

The Hawai‘i Rainfall Analysis and Mapping Application (HI-RAMA): Decision Support and Data Visualization for Statewide Rainfall Data

JARED H. MCLEAN, University of Hawai‘i - System, USA

SEAN B. CLEVELAND, University of Hawai‘i - System, USA

MATTHEW LUCAS, University of Hawai‘i - Mānoa, USA

RYAN LONGMAN, University of Hawai‘i - Mānoa, USA

THOMAS W. GIAMBELLUCA, Water Resource Research Center; University of Hawai‘i - Mānoa, USA

JASON LEIGH, University of Hawai‘i - Mānoa, USA

GWEN A. JACOBS, University of Hawai‘i - System, USA

This paper discusses the design and implementation of the Hawai‘i Rainfall Analysis and Mapping Application (HI-RAMA) decision support tool. HI-RAMA provides researchers and community stakeholders interactive access to and visualization of hosted historical and near-real-time monthly rainfall maps and aggregated rainfall station observational data for the State of Hawai‘i. The University of Hawai‘i Information Technology Services Cyberinfrastructure team in partnership with members of the Hawai‘i Established Program to Stimulate Competitive Research (EPSCoR) ‘Ike Wai project team developed this application as part of the ‘Ike Wai Gateway to support water sustainability research for the state of Hawai‘i. This tool is designed to provide user-friendly access to information that can reveal the impacts of climate changes related to precipitation so users can make data-driven decisions.

CCS Concepts: • **Software and its engineering** → **Designing software**.

Additional Key Words and Phrases: rainfall, data visualization, web interface

ACM Reference Format:

Jared H. McLean, Sean B. Cleveland, Matthew Lucas, Ryan Longman, Thomas W. Giambelluca, Jason Leigh, and Gwen A. Jacobs. 2020. The Hawai‘i Rainfall Analysis and Mapping Application (HI-RAMA): Decision Support and Data Visualization for Statewide Rainfall Data. In *Practice and Experience in Advanced Research Computing (PEARC ’20)*, July 26–30, 2020, Portland, OR, USA. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3311790.3396668>

1 INTRODUCTION

A sustainable supply of clean water is essential for all aspects of life in Hawai‘i and there are multiple threats to maintaining this resource for future generations. Climate change is decreasing rainfall and increasing sea levels throughout the islands and population growth, urban development, agricultural use continue to increase, bringing into question what is sustainable. Knowledge of mean rainfall patterns in Hawai‘i are critically important in support of the research and understanding of watershed modeling as well as for the management and protection of groundwater and surface water resources. Further, rainfall information can inform methods to control and eradicate invasive species, methods to protect and restore native ecosystems, and planning for the effects of global climate change.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2020 Copyright held by the owner/author(s).

Manuscript submitted to ACM

The Hawai‘i Established Program to Stimulate Competitive Research (EPSCoR) ‘Ike Wai project in partnership with the University of Hawai‘i Water Resource Research Center (WRCC) has spent the last several years developing gridded monthly rainfall products. The gridded month rainfall products are highly useful for water resources analysis and planning, ecological modeling, and other applications. Prior to this work, maps were produced at irregular intervals of at least several years, and maps for recent months were generally not available. By automating and streamlining data acquisition, quality assurance and quality control (QA/QC), gap filling, interpolation, and dissemination, this initiative intends to produce results in regularly updated maps with the goal of releasing a preliminary product within one month of the end of the most recent month. Having more up-to-date analyses available will potentially facilitate additional applications for these datasets, such as drought planning and wildfire management.

In addition, to the generation of these rainfall products the dissemination and visualization of these products is crucial for engaging stakeholders and community members in utilizing this data. The Hawai‘i Rainfall Analysis and Mapping Application (HI-RAMA) is a publicly available tool to facilitate rapid and easy, web-based access to current and legacy monthly rainfall products. In this paper we introduce the HI-RAMA and its design and implementation related to data and visualization.

2 BACKGROUND

2.1 ‘Ike Wai

The current Hawai‘i EPSCoR Track 1 project, named ‘Ike Wai, has a mission to ensure Hawai‘i’s future water security through an integrated program of research, education, community engagement, and decision support. Its vision is to promote water resource management in Hawai‘i that is sustainable, responsible, and data driven. Scientific, cultural, and social dimensions to the problem of water security are integrated in a transparent, stakeholder-driven, and rigorous water research enterprise. In order to support this mission a robust cyberinfrastructure solution such as a science gateway was required.

