

01 Jan 2015

## Data-enabled Field Experiment Planning, Management, And Research Using Cyberinfrastructure

Ibrahim Demir

Helen Conover

Witold F. Krajewski

Bong Chul Seo

*Missouri University of Science and Technology*, [bongchul.seo@mst.edu](mailto:bongchul.seo@mst.edu)

*et. al.* For a complete list of authors, see [https://scholarsmine.mst.edu/civarc\\_enveng\\_facwork/2653](https://scholarsmine.mst.edu/civarc_enveng_facwork/2653)

Follow this and additional works at: [https://scholarsmine.mst.edu/civarc\\_enveng\\_facwork](https://scholarsmine.mst.edu/civarc_enveng_facwork)



Part of the [Civil and Environmental Engineering Commons](#)

---

### Recommended Citation

I. Demir et al., "Data-enabled Field Experiment Planning, Management, And Research Using Cyberinfrastructure," *Journal of Hydrometeorology*, vol. 16, no. 3, pp. 1155 - 1170, American Meteorological Society, Jan 2015.

The definitive version is available at <https://doi.org/10.1175/JHM-D-14-0163.1>

This Article - Journal is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Civil, Architectural and Environmental Engineering Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact [scholarsmine@mst.edu](mailto:scholarsmine@mst.edu).



## Data-Enabled Field Experiment Planning, Management, and Research Using Cyberinfrastructure

IBRAHIM DEMIR

*IIHR–Hydroscience and Engineering, The University of Iowa, Iowa City, Iowa*

HELEN CONOVER

*Information Technology and Systems Center, University of Alabama in Huntsville, Huntsville, Alabama*

WITOLD F. KRAJEWSKI, BONG-CHUL SEO, AND RADOSŁAW GOSKA

*IIHR–Hydroscience and Engineering, The University of Iowa, Iowa City, Iowa*

YUBIN HE, MICHAEL F. MCENIRY, AND SARA J. GRAVES

*Information Technology and Systems Center, University of Alabama in Huntsville, Huntsville, Alabama*

WALTER PETERSEN

*Office of Field Support, NASA Goddard Space Flight Center, Wallops Island, Virginia*

(Manuscript received 28 August 2014, in final form 30 January 2015)

### ABSTRACT

In the spring of 2013, NASA conducted a field campaign known as Iowa Flood Studies (IFloodS) as part of the Ground Validation (GV) program for the Global Precipitation Measurement (GPM) mission. The purpose of IFloodS was to enhance the understanding of flood-related, space-based observations of precipitation processes in events that transpire worldwide. NASA used a number of scientific instruments such as ground-based weather radars, rain and soil moisture gauges, stream gauges, and disdrometers to monitor rainfall events in Iowa. This article presents the cyberinfrastructure tools and systems that supported the planning, reporting, and management of the field campaign and that allow these data and models to be accessed, evaluated, and shared for research. The authors describe the collaborative informatics tools, which are suitable for the network design, that were used to select the locations in which to place the instruments. How the authors used information technology tools for instrument monitoring, data acquisition, and visualizations after deploying the instruments and how they used a different set of tools to support data analysis and modeling after the campaign are also explained. All data collected during the campaign are available through the Global Hydrology Resource Center (GHRC), a NASA Distributed Active Archive Center (DAAC).

### 1. Introduction

The Iowa Flood Studies (IFloodS) campaign collected massive amounts of precipitation and streamflow data to provide a robust reference for investigations of space-based products and their ability to drive hydrologic

models used for flood prediction in real time. Of particular interest was the evaluation of the capacities of the international Global Precipitation Measurement (GPM) satellite mission (Schwaller and Morris 2011; Hou et al. 2014; Tapiador et al. 2012) to provide global predictions of floods and other precipitation-induced natural hazards. The precipitation data allow us to evaluate (validate) space-based rainfall estimates across several spatial scales, while multiple stream gauges allow us to evaluate the predictive abilities of the hydrologic models used to convert rainfall into runoff and to forecast discharge and

---

*Corresponding author address:* Ibrahim Demir, Iowa Flood Center, IIHR–Hydroscience and Engineering, 207 C. Maxwell Stanley Hydraulics Laboratory, Iowa City, IA 52242.  
E-mail: ibrahim-demir@uiowa.edu

TABLE 1. Summary of GPM GV field campaigns supported by GHRC.

GPM GV field campaigns		
Fall 2010	Light Precipitation Validation Experiment (LPVEx)	Finland
Apr–Jun 2011	Midlatitude Continental Convective Clouds Experiment (MC3E)	North-central Oklahoma
Jan–Feb 2012	GPM Cold-Season Precipitation Experiment (GCPEX)	Ontario, Canada
Apr–Jul 2013	IFloodS	Northeastern Iowa
May–Jun 2014	Integrated Precipitation and Hydrology Experiment (IPHEX)	North Carolina, Appalachians/Piedmont region
Nov 2015–Jan 2016	Olympic Mountain Experiment (OLYMPEX)	Washington, Olympic Peninsula

flooding. Petersen and Krajewski (2013) provide an overview of the IFloodS science objectives.

The deployment of multiple weather radars as well as numerous disdrometers and rain gauges, soil moisture sensors, and streamflow and stream-stage measuring devices while engaging in real-time data acquisition and ingestion into models requires an appropriate cyberinfrastructure. The term cyberinfrastructure, as defined in the National Science Foundation's (NSF) landmark Atkins report (Atkins et al. 2003), encompasses sensors, communication, computers, models, and people (Newman et al. 2003; Bottum et al. 2008). Off-the-shelf comprehensive solutions for such an infrastructure are not available, in part because of the fast pace of technological advances related to communication and computation. The limited financial resources available to research groups within academia and federal agencies prevent the rapid integration of the top advances from all relevant fields, including sensor (instrumentation) and modeling development. Nevertheless, widespread access to personal computers and the Internet allows unprecedented integration of sensors, databases, computational clusters, and a variety of complex models in order to address research questions in near-real time. However, site-specific limitations may prevent the implementation of solutions that are otherwise readily available.

In IFloodS, campaign participants faced the challenge of deploying six scanning weather radars, 20 optical disdrometers, about 100 tipping-bucket rain gauges, 25 soil moisture stations, and 100 stream gauges alongside a few other specialized instruments that were distributed over a region of about 50 000 km<sup>2</sup>. Many of these instruments reported in real time or with only small (minutes) delays and deposited their data in relational databases that were relayed to other centers and users. Computational models were fed with input data from these databases, as well as satellite observations and other products, to provide output that was ready for inspection by the researchers.

Crews of technical staff, assisted by students, frequently monitored the deployed instruments and

maintained them on site. The maintenance is more cost effective and practical if the instruments can be monitored remotely and only checked on site when a problem is detected. Data transmission during IFloodS used a hybrid system with specific solutions that depended on site-specific circumstances. Solutions included cell phone-based data transmission, radio links, digital subscriber line (DSL) service, and fiber optics-based fast Internet networks.

In this paper, we limit our discussion to the data management aspects of the IFloodS campaign cyberinfrastructure. Other aspects, such as selection of the sensors and their deployment, data transmission modes, relevant atmospheric and hydrologic models, and data analyses are described in other papers included in this special collection. The GPM Ground Validation (GV) program has conducted a series of field experiments approximately once per year since 2010, as shown in Table 1. Informatics experts at the Global Hydrology Resource Center (GHRC) Distributed Active Archive Center (DAAC), a partnership between NASA's Marshall Space Flight Center and the University of Alabama in Huntsville (UAH), have provided the GPM GV program with collaboration tools in order to facilitate the following:

- 1) the exchange of planning information before each field campaign;
- 2) collection of all data, mission science, project and instrument status reports, and other information during the campaign; and
- 3) the provision of long-term archive and distribution of post-campaign data and information.

For IFloodS, the same group provided similar data- and information-sharing capabilities and partnered with the Iowa Flood Center (IFC) at the University of Iowa to provide additional tools for this and future campaigns. The Iowa legislature established the IFC in 2009, following a devastating flood in 2008 that caused multibillion dollar losses to the state, with the charge of conducting applied research to help the state mitigate future flood disasters. The Iowa group has experience

developing relevant information systems that support real-time data acquisition and forecasting for the public as well as for the scientific community. Primary examples include the Iowa Flood Information System (Demir and Krajewski 2013) and the Hydro-NEXRAD system (Krajewski et al. 2011, 2013; Kruger et al. 2011; Seo et al. 2011). These two groups exchanged and developed new tools and ideas, many of which we describe in this paper.

We organized the paper as follows. Section 2 documents the main IFloodS campaign collaboration portal, which is the central site for the exchange of data and information among the team, and describes how it was used before and during the field campaign as well as for the subsequent distribution of quality-controlled IFloodS data to the public. Section 3 discusses the IFloodS Planning Platform, an interactive tool for campaign planning and instrument placement using a web-based mapping environment. Section 4 describes the IFloodS Information System for accessing, examining, and visualizing instrument data for further analysis and modeling.

## 2. IFloodS campaign collaboration portal

The central site for exchanging data and information among the team, both prior to and during the experiment, was the IFloodS collaboration portal that is hosted at the GHRC DAAC. The GHRC, which is a NASA Earth science data center, serves as the primary data archive for GPM GV field campaigns. GHRC is a full-service science data center that can ingest, process, archive, catalog, document, and distribute data in addition to offering user services. Processing varies from the execution of product generation software that is supplied by instrument scientists to the conversion of ASCII or binary files into a standard format such as Hierarchical Data Format—Earth Observing System (HDF-EOS) or netCDF. File-level metadata such as spatial and temporal bounds, file size, and checksum are generated and inserted into the GHRC metadata catalog as the data files are acquired or produced.

