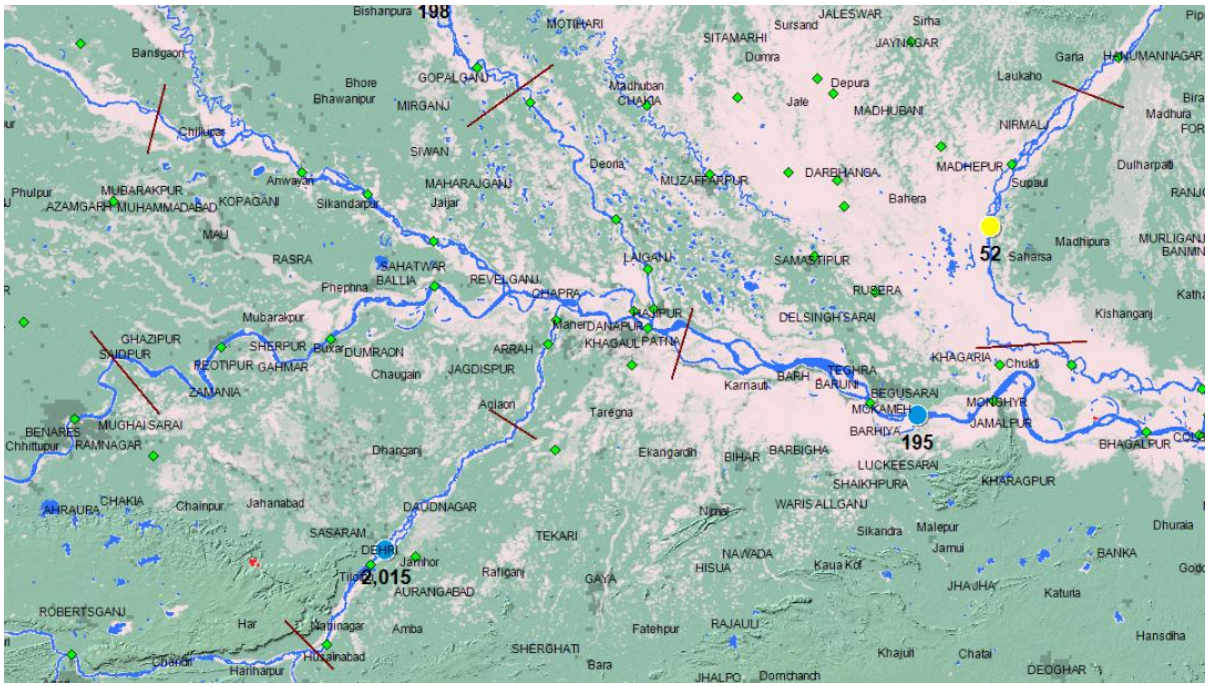


Figure 45: Ganges River, no flooding, light pink are previously flooded areas, 1999-present. River Watch measurement sites (large colored dots) and NCAR forecast points (small green dots) are shown.

Brahmaputra River, flooding is underway (red), as shown below.

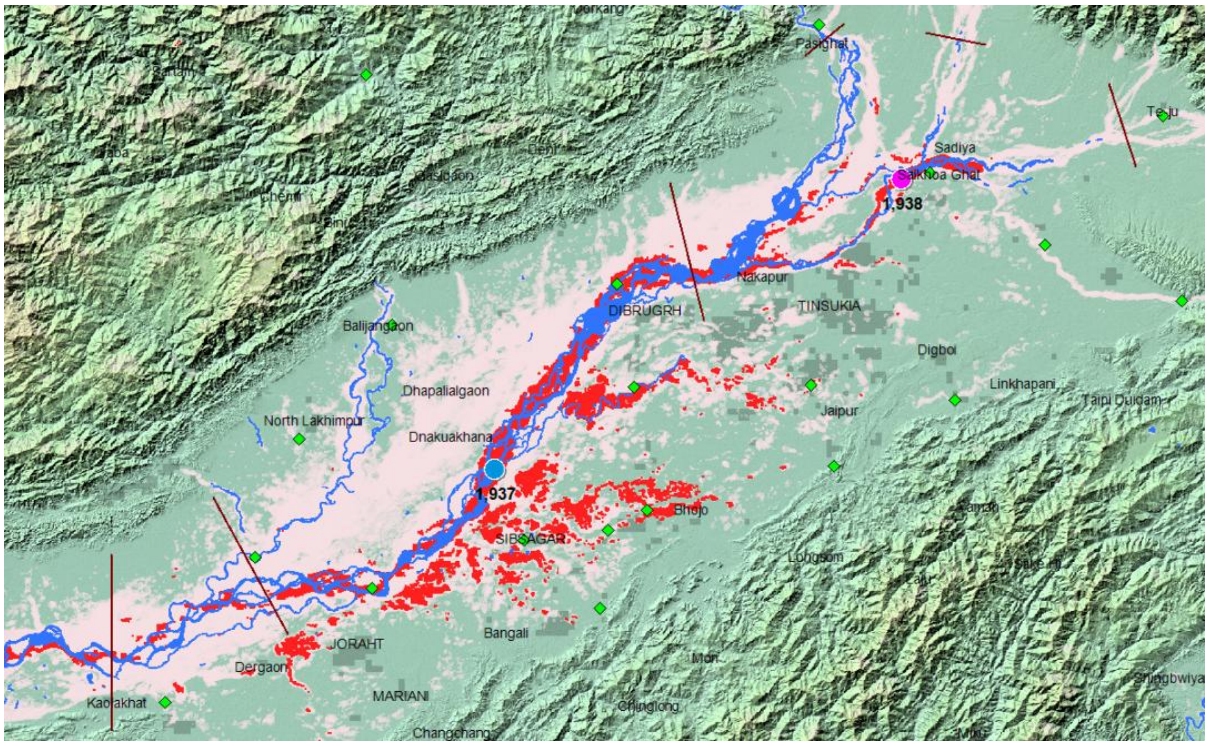


Figure 46: Brahmaputra River, flooding is underway (red).

River Watch Site 1938 is also measuring the flooding (below).

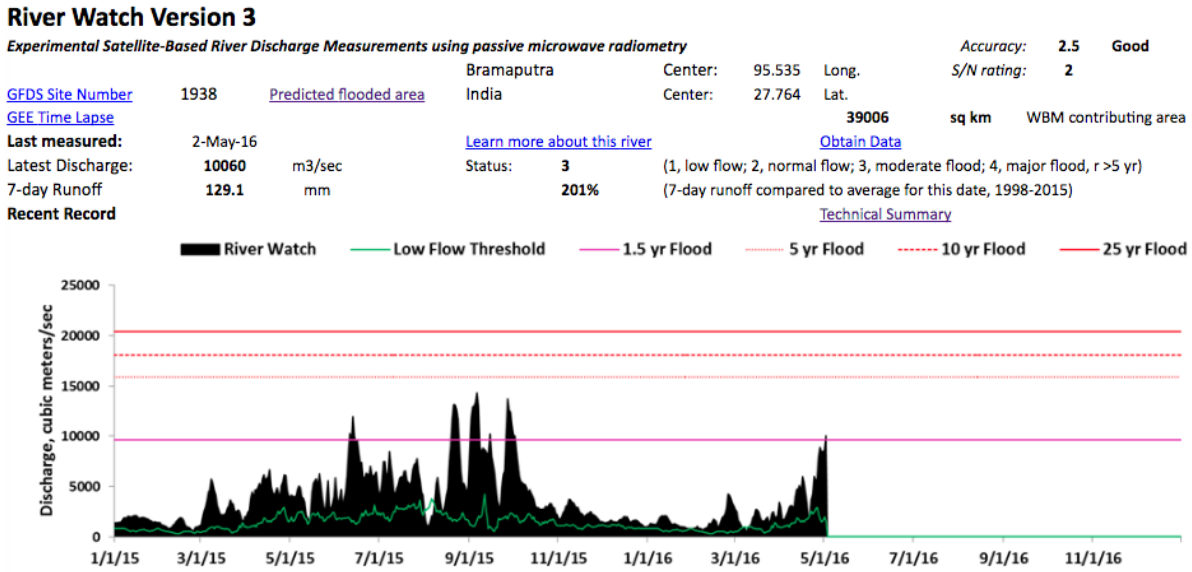
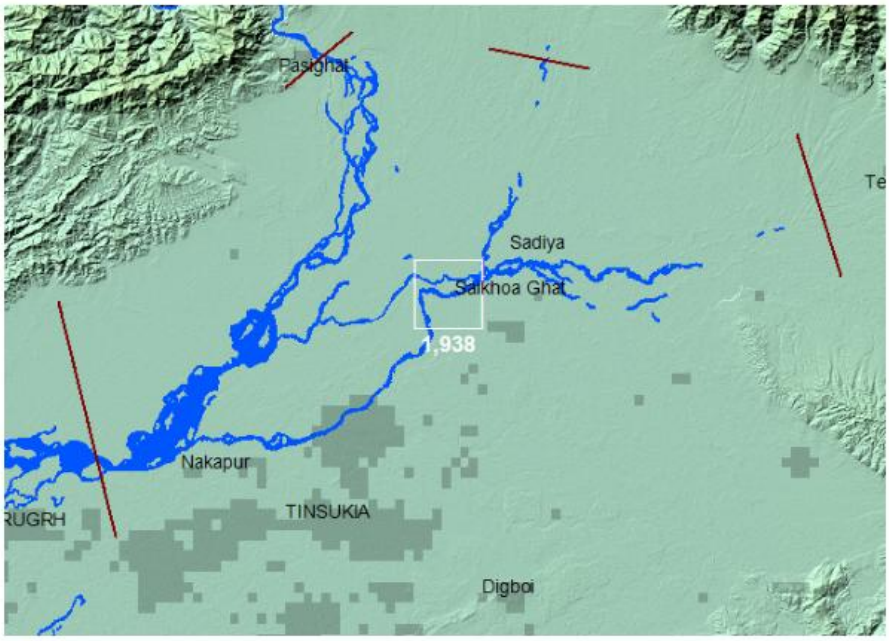


Figure 47: River Watch Site 1938 is also measuring the flooding.

Below we show pilot Predicted Flooded Area Display, Site 1938, online at: http://floodobservatory.colorado.edu/PredictedFloodedArea/1938.html. The display is based on forecasts from the WorldBank-SAWI-NCAR system furnished to the Dartmouth Flood Observatory, DFO) on May 1, 2016. These two displays below show normal flow conditions and 2-yr return level inundation levels.

Flooded area for Normal Flow, Winter (936 m3/sec, observed February 11-22, 2000)



Geotif version

Google Earth kmz version

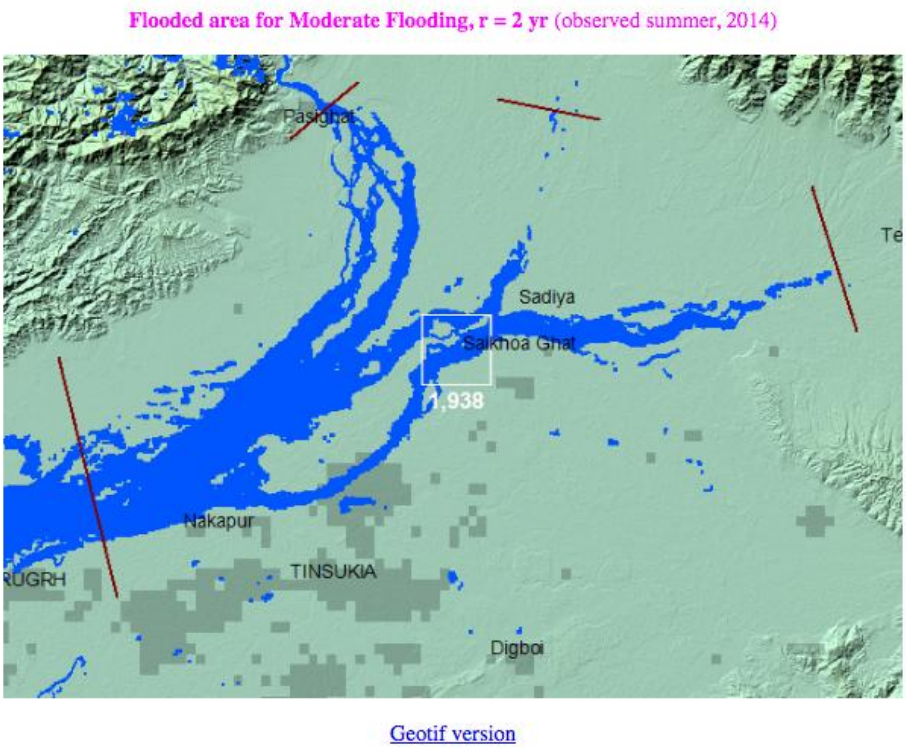


Figure 48: Predicted Flooded Area Display, Site 1938 based on forecasts from the WorldBank-SAWI-NCAR system furnished to the Dartmouth Flood Observatory, DFO) on May 1, 2016, showing normal flow conditions and 2-yr return level inundation levels.

Ongoing Work To Be Done at DFO

- **18** River Watch measurement sites are already established in the Ganges Brahmaputra. Most need to be upgraded from version 2 or version 3 to version 3.1.
- **~ 10** new sites need to be established where observational data are most sparse.
- **6** Predicted Flooded Area displays have now been produced for matching River Watch measurement sites. Eventually a total of 28 should be produced and published online.
- Code for ingesting the NCAR prediction information has been built by Albert Kettner at DFO. Incorporation into the River Watch site displays must now be accomplished.

- Automated updating on at least daily basis must be established for these new prediction displays The Flood Awareness System; other map view displays should be augmented to show predicted as well as current conditions.

Appendix C – Simplified Concept of River Flow Predictability

In this appendix we provide a brief overview of the basis for predictability in water flow, and thus flooding, in a river catchment. A simple conceptualization of water transport through a watershed is shown in Figure 1, which shows a hydrological isochrone (often called "isoline") map of the Brahmaputra and Ganges basins. The isochrones represent the average time it takes a parcel of water, located at some location within the basin, to reach the outflow point of the two rivers out of India and into Bangladesh. Watershed hydrologic time scales related to these isochrones are the *time of concentration* (time needed for water to flow from the most remote point in a watershed to the watershed outlet) and watershed lag time (time from the center of mass of excess rainfall to the center of mass of the direct runoff hydrograph). The hydrological time scale of the Ganges and the Brahmaputra are roughly 1 to 3 weeks (with the Brahmaputra having the shorter response times relative to the Ganges, and the ambiguity due to whether one is considering direct runoff, baseflow, or in-stream flow time scales of water migration through the system – in Figure 1 we show estimates of transport times averaged over the seasonal cycle derived from routing of subcatchment hydrologic flows).

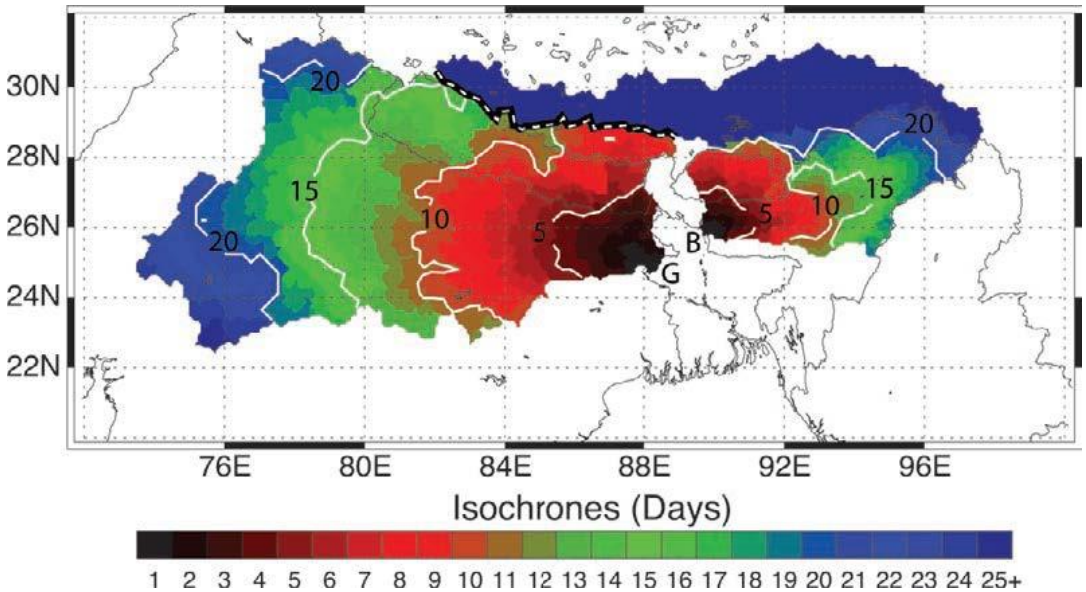


Figure 1: Isochrones (days) for the Brahmaputra and Ganges basins. The contours provide an estimate of the time it takes for water in a particular location in the basin to pass from the headwaters to the outflow points into Bangladesh ("G" for the Ganges, and "B" for the Brahmaputra). Isochrones were constructed using a flow routing algorithm and digital elevation data (details in text). The separation of the two basins is shown as a thick dashed white lines. Isochrones for 5, 10, 15 and 20 days are shown as solid white lines.