Science gateways and science gateway frameworks have become popular for providing research infrastructure and are being supported by organizations such as the Extreme Science and Engineering Discovery Environment (XSEDE) [8] and the Science Gateways Community Institute (SGCI) [12, 13, 28]. Several efforts to provide “as a service” gateways such as WS-PGRADE/gUSE project [17], the HUBzero project [22], Apache Airavata [19, 24], Galaxy [4], and Globus [5] have a cloud-style model for managing data transfers and sharing data. Existing hydrology gateways and tools such as the CUAHSI Hydrological Information System (HIS) [27], Hydroshare [26], and Virtual Observatory and Ecological Informatics System (VOEIS) [20] provide exceptional server side support, data entry, and quality control client tools. However, existing water data gateways lack tools for managing and validating multi-domain metadata, integration for multiple data stores, and non-integration of compute resources or flexibility in access and selection of those resources.

To address these challenges Tapis [25], built on Agave [9, 10], was chosen as the framework to build the ‘Ike Wai gateway due the wide range of supported functionality that enable supporting the above needs and UH’s partnership with the Texas Advanced Compute Center (TACC) that support it. Tapis is an open source, scientific API middleware for hybrid cloud computing and data management that is powering a number of current community science gateways including CyVerse [15, 23], which provides life scientists with powerful computational infrastructure to handle huge datasets and complex analyses that enable data-driven discovery.

The ‘Ike Wai gateway [7] supports research in hydrology and water management and provides tools to address questions of water sustainability in Hawai‘i. The gateway provides centralized web-based user interfaces and RESTful

APIs for multi-domain data management, computation, analysis, and visualization tools. These tools provide reproducible science, modeling, data discovery, and decision support for the 'Ike Wai research team and wider Hawai'i hydrology community. By leveraging the Tapis framework [25] UH has constructed a gateway that ties data management and advanced computing resources together to support diverse research domains including microbiology, geochemistry, geophysics, economics, and humanities.

The HI-RAMA tool utilizes the 'Ike Wai Gateway cyberinfrastructure and APIs for implementation of data and metadata storage, management and retrieval.

2.2 Precipitation Data in Hawai'i

Monthly and daily rainfall data hosted within the 'Ike Wai Gateway, is obtained from a number of sources (see Longman, et al. [18] for a comprehensive description of the climate data networks in Hawai'i). The near-real-time data utilized in some of the rainfall data products are obtained from several national, online data repositories including; the Hydrometeorological Automated Data System (HADS) [1], National Centers for Environmental Information (NCEI) [2] and the Soil Climate Analysis Network (SCAN) [3]. Additional station data are obtained from various local networks either through telemetry or via manually upload to the 'Ike Wai gateway.

For rainfall analysis, all data are quality controlled at the daily time step. First, all data are put into a standard format (SI units, time-stamp, missing value place holder, and time step if necessary). Next, data are screened, and subsequently flagged for extraneous outlying values. Then, gaps are filled using either a normal ratio or inverse distance weighting approach depending on the length of record at an individual station [18]. Finally daily data are aggregated to the monthly time step.

Monthly rainfall data are used to create gridded monthly rainfall maps for the State of Hawai'i at a 250m x 250m resolution using a geo-statistical Ordinary Kriging (OK) method combined with a climatologically aided interpolation (CAI) approach that combines long-term climatologies with monthly station data [11]. Frazier et al., 2016 has shown that OK combined with the CAI is the most effective method for mapping monthly rainfall in Hawai'i. The main advantage of the CAI approach is that the climatology maps [14] have more spatial information built into them because they were created using a denser network of stations and more sophisticated mapping approach.

At the start of each month, the rainfall analysis workflow (Figure 1) produces a rainfall map for the seven major islands Kaua'i, O'ahu, Moloka'i, Lana'i, Maui, Kaho'olawe, and Hawai'i (the island of Ni'ihau is not considered in this study because no rainfall data is available there). For every monthly rainfall map, corresponding uncertainty maps including cross standard error maps and cross validation statistics are calculated.

The 'Ike Wai Gateway hosts a number of rainfall products available in several user friendly and downloadable formats. Daily rainfall time series' are available from 1990 to present-day and monthly time series and maps are available from 1920 to present-day.

2.2.1 Datasets and Query Tools. HI-RAMA users are able to query data based on parameters including, variable period-of-record, geographic coordinates, station name, elevation, or available record. The website will host an interactive map (Figure 3) allowing users to plot any month in the time series, zoom in on areas of particular interest, navigate to specific locations with the help of a choice of different base maps, and click on any location to get the monthly rainfall. When individual station points are selected, the accompanying meta-data including the availability of data will be displayed. Data tables will be available for download as well as customized data extraction from the entire gridded time series for predetermined point location or defined region of interest from the entire gridded time series.