For each GPM GV campaign, GHRC developers and managers work closely with GPM GV coordinators to prepare for and support field campaign activities, and they provide mission coordination tools and near-real-time data prior to and during the campaign. The IFloodS collaboration portal, shown in Fig. 1, provided a central reference point for team members to access the mission schedule, the plan of the day, weather forecasts, current instrument status, and quick-look imagery. Daily mission science reports presented an overview of goals and accomplishments. The portal was designed as an internal communication tool to enable collaboration

among the participating scientists, and as such, access was restricted to team members. General project information was made available through public-facing websites.

The IFloodS portal's addition of a map-based data visualization capability offered a significant improvement over previous GPM GV field campaigns. This map feature supported a display of various types of spatial information, including IFloodS instrument locations and coverage areas, rivers and watershed boundaries, and roads and political boundaries. Investigators could overlay this base map with near-real-time imagery from IFloodS instruments and ancillary datasets. Near-real-time data acquisition and display supported mission planning and situational awareness among the team and also provided for a comparison of coincident observations from multiple instruments.

The portal also provided a gateway for collecting data from the scientists into the GHRC for long-term archive, as it is NASA's responsibility to curate scientific data that are collected in support of its mission. Maintaining all of the data at GHRC allows for easy access to and long-term preservation of these unique data collections. At the GHRC, data from successive field campaigns are tied together through common procedures, consistent metadata, and discovery and archival systems, which make it easy to access data from instruments that have been employed across several missions. These data are also valuable when preparing for new field campaigns.

### a. IFloodS portal architecture

GHRC's web-based field campaign coordination tools have evolved over the years from hand-coded HTML and web forms to collaboration portals built on content management systems (CMSs) including Plone and WordPress. The latest generation, implemented for IFloodS, was built on the Drupal CMS (Ramachandran et al. 2012) and incorporated spatial data infrastructure from ORION, a customizable situational awareness framework developed by UAH (Ramachandran et al. 2013). The ORION framework contains open-source components such as GeoServer and Thematic Real-Time Environmental Distributed Data Services (THREDDS) Data Server and utilizes well-established interoperability standards to provide a robust architecture for serving and managing curated environmental data and information. ORION supports real-time/near-real-time data and information acquisition and displays and stores comprehensive metadata to facilitate the search, evaluation, and interpretation of these datasets. ORION integrates and synthesizes these real-time data feeds into interactive maps, which yields quick

The screenshot shows the IFloodS collaboration portal dashboard. At the top, there is a header with the NASA and Iowa Flood Center logos, a search bar, and a navigation menu. The main content area is divided into several sections:

- Home:** Contains a 'Daily Events Schedule' and a 'Plan of the Day' for June 15, 2013. The 'Plan of the Day' section includes a title, date, and a discussion about the day's activities and goals.
- Latest Instrument Status Reports:** Lists various instruments and their status, including '2D Video Diadrometer', 'Autonomous Parsivel Units (APU)', 'Micro Rain Radar (MRR) Diadrometer', 'Rain Gauges - IRC', and 'Rain Gauges - NASA'. Each report includes an update date.
- Latest Weather Forecast:** Provides a summary of the weather forecast for June 15, 2013, mentioning scattered convective rainfall and thunderstorms.
- Right-hand Sidebar:** Contains several sections:
  - Plan of the Day:** Lists daily plans for various dates in June 2013.
  - Hydro Forecasts:** Lists hydro forecasts for various dates in June 2013.
  - Weather Forecasts:** Lists weather briefings and discussions for various dates in June 2013.
  - Mission Scientist Reports:** Lists mission scientist reports for various dates in June 2013.
  - Current Instrument Scientist Reports:** Lists current instrument scientist reports, including 'MRRs Green' and '2DVDs Yellow'.

FIG. 1. IFloodS collaboration portal dashboard including daily schedule and plan of the day (center left), instrument status report list (center right), latest forecast summary (bottom center), and links to other reports (right).

visualization of the situation and improved coordination among responders and stakeholders.

Figure 2 shows the high-level architecture of the IFloodS collaboration portal in the overall context of GHRC systems. Distributed data sources (Fig. 2a) include IFloodS instruments, operational systems such as National Weather Service (NWS) radars, and research data from satellites and other sources. The GHRC infrastructure (Fig. 2b) provides core components and services for data ingest, access, and long-term data management. The IFloodS portal (Fig. 2c) augments this core with specialized features that are required for field campaign coordination. The system's core components include GeoServer for managing and serving spatial data and Drupal for all web functions and associated information. The Drupal "profile," or suite of modules defined in ORION for Earth science data systems, includes a web geographic information system (GIS)

with an associated catalog of map layers for display in the GIS as well as a set of collaboration tools.

### b. Drupal CMS

The IFloodS portal was built on version 7 of the Drupal CMS platform (<http://drupal.org>). Drupal is an open-source web software suite with thousands of modules that were created and maintained by an active community of developers. The Drupal core includes a library of common functions and modules that provide basic features like user management, access control, session management, etc. (VanDyk and Westgate 2007). These core functions enable collaboration among team members with respect to sharing data and documents, calendar, weather forecasts, and status reports. For IFloodS, Drupal's user management and access control features restricted portal access to team members during the campaign. While every team member was able to

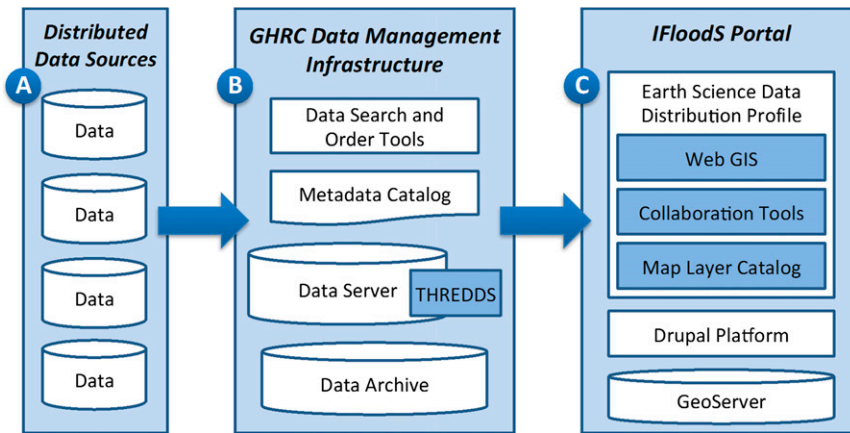


FIG. 2. High-level architecture of IFloodS collaboration portal in the overall context of GHRC systems.

access any material on the portal, including field campaign planning documents, quick-look imagery from instruments, and experimental data products, user roles provided customized access to the site’s reporting functions for different team members (mission manager, project scientists, instrument team members, and weather forecasters).

A Drupal website can support a variety of custom “content types” or document templates around which the site is organized. Key content types defined for the IFloodS portal were the various reports, and common features included title, author, date submitted, and the main body of text. We could insert images into the reports and attach additional documents. Weather forecasts included separate text fields for different forecast periods (current conditions, forecast for next 12 h, next day, etc).

We used taxonomies to define semantic relationships among concepts used in the site and to tag different pages for efficient searches. We encoded relationships among the different IFloodS instruments and other data sources in a two-level taxonomy, with higher-level categories used to group the different radars, other ground instruments, model results, and satellite products. We developed custom software around the instrument taxonomy to display instrument status reports as a hierarchical list, and users were able to hide or expand the list of instrument reports in each group. Custom software also provided a color-coded (green, yellow, red) indicator for instrument status reports so that users could quickly see which instruments are working nominally. Figure 1 shows the portal home screen, including the instrument status report list (right center). Color blocks around the instrument names indicate that the 2D video disdrometer and X-band polarimetric radar 2 (XPOL-2) have questionable statuses, while all other instruments are shown to be operating nominally. A red status code

would indicate a completely inoperable instrument. Instrument teams provided status details in the full report, which can be accessed from this list.

*c. Spatial data handling and visualization*

The IFloodS portal incorporates spatial data infrastructure from the ORION framework. ORION’S web-based GIS tools, in turn, are based on the OpenGeo Suite. We chose this software stack because it is widely used, well documented, open source, and supports Open Geospatial Consortium (OGC) standards for mapping applications. Components of the OpenGeo Suite used in ORION and the IFloodS portal included the GeoExplorer mapping application, OpenLayers map rendering library, GeoServer spatial data handler, and PostGIS spatial database for vector data storage. In addition, we used the Unidata’s THREDDS Data Server to provide data access via OGC Web Map Service (WMS) for gridded data in standard formats (netCDF or HDF). These mapping functions were built on existing software components but were packaged as part of the familiar collaboration portal, with data layers relevant to and collected during the campaign.

GeoExplorer includes an integrated mapping library known as OpenLayers that was implemented in open source JavaScript. OpenLayers provides the web GIS display of any of the data available through the IFloodS GeoServer as well as from other sources that support standard interfaces such as WMS. For three-dimensional map displays, the IFloodS portal also includes a Google Earth plug-in. Unfortunately, this plug-in is not supported by all web browsers.