For our purpose, the important concept is that upstream observations (whether of rainfall or river discharge) inherently can provide information on future river flows that will occur lower down in the catchment up to the *time of concentration* time-scale. To elucidate these concepts further using a simplified context of precipitation-streamflow relationships, consider a simple linear basin with equal area sub-basins (Figure 2), each having a flow-through time of one day. The goal in this exercise is to forecast the flow out of the M^{th} sub-basin I days ahead. The discharge out of each sub-basin n, has two contributing parts: the local runoff R_n which depends on area integrated precipitation, and Q_{n-1}, the discharge from the adjacent upstream sub-basin n-1. In any sub-basin n:

$$Q_n(i) = R_n(i) + Q_{n-1}(i-1) \tag{1}$$

where the subscript i refers to the time in days. The local runoff, assuming that precipitation within the basin takes half a day to flow through, is given by:

$$R_n(i) = \int_A \frac{P_n(i) + P_n(i-1)}{2} dA \tag{2}$$

where A is the area of the sub-basin. To forecast the discharge out of the last sub-basin (i.e., $n = M$) I days in the future (for $I < M$) using available precipitation forecasts, we have

$$Q_M(I) = \left[\sum_{i=0}^{I-1} R_{M-i}(I-i)\right] + Q_{M-I}(0) \tag{3}$$

Note that if no discharge observations are available (or $I > M-1$), $Q_M(I)$ can still be determined only by rainfall observations and forecasts.

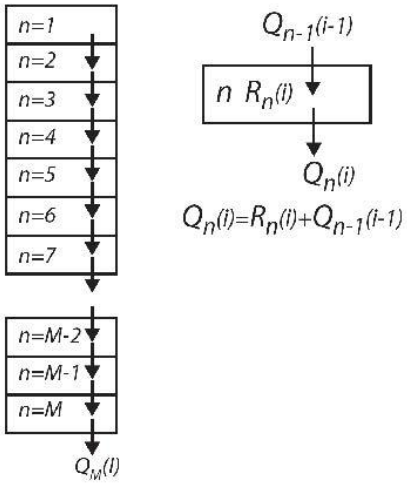


Figure 2: Simple model of a linear river basin containing M sub-basins in which the transit flow time is 1 day. Thus if M =10, the flow- through period of water entering the basin at the headwaters (n =1) and leaving the last sub- basin would be 10 days. The flow out of some sub-basin n at time i ($Q_n(i)$) is determined by the flow into the sub-basin ($Q_{n-1(i-1)}$) plus the rainfall occurring in sub-basin n during the (i-*th*) 24 hour period. We are interesting in determining the flow through some *nth* box along the Ganges and Brahmaputra. Depending on the flow-through time (*time of concentration*) of the basin upstream of the *nth* location and the length of the forecast lead-time, *in situ* measured (discharge and rainfall) data and forecast rainfall contribute in

differing proportions.

We now pursue this simplified analogy further, considering the case where the *M* subcatchments of Figure 2 each contribute, on average, the same proportion to the overall river flow, where *M=10* for simplicity. The cost in not having discharge data for forecasts horizons less than the hydrological time scale *M* (*time of concentration*) can be assessed as follows, assuming a perfect precipitation forecast:

(i) *One-day precipitation forecasts available*: For $I = 1$, then $Q_M(1) = R_M(1) + Q_{M-1}(0)$ from Equation 3. In this limiting case, 10% of the information for $Q_M(1)$ would come from knowledge of the precipitation (from Equation 2, we have 5% each from the observed and forecasted precipitation) and 90% from the discharge observed from the upstream basin Q_{M-1} (since, on average, 90% of the water originates above this location).

(ii) *Two-day precipitation forecasts available*: For I =2, then $Q_M(2) = R_M(2) + Q_{M-1}(1) = R_M(2) + R_{M-1}(1) + Q_{M-2}(0)$ from Equation 3. For this case, 20% of the information for $Q_M(2)$ comes from precipitation (from Equation 2, we have 5% each from the observed and 2-day forecasted precipitation, and 10% from the 1-day forecasted rainfall) and 80% from the discharge passing from the second upstream catchment area.

(iii) *M-day precipitation forecast available*: For *I=M=10*, the forecast is the same length as the hydrological time scale of the basin and hence the basin hydrology is determined completely by precipitation (from Equation 2, we have 5% each from the observed and 10-day forecasted precipitation, and 10% each from the 1- to 9-day forecasted rainfall).

(iv) *General case, with no discharge observations available*: For an *Ith*-day forecast, 5% of the flow is determined by the *Ith*-day lead-time precipitation forecast, 10% each from the (*I-1*)-, (*I-2*)-, …, 1-day forecasts, and the remainder ([*105 – (10 X I)*]%) from rainfall observations.

From this "thought experiment", we can draw the following conclusions:

- Discharge observations (from which satellite altimetry or river width measurements are part of) have no utility for forecasts beyond the time-of-concentration (*M*th time scale) at the location in the river basin of interest;
- If no discharge observations are available, precipitation observations (and hydrologic modeling) can be used in their place;
- However, typically the errors in runoff derived from rainfall-runoff modeling are significantly greater than river observations (which includes satellite altimetry, but not necessarily river-width observations);
- For any river discharge forecast, part of the discharge can only be derived from rainfall observations and/or forecasts;

- Given imperfect rainfall observations, forecasts, and hydrologic models, there is always added utility in incorporating "skillful" river discharge observations for forecast lead-times less than the *time-of-concentration*, and locations below the upper catchment headwaters.

Although this simple model made a number of broad simplifications (e.g. equal travel-through times, equal discharge contributions for each subcatchment, etc.), these concepts can be generalized and provide insight into the sources of predictability of our specific catchments, which we discuss in the following sections.

Appendix C – India River Stage Gauge Quality Control Procedures

Raw river level gauge data can occasionally have errors, such as those shown in Fig. 1. Therefore, it is important to have quality control (QC) procedures in place, to flag bad data, as well as identify suspicious data. We employed several quality control procedures on several months of data from over 300 river level gauges in India. The data from each station were used to develop quality control thresholds for its own data; this assumes that there is a sufficient amount of data (e.g., several months), and that most data at each station are accurate. The quality control procedures were used to identify the following common errors: 1) multiple observations at the exact same time, 2) data points that are exceptionally larger or smaller than nearly all other observations at that station, 3) excessively rapid changes in river level data, and 4) river levels that are indicated to be constant for too long that are likely the result of a stuck gauge. Due to the large amount of data to be QCed, it was important that the procedures be designed to be very computationally efficient. Different QC values are used to identify values that fail each procedure, while a QC value of 100 indicates highly probable good data. Since data are compared with each other for each QC procedure, data that fail one procedure are not considered for comparison in future procedures; likewise, each procedure rejects erroneous data one-by-one, from the most extreme to the least, so that nearest-neighbor procedures do not falsely identify good data. However, the raw data value is always retained, accompanied by its appropriate QC flag value. The details of each procedure are laid out below.

The first procedure determines if the raw data are actual levels and dates, and not seeming nonsense. Any levels and/or dates that are unidentifiable are marked with the QC value of -1 for missing (Table 1.)

QC score	Meaning
100	Good
0-99	questionable rate of change or possible stuck sensor

-1	Missing
-2	Isolated
-3	too high/low
-4	extreme rate of change
-5	stuck sensor
-11	Duplicate

Table 4: QC score values, and their meaning.

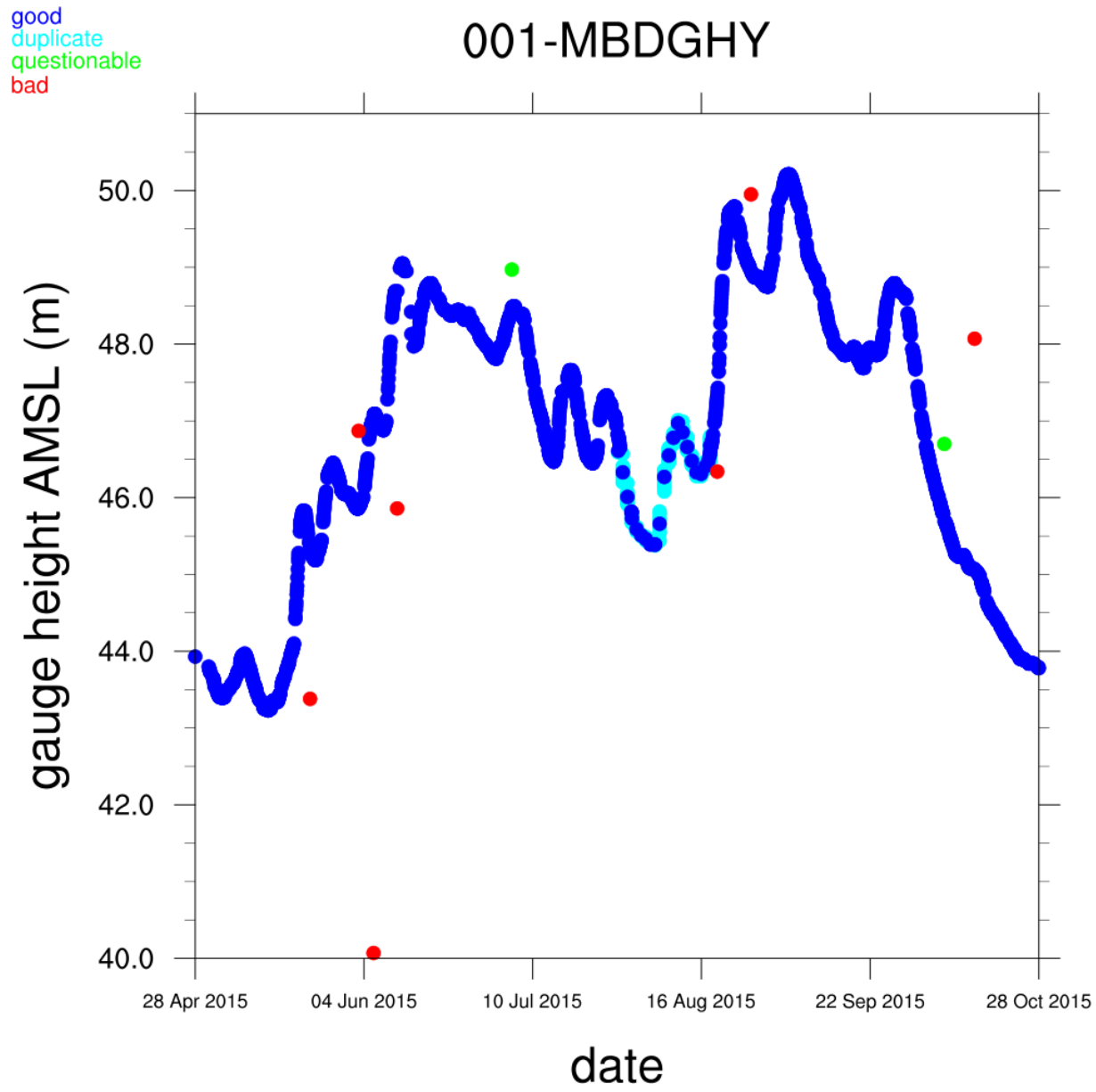


Figure 49: River level data from station 001-MBDGHY. The dot color identifies the data as: (blue) likely valid river levels, (red) very likely erroneous data, (green) questionable data, and (cyan) data observed at the exact same time.

The second procedure identifies all observations at the exact same time. If they all have identical river levels, only one is retained. Of those that share a common time, but are not identical, there are two tests employed. If there are only two different observations at the exact same time, the one with the greater difference with its other nearest neighbor in time is identified with a QC value of -11. If there are three or more different river levels sharing the same time stamp, the median is calculated, and all observations are marked with a QC value of -11, except for the level closest to the median. Examples of multiple observations identified at the exact same time are indicated by cyan dots on Fig. 1.

The third procedure identifies data points with no other nearby data points in time. If a value has no other valid value within 10 days, it is marked by a QC value of -2.

The fourth procedure identifies gauge levels that are exceptionally lower, or higher, than nearly all other levels. To do this, the 10th and 90th percentiles are calculated. The Xth percentile is defined as the level at which only X% of the river levels are lower than that value. Then, the difference, ΔL_{10th}^{90th}, between the 10th and 90th percentiles is calculated. If a gauge level is greater than $L_{90th} + c_H * \Delta L_{10th}^{90th}$, or lower than $L_{10th} - c_L * \Delta L_{10th}^{90th}$, it is marked with a QC value of -3. The constants c_H and c_L were optimized so that they rejected most excessively high and low values, while not falsely identifying any valid values. These optimized values were found to be $c_H = 4.1$ for and $c_L = 1.3$. Examples of river gauge levels identified by this procedure are shown in Fig. 2.

The fifth procedure identifies data with excessively rapid time rates of change on either side, namely sharp spikes and dips in the data. This is done by calculating the time rate of change, both before and after each data point, and then summing the two together: $\left(\sum \left|\frac{dL}{dT}\right|\right)$. Next, the 95th percentile, is calculated from all these values for each station. If any $\sum \left|\frac{dL}{dT}\right|$ exceeds $c_T * \sum \left|\frac{dL}{dT}\right|_{95th}$, it is marked with a QC value of -4. The value of c_T was optimized at 16 so that it rejected most bad data,

yet did not reject hardly any good data. Examples of data identified by this QC procedure are shown by red dots on Fig. 1. There remained, however, additional minor errors that were not flagged by this procedure, such as some of those shown in green dots on Fig. 1. Therefore, another smaller constant, c_B, was calculated to identify data points with rapid changes on either side; if their value $\sum\left|\frac{dL}{dT}\right|$ exceeded c_B they were given a QC score of 0-99, according to the following equation: $100 - 100 * \frac{\sum\left|\frac{dL}{dT}\right| - c_B * \sum\left|\frac{dL}{dT}\right|_{95th}}{c_T * \sum\left|\frac{dL}{dT}\right|_{95th} - c_B * \sum\left|\frac{dL}{dT}\right|_{95th}}$. The higher the score, the more probable it is that the data is good, while the lower scores indicate a high probability it is bad. c_B was optimized at 5 to identify nearly all visually bad gauge level spikes. Examples of gauge levels identified by a score of 0-99 are indicated by green dots in Fig. 1.

Lastly, the sixth procedure identifies river gauge levels that have been constant for an unusual length of time, indicative of a possible stuck gauge. Since the temporal resolution of the river level data varies substantially with time and location, a count of successive constant observations (s), of any time length, were used to identify possible stuck gauges. A lower threshold s_B was used to identify a possible stuck gauge, while an upper threshold s_T was used to indicate a nearly certain stuck gauge. All level data constant for greater than s_T consecutive observations are given a QC score of -5, while all data constant between s_B and s_T time steps were given a QC score between 0-99, according to the equation: $100 - 100 * \frac{s - s_B}{s_T - s_B}$. All river level data in a string of constant values are given the same QC score, from the first to the last.

The next steps to implement in this QC system are: 1) collect at least one year of river gauge level data, in order to capture at least one full monsoon cycle; this will allow for re-calculation of better thresholds for each station. 2) Determine a way to reject "spikes" of data that are more than a single point. 3) Identify spikes of bad data with missing temporal neighbors. There are certainly other QC procedures that can also be developed and implemented on these data.