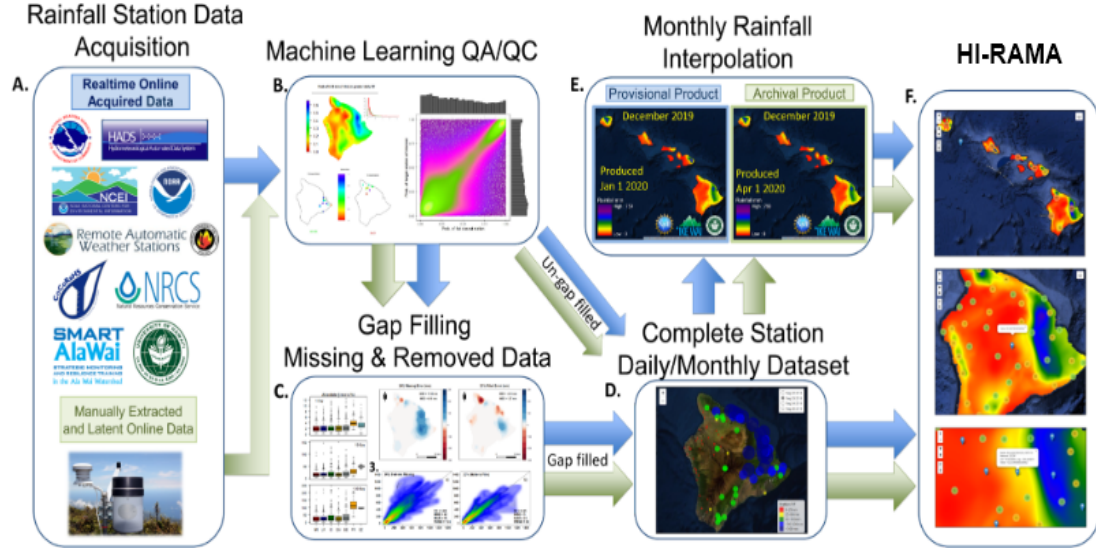


Fig. 1. Automated Hawai'i Statewide Rainfall Data Mapping Workflow A) Rainfall station data is collected from multiple sources and processed to daily totals with online resources to make a provisional product. Additionally, after some delay, manually collected rainfall is used to produce an archival product. B) A custom, random-forest machine learning algorithm uses the 10 closest stations to estimate if an observed rainfall value is true or a data collection error. C) Data-driven gap filling of removed and missing values using best available methods to fill missing station values. D) Daily processed, raw, and gap filled station data is produced. E) Gap filled monthly station values are used to create 250m x 250m resolution rasters of monthly rainfall amount, anomaly, and respective standard errors. F) HI-RAMA serves as the user interface for data interaction, query, and download.

3 IMPLEMENTATION

This project builds on lessons learned from developing The 'Ike Wai Hawai'i Groundwater Recharge Tool [21] to create a responsive, interactive web interface for visualizing the collected rainfall data and facilitating data retrieval for user analysis.

3.1 Architecture

The HI-RAMA web interface is built on the Angular 7 web application framework. Angular is a TypeScript – a typed superset of JavaScript – based framework for building scalable, modular web applications. This framework allows for application functions to be separated out into logical components for better extensibility and stability while developing features.

For generating maps that are used by the application for visualizing the geo-spatial rainfall data, Leaflet is used. Leaflet is a JavaScript library for creating interactive maps and has an extensive API for placing map points, generating layers for viewing geolocated data, and getting user input map locations.

Storage of the rainfall dataset for retrieval by the web interface is handled by leveraging the Tapis metadata API [25]. Rainfall data can be processed into JSON documents and added to a metadata store. JSON documents are loaded into

the 'Ike Wai Gateway and stored in a Mongo Database which allows easy data querying and retrieval via RESTful API requests to the Tapis metadata API endpoint.

3.2 Data Import and Storage

After initial processing, rainfall data is provided as a set of rainfall station metadata, rainfall station values for each daily time step, and the interpolated raster data for the state at each time step. This data is processed into a set of JSON documents that can be loaded into the Tapis Metadata for storage and can then be queried by the HI-RAMA web interface. The Agave command line interface is used to handle authentication and API request generation to the Tapis metadata endpoint for easily inserting JSON documents into the metadata store.

Rainfall station values and metadata are decoupled and loaded as separate JSON documents, such that a single instance of each station's metadata is stored and the value documents contain the unique identifier for the station which can be used to reference the station's metadata document. Decoupling the rainfall station information into separate metadata and value documents reduces the amount of redundant information that needs to be retrieved for each time frame.

3.3 Application Design

Once the raw rainfall data is processed and stored in the database, the web application provides users with an interactive interface to retrieve and visualize the rainfall data rasters and station metadata for a defined time period. Figure 2 shows the primary set of organizational components defined by the web application to handle and display the requested set of rainfall data. These organizational structures are defined to provide isolation of different application functions, preventing cascading issues when making updates and modifications that do not affect the main data model.