GeoServer manages the geographic data and imagery for the IFloodS Portal’s web GIS. It is a Java-based open source software that is distributed as part of the OpenGeo Suite that provides services for creating and editing

maps as well as features that can be used for visualizing, analyzing, and editing geospatial data from spatial data sources using open standards. Designed for interoperability, GeoServer publishes data from any major spatial data source, including PostGIS databases, and provides implementations of OGC services such as WMS, Web Feature Service (WFS), Web Coverage Service (WCS), and Keyhole Markup Language (KML). It also provides various representational state transfer (REST) application programming interfaces (APIs) to interact with spatial data. Furthermore, client applications can request and query information via standards-based services and utilize a presentation layer via styling of the data using other OGC standard-styled layer descriptors (SLDs). Additional features of GeoServer include support for many back-end data formats (ArcSDE, Oracle Spatial, DB2, Microsoft SQL Server, Shapefile, GeoTIFF, and many more), multiple output formats (ESRI Shapefiles, KML, GML, GeoJSON, PNG, JPEG, TIFF, SVG, PDF, GeoRSS), fully featured web administration interface and REST API for easy configuration, dynamic reprojection of spatial data, and a configurable role-based security subsystem based on Spring Security.

PostGIS is an open source spatial database (an extension of PostgreSQL) that is optimized to store and query spatial data such as points, lines, and polygons. While typical relational databases can understand various numeric and character types of data, we need additional functionality to process such spatial data types as geometries or features. Spatial databases can perform a wide variety of spatial operations that include spatial measurements such as the distance between points or polygon area, spatial functions such as transforming features to create new ones like overlapping instrument fields of view, and Boolean queries such as “are there any heavy rainfall sensitive areas within a mile of the rain?” PostGIS adds functions, operators, and index enhancements to these spatial types that augment the power of the core PostgreSQL database management system. It is also possible to combine information from nonspatial sources with PostGIS spatial data and expose the aggregated information using GeoServer.

During the IFloodS campaign, GHRC brought in over 50 data products in near-real time, ranging from operational weather data (NWS NEXRAD radar and USGS stream gauges) to satellite imagery and derived products to quick-look data from instruments deployed by NASA, the University of Iowa, and other IFloodS investigators. The ORION-IFloodS's infrastructure handled the different types of data in different ways. Many of these data were provided in a form that was

suitable for display as map layers in ORION's web GIS (Fig. 3). These included the following:

- 1) Point data (e.g., from rain gauges or stream gauges). These data were generally acquired in XML or another structured ASCII format and were loaded directly into a PostGIS database for access via the GeoServer. For example, the USGS stream gauge data were a time series of several measured parameters (height, velocity, etc.) for each gauge location. Each parameter was cataloged as a separate map layer as well as a layer showing the age of the most recent update for the various sites. Each such layer is a database view into an underlying table holding all the stream data.
- 2) Raster imagery (e.g., flood maps and other satellite products provided as PNG or JPG files). Typically, all of the images in each dataset were in a standard map projection with the same spatial extent. These data were stored on the GHRC file system with locally generated KML files.
- 3) Satellite-derived rainfall and other gridded science products [e.g., Climate Prediction Center morphing technique (CMORPH; Joyce et al. 2004) and PERSIANN (Hsu et al. 1997)]. These data were translated to netCDF on ingest. We used a THREDDS Data Server to render imagery on demand via a WMS.

All of these map-ready data were registered in the portal's map layers catalog. The layer catalog was implemented as Drupal content populated with uploaded data files (KML) or linked files or services (WMS) and a title and brief description for each map layer. The OpenLayers library interpreted the map layers in the catalog. Some of the near-real-time imagery included borders, text, and other annotations as part of the image and were consequently more suited for visualization as self-contained images rather than as map layers. These images were accessible through a separate image viewer on the portal, and the most recent images were frequently offered as an animation.

### 3. IFloodS Planning Platform

The IFloodS Planning Platform was developed at the Iowa Flood Center at the University of Iowa and hosted at the University of Iowa Data Center. The Planning Platform is an interactive web application with GIS capabilities, developed to support campaign planning, the design of environmental monitoring systems, and instrument placement (Fig. 4). Core functionality specifically developed for the platform includes the following:

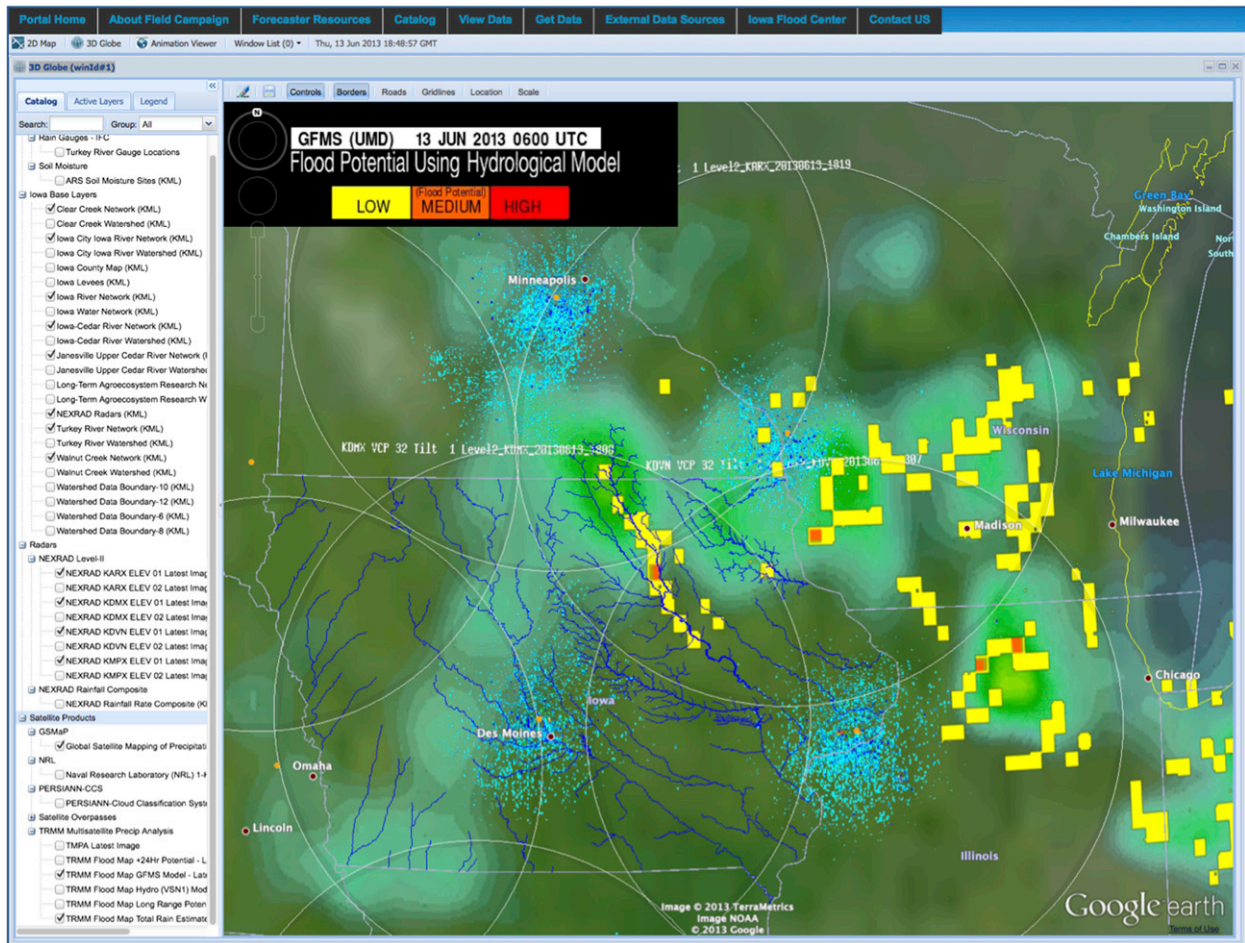


FIG. 3. IFloodS portal map display with base layer from a global satellite map of precipitation, with overlays showing cities, state borders, and Iowa rivers. Large circles indicate ranges of NWS radars, with each radar reporting light precipitation. Red and yellow squares indicate flood potential calculated using a hydrological model and precipitation estimates.

- 1) mapping tools to decide placement of the instruments and access information for public/private land within the study area;
- 2) adding and removing in situ (point; e.g., rain gauges, disdrometers) instruments as well as instruments with spatial coverage (e.g., research weather radars) to the interface;
- 3) overlaying GIS layers of interest (e.g., watersheds, existing sensor networks and instruments, and wind farms); and
- 4) generating reports of instrument placement for sharing and discussing with other researchers.

The Planning Platform allows users to navigate around the proposed sites in the interactive map interface and to explore various sampling schemes in the context of other instruments, mainly weather radars and the coverage they provide. We made numerous map layers available, including watershed boundaries and the stream drainage

networks. To provide a historical context for studies, we made various precipitation datasets from weather radars, satellites, and rain gauges, which have been collected over the past several years and processed to support validation and flood-related rainfall–runoff modeling studies, available for browsing. Table 2 lists these historical datasets that are included in the IFloodS Planning Platform and Information System (IFloodS IS). These heterogeneous datasets have different temporal and spatial resolutions as well as different uncertainty characteristics. These differences provide benefits for product validation using multiscale data as well as for hydrologic modeling where different models require different scales of rainfall inputs.