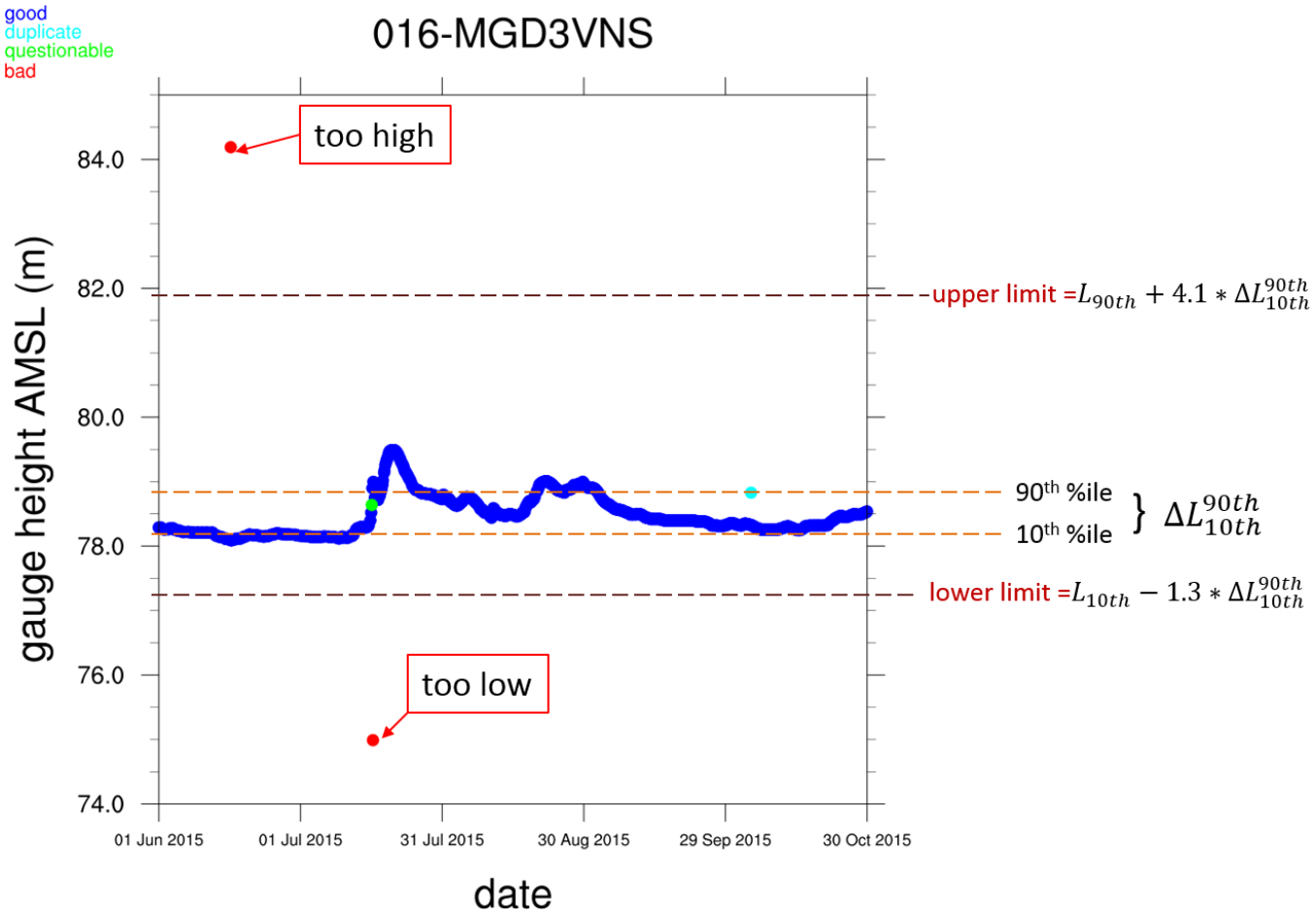


Figure 50: River level data from station 016-MGD3VNS, showing examples of erroneous extremely high and low river gauge values.

Appendix B – Rating Curve App highlighting results

This web application (built using Shiny, a web application framework for the R statistical computing language) displays interactively results from our rating curve fitting (described above), allowing the user to see the stage and discharge data used to fit the rating curves, the rating curve, rating curve fitted parameters, and skill score. Please see: http://indiawbg.rap.ucar.edu/Ratefit/

This app offers practitioners a way of viewing the rating curve fits to ensure that values and fits seem reasonable. For a given selected location, the default view shows the stage and discharge at an optimal lag as defined by the best rating curve fit. Users can vary this lag to see the effect on the rating curve fit.

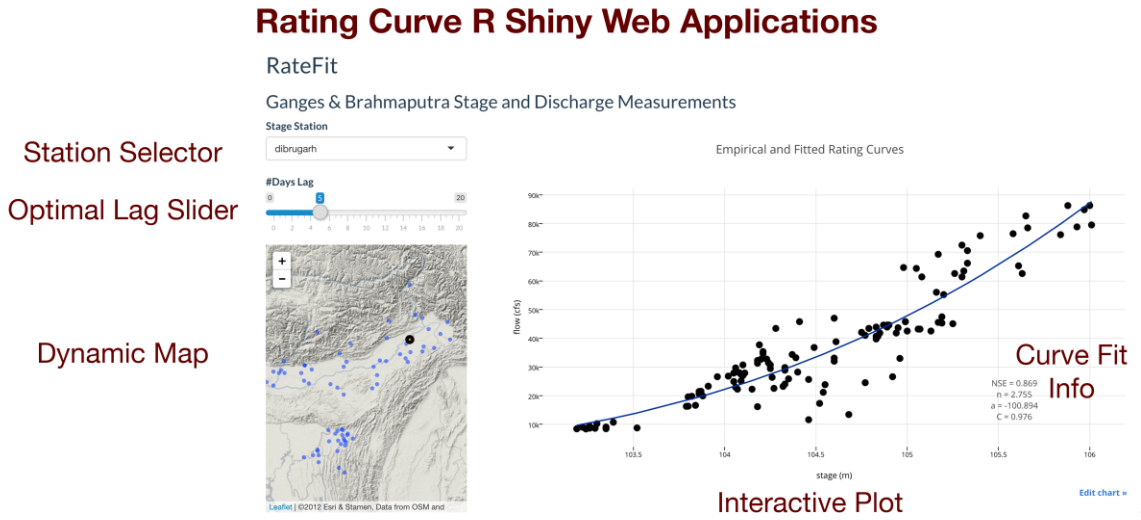


Figure 51: this figure shows the features of the App: station selector, optimal lag slider, dynamic map, interactive plot, and curve fit information.

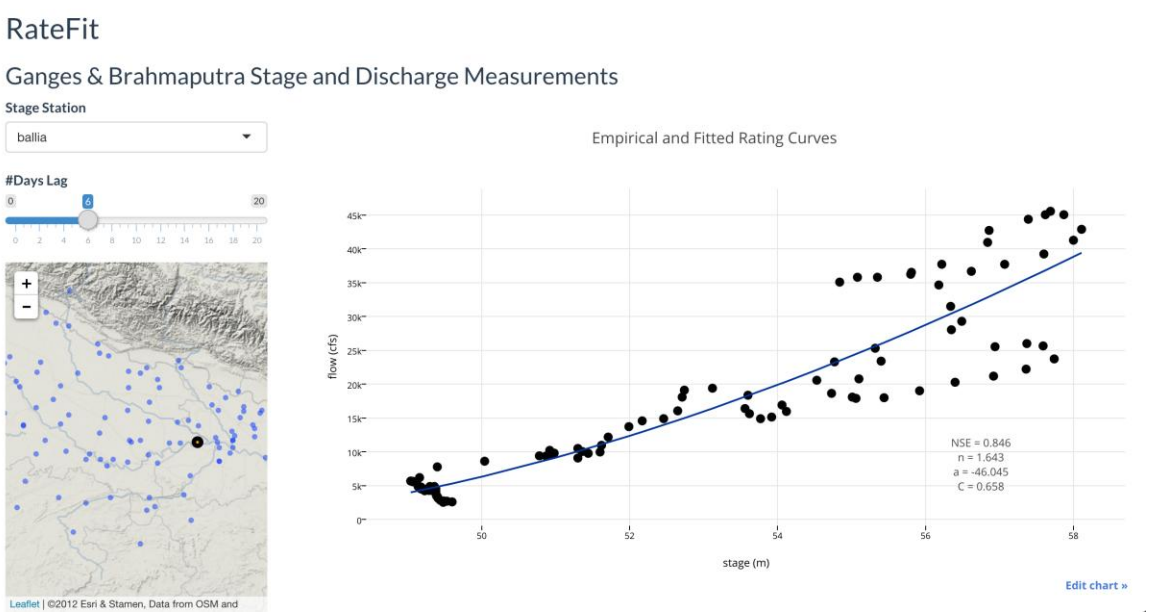


Figure 52: an example of the App showing results for a fit for a Ganges river point.

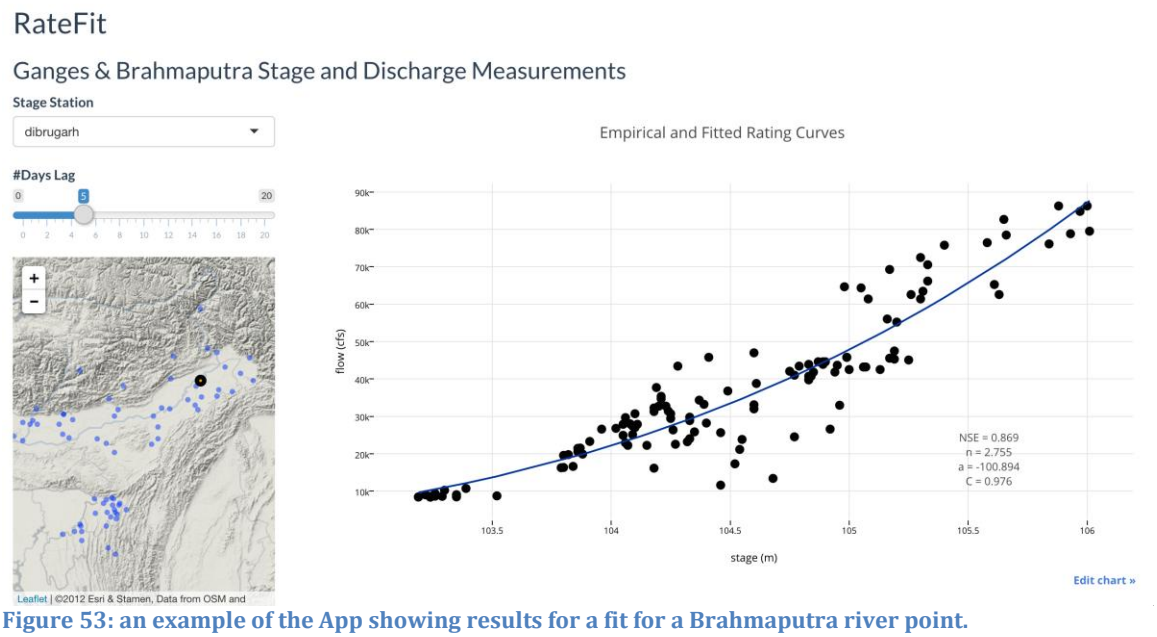


Figure 53: an example of the App showing results for a fit for a Brahmaputra river point.

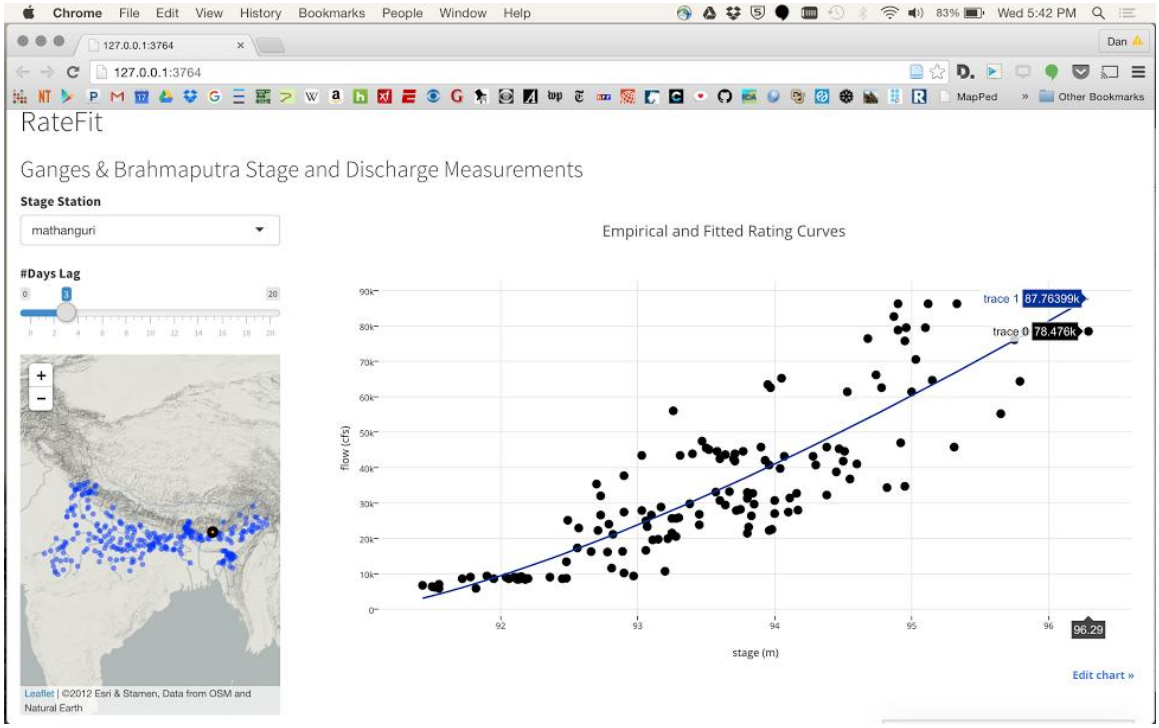


Figure 54: an example showing an actual screen shot of the App, showing that a click on a dot also provides the corresponding stage and discharge values.

Appendix E – Merged Satellite Precipitation versus Rain Gauge Time-Series Accuracy

In this appendix we present daily sub-basin time-series correlation analysis comparing our merged satellite precipitation product (combining NASA TRMM, NOAA CMORPH, and JAXA GSMaP) with our daily 0.5X0.5 gridded rain gauge product. This analyses was done for July. Our findings in the figures below show that we find correlations between these products of between 0.6-0.7; which although not extremely high, we argue they do prove a strong-enough relationship between these products to affirm the use of satellite precipitation for river discharge forecasting over the Indian subcontinent. Also, histogram comparisons show similar results (Yamuna Basin below), but further analysis of the Meghna suggests that the gauge data has proximately 3x as many days with zero or very low precipitation, possibly suggesting that rainfall events are being missed to lack of coverage.

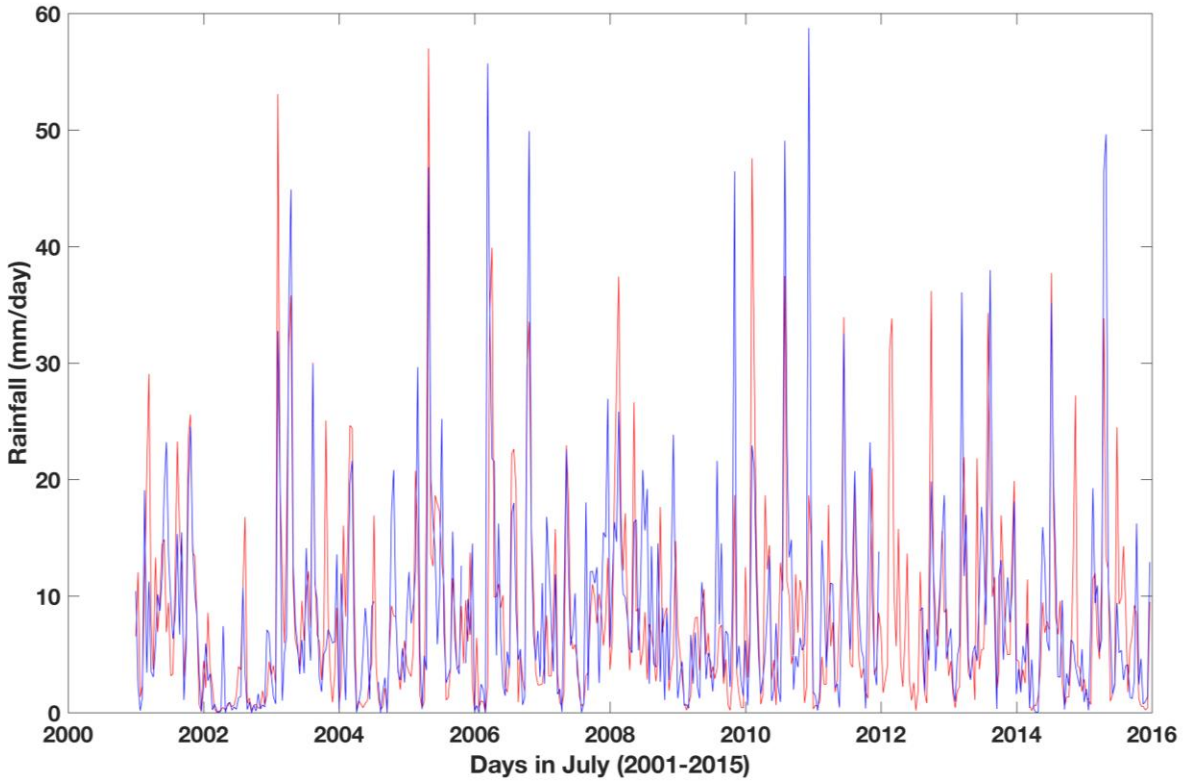


Figure 55: Comparison of time series of rainfall for days in July in Yamuna basin. Blue: Rain gauge Red: Merged Satellite. Correlation coefficient between the two time series: 0.67.