3.4 Data Handling

At application initialization, the most recent set of rainfall data is loaded into the application, and metadata for all rainfall stations is loaded and stored by the data manager. Station metadata is pre-loaded into the application during initialization. Although not all of the station metadata will be in use for every dataset due to differences in active periods between rainfall stations, the relatively few rainfall stations allows for this information to be stored without much overhead and allows for quick retrieval of station metadata between datasets. On loading a new dataset into the visualization, the relevant station metadata can be quickly referenced and displayed.

The data processor component handles the conversion of any incoming or outgoing data. Modules for loading local data from application assets or for retrieving data loaded into the Tapis metadata store are defined to handle the actual data retrieval/querying and the data manager handles caching and application storage, but this component acts as a filter to maintain consistency for handling multiple formats. Additionally, having a single, standardized internal data representation allows for easy adaptation to any additional data formats that may be desired. All that is required to add a format definition is a transformation function in this component, and the rest of the application will handle the data without any changes.

Incoming data are standardized into an internally defined raster format for map-wide value grids, or a rainfall station information format which contains the set of available rainfall station values composed with their respective metadata. Initial raster values are accepted in geoTIFF format or as JSON documents containing a value grid and the necessary information to construct the data on a map – e.g. lower left coordinates, grid cell dimensions, and grid width and height. Since each data grid for the application is expected to be identically positioned and sized, once initialized only the

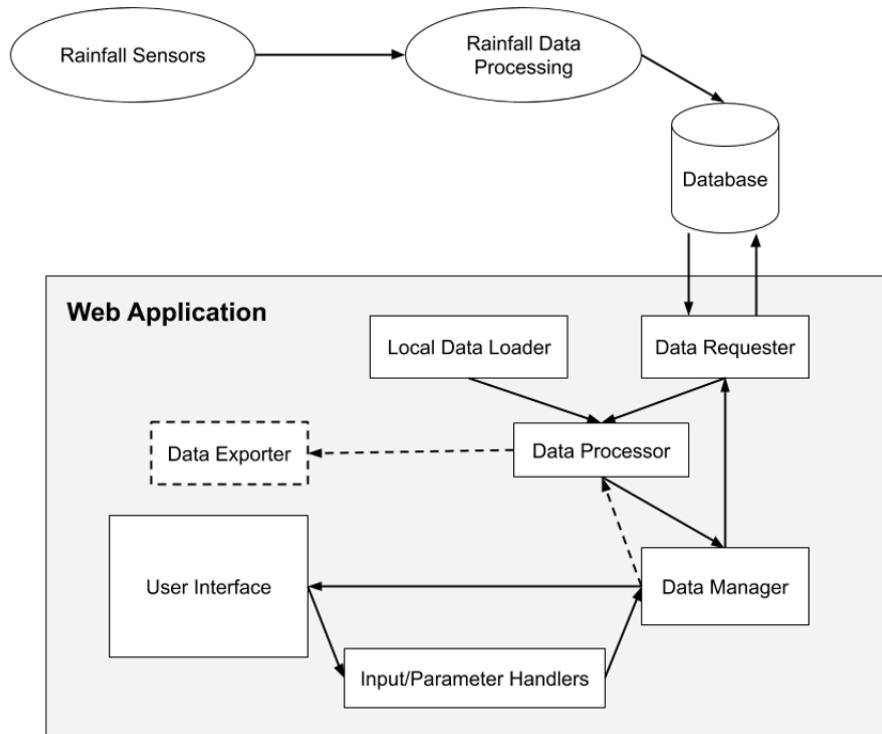


Fig. 2. The flow of data through HI-RAMA and its primary data handling components.

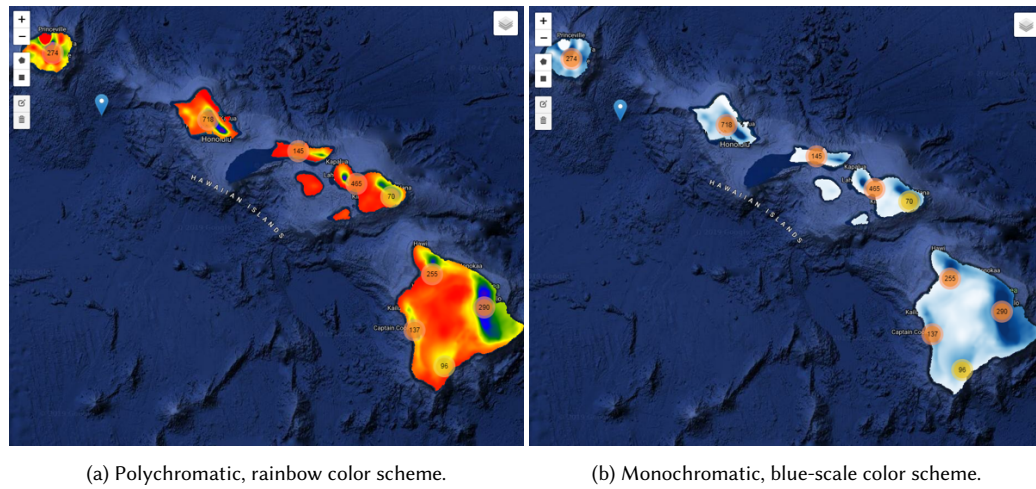


Fig. 3. The HI-RAMA interactive map interface. Two color schemes are provided for visualizing rasterized rainfall values.

actual value array is required to update the application's data raster. This prevents redundant positional information from being loaded for every raster. On data export, the internal storage format can be serialized into JSON documents or geoTIFFs.