We developed the Planning Platform based on the existing resources at the Iowa Flood Center, specifically, the Iowa Flood Information System (IFIS). IFIS is a one-stop web platform for community-based information such as current flood conditions; forecasts;



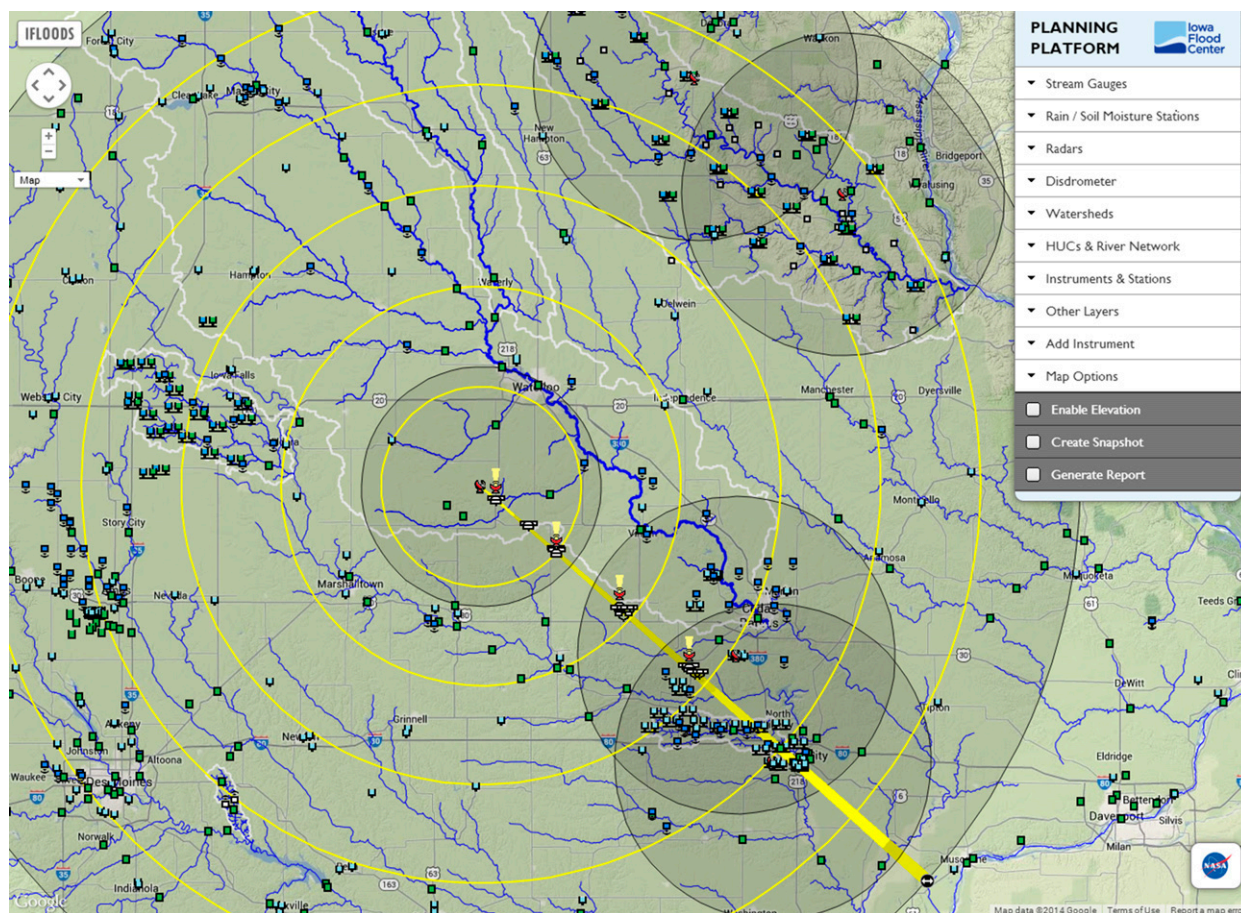


FIG. 4. Screenshot of the IFloodS Planning Platform. The symbols included on the map correspond to double rain gauge platforms (two light blue buckets), standard (single) gauge (one light blue bucket), platforms with soil moisture sensors (blue and green buckets), vertically pointing radars (red with yellow funnels), stream-stage sensors (dark blue squares), and standard stream gauges (green squares). All are shown in the context of NASA S-band dual-polarimetric radar (NPOL) radar range, and the other circles represent the XPOL radars, river network, and basin boundaries.

visualizations; inundation maps; and flood-related data, information, and applications (Demir and Krajewski 2013), many of which rely on the distributed hydrologic model. The IFIS helps communities make better-informed decisions with respect to the occurrence of floods and provides advance warning to communities to help them minimize flood damages. The IFIS is widely used by the general public in the Midwest, with more than 100 000 users, and is a key source of information for many newspapers and TV stations in Iowa. The IFIS also integrates the developments of and collaborations among experts in engineering, hydrology, informatics, mathematics, and communication science.

The IFIS includes several elements (e.g., data resources and visualization libraries) that are highly pertinent to the IFloodS' objectives. These include digital representation of the basin boundaries, the stream and river network, flood inundation maps, and the locations

of the local stream, rain, and soil moisture sensors. All of these elements, together with the locations and coverage of the area by the NWS weather radars [Weather Surveillance Radar-1988 Doppler (WSR-88D)] and the local network of roads and communities, constitute the key context for siting other instruments brought to Iowa for the campaign. Figure 5 depicts the high-level architecture of the IFIS that is pertinent to IFloodS. IFIS infrastructure consists of a hybrid of file and compute servers, a central database server, and a high-performance computing (HPC) cluster. The rainfall subsystem runs at the University of Iowa Data Center and generates rainfall maps for precipitation intensity, daily accumulation (up to the past 2 weeks), and cumulative rainfall (1–15 days). The flood forecasting model runs on the HPC cluster and feeds model results to the central database server for easy access and visualization. Data collected from sensors were stored in the

TABLE 2. Historical datasets included in the IFloodS Planning Platform and IS.

TRMM Multisatellite Precipitation Analysis (TMPA)	Xue et al. (2013)
Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN)	Hsu et al. (1997)
Climate Prediction Center morphing technique (CMORPH) Stage IV	Joyce et al. (2004) Lin and Mitchell (2005)
Hydro-NEXRAD	Krajewski et al. (2011)
National Mosaic and Multi-Sensor QPE (NMQ)	Zhang et al. (2011)
Iowa Flood Center (IFC) Stream Sensors	
Automated Surface Observing System (ASOS)	Clark (1995)
NWS Cooperative Observer (COOP) program	
IFC research network for rain gauge data	

central database server for long-term data access. IFIS ingests short-term data to the IFIS database for quick access by federal agencies and weather services. We integrated all of the data sources, rainfall maps, and custom map layers from IFIS into the IFloodS Planning Platform and IFloodS IS.

To help scientists design a monitoring system, we created the capability to place observational points that are denoted with different symbols to help distinguish the functionality of instruments. For example, rain gauge platforms with two side-by-side tipping-bucket gauges are denoted by a double-bucket blue symbol, and similar platforms with added soil moisture sensors are shown as one blue and one green bucket. Disdrometers and vertically pointing radars have their own symbols. When placing weather radars, it is important to consider potential beam blockage by nearby (trees, buildings, etc.) and distant (mountains or hills, wind farms, etc.) objects. It is also important to mark the anticipated maximum range for data collection and for sectors and directions of particular interest. Because of the relatively simple (versus complex) terrain in Iowa, we did not use sophisticated tools to simulate potential beam blockage (e.g., Kucera et al. 2004; Krajewski et al. 2006; Lang et al. 2009). Instead, we used Google Earth- and Google Maps-enabled elevation readings and Planning Platform tools in the initial stages of the radar site selection. We based our final selection on an on-site inspection.

The Planning Platform tools for radar site selection included range adjustment (circle radius) and sector selection that were connected to a radar icon that a user could “drop” on a Google map. Because one of the objectives of the campaign included studies of the range’s effect on estimating rainfall, the tools included a single ray with range gates marked approximately to scale. Users could “grab” and move the ray in any desired direction, and the gate markings allowed the siting of in situ instrument clusters (disdrometers and rain gauges).

To facilitate sharing a proposed placement for a given instrument or a set of instruments, the Platform allows

users to save the locations by taking a “snapshot” and generating a report that can be transmitted to another user for inspection. The Snapshot feature allows users to create a system state with custom added instruments and share this state with other researchers as a URL. This helps the decision-making process and collaboration among team members with respect to instrument placement by allowing them to use a custom-generated URL to share and modify the proposed campaign. The Snapshot generator saves the location of individual or grouped instruments, the range and location of radars, the zoom level, and the region of interest. The “Generate Report” function allows users to create a report that indicates the latitude and longitude of the proposed instrument placements, which consequently helps team members share the coordinates of the instruments as a report.

#### 4. IFloodS Information System

The IFloodS IS was developed at the Iowa Flood Center as a web-based platform for accessing, examining, and visualizing instrument data for further analysis and modeling (Fig. 6). IFloodS IS, which is hosted at the University of Iowa Data Center, utilizes some of the data resources and built-in visualization libraries of IFC’s IFIS. An enormous volume of real-time data from a variety of sensors (radar, rain gauges, stream-stage sensors, USGS gauges, etc.) was integrated from IFIS and the field campaign into the information system for visualization, analysis, and examination of data for further research. We describe the data management and scientific visualization components of IFloodS IS in the following sections.