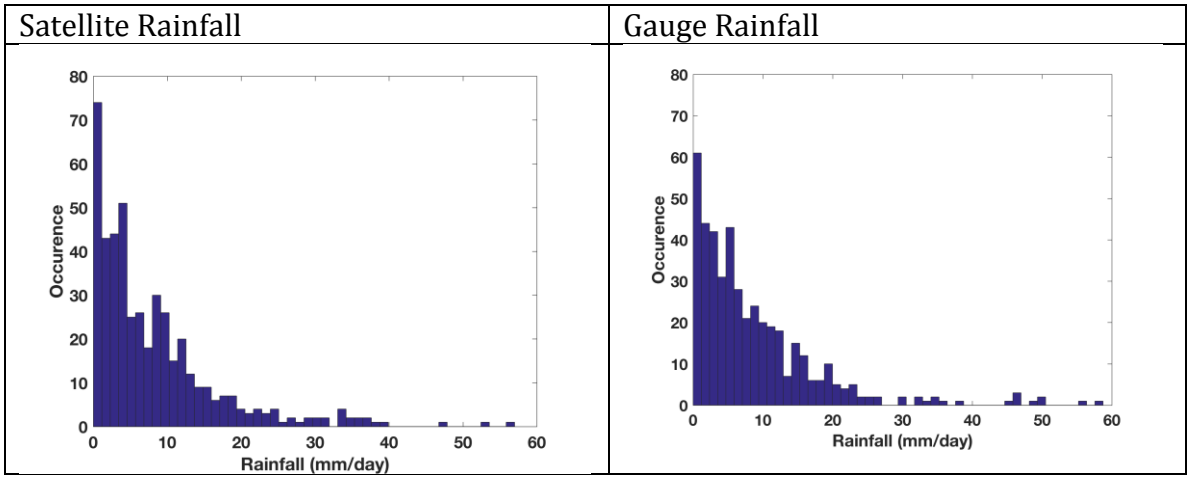


Figure 56: Comparison of rainfall histograms of July rainfall in the Yamuna Basin.

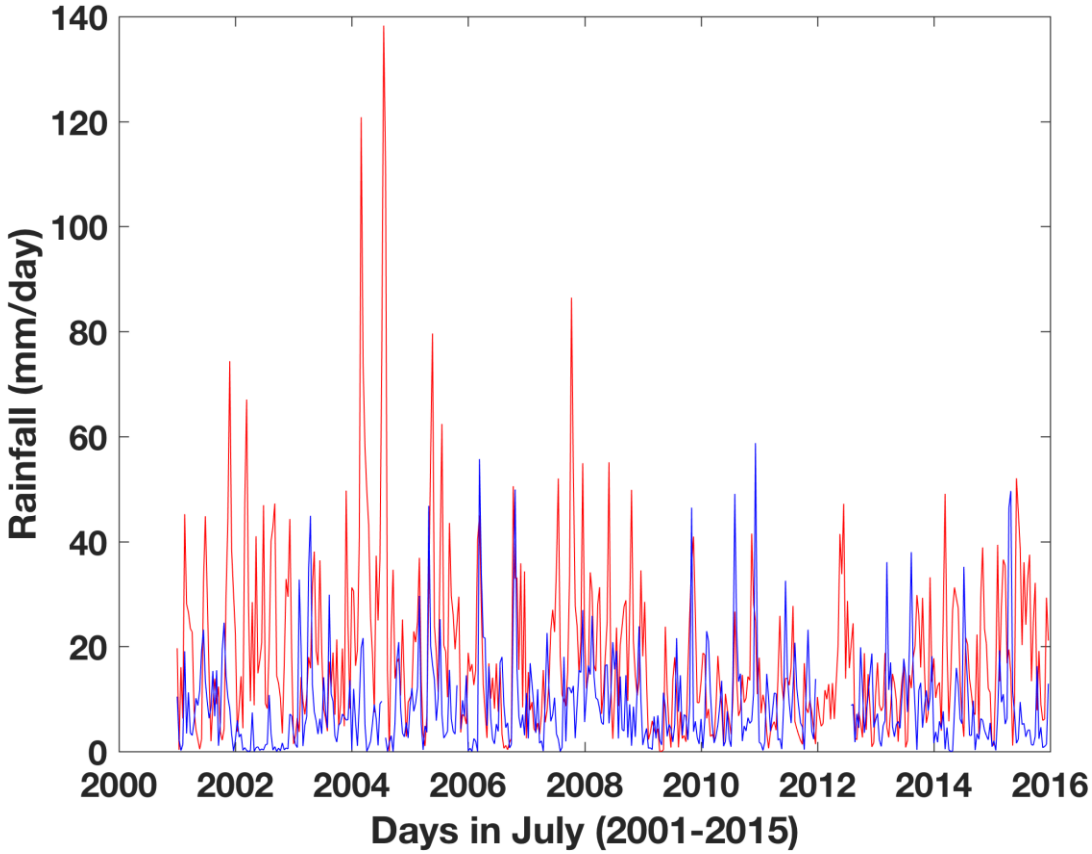


Figure 57: Comparison of time series of rainfall for days in July in Meghna basin. Blue: Rain gauge Red: Merged Satellite. Correlation coefficient between the two time series: 0.58.

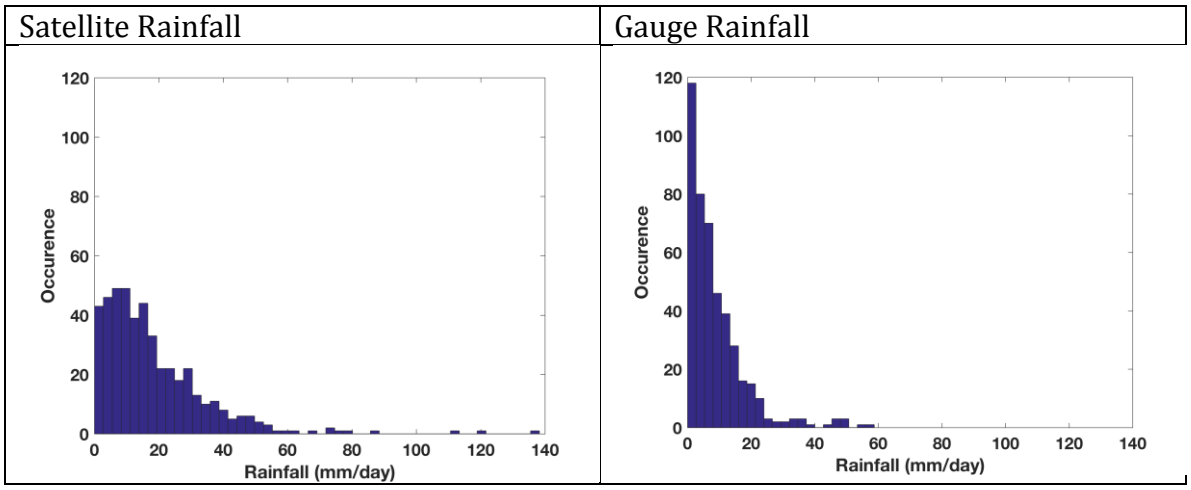


Figure 58: Comparison of rainfall histograms of July rainfall in the Meghna Basin.

Appendix D – Additional River Altimetry Data Examples and Plots

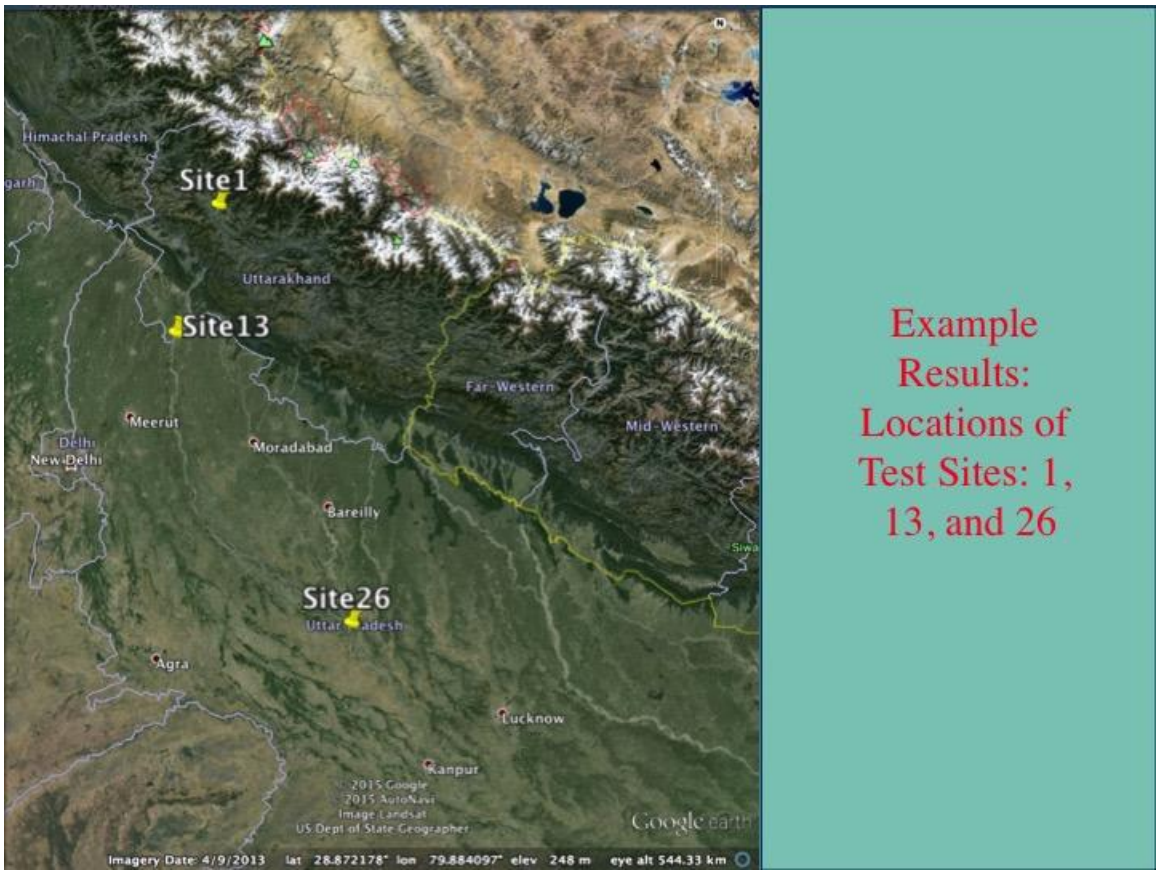


Figure 59: examples of locations of three altimetry sites high up in the Ganges catchment.

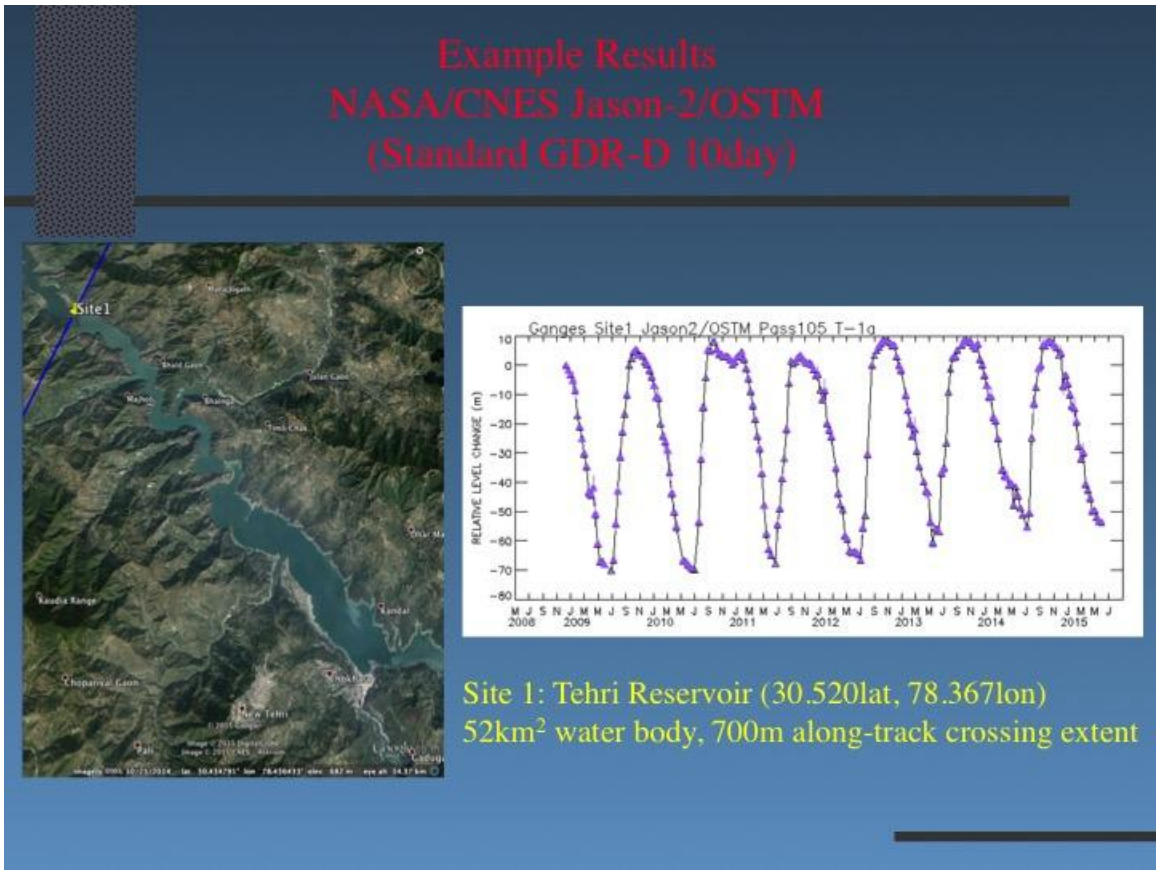


Figure 60: Time series of site 1.

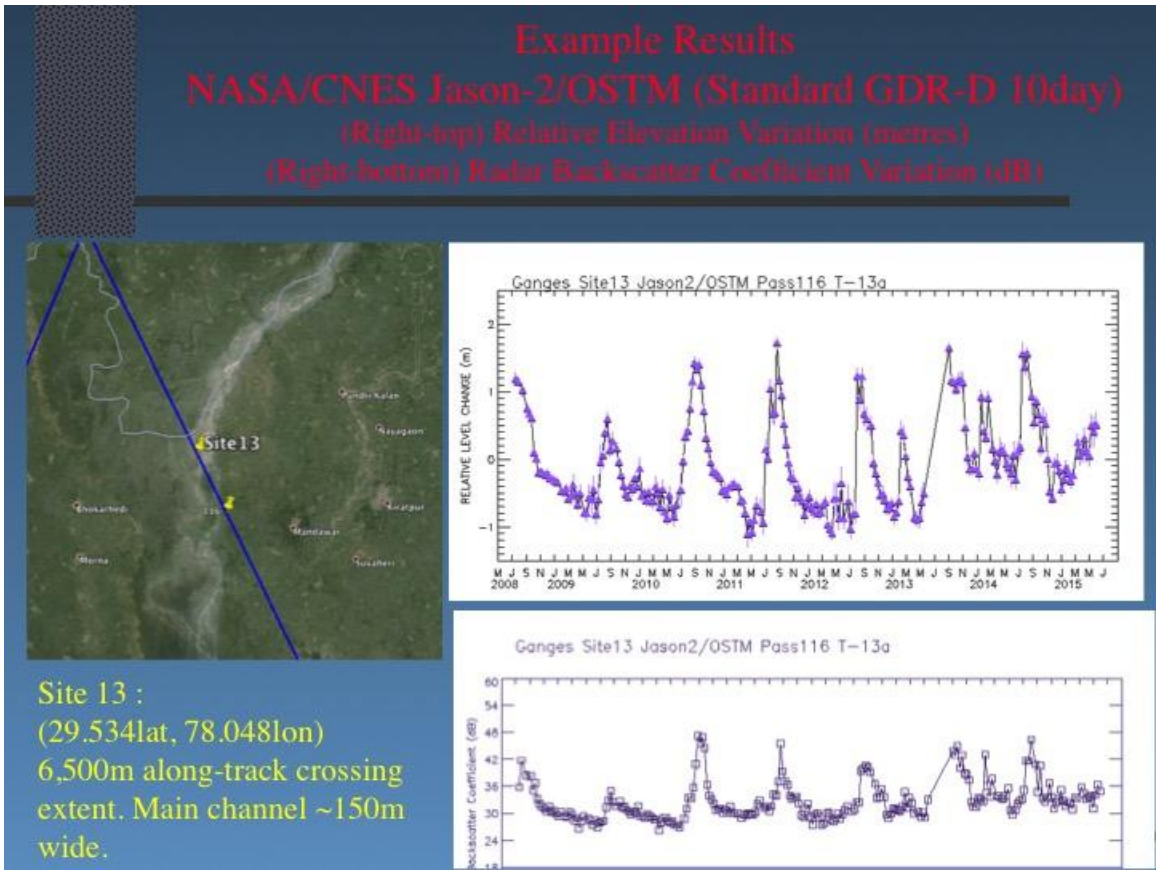


Figure 61: Time series of site 13.