User input parameters, such as date range selection, are linked to a set of input handlers that process the user input and pass it to the data manager. The data manager coordinates retrieval and caching of data being used by the application. When a new set of data or data view is requested by the user interface, the data manager will check if the requested dataset is already loaded into storage and trigger a request for the data if not available. Once the required data is returned by the data processor, the information will be cached and pushed to the user interface. On the user interface receiving a set of data, the values are loaded into the visualization to update the user's view.

In the case that a very large range of data is requested by the user, the data manager will also handle breaking the request into chunks for retrieval. Since the visualization only displays a specific time instance at a given moment, requests to load in large time ranges can be broken up and different time frames requested as they need to be displayed. The main issue with this approach is that it does not allow metrics for large time ranges to be computed and displayed, since not all of the data is loaded into the application. To work around this, values for data subsets, such as values for a specific selected rainfall station, can be loaded in as needed — e.g. when a user selects a rainfall station — and cached separately. This set of values is much smaller than the set required for the full visualization and can be stored without a major impact on performance.

3.4.1 Performance and Data-Processing Extensions. For aggregate metrics over the entire data set, some values can be pre-computed and stored along with the full value rasters and rainfall station data. This pre-aggregated data could be loaded without excessive overhead. Additionally, some metric computation functionality could potentially be implemented server-side so the values could be retrieved directly without requiring the full set of data to be loaded into the application.

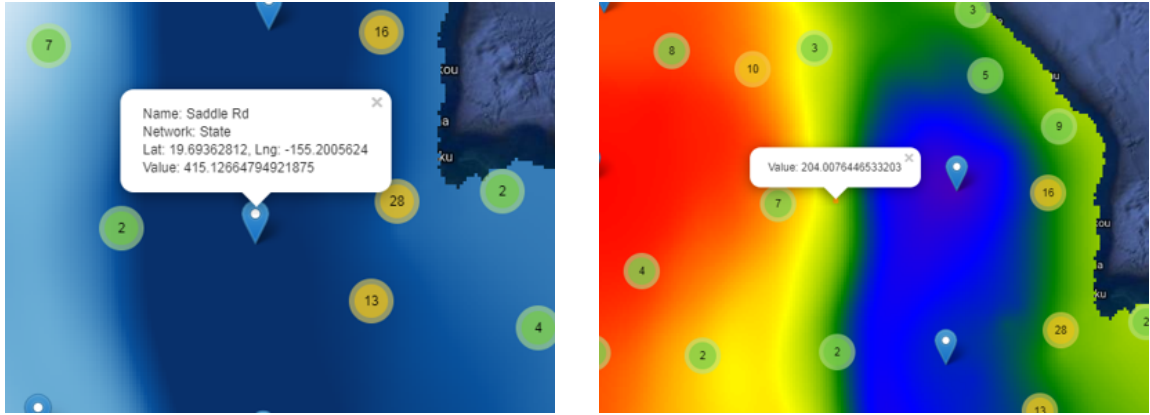
GeoTIFFs for the raster data should be stored in the database alongside the raw raster data for retrieval to avoid the extra processing overhead and potential loss of data from converting the raster data back into a geoTIFF on export. Additional issues arise for exporting very large sets of data — for example, rainfall data over a large time range — so data exportation will eventually have the option to delegate to a server side operation with the ability to email links to results to a user after processing is complete.

3.5 Map Generation

An interactive map is central to the data visualization component of the application. Raster data loaded into the visualization is used to generate a layer in the Leaflet map instance. Layers are produced using a custom Leaflet layering system extending Leaflet's `GridLayer` — a generic layer template that can be extended to create arbitrary tiled HTML layers. This allows each raster cell to be colored relative to its value and displayed on the map in its respective location.

A set of functions are defined for quickly converting longitude, latitude positional data into the corresponding position within the raster grid by rounding down to the nearest raster cell width and height unit. Given an efficient color lookup algorithm, this allows the map tiles for the raster layers to be generated quickly, resulting in a responsive interface when interacting with the map.

After the map layers are generated, the raster data objects remain exposed to the application to allow the raster values to be swapped out on the fly for reuse. Resultantly, the map layers only have to be created once at application



(a) Map markers represent the various rainfall stations around the state. Clicking on a marker displays information about the rainfall station and the rainfall value it detected for the currently displayed time step.