##### a. Data management

IFloodS IS relies on the IFIS’s utilization of a hybrid of file and compute servers, including an HPC cluster, codes in different languages, data streams and web services, databases, scripts, and visualizations. The IFIS processes raw data ( $50 \text{ GB day}^{-1}$ ) from NEXRAD radars, creates rainfall maps ( $3 \text{ GB day}^{-1}$ ) every 5 min, and integrates

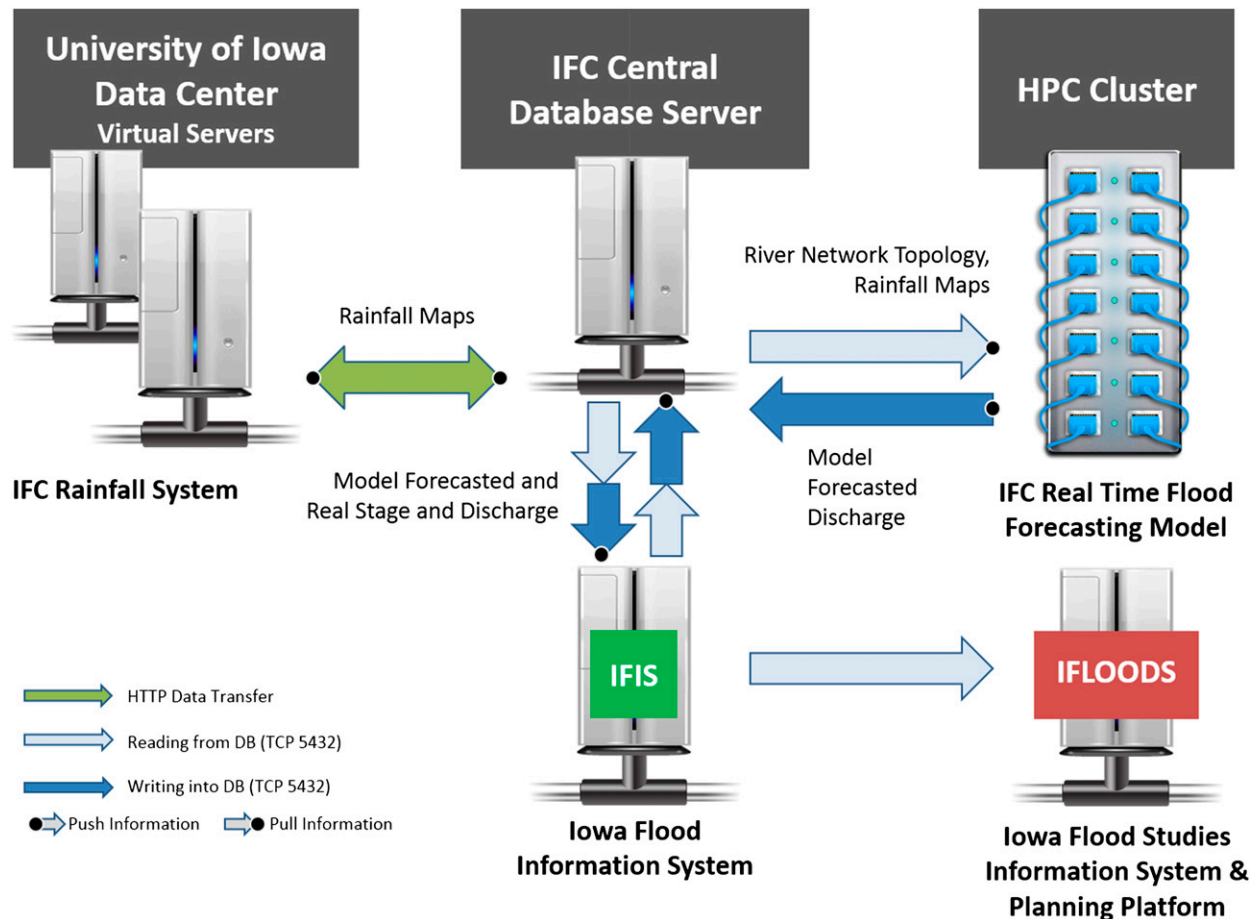


FIG. 5. High-level architecture of the IFIS pertinent to IFloodS.

real-time data from over 600 sensors in Iowa. IFloodS IS integrates data and resources, stored in various database systems (e.g., PostgreSQL, MySQL, Google Fusion), directly from the IFIS infrastructure in order to provide comprehensive data and services. The IFloodS IS client side was developed using HTML, JavaScript, Google Maps API, and many graphics libraries. The server side application in the system is based on PHP, PostgreSQL, and a file server for rainfall data and maps. We organized all datasets in a database and developed a comprehensive information system to share and visualize large-scale spatial and temporal datasets in a web browser using the latest web technologies. The information system allows users to easily locate significant events based on metadata that contain, in addition to standard data descriptions, precomputed data characteristics that might be of interest to the researchers.

#### b. Scientific visualization

IFloodS IS provides the location and properties of existing and new instruments that we deployed during

the IFloodS field campaign as well as information on public and private land and numerous geospatial layers for watersheds. We developed custom scientific visualization tools for each instrument type, which allows researchers to examine and download data for further analysis and modeling. IFloodS IS provides time series graphs of stream and rain gauges and soil moisture sensors. The system utilized a graphical processing unit (GPU) for processing and visualizing large-scale data that we collected using radars. Custom visualization interfaces provide a web-based interactive visualization of raw and processed sensor data for disdrometers and soil moisture sensors. IFloodS IS includes visualization and animation for various radar (NEXRAD, Q2) and satellite-based (PERSIANN and Hydro-Estimator) rainfall products. The interface allows users to select a date from the campaign period and load an animation of the rainfall map for the entire Iowa domain. The rainfall map archive has over 25 000 images for each product, which allows a smooth animation with 5-min intervals.

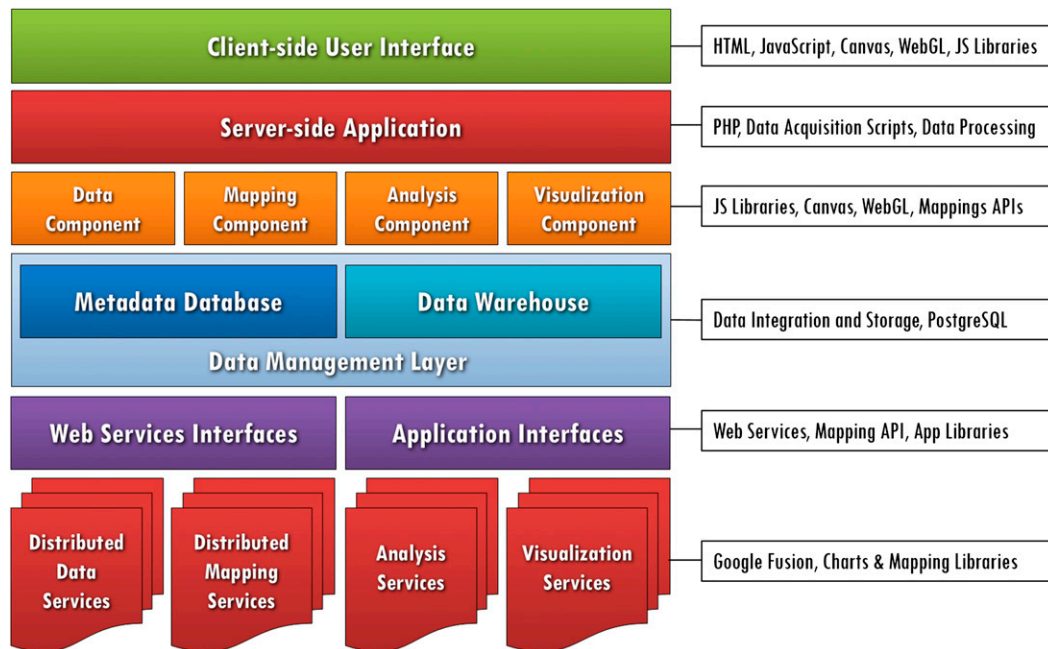


FIG. 6. Architecture of IFloodS IS and Planning Platform.

The performance of these web-based systems has increased significantly in recent years because of the advancements in Internet technologies and standards, speed-ups in scripting languages, integration, and direct access to graphics processing units (GPUs) from web browsers. Web-based scripting languages like JavaScript perform at speeds close to native C++ code with the help of new libraries (e.g., ASM.js) and GPU. The IFloodS IS includes three comprehensive visualization systems that utilize new web standards and GPU for XPOL and WSR-88D Level II data as well as radar and satellite-based rainfall products.

### 1) DISDRMETER DATA VISUALIZATION

The IFloodS Information System provides a custom visualization interface (Fig. 7) for disdrometer data. At this point, the interface is limited to displaying data from one-dimensional disdrometers that report collected data in the form of an array of drop numbers for different size and terminal velocity categories. We used the velocity data to estimate rainfall rate. We developed an original interface using HTML, JavaScript, CSS, and Canvas, which allows us full control of the design and customization of functionality and “look.” Compared to traditional approaches for visualizing time series data, our custom visualization interface for disdrometer data allows researchers to find a day of interest by examining metadata, to interact with the high-resolution data, and to read individual values from the graph. The interface

allows users to visualize rain rate ( $\text{mm h}^{-1}$ ), reflectivity (dBZ), and drop size distribution at 1-min resolution for a 24-h time period of their choice. The calendar component of the interface allows users to select a day of interest based on precalculated metadata (days with total rainfall amounts). The time slider allows the inspection of data over the selected day with 1-min resolution. The rainfall characteristics (variables) shown can be read from a label (not shown) that appears as the cursor hovers over the minute bar of interest.