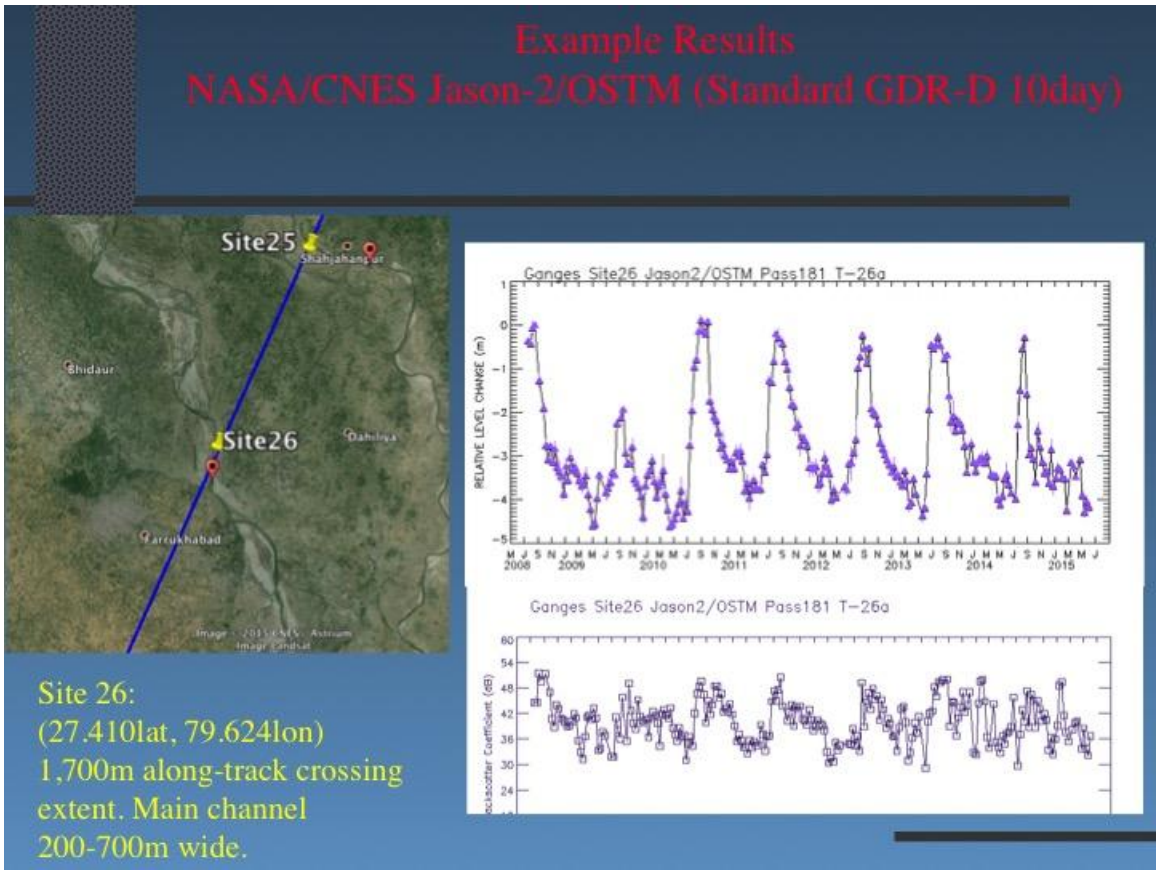


Figure 62: Time series of site 26.

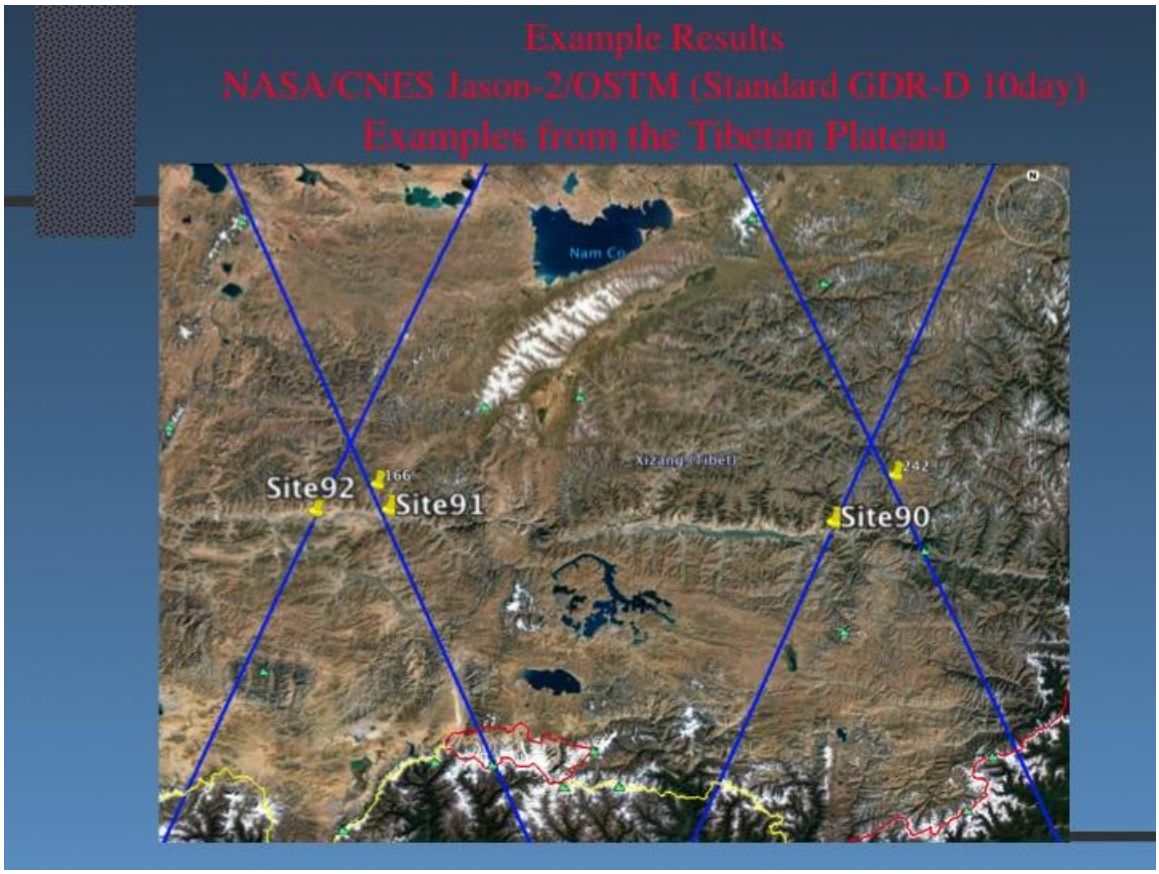


Figure 63: Sites 90, 91, and 92 high up in the Brahmaputra catchment.

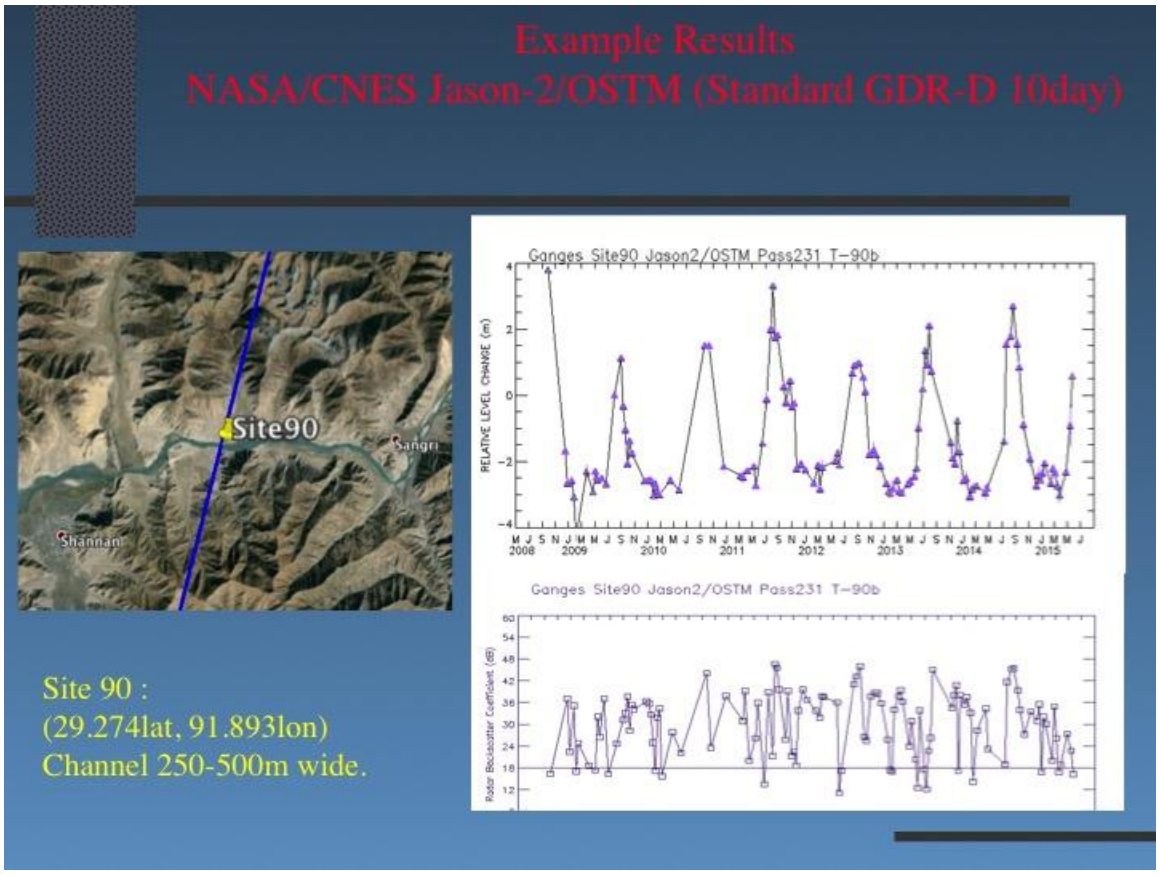


Figure 64: Time series of site 90 in the Brahmaputra.

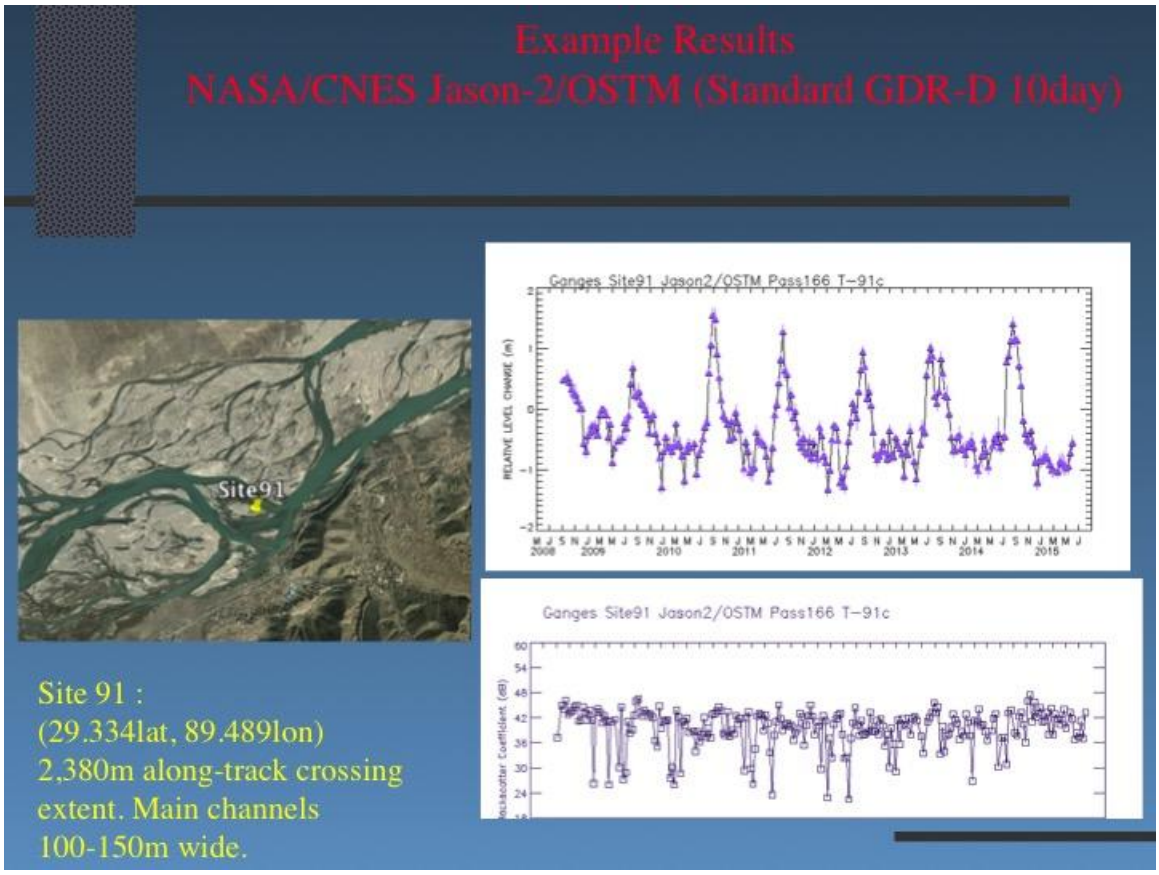


Figure 65: Time series of site 91 in the Brahmaputra.

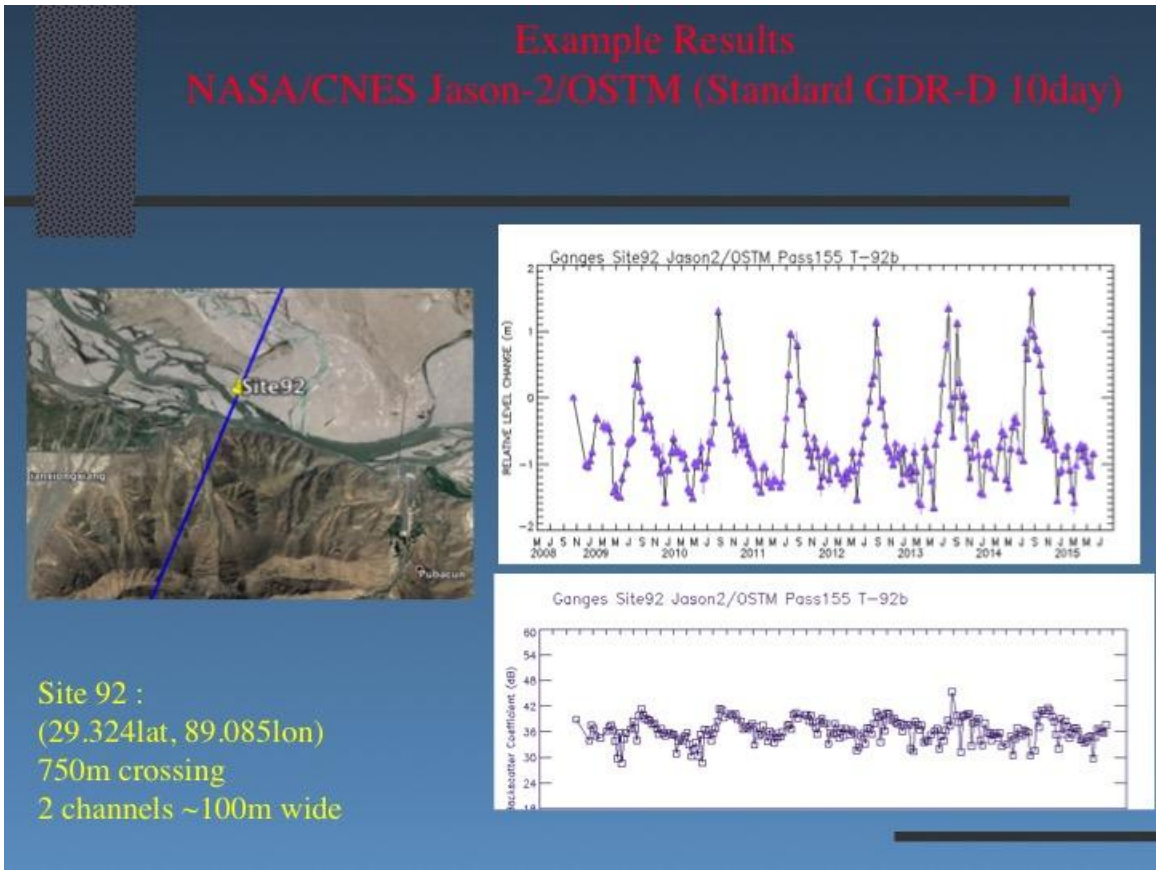


Figure 66: Time series of site 92 in the Brahmaputra.