(b) Hovering the cursor over an area on the map highlights the extent of the raster cell under the cursor and displays the rainfall value for that cell.

Fig. 4. Map data display functionality

initialization; changes made to each visualization layer can be inserted as the data is received and are reflected immediately by redrawing the layer.

Two different color schemes are available for coloring the map values (figure 3). A polychromatic scheme with a rainbow color progression – red, orange, yellow, green, blue, purple – is provided for user comfort due to its prevalence in environmental sciences; however, this color scheme can be problematic for visualizations due to the progression of relative luminance within the color scheme. The intermediary colors in this scheme have a higher relative luminance than that of the extreme values, which can result in perceptual artifacts. Bands of moderate values stand out from the more extreme values and can draw unwarranted attention [6, 21].

A monochromatic color scheme is also provided to offset this disadvantage and provide a more consistent visualization option. Monochromatic color scaling provides a consistent relative luminosity progression and can be converted to greyscale without loss of data. This color scheme uses a white to blue color progression based on a single hue color scale provided by ColorBrewer2, a tool for selecting color schemes for cartographic visualizations [16].

To generate a continuous gradient between colors, chroma.js, a JavaScript library for generating color palettes, was used. This library has scale and domain functions which can generate a function for mapping values to a continuous color scheme on a defined value domain. Since efficiency is important for maintaining a responsive map when data layers are redrawn, this function is used to generate a static array of colors at small value steps – the domain is broken into 450 steps, allowing for constant time color lookup with minimal loss of quality.

Rainfall station metadata and values that are available for the currently viewed time are placed as markers on the interactive map. To avoid loading too many markers at a time, the Leaflet marker cluster plugin is used. This plugin creates clusters of similarly located markers added to the map and displays the number of sites within the cluster. Hovering over a cluster also shows a convex shell for the area the cluster covers. Single rainfall stations can be clicked to view information about the site on a map popup (figure 4a). This includes some of the site’s metadata and its rainfall value. Additionally, when moving the mouse over the map, hovering over an area will display the raster data value at

that location and highlight the extent of the data cell under the cursor. This allows for users to quickly identify more specific values for the map at large (figure 4b).

4 CONCLUSION

The HI-RAMA decision support tools offers improved accessibility of statewide rainfall data for Hawai'i and allows for the visualization and exploration of the data in a user-friendly and responsive interface. This should allow for improvements in the ability to analyze rainfall data and facilitate its use in decision making and additional research products. Improving the ability to build off previously collected data allows for more advanced developments with lower initial overhead. Additionally, this tool could provide similar functionality as a user endpoint for other data workflows. Structuring the application in such a way as to allow easy integration with other, similarly composed datasets provides potential extensions to other climatological data with the same benefits.

5 FUTURE WORK

The HI-RAMA will be integrated into a state of the art climate gateway for data retrieval, quality control, and near-real-time rainfall mapping with additional climate variables. Near-surface air temperature, relative humidity, wind speed, and many others will be available through this gateway. For the first time ever, researchers will have access to all of the available climate data up to the present day. Additional mapping efforts will include the development of near-real-time monthly temperature maps and eventually maps of both rainfall and temperature at the near-real-time daily time step. The creation and ongoing production of rainfall estimates will facilitate the potential for further study into integration and application of this and other climate products into additional components of the 'Ike Wai project, such as economics, conservation planning, and groundwater balances and discharge.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation Office of Advanced CyberInfrastructure [grant numbers 1931439 and 1931575]. Support for the Hawai'i EPSCoR Program is provided by the National Science Foundation's Research Infrastructure Improvement (RII) Track-1: 'Ike Wai: Securing Hawai'i's Water Future Award # OIA-1557349. Collaborative Research: SS2-SSI: The Agave Platform: An Open Science-As-A-Service Cloud Platform for Reproducible Science NSF ACI #145041

We thank, Kevin Kodoma from the National Weather Service. Dr. YinPhan Tsang, Yu-Fen Huang and Dr. Creighton Litton from the Department of Natural Resource Management, Dr. Paul Krushelnycky from the Department of Plant and Environmental Protection Sciences, Ryan Mudd from the Department of Geography and Environment, Dr. Brian Glazer and Cori Yap from the Department of Oceanography. Andrea Buckman, from the Leeward Haleakala Watershed Restoration Partnership. Mark Wasser and Ryan Monelo from the National Park Service, Dr. Lucas Fortini, from the United States Geological Survey, Pacific Island Ecosystem Research Center. Dr. Susan Cordell and Robert Hegemann from the Forest Service, Pacific Southwest Research Station. Dr. Safeeq Khan, University of California Division of Agriculture and Natural Resources, Ann Prouty, Ulupalakua Ranch and citizen scientists Neal Fujii, and Glenn Bauer.