### 2) XPOL RADAR DATA VISUALIZATION

XPOL radars collect a significant amount of data during each scan and generate volumetric data with various elevation scans, sectorial readings, and scanning modes. A common approach to sharing and visualizing research radar data during or shortly after field campaigns entails static images, which provide limited interactivity and spatial context. Map-based visualization of the radar data yields an accurate and high-resolution projection of data onto an interactive map environment. We developed an XPOL radar visualization interface (Fig. 8) in IFloodS IS that allows users to select an XPOL radar, choose a date and hour of interest from the calendar and time components, select the scan mode and variable, and browse through different elevations and scans using navigation controls. The time component shows the minimum, maximum, and average values for rainfall intensity and makes it easy for

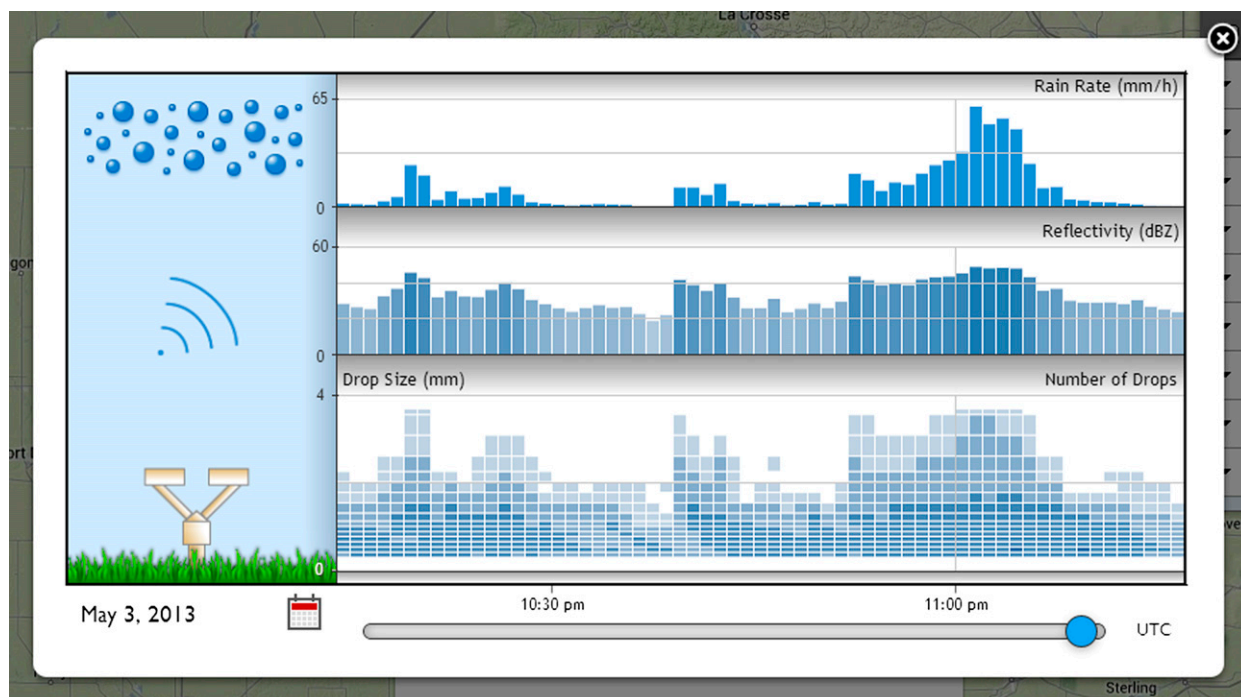


FIG. 7. Disdrometer data visualization interface.

researchers to find a scan of interest. The interface utilizes the WebGL graphic library to visualize data on Google Maps and allows browsing data in various zoom levels. The tool is not a replacement for a full suite of radar data analysis software (e.g., RSL, PPI-MMM, SPRINT, CEDRIC, GEMINI, Solo 3, and TREC), but it is a flexible method to quickly view full-resolution data in an easily customizable geographic context.

### 3) RAINFALL DATA BROWSER

IFloodS participants collected a variety of radar and satellite rainfall data in order to facilitate hydrologic analyses using high-resolution rainfall estimates. To support the IFloodS experiment, the Iowa Flood Center developed an interactive radar-satellite rainfall browser (Fig. 9) that allows researchers to examine, visualize, and download historical radar- and satellite-based rainfall products in an interactive web-based environment. The browser covers the period from 2002 to 2012 and enables users to identify significant rainfall events for their own purposes (e.g., flooding). We ingested Stage IV (Lin and Mitchell 2005), IFC (HUC 0708 and Flood 2008; Krajewski et al. 2011, 2013; Seo et al. 2013), and Q2 (Zhang et al. 2011) radar-rainfall and CMORPH (Joyce et al. 2004), PERSIANN (Hsu et al. 1997), and TMPA 3B42, version 7 (Xue et al. 2013), satellite-rainfall products into a relational database that provides faster access to the data at users' requests. Since

the use of a static image overlay for the display of rainfall maps often leads to a map distortion or projection error (especially at higher zoom levels), we utilize JavaScript, WebGL, and GPU to directly draw individual gridded rainfall cells on a map environment. This provides an accurate projection of rainfall data onto a correct spatial grid. The calendar and time components allow researchers to examine daily and hourly metadata values (maximum, average, and coverage percentage) and help identify rainfall events. Our visualization interface provides various map layers such as IFloodS's radar sites, wind farm locations with all single turbine towers, river network and watershed boundaries, etc. The data download panel allows users to download rainfall data for a selected product for a point, region, whole map, and selected time range.

### 4) WSR-88D (LEVEL II) DATA BROWSER

Since weather radars can provide crucial information on quantitative rainfall characteristics such as raindrop size, shape, and concentration, we developed an interactive browser (Fig. 10) to allow researchers to access and navigate the base data of WSR-88D radars for the months of April-June 2013. We acquired Level II data from the four WSR-88D radars (KARX in La Crosse, Wisconsin; KDMX in Des Moines, Iowa; KDVN in Davenport, Iowa; and KMPX in Minneapolis, Minnesota) that are located in the IFloodS domain through

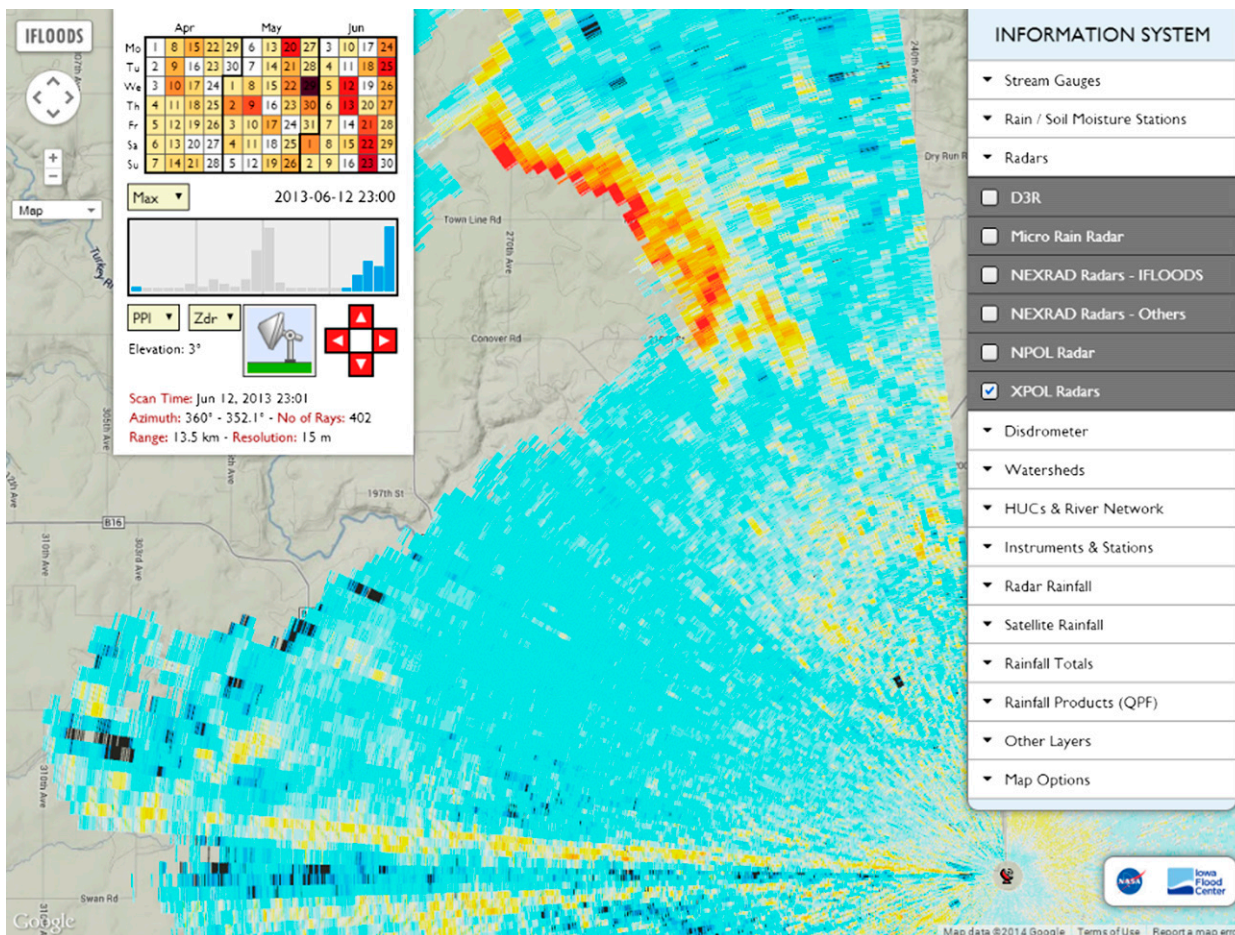


FIG. 8. XPOL radar data visualization interface. The display “mimics” a real-time display of data stored in a database. The user selects the day and hour based on precalculated metadata that characterize the collected data. Then the user selects and can browse the scan type and the variable of interest.

