

Using TEEHR Dashboards to Explore Hydrologic Model Performance

Hydroinformatics

Friday May 29th Session 1 (10:30AM)

DevCon 2026

This project received funding under award NA22NWS4320003 from NOAA Cooperative Institute Program. The statements, findings, conclusions and recommendations are those of the author(s) and do not necessarily reflect the views of NOAA.



The RTI International logo is displayed on a white rectangular background. It features a stylized blue square icon to the left of the letters "RTI" in a large, bold, blue serif font. Below "RTI", the word "INTERNATIONAL" is written in a smaller, blue, all-caps sans-serif font.

Team TEEHR



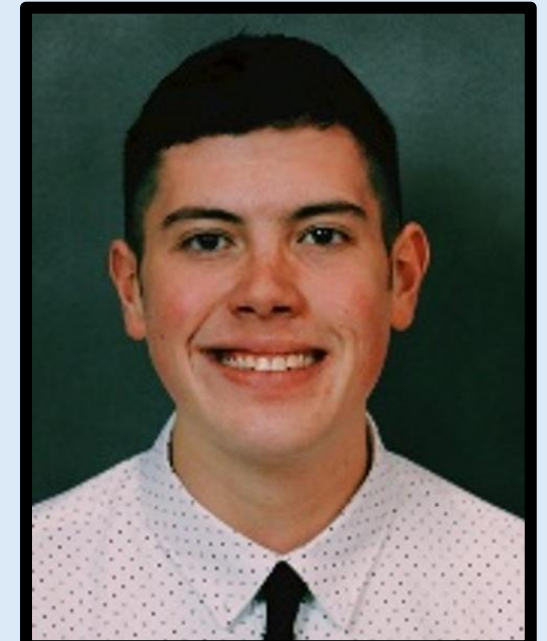
Matthew Denno
(PI - RTI)



Katie van Werkhoven
(co-PI - RTI)



Sam Lamont
(Developer - RTI)



Sam Landsteiner
(Developer - RTI)

Workshop Outline