REFERENCES

- [1] [n.d.]. *Hydrometeorological Automated Data System*. Retrieved Feb 17, 2020 from <https://hads.ncep.noaa.gov/>
- [2] [n.d.]. *National Centers for Environmental Information*. Retrieved Feb 17, 2020 from <https://www.ncei.noaa.gov/>
- [3] [n.d.]. *Soil Climate Analysis Network*. Retrieved Feb 17, 2020 from <https://www.wcc.nrcs.usda.gov/scan/>

- [4] Enis Afgan, Jeremy Goecks, Dannon Baker, Nate Coraor, Anton Nekrutenko, and James Taylor. 2011. Galaxy: A Gateway to Tools in e-Science. *Guide to e-Science* (April 2011), 145–177. https://doi.org/10.1007/978-0-85729-439-5_6
- [5] Rachana Ananthakrishnan, Kyle Chard, Ian Foster, and Steven Tuecke. 2014. Globus platform-as-a-service for collaborative science applications. *Concurrency and Computation: Practice and Experience* 27, 2 (March 2014), 290–305. <https://doi.org/10.1002/cpe.3262>
- [6] David Borland and Russell M. Taylor II. 2007. Rainbow Color Map (Still) Considered Harmful. *IEEE Computer Graphics and Applications* 27, 2 (March–April 2007), 14–17. <https://doi.org/10.1109/MCG.2007.323435>
- [7] Sean B. Cleveland et al. 2018. The ‘Ike Wai Gateway- A Science Gateway For The Water Future of Hawai‘i’. Science Gateways Community Institute, Proceedings of Science Gateways 2018, Austin TX, USA September 2018.
- [8] Maytal Dahan, John Towns, Tim Cockerill, Maytal Dahan, Ian Foster, Kelly Gaither, Andrew Grimshaw, Victor Hazlewood, Scott Lathrop, Dave Lifka, Gregory Peterson, Ralph Roskies, J. Scott, and Nancy Wilkins-Diehr. 2014. XSEDE: Accelerating Scientific Discovery. *Computing in Science & Engineering* 16, 5 (September 2014), 62–74. <https://doi.org/10.1109/MCSE.2014.80>
- [9] Rion Dooley, John Fonger, Gwen Jacobs, and Steven Brandt. 2018. The Agave Platform: An Open Science-As-A-Service Cloud Platform for Reproducible Science. SGCI, Gateways 2017. <https://doi.org/10.6084/m9.figshare.4675765.v2>
- [10] Rion Dooley, Matthew Vaughn, Dan Stanzione, and Steve Terry. 2012. Software-as-a-service: the iPlant foundation API. IEEE, 5th Workshop on Many-Task Computing on Grids and Supercomputers.
- [11] Abby Frazier, Thomas Giambelluca, Henry Diaz, and Heidi Needham. 2016. Comparison of geostatistical approaches to spatially interpolate month-year rainfall for the Hawaiian Islands. *International Journal of Climatology* 36, 3 (2016), 1459–1470. <https://doi.org/10.1002/joc.4437> arXiv:<https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.4437>
- [12] Sandra Gesing, Katherine Lawrence, Maytal Dahan, Marlon E. Pierce, Nancy Wilkins-Diehr, and Michael Zentner. 2019. Science gateways: Sustainability via on-campus teams. *Future Generation Computer Systems* 94 (2019), 97–102.
- [13] Sandra Gesing, Nancy Wilkins-Diehr, Maytal Dahan, Katherine Lawrence, Michael Zentner, Marlon Pierce, Linda Hayden, and Suresh Marru. 2017. Science gateways: the long road to the birth of an institute. Proceedings of the 50th Hawaii International Conference on System Sciences. <https://doi.org/10.24251/HICSS.2017.755>
- [14] Thomas Giambelluca, Qi Chen, Abby Frazier, Jonathan Price, Yi-Leng Chen, Pao-Shin Chu, Jon Eischeid, and Donna Delparte. 2013. Online Rainfall Atlas of Hawai‘i. *Bulletin of the American Meteorological Society* 94, 3 (March 2013), 313–316. <https://doi.org/10.1175/BAMS-D-11-00228.1>
- [15] Stephen Goff, Matthew Vaughn, Sheldon McKay, Eric Lyons, Ann Stapleton, Damian Gessler, Naim Matasci, Liya Wang, Matthew Hanlon, Andrew Lenards, Andy Muir, Nirav Merchant, S. Lowry, Stephen Mock, Matthew Helmke, Adam Kubach, Martha Narro, Nicole Hopkins, David Micklos, and Dan Stanzione. 2011. The iPlant Collaborative: Cyberinfrastructure for Plant Biology. *Frontiers in Plant Science* 2, 34 (July 2011). <https://doi.org/10.3389/fpls.2011.00034>
- [16] Mark Harrower and Cynthia A. Brewer. 2003. ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. *The Cartographic Journal* 40, 1 (June 2003), 27–37. <https://doi.org/10.1559/152304003100010929>