Unidata Local Data Manager (LDM) and Internet Data Distribution software (IDD; see, e.g., Sherretz and Fulker 1988; Fulker et al. 1997) in real time [for the quality control of Level II data, refer to Krajewski et al. (2013)]. Upon the reception of the Level II data in the local downstream LDM in IFC, all variables (horizontal reflectivity  $Z_h$ , differential reflectivity  $Z_{dr}$ , copolar correlation  $RHO$ , and differential phase  $PHI_{dp}$ ) in Level II were ingested into a relational database and organized by radar elevation angle, azimuth, and range. Pre-computed metadata data using coverage thresholds for  $Z_h$  and  $Z_{dr}$  enables a search for rain event data and the visual investigation of the retrieved data.

We utilize JavaScript, Canvas, and Google Maps API to visualize the Level II variables directly on a polar-based map environment. The developed radar data browser provides unique capabilities (e.g., showing vertical structure of observed variables at a specified azimuth angle) as well as similar functionalities

compared to a data file-based platform (e.g., NOAA’s Weather and Climate Toolkit; [www.ncdc.noaa.gov/wct/](http://www.ncdc.noaa.gov/wct/)). The interface allows users to select one of the four radars, the observed variables, and a time of interest from the control panel. The map environment allows users to zoom in and navigate on a selected region and to visualize various elevation and azimuth angles.

### 5. Discussion and conclusions

This article presents the cyberinfrastructure tools and systems necessary to support the planning, reporting, and management of the field experiment and to access and share data and models for assisting research. We described a set of informational technology tools that we developed for use before, during, and after the IFloodS field campaign. Here, we share a few reflections on our efforts.

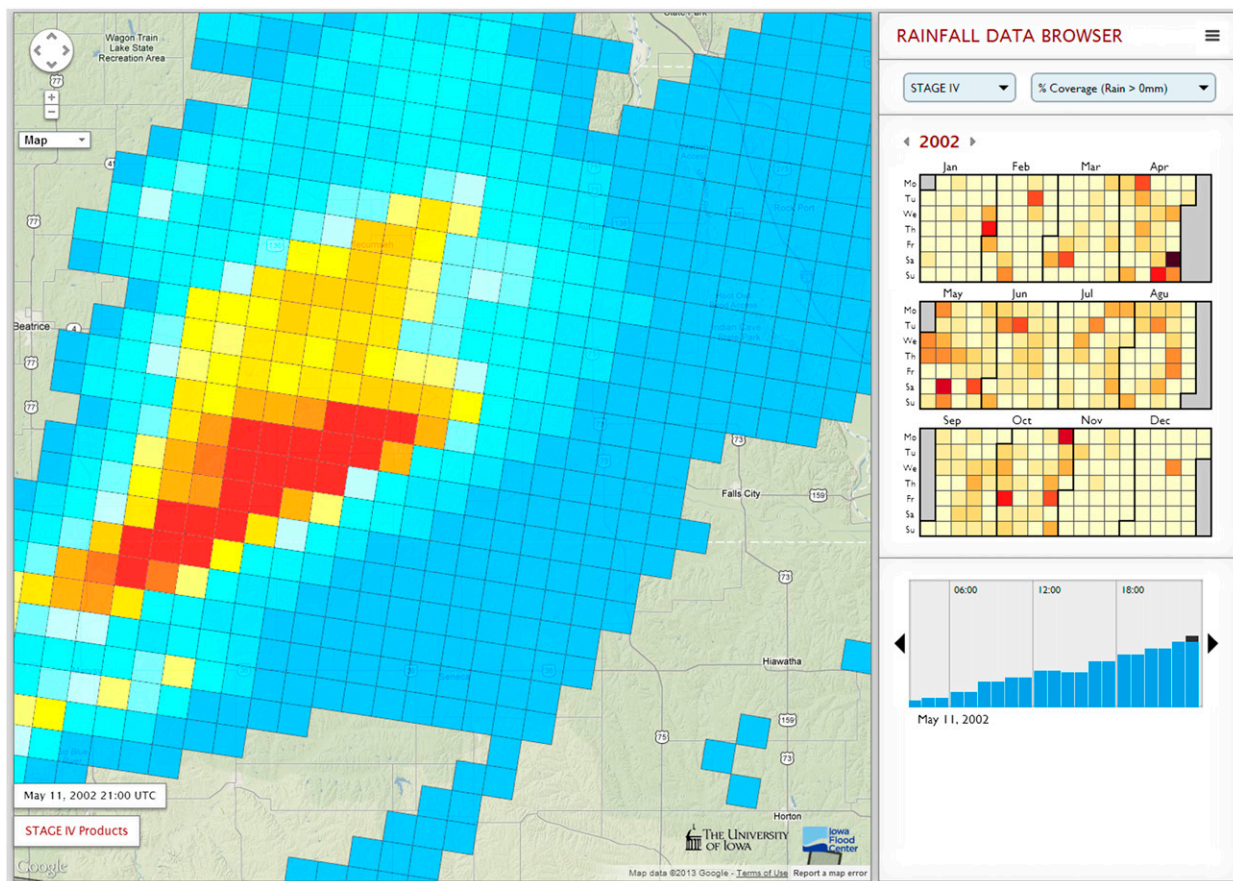


FIG. 9. Interactive rainfall data browser and visualization interface.

First, the fast pace of IT development is both a major opportunity as well as a major challenge. Because of the roughly annual frequency of field campaigns and their varied scope and objectives, it is difficult to develop a set of tools that can be used and reused for a number of years. Therefore, to take advantage of the new developments, domain scientists and IT professionals need to collaborate. It is the responsibility of the domain scientists to constantly assess the current tools and capabilities and request from IT support professionals new means of collecting, organizing, serving, and visualizing scientific data coming both from the field and from mathematical models. Failure to do this will delay progress, as IT professionals do not and cannot know the scientific analyses, questions, and hypotheses.

Our second conclusion is that efficient and effective systems necessitate a hybrid of different technologies and tools. Attempts to develop a tool that can do it all are bound to result in software “monsters” that do not do anything particularly well. These tools also become quickly outdated because it takes a long time, much

effort, and great expense to develop them. Examples include database schemas that can accommodate any sensor data, visualization that can display any type of data, and information systems that can handle any type of data. For IFloodS, we implemented a series of specialized tools, described in sections 2–4, and linked them all through the collaboration portal for ease of access.

The development of hybrid systems requires a comprehensive set of skills on the part of developers. It is unlikely that a small university research group would have all such required skills. This underscores the need for collaboration with computer scientists, which represents other challenges such as the need to learn each other’s technical jargon in order to clearly communicate a scientific vision and the technological limitations of the day.

We would like to close by expressing hope that the GPM research community will embrace informatics technology in order to accelerate their own research and to facilitate the communication of our research results to the general public.