Appendix A – Quantile Regression

Quantile regression (QR) has been little used in the atmospheric and hydrologic communities. We introduced this approach for this project, relaying on it heavily to generate ensemble forecasts and forecast error corrections that provide reliable probabilities of what values the river flow will exceedance on a daily basis (such that "flat histograms" of the forecasts are produced – described further in an appendix below).

Introduced in 1978 (Koenker and Bassett, 1978), quantile regression (QR) is an absolute error estimator that can conditionally fit specific quantiles of the regressand distribution (beyond just the mean or median), which does not rely on parametric assumptions of how either the regressand or residuals are distributed. And by virtue of being an ℓ_1-method, the conditional fit is less sensitive to outliers than square error estimators (Koenker and Portnoy 1997). See also Bremnes (2004) for an application of QR to calibrating weather variable output. Specifically, let $\{y_i\}$ represent a set of observations of the regressand y of interest, and $\{\mathbf{x}_i\}$ an associated set of predictor values. Analogous to standard linear regression, a linear function of $\mathbf{x}$ can be used to estimate to a specific quantile q_θ of y

$$q_\theta(\mathbf{x}_i;\boldsymbol{\beta}) = \beta_0 + \sum_{k=1}^{n} \beta_k x_{ik} + r_i \qquad (1)$$

with residuals $r_i = y_i - q_\theta(\mathbf{x}_i;\boldsymbol{\beta})$ and $q \hat{\imath}\ (0,1)$. However, instead of minimizing the squared residuals as with standard linear regression, in QR a weighted iterative minimization of $\{r_i\}$ is performed over $\boldsymbol{\beta}$:

$$\min \sum_{i=1}^{n} \rho_\theta(r_i) = \arg\min_{\beta} \sum_{i=1}^{n} \rho_\theta(y_i - q_\theta(\mathbf{x}_i;\boldsymbol{\beta})) \qquad (2)$$

with weighting function

$$r_q(u) = \begin{cases} qu & u \geq 0 \\ (q-1)u & u < 0 \end{cases} . \qquad (3)$$

In addition to the benefit of being less sensitive to outliers compared to standard linear regression, QR optimally determines the relationship of a regressor set on specific quantiles of the regressor, with no parametric assumptions. This point can be seen in the figure below where, by way of example, we have applied Eqs. (1) – (3) for persistence (previous day's temperature) as a forecast of 24hr July temperature (Salt Lake City airport 1979 to 2001; discussed below) for the 0.1, 0.5 (median), and 0.9 quantiles. The red line is the fit for the central tendency (mean) using standard linear regression; the middle black line is the fit of the median ($q_{0.5}$); upper black line is for $q_{0.9}$; and the lower black for $q_{0.1}$. Notice the similarity but noticeable divergence of the median and mean fits for larger temperatures. Notice

also the heteroscedastic behavior of the persistence QR fitting, which is seen by the convergence of the 0.1 and 0.9 quantile lines for higher temperatures. Physically, such behavior can be justified from higher temperatures being typically associated with high pressure anomalies, which have longer persistence than cold front passages.

Another powerful benefit of QR is that the cost function Eq (3) is precisely associated with the creation of a flat rank histogram verification, which itself is a necessary requirement for a calibrated ensemble forecast. Here, we implicitly define a "calibrated ensemble" as one being equivalent to a random draw from an underlying (but typically unknown) probability distribution function. In addition, although QR constrains the resultant quantile estimators to satisfying this requirement, at the same time it also constrains the estimators to optimal "sharpness" (Wilks 1995; i.e. creating "narrow" forecast PDFs as compared to a purely climatological distribution).

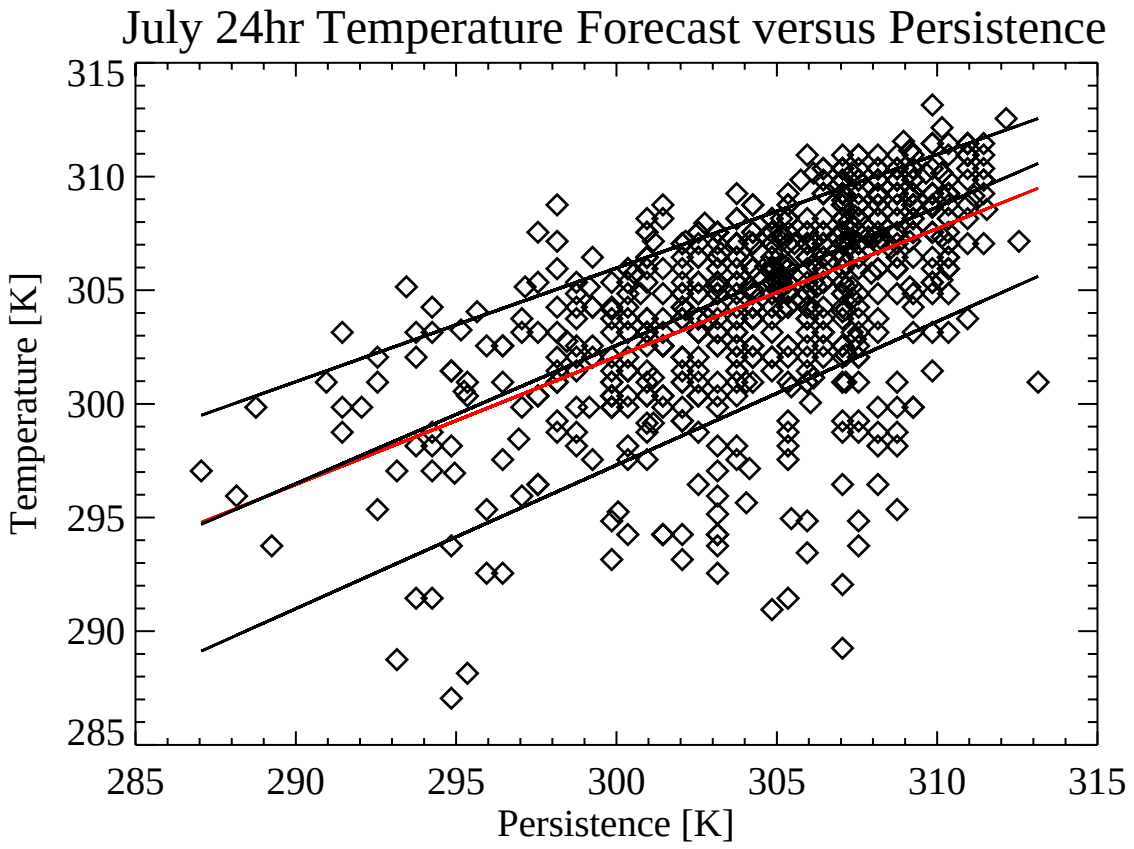


Figure 67: Previous day's temperature (persistence) used as a forecast of 24hr temperature, Salt Lake City airport data, 1979 to 2001. Red line is the fit for the central tendency (mean) using standard linear linear regression; middle black line is the fit of the median (0.5 quantile); upper black 0.9 quantile; lower black 0.1 quantile. Notice the similarity but noticeable divergence of the median and mean fits for larger temperatures. Notice also the heteroscedastic behavior of the persistence fitting, which is seen by the convergence of the 0.1 and 0.9 quantile lines for higher temperatures.

Appendix B – Quantile-to-Quantile Mapping

Generally, all practical hydrologic models require some form of calibration arising from reasons such as incomplete knowledge of watershed properties and the necessary parameterization of transport at unresolved scales. Through the hydrologic model parameter calibration process, some of the additional errors in the watershed inputs (rainfall, in particular) can also be implicitly reduced (e.g. runoff errors arising from an over-bias in rainfall forcing can be implicitly minimized by increasing the parameterization of evapotranspiration). However, such bias reduction through calibration does not occur if there are relative errors between more than one input of the same physical quantity. In particular, using both weather forecasts and observationally-based estimates of precipitation concurrently to generate unbiased discharge forecasts, requires that these two data sources maintain their statistical similarity such that there are no relative biases between their statistical moments (mean, variance, and skewness, in particular). Note that quantile regression (discussed above) could also provide the required bias reduction on a forecast-day-by-forecast-day bases. However, because we feel that the merged satellite precipitation product (i.e. "observations") inherently have their own error, we hypothesize it is more optimal to only utilize a more "global" statistical adjustment of the TIGGE precipitation forecasts to require them to only be drawn from the same probability distribution function as the observations. Quantile-to-quantile (q-to-q) mapping described next is designed to exactly do this.

To minimize systematic differences in the hydrologic model inputs that can lead to systematic hydrologic forecasting errors (in our case, satellite precipitation estimates [which were used to calibrate the hydrologic models] and the THORPEX-TIGGE ensemble precipitation forecasts which drive the hydrologic model forward in time), a quantile-to-quantile (q-to-q) mapping technique was implemented in this project. As applied here, the q-to-q technique forces each TIGGE ensemble precipitation forecast catchment-averaged over each basin we are forecasting for (input E, Fig 3) to be statistically sampled from the cumulative distribution function (CDF) of the associated observational historical record (i.e. satellite-precipiation estimates) catchment-integrated for the same basin (input S Fig 2). In essence, this converts the numerical weather prediction product from a physical variable to a probability forecast by performing a bias correction at each quantile of the forecast CDF. The manner in which the technique is applied to the TIGGE medium-range forecasts is described below in more detail.

Calculate the monsoon season "climatological" empirical CDF's for both the observations and the forecast model catchment-integrated valuesfor basin *i*. The observational (*o*) climatological CDF for basin *i* and precipitation amount *y*, $C_o^i(y)$, is derived from the merged satellite data (available from 2002 onwards). The forecast (*M*) CDF's at lead-time *f* and basin *i*, $C_M^{i,f}(y)$, are derived from stored forecasts

(available 2014 onwards), which are first catchment-integrated over the same basin, and the CDF computation is done independently for each forecast lead-time *f* (24 hour, 48 hour, etc.). Formally, these CDF's are given by

$$C_o^i(x) = \mathrm{P}\{X \leq x\}, \tag{4}$$

and

$$C_M^{i,f}(y) = \mathrm{P}\{Y \leq y\}, \tag{5}$$

where P{α} denotes the probability of the event α, and *x* (*y*) a specific value of the generic random variable *X* (*Y*) representing the observed (forecast) weather variable of interest (here, precipitation only). Using these CDF's, a direct quantile-by-quantile bias correction is made to each basin *i*, lead-time *f*, and precipitation ensemble member *fcst* of the forecast, $y = p_{fcst}^{i,f}$, of the forecast by simply setting (4) equal to (5), and determining the value of $x = p_{adj}$ that satisfies the equation. In this manner, the "observation-space" quantile is matched with the "forecast model-space" quantile as shown in the figure below, and p_{adj} is used in replace of $p_{fcst}^{i,f}$ in the hydrometeorological application.

Note that this technique ensures that the forecasts produce the same climatological number of "no rain" events as observed. It also preserves the spatial and temporal (ranked) covariances of the weather variable forecast fields (as generated by the numerical weather prediction model – although not so important for our application of catchment-integrated values).

If there is no contemporaneous forecast and observational data, we suggest that the q-to-q technique may be the best one can do in bringing forecasts into statistical similarity with the observations. However, although the mapping improves the forecast bias at each quantile by steering the forecasts into having the same frequency of events as the observation climatological record, this process still does not necessarily ensure that the forecasts' skill will be increased. Such skill degradation can (arbitrarily) arise because the technique does not directly improve the forecast model's reliability by neglecting the conditional relationship between observations and forecasts.

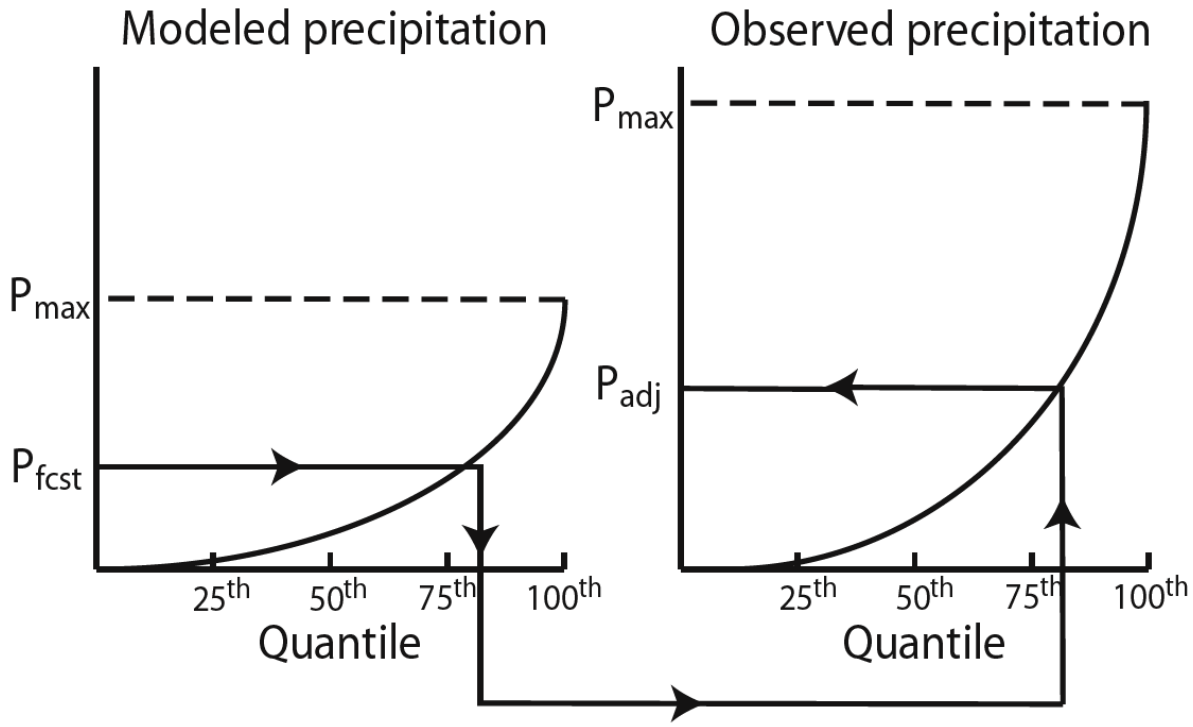


Figure 68: The quantile-to-quantile (q-to-q) correction system. Both modeled and observed precipitation are binned into quantiles. The model precipitation is mapped onto the observed precipitation fields by setting respective modeled quantiles to observed quantiles. In the figure, the forecast precipitation (p_{fcst}) of the 80th quantile is set to observed precipitation (p_{adj}) 80th quantile. The method ensures that the forecasts produce the same number of "no rain" events as the observations.

Appendix D – WRF Application for Flood Forecasting

In this appendix we discuss further the preliminary exploration of the skill of employing the NCAR Weather Research and Forecasting (WRF) model over the Ganges and Brahmaputra domains (with special focus on the northeast Haor regions of Bangladesh) to capture small scale "flash flood" events that would be at smaller resolution than the medium-range TIGGE ensemble NWP forecasts that are used in this project, in collaboration with our colleague Prof. Saiful Islam at the Bangladesh University of Engineering and Technology (BUET).

The northeast Haor regions of Bangladesh are typical lowland which experiences flash flood every year. If flash flood occurs before the harvesting of the only crop (Boro rice), it becomes critical. According to the officials of the Department of Agricultural Extension, Boro rice crop on over 22, 000 hectares out of 4,00,000 hectares was damaged by flash floods this year. Providing accurate early warning of the flash with sufficient lead time is essential for that region. At present, the early warning and flash flood mechanism of that region is neither efficient nor timely. Here we include a recent newspaper articles discussing some of the importance of improved flash flood prediction: http://bangladeshchronicle.net/2016/04/rice-crop-damaged-by-flash-floods-in-haor-belt/.

Below we show the model domain of the WRF application.

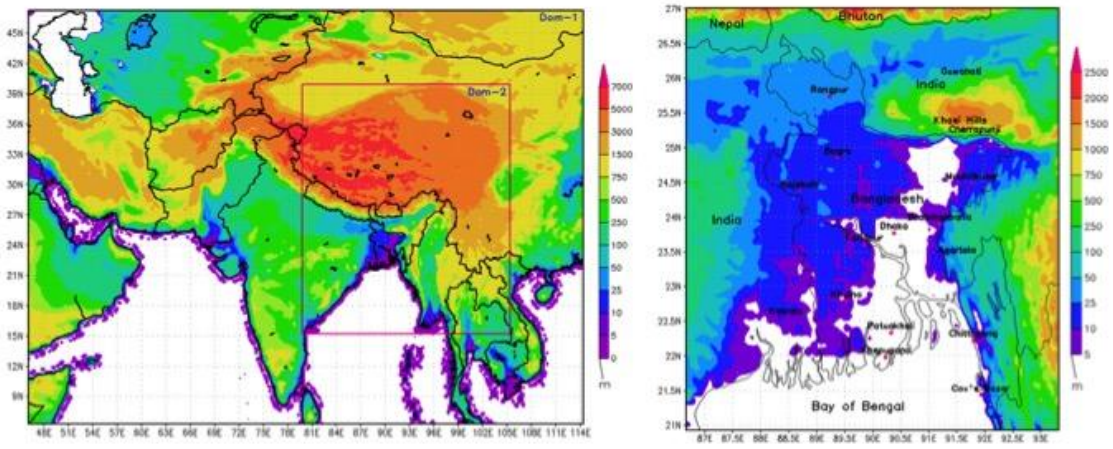


Figure 69: WRF model domain used for rainfall forecasting to improve flash flood forecasting of the north east Haor regions of Bangladesh

Below for reference we provide the model parameters used in the WRF setup.