- [17] Peter Kacsuk, Zoltan Farkas, Miklos Kozlovsky, Gabor Hermann, Akos Balaskó, K. Karoczkai, and Is Marton. 2012. WS-PGRADE/gUSE Generic DCI Gateway Framework for a Large Variety of User Communities. *Journal of Grid Computing* 10 (December 2012), 1–30. <https://doi.org/10.1007/s10723-012-9240-5>
- [18] Ryan J. Longman, Thomas W. Giambelluca, Michael A. Nullet, Abby G Frazier, Kevin Kodama, Shelley D. Crausbay, Paul D. Krushelnycky, Susan Cordell, Martyn Clark, Andy Newman, and Jeffrey R. Arnold. 2018. Compilation of climate data from heterogeneous networks across the Hawaiian Islands. *Scientific Data* 5, 180012 (February 2018). <https://doi.org/10.1038/sdata.2018.12>
- [19] Suresh Marru, Lahiru Gunathilake, Chathura Herath, Patanachai Tangchaisin, Marlon Pierce, Chris Mattmann, Raminderjeet Singh, Thilina Gunarathne, Eran Chinthaka, Ross Gardler, Aleksander Slominski, Ate Douma, Srinath Perera, and Sanjiva Weerawarana. 2011. Apache airavata: a framework for distributed applications and computational workflows. Proceedings of the 2011 ACM workshop on Gateway computing environments, 21–28.
- [20] Seth J.K. Mason, Sean B. Cleveland, Pol Llovet, Clemente Izurieta, and Geoffrey C. Poole. 2014. A centralized tool for managing, archiving, and serving point-in-time data in ecological research laboratories. *Environmental Modeling & Software* 51, 59–69. <https://doi.org/10.1016/j.envsoft.2013.09.008>
- [21] Jared McLean, Sean Cleveland, Kolja Rotzoll, Scot Izuka, Jason Leigh, Gwen Jacobs, and Ryan Theriot. 2019. The ‘Ike Wai Hawai‘i Groundwater Recharge Tool. online. Proceedings of Gateways 2019. <https://doi.org/10.17605/OSF.IO/TH9W2>
- [22] Michael McLennan and Rick Kennell. 2010. HUBzero: A Platform for Dissemination and Collaboration in Computational Science and Engineering. *Computing in Science & Engineering*. 12, 2 (May 2010), 48–53.
- [23] Nirav Merchant, Eric Lyons, Stephen Goff, Matthew Vaughn, Doreen Ware, David Micklos, and Parker Antin. 2016. The iPlant Collaborative: Cyberinfrastructure for Enabling Data to Discovery for the Life Sciences. *PLoS Biology* 14, 1 (January 2016). <https://doi.org/10.1371/journal.pbio.1002342>
- [24] Marlon Pierce, Suresh Marru, Lahiru Gunathilake, Don Wijeratne, Raminderjeet Singh, Chathuri Peli Kankanamalage, Shameera Ratnayaka, and Sudhakar Pamidighantam. 2015. Apache Airavata: design and directions of a science gateway framework. *Concurrency and Computation: Practice and Experience* 27, 16 (May 2015), 4282–4291.
- [25] Joe Stubbs, Richard Cardone, Mike Packard, Anagha Jamthe, Smruti Padhy, Steve Terry, Julia Looney, Joseph Meiring, Steve Black, Maytal Dahan, Sean Cleveland, and Gwen Jacobs. 2020. Tapis: An API Platform for Reproducible, Distributed Computational Research. (2020). submitted.

- [26] David Tarboton, Ray Idaszak, Jeffery Horsburgh, Jeff Heard, Daniel Ames, Jonathan Goodall, L. Band, Venkatesh Merwade, Alva Couch, Jennifer Arrigo, Richard Hooper, David Valentine, and David Maidment. 2014. HydroShare: Advancing Collaboration through Hydrologic Data and Model Sharing. In *Proceedings of the 7th International Congress on Environmental Modelling and Software, San Diego, California, USA, International Environmental Modelling and Software Society (iEMSs)*. <https://doi.org/10.13140/2.1.4431.6801>
- [27] David G. Tarboton, Jeffery S. Horsburgh, David R. Maidment, Timothy L. Whiteaker, Ilya Zaslavsky, Michael Piasecki, Jonathan L. Goodall, David L. Valentine, and Thomas Whitenack. 2009. Development of a community hydrologic information system. In *Proceedings of 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation*. 988–994.
- [28] Nancy Wilkins-Diehr and T. Daniel Crawford. 2018. The science gateways community institute and the molecular sciences software institute. *Computing in Science & Engineering* 5, 20 (2018), 26–38.