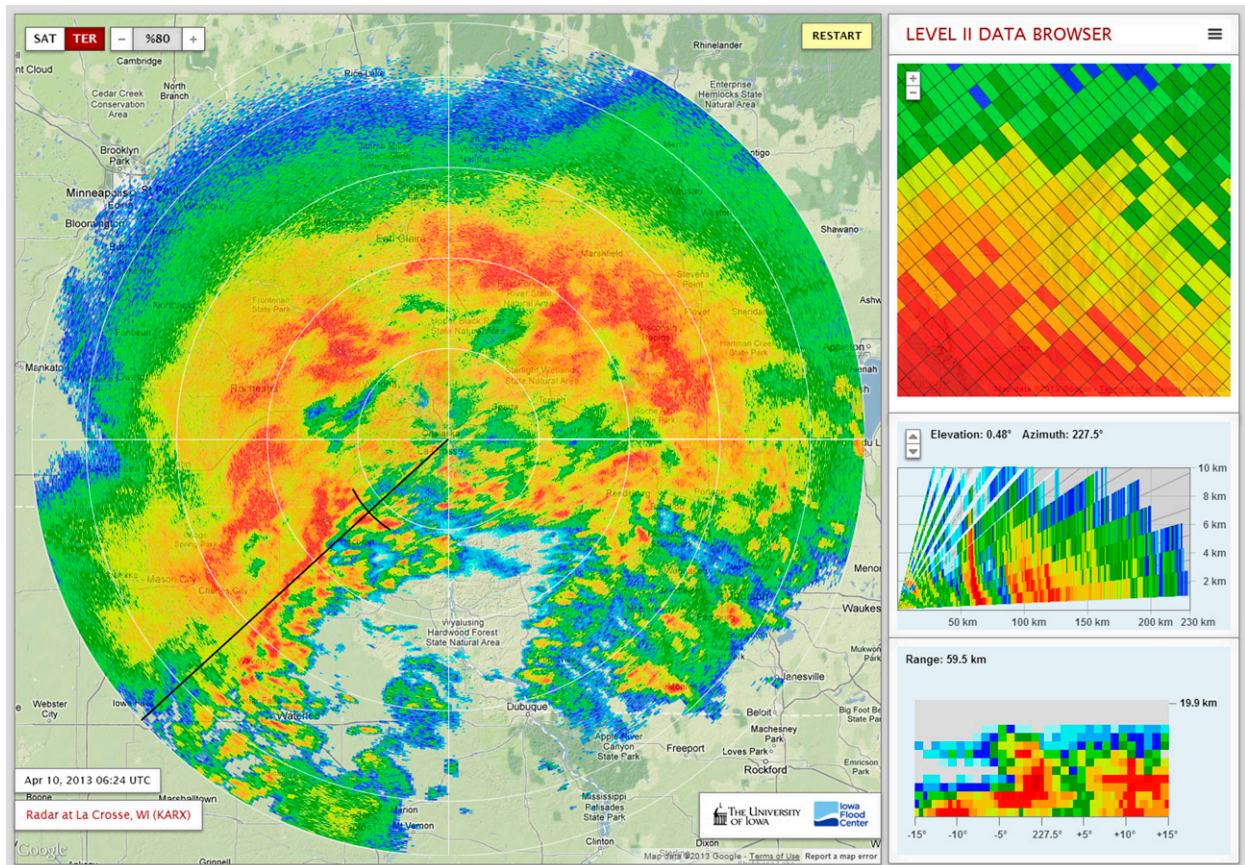


FIG. 10. WSR-88D (Level II) data browser and visualization interface.

### 6. Challenges and future work

With the advancement in Internet and communication technologies, environmental information systems are moving from desktop-based systems to the web. New web standards and libraries (WebGL, asm.js, etc.) are improving the performance and capabilities of data transfers, which allows visualizations with direct access to graphic cards and the running of scripts as fast as native code. However, there are still limitations involved in transferring and handling large-scale data on the web. One challenge is the speed of adaptation of these technologies by different platforms and web browsers. Limited publications and information is another challenge related to the widespread adaptation of these technologies. While pushing the limits of these web technologies, we are working to develop more capable web systems with rich graphics, interactivity, and performance, which will allow us to create virtual, augmented, and immersive reality applications for research and education.

*Acknowledgments.* Funding for this work was provided by NASA’s Cooperative Agreement NNM11AA01A

and Grants NNX13AD83G and NNX13AG94G, the Iowa Flood Center, and the National Science Foundation Award 1327830. We also gratefully acknowledge funding and management support from the NASA GPM and Precipitation Measurement Mission programs.

### REFERENCES

Atkins, D. E., and Coauthors, 2003: Revolutionizing science and engineering through cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. National Science Foundation, 84 pp. [Available online at [www.nsf.gov/cise/sci/reports/atkins.pdf](http://www.nsf.gov/cise/sci/reports/atkins.pdf).]

Bottum, J. R., J. F. Davis, P. M. Siegel, B. Wheeler, and D. G. Oblinger, 2008: Cyberinfrastructure: In tune for the future. *EDUCAUSE Rev.*, **43** (4), 11–17. [Available online at <https://net.educause.edu/ir/library/pdf/ERM0840.pdf>.]

Clark, P., 1995: Automated surface observations, new challenges - new tools. Preprints, *Sixth Conf. on Aviation Weather Systems*, Dallas, TX, Amer. Meteor. Soc., 445–450.

Demir, I., and W. F. Krajewski, 2013: Towards an integrated Flood Information System: Centralized data access, analysis, and visualization. *Environ. Modell. Software*, **50**, 77–84, doi:10.1016/j.envsoft.2013.08.009.

Fulker, D., S. Bates, and C. Jacobs, 1997: Unidata: A virtual community sharing resources via technological infrastructure. *Bull.*



- Amer. Meteor. Soc.*, **78**, 457–468, doi:10.1175/1520-0477(1997)078<0457:UAVCSR>2.0.CO;2.
- Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement (GPM) mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, doi:10.1175/BAMS-D-13-00164.1.
- Hsu, K., X. Gao, S. Sorooshian, and H. Gupta, 1997: Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteor.*, **36**, 1176–1190, doi:10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2.
- 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, doi:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.
- Krajewski, W. F., A. Ntelekos, and R. Goska, 2006: A GIS based methodology for the assessment of weather radar beam blockage in mountainous regions: Two examples from the U.S. NEXRAD network. *Comput. Geosci.*, **32**, 283–302, doi:10.1016/j.cageo.2005.06.024.
- , and Coauthors, 2011: Towards better utilization of NEXRAD data in hydrology: An overview of Hydro-NEXRAD. *J. Hydroinf.*, **13**, 255–266, doi:10.2166/hydro.2010.056.
- , A. Kruger, S. Singh, B.-C. Seo, and J. A. Smith, 2013: Hydro-NEXRAD-2: Real time access to customized radar-rainfall for hydrologic applications. *J. Hydroinf.*, **15**, 580–590, doi:10.2166/hydro.2012.227.
- Kruger, A., W. F. Krajewski, P. Domaszczynski, and J. A. Smith, 2011: Hydro-NEXRAD: Metadata computation and use. *J. Hydroinf.*, **13**, 267–276, doi:10.2166/hydro.2010.057.
- Kucera, P. A., W. F. Krajewski, and C. B. Young, 2004: Radar beam occultation studies using GIS and DEM technology: An example study of Guam. *J. Atmos. Oceanic Technol.*, **21**, 995–1006, doi:10.1175/1520-0426(2004)021<0995:RBOSUG>2.0.CO;2.
- Lang, T. J., S. W. Nesbitt, and L. D. Carey, 2009: On the correction of partial beam blockage in polarimetric radar data. *J. Atmos. Oceanic Technol.*, **26**, 943–957, doi:10.1175/2008JTECHA1133.1.
- Lin, Y., and K. E. Mitchell, 2005: The NCEP Stage II/IV hourly precipitation analyses: Development and applications. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., 1.2. [Available online at [https://ams.confex.com/ams/Annual2005/techprogram/paper\\_83847.htm](https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm).]
- Newman, H. B., M. H. Ellisman, and J. A. Orcutt, 2003: Data-intensive e-science frontier research in the coming decade. *Commun. ACM*, **46** (11), 68–77, doi:10.1145/948383.948411.
- Petersen, W. A., and W. Krajewski, 2013: Status update on the GPM ground validation Iowa Flood Studies (IFloodS) field experiment. *Extended Abstracts, European Geosciences Union General Assembly 2013, Geophys. Res. Abstracts*, **15**, Vienna, Austria. [Available online at <http://meetingorganizer.copernicus.org/EGU2013/EGU2013-13345.pdf>.]
- Ramachandran, R., M. Maskey, A. Kulkarni, H. Conover, U. S. Nair, and S. Movva, 2012: Talkoot: Software tool to create laboratories for Earth science. *Earth Sci. Inf.*, **5**, 33–41, doi:10.1007/s12145-012-0094-y.
- , A. Kulkarni, M. McEniry, A. Lin, S. Tanner, and S. Graves, 2013: Fostering national and international collaborations for Arctic resources using a virtual collaborator. *29th Conf. on Environmental Information Processing Technologies*, Austin, TX, Amer. Meteor. Soc., 2. [Available online at <https://ams.confex.com/ams/93Annual/webprogram/Paper214106.html>.]
- Schwaller, M. R., and K. R. Morris, 2011: A ground validation network for the Global Precipitation Measurement Mission. *J. Atmos. Oceanic Technol.*, **28**, 301–319, doi:10.1175/2010JTECHA1403.1.
- Seo, B.-C., W. F. Krajewski, A. Kruger, P. Domaszczynski, J. A. Smith, and M. Steiner, 2011: Radar-rainfall estimation algorithms of Hydro-NEXRAD. *J. Hydroinf.*, **13**, 277–291, doi:10.2166/hydro.2010.003.
- , L. K. Cunha, and W. F. Krajewski, 2013: Uncertainty in radar-rainfall composite and its impact on hydrologic prediction for the eastern Iowa flood of 2008. *Water Resour. Res.*, **49**, 2747–2764, doi:10.1002/wrcr.20244.
- Sherretz, L. A., and D. W. Fulker, 1988: Unidata: Enabling universities to acquire and analyze scientific data. *Bull. Amer. Meteor. Soc.*, **69**, 373–376, doi:10.1175/1520-0477(1988)069<0373:UEUTAA>2.0.CO;2.
- Tapiador, F. J., and Coauthors, 2012: Global precipitation measurement: Methods, datasets and applications. *Atmos. Res.*, **104–105**, 70–97, doi:10.1016/j.atmosres.2011.10.021.
- VanDyk, J. K., and M. Westgate, 2007: *Pro Drupal Development*. Apress, 428 pp.
- Xue, X., Y. Hong, A. S. Limaye, J. J. Gourley, G. J. Huffman, S. I. Khan, C. Dorji, and S. Chen, 2013: Statistical and hydrological evaluation of TRMM-based multi-satellite precipitation analysis over the Wangchu basin of Bhutan: Are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins? *J. Hydrol.*, **499**, 91–99, doi:10.1016/j.jhydrol.2013.06.042.
- Zhang, J., and Coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) system: Description, results, and future plans. *Bull. Amer. Meteor. Soc.*, **92**, 1321–1338, doi:10.1175/2011BAMS-D-11-00047.1.