Model Features	Configurations
Horizontal Resolution	18 km and 6 km
Vertical Levels	30
Topography	USGS
Time Integration	Semi Implicit
Vertical Differencing	Arakawa's Energy Conserving Scheme
Time Filtering	Robert's Method
Horizontal Diffusion	2nd order over Quasi-pressure, surface, scale selective
Convection	KF
PBL	YSU
Cloud Microphysics	WSM6
Surface Layer	Monin-Obukhov
Radiation	RRTM (LW), SW (Dudhia 1989)
Gravity Wave Drag	No
Land Surface Processes	Unified NOAH Land Surface Model

Table 5: WRF model parameters for simulating flash flood event on 18 April 2010

In the next two figures below we show the TRMM 3-hourly precipitation we use as comparison to the WRF model simulations for the severe rainfall event of April 17 and 18. The first figure is for April 17th, and the second for April 18th.

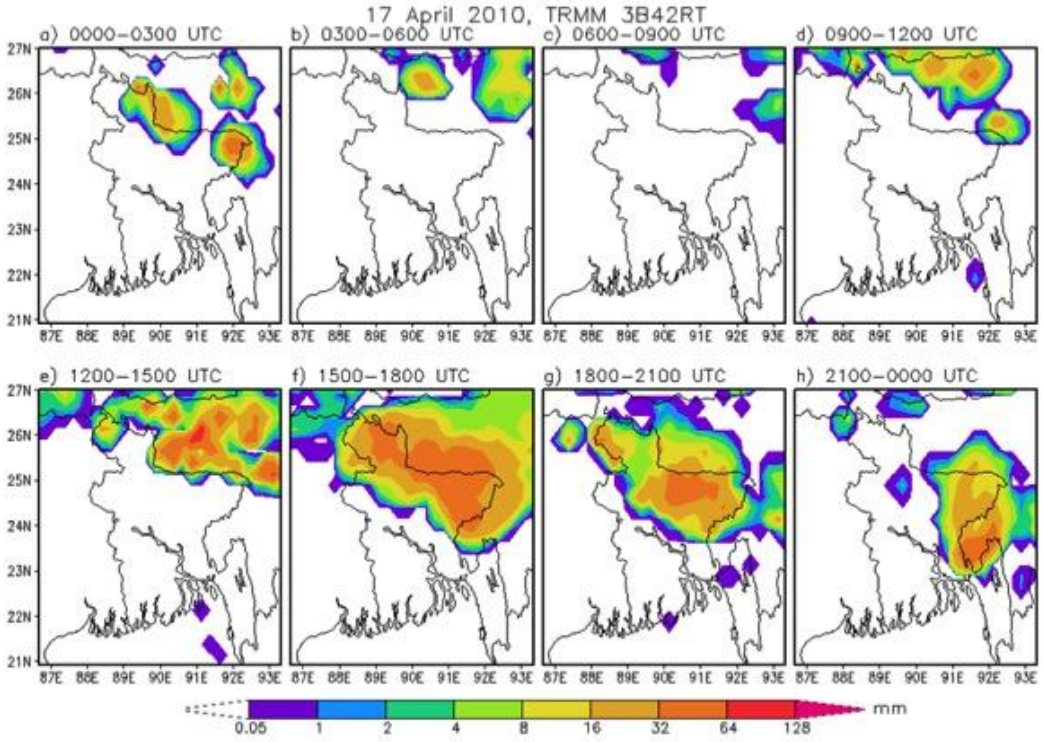


Figure 70: TRMM derived 3 hourly rainfall on 17 April 2010

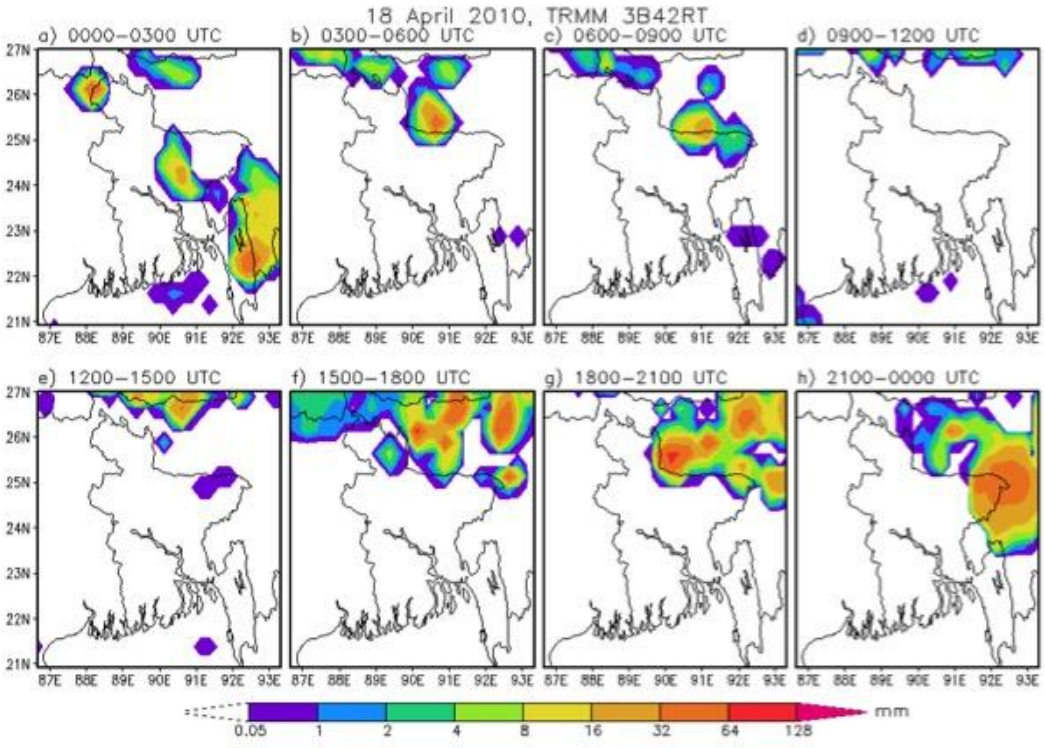


Figure 71: TRMM derived 3 hourly rainfall on 18 April 2010

In the figure below, we show the WRF simulations for the same April 17-18th event, by way of comparison.

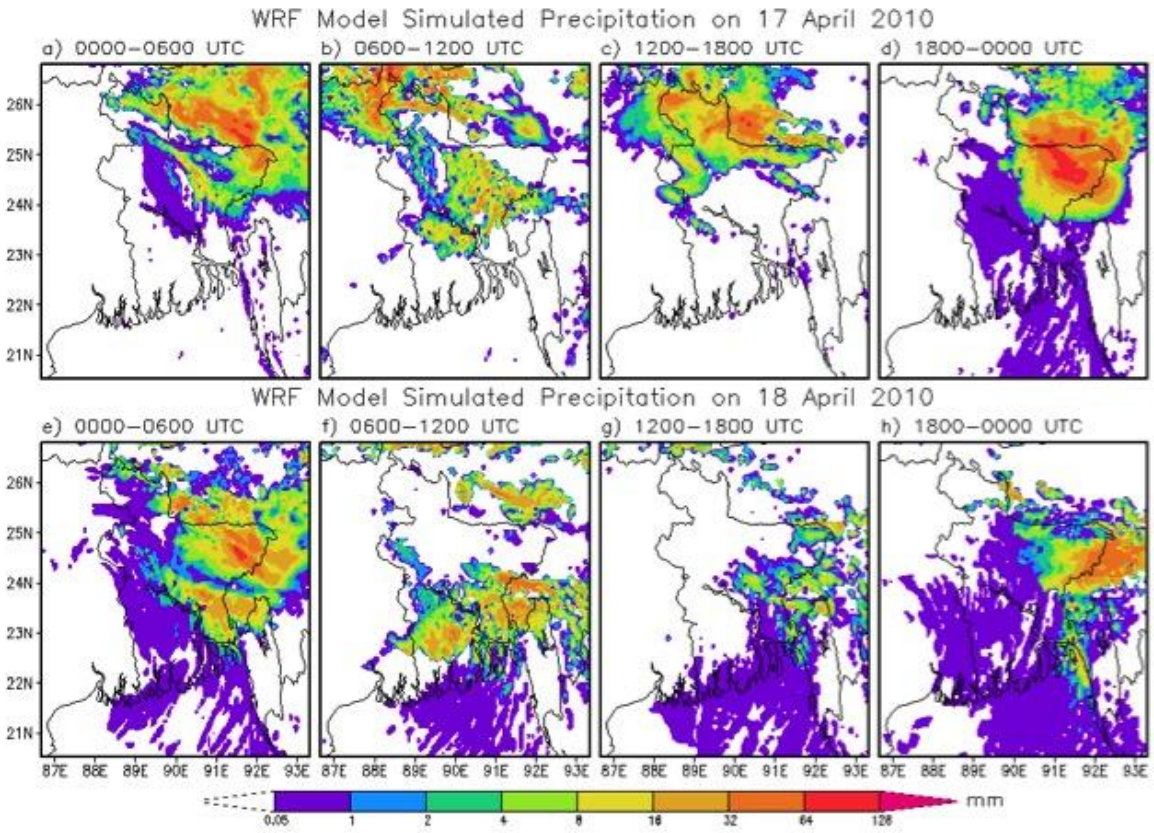


Figure 72: WRF simulated 6 hourly rainfall on 17 and 18 April 2010

In the table below, we show a statistical analysis of this comparison between TRMM and WRF simulations, showing the overall similarity of the products, albeit with different scale.

Rainfall		AWS	TRMM	Model D1	Model D2
24 hour rainfall (mm)		152.5	100.44	97.89	93.92
3 hourly rainfall (mm)	Mean Difference	37.56	19.05	28.31	28.58
	RMSD	48.45	27.37	36.99	37.17
	NRMSD	0.74	0.42	0.57	0.57

Table 6: Statistical Analysis -- The mean difference, RMSD and NRMSD of the rainfall is comparatively more (wet bias) for the model compared to rain gauge observation (Table). It should be noted that, rainfall measured by BMD was 161mm on 17 April 2010.

Future Improvements

In addition to further comparisons with TRMM for further extreme rainfall events, future work will also compare the WRF simulations with the Bangladesh Meteorology Department's S-band weather radar (~10 cm wave length) at Dhaka (90.4° N, 23.7° E), running since 2000. It has a maximum scan radius of about 400 km horizontally and effective scan radius of about 250 km. The radar collects reflectivity data, which are stored in six rain status categories. The BMD Dhaka radar collects reflectivity (dBZ) data and automatically converts it to the precipitation rate (mm h^{-1}). The figure below shows the scan radius of this radar.

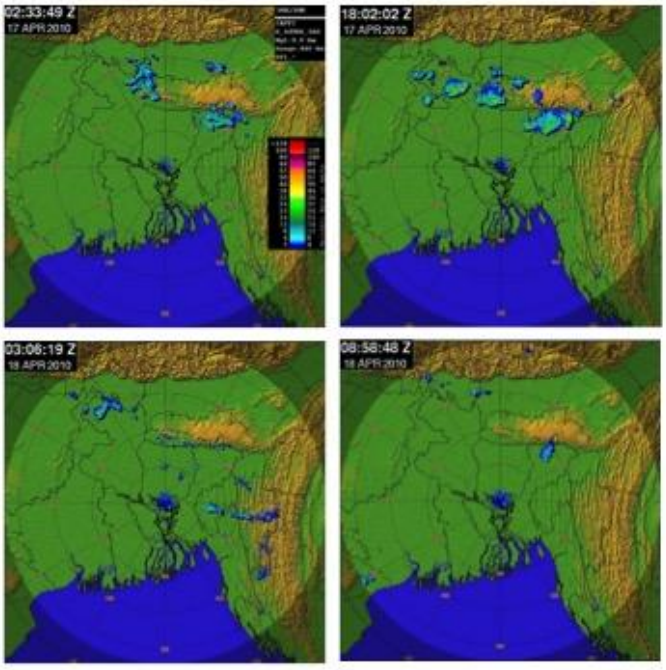


Figure 73: Planning to validate with Radar Driven Rain Rate.

Conclusions

- This study has investigated the ability of a cloud-resolving WRF to reproduce the convective cells associated with a heavy rainfall event over the Haor region of Bangladesh.
- The ability of WRF model with a highest resolution of 18 and 6 km horizontal grid spacing to predict heavy rainfall near Sylhet in the northeastern part of Bangladesh on 17 and 18 April 2010 was evaluated.
- The model underestimated the strength of the storm in general in term of the rainfall. The simulated rainfall is usually about 93.92 mm day^{-1}, for outer domain and 97.89 mm day^{-1} for inner domain but observed amount is 161

mm day^{-1}. TRMM retrieved 24 hours rain amount is 100 mm which is also less compared to rain gauge observation.

Appendix C – GIS-based Rain Gauge Siting Tool

Overview

Accurate and reliable real-time monitoring and dissemination of observations of surface weather conditions is critical for a variety of societal applications. Applications that provide local and regional information about temperature, precipitation, moisture, and winds, for example, are important for agriculture, water resource monitoring, health, and monitoring of hazard weather conditions. These observations can be used to provide early warning alerts for extreme weather events such as flash floods and dust storms. The resulting data can be used by climate services to provide guidance on longer-term phenomena such as droughts and prolonged heat waves. When coupled with numerical weather predication (NWP) systems, the observations can improve the model forecasts to improve short-term (1-3 day) to seasonal forecasts (1-3 months). These data can provide verification information about NWP forecasts and remote sensing observations.

Background

To adequately capture the local spatial and temporal weather phenomena, an observation network has to be well designed. The design of the observation network requires evaluation of existing observations to quantify the variability of weather conditions, which inform the spatial and temporal monitoring requirements. The network design also depends on available resources (number of available instruments), land-surface/land-use conditions (rough/flat terrain, woodland/farmland), available infrastructure (roads, secure locations, communication networks), and location of population centers, rivers, or other critical monitoring regions.

A GIS framework is well suited to optimally design the network. A GIS network design tool will ingest all relevant and available geographical information, and objectively locate possible weather sites based on spatial design requirements.

An ideal observation network is designed to adequately observe the natural variability of a specific weather observable such as precipitation. Spatial correlation functions (Ciach and Krajewski 1999; Habib and Krajewski 2002, Krajewski et al. 2003) could be used to represent the spatial variability. These spatial correlation functions can also be used to spatially extrapolate the observations in other applications (NWP, hydrologic prediction, etc.). An example of one such spatial correlation function designed for a precipitation network (Senegal) is shown in Figure 74, giving a sense of the spatial scale and thus gage siting density required to adequately capture rainfall variability and magnitude for this region. However, another alternative to developing spatial correlation functions would be to rely on historical records of estimated weather fields, and sample over all possible

combinations of gage siting locations to minimize representative errors. It is this latter approach we utilize, where we rely on satellite precipitation estimates to provide approximate representations of rainfall magnitude, variability, and spatial and temporal covariance.

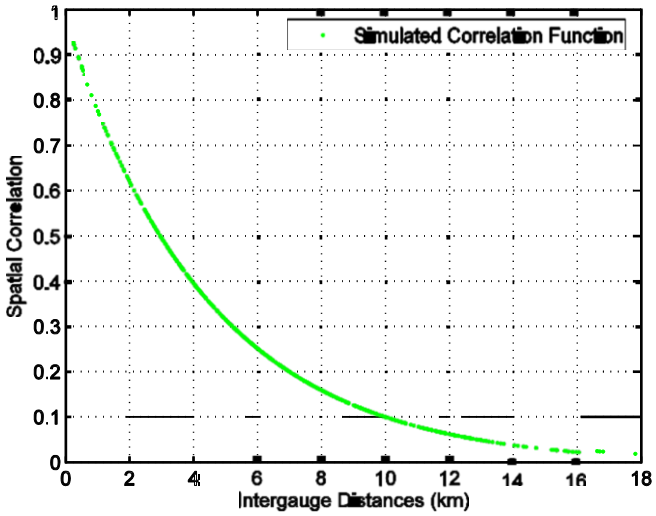


Figure 74: Example spatial correlation function used to optimize siting of precipitation gauges.

Design Considerations

Once an initial network has been designed using a GIS network design tool specifying general locations, site surveys would need to be conducted to determine the specific and proper location of each instrument. Ideally, the sites would be placed in a secure location such as government facilities, radio stations, or primary or secondary schools. Sites must be selected away from obstructions such as trees, buildings, fences, etc., identifying the required number of sites during the survey. Weather service staff could participate in the survey, since their participation would provide them with the opportunity to connect with local residents to explain the value of observations, and also give them hands-on experience with siting the stations. More specific technical design considerations are given below:

Required functionality:

- Clip to region of interest (river basin, state, etc.) if desired
- Mask by population, flood risk, slope, or some combination, if desired
- Determine some threshold for masking rainfall (80th percentile e.g.) and keep only days where at least one of the cells exceeds the threshold
- Calculate a region-total time series
- Restrict potential station locations by road network, accessibility, etc. if desired

- Through a skill measure (e.g. RMSE), determine the best location to put stations to best capture the region-total time series - done by either specifying the number of stations desired, or some level of acceptable uncertainty.

Required inputs:

- Precipitation products
- Shape files of desired regions - river basins and states
- Population, DEM slope, flood risk - to mask rainfall product with
- Road networks, or other station location restriction data

User inputs:

- Number and which precipitation products desired (e.g. TRMM, CMORPH, etc.)
- Region of interest
- Mask layers (if any)
- Precipitation threshold (if any)
- Number of sites or acceptable uncertainty
- Station location mask (if any).

An example of a prototype GIS network design is shown in Figure 75 for optimal placement of a single rain gage over Uttar Pradesh (derived using historic TRMM rainfall fields) to represent the rainfall over the whole state. The dot shows the optimal location; with the colors in each box showing the expected errors one would see at that respective location by relying on that single gage site to represent the rainfall over the whole state.

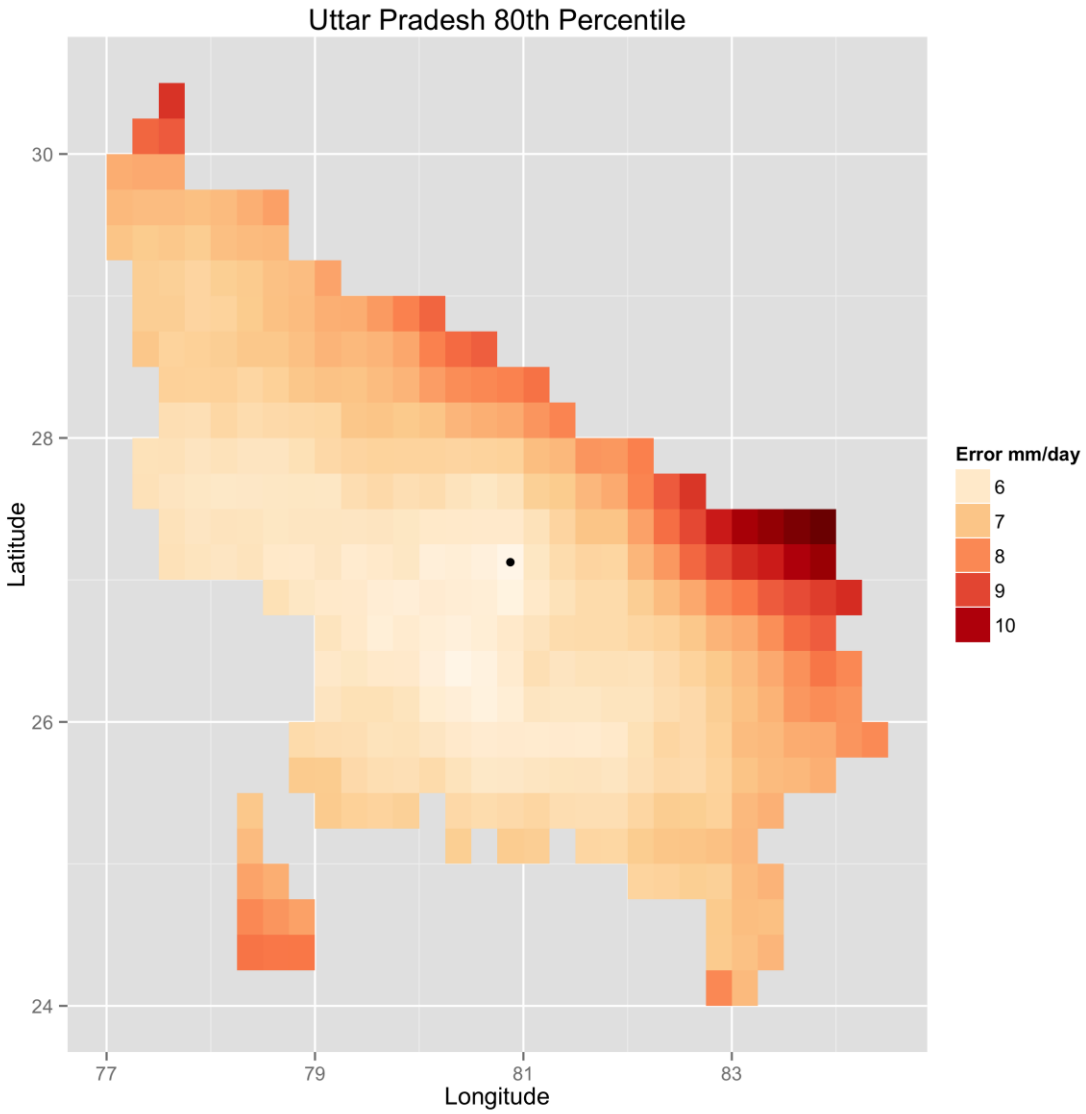


Figure 75: optimal siting of a single gage over Uttar Pradesh using historic TRMM rainfall fields, shown by the dot. Colors show the expected errors at that respective location in using this single site to represent the whole state's rainfall amount.

References

Amarnath, G., Umer, Y. M., Alahacoon, N., and Inada, Y.: Modelling the flood-risk extent using LISFLOOD-FP in a complex watershed: case study of Mundeni Aru River Basin, Sri Lanka, Proc. IAHS, 370, 131-138, doi:10.5194/piahs-370-131-2015, 2015.

Arulampalam, M. S., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian Tracking. *IEEE Transactions on Signal Processing, 50*, 174-188.

Biancamaria, S., F. Hossain, and D. P. Lettenmaier, 2011: Forecasting transboundary river water elevations from space: RIVER WATER HEIGHT FORECAST FROM SPACE. *Geophysical Research Letters*, **38**, n/a–n/a, doi:10.1029/2011GL047290.

Brakenridge, G. R., Nghiem, S. V., Anderson, E., & Chien, S. (2005). Space-based measurement of river runoff. *EOS, Transactions of the American Geophysical Union, 86*(19), 185-188.

Brakenridge, G. R., Nghiem, S. V., Anderson, E., & Mic, R. (2007). Orbital microwave measurement of river discharge and ice status. *Water Resources Research, 43*, W04405. doi: 10.1029/2006WR005238

Bremnes, J. B., 2004: Probabilistic forecasts of precipitation in terms of quantiles using NWP model output. *Mon. Wea. Rev.,* **132**, 338–347.

CEGIS. (2006). Sustainable end-to-end climate/flood forecast application through pilot projects showing measurable improvements. *CEGIS Base Line Report, 78 pp*.

Ciach, G. J., and W. F. Krajewski, 1999: On the estimation of radar rainfall error variance. *Advances in Water Resources*, **22**, 585–595.

Clark, M. P., Rupp, D. E., Woods, R. A., Zheng, X., Ibbitt, R. P., Slater, A. G., Schmidt, J., Uddstrom, M. J. (2008). Hydrocological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. *Advanced Water Resource, 31*, 1309-1324.

De Groeve, T., R. G. Brakenridge, and S. Paris, 2015: Global Flood Detection System: Data Product Specifications. https://floodobservatory.colorado.edu/Technical%20Note%20GFDS%20Data%20Products%20v1.pdf (Accessed February 10, 2016).

Habib, E., and W. F. Krajewski, 2002: Uncertainty analysis of the TRMM ground-validation radar-rainfall products: Application to the TEFLUN-B field campaign. *Journal of Applied Meteorology*, **41**, 558–572.

Hamill, T. M., 2012: Verification of TIGGE multimodel and ECMWF reforecast-calibrated probabilistic precipitation forecasts over the contiguous United States*. *Monthly Weather Review*, **140**, 2232–2252.

Hirpa, F. A., T. M. Hopson, T. De Groeve, G. R. Brakenridge, M. Gebremichael, and P. J. Restrepo, 2013: Upstream satellite remote sensing for river discharge forecasting: Application to major rivers in South Asia. *Remote Sensing of Environment*, **131**, 140–151, doi:10.1016/j.rse.2012.11.013.

Hopson, T. M., and P. J. Webster, 2010: A 1–10-Day Ensemble Forecasting Scheme for the Major River Basins of Bangladesh: Forecasting Severe Floods of 2003–07*. *Journal of Hydrometeorology*, **11**, 618–641, doi:10.1175/2009JHM1006.1.

Hopson, T. M., 2005: Operational Flood-Forecasting for Bangladesh. Ph.D. Thesis, Univ. of Colorado, 225 pp.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., . . . Stocker, E. F. (2007). The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology, 8*(1), 38-55.

Huffman, G. J., R. F. Adler, S. Curtis, D. T. Bolvin, and E. J.Nelkin (2005). Global rainfall analyses at monthly and 3-hr time scales. Measuring Precipitation from Space: EURAINSAT and the Future, V. Levizzani, P. Bauer, and J. F. Turk, Eds., Springer, 722 pp.

IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

Jongman, B., H. C. Winsemius, J. C. J. H. Aerts, E. Coughlan de Perez, M. K. van Aalst, W. Kron, and P. J. Ward, 2015: Declining vulnerability to river floods and the global benefits of adaptation. *Proceedings of the National Academy of Sciences*, **112**, E2271–E2280, doi:10.1073/pnas.1414439112.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeor., 5, 487–503.

Kleuskens, M., Westerhoff, R. S., & Huizinga, J. (2011). *Operational flood mapping: A pilot study in the Mekong Area*. Paper presented at the Proceedings of the International Symposium of Remote Sensing of the Environment, Sydney, Australia.

Krajewski, W. F., G. J. Ciach, and E. Habib, 2003: An analysis of small-scale rainfall variability in different climatic regimes. *Hydrological Sciences Journal*, **48**, 151–162, doi:10.1623/hysj.48.2.151.44694.

Koenker, R. and W. Portnoy, 1997: The Gaussian Hare and the Laplacean Tortoise: Computability of Squared-error vs. Absolute Error Estimators. Statistical Science, 12, 279-300.

Koenker, R. W. and G. W. Bassett, 1978: Regression Quantiles. Econometrica, 46, 33-50.

Lee, H. S., Seo, D. J., Lui, Y., Koren, V., McKee, P., & Corby, R. (2012). Variatinal assimilation of streamflow into operational ditributed hydologic models: Effect of spatiotemporal scale adjustment. *Hydrology and Earth System Sciences, 16*, 2233-2251. doi: 10.2194/hess-16-2233-2012

Mitra, A. K., G. R. Iyengar, V. R. Durai, J. Sanjay, T. N. Krishnamurti, A. Mishra, and D. R. Sikka, 2011: Experimental real-time multi-model ensemble (MME) prediction of rainfall during monsoon 2008: large-scale medium-range aspects. *Journal of earth system science*, **120**, 27–52.

Montanari, M., Hostache, R., Matgen, P., Schumann, G., Pfister, L., & Hoffman, L. (2009). Calibration and sequential updating of a coupled hydrolic-hydrualic model using remote sensing-derived water stages. *Hydrology and Earth System Sciences, 13*, 367-380. doi: 10.5194/hess-13-367-2009

Moradkhani, H., Hsu, K., Gupta, H. V., & Sorooshian, S. (2005). Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using particle filter. *Water Resources Research, 41*, W05012. doi: 10.1029/2004WR003604

Mylne, K. R., 1999: The use of forecast value calculations for optimal decision making using probability forecasts. Preprints, 17th Conf. on Weather Analysis and Forecasting. Denver, CO, Amer. Meteor. Soc., 235-239.

Pandey, R., and G. Amarnath, 2015: The potential of satellite radar altimetry in flood forecasting: concept and implementation for the Niger-Benue river basin. *Proceedings of the International Association of Hydrological Sciences*, **370**, 223–227, doi:10.5194/piahs-370-223-2015.

Prakash, S., A. K. Mitra, I. M. Momin, D. S. Pai, E. N. Rajagopal, and S. Basu, 2015: Comparison of TMPA-3B42 Versions 6 and 7 Precipitation Products with Gauge-Based Data over India for the Southwest Monsoon Period. *Journal of Hydrometeorology*, **16**, 346–362, doi:10.1175/JHM-D-14-0024.1.

Ranade, A., A. K. Mitra, N. Singh, and S. Basu, 2014: A verification of spatio-temporal monsoon rainfall variability across Indian region using NWP model output. *Meteorology and Atmospheric Physics*, **125**, 43–61, doi:10.1007/s00703-014-0317-5.

Revilla-Romero, B., H. E. Beck, P. Burek, P. Salamon, A. de Roo, and J. Thielen, 2015: Filling the gaps: Calibrating a rainfall-runoff model using satellite-derived surface water extent. *Remote Sensing of Environment*, **171**, 118–131, doi:10.1016/j.rse.2015.10.022.

Richardson, D. S., 2000: Skill and economic value of the ECMWF ensemble prediction system. Quart. J. Roy. Meteor. Soc., 126, 649-668.

Rosenberg, E. A., A. W. Wood, and A. C. Steinemann, 2013: Informing Hydrometric Network Design for Statistical Seasonal Streamflow Forecasts. *Journal of Hydrometeorology*, **14**, 1587–1604, doi:10.1175/JHM-D-12-0136.1.

Seo, D. J., Cajina, L., Corby, R., & Howieson, T. (2009). Automatic state updating for operational streamflow forecasting via variational data assimilation. *Journal of Hydrology, 367*, 255-275.

Siddique-E-Akbor, A. H. M., F. Hossain, H. Lee, and C. K. Shum, 2011: Inter-comparison study of water level estimates derived from hydrodynamic–hydrologic model and satellite altimetry for a complex deltaic environment. *Remote Sensing of Environment*, **115**, 1522–1531, doi:10.1016/j.rse.2011.02.011.

Song, X. Y., & Lee, S. Y. (2004). Bayesian analysis of two-level nonlinear structural equation models with continuous and polytomous data. *British Journal of Mathematical and Statistical Psychology, 57*, 29-52.

Sunilkumar, K., T. Narayana Rao, K. Saikranthi, and M. Purnachandra Rao, 2015: Comprehensive evaluation of multisatellite precipitation estimates over India using

gridded rainfall data: EVALUATION OF MPES OVER INDIA. *Journal of Geophysical Research: Atmospheres*, **120**, 8987–9005, doi:10.1002/2015JD023437.

Syvitski, J. P. M., & Brakenridge, G. R. (2013). Causation and avoidance of catastrophic flooding along the Indus river, Pakistan. *GSA Today, 23*(1). doi: 10.1130/GSATG1165A.1131

Webster, P. J. (2013). Improve weather forecasts for the developing world. *Nature, 493*, 17-19.

Webster, P. J., Jian, J., Hopson, T. M., Hoyos, C. D., Agudelo, P., Chang, H.-R., . . . Subbiah, A. R. (2010). Extended-range probabilistic forecasts of Ganges and Brahmaputa floods in Bangladesh. *Bulletin of the American Meteorological Society, 91*, 1493-1514.

Webster, P. J., T. Hopson, C. Hoyos, A. Subbiah, H-. R. Chang, R. Grossman, 2006: A three-tier overlapping prediction scheme: Tools for strategic and tactical decisions in the developing world. In Predictability of Weather and Climate, Ed. T. N. Palmer, Cambridge University Press.

Weerts, A. H., & El Serafy, G. Y. H. (2006). Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models. *Water Resources Research, 42*, W09403. doi: 10.1029/2005WR004093

Westerhoff, R. S., Huizinga, J., Kleuskens, M., Burren, R., & Casey, S. (2010). *Operational satellite-based flood mapping using the Delft-FEWS System.* In Proceedings of the ESA Living Planet Symposium, Bergen, Norway, June 28-July 2, 2010. Retrieved Jan 10, 2013, from http:/kennisonline.deltares.nl/product/22381

Wilks, D. S., 2001: A skill score based on economic value for probability forecasts. Meteor. Appl., 8, 209-219.

Wilks, D. S., 1995: Statistical Methods in the Atmospheric Sciences. Cambridge Press, 467 pp.

Xie P. P., B. Rudolf, U. Schneider, and P.A. Arkin, 1996: Gauge-based monthly analysis of global land precipitation from 1971 to 1994. *J. Geophys. Res. - Atmos.,* 101 (D14), 19023-19034.

Zhu, Y., Z. Toth, R. Wobus, D. Richardson, K. Mylne, 2002: The economic value of ensemble-based weather forecasts. Bull. Amer. Meteor. Soc., 83, 73-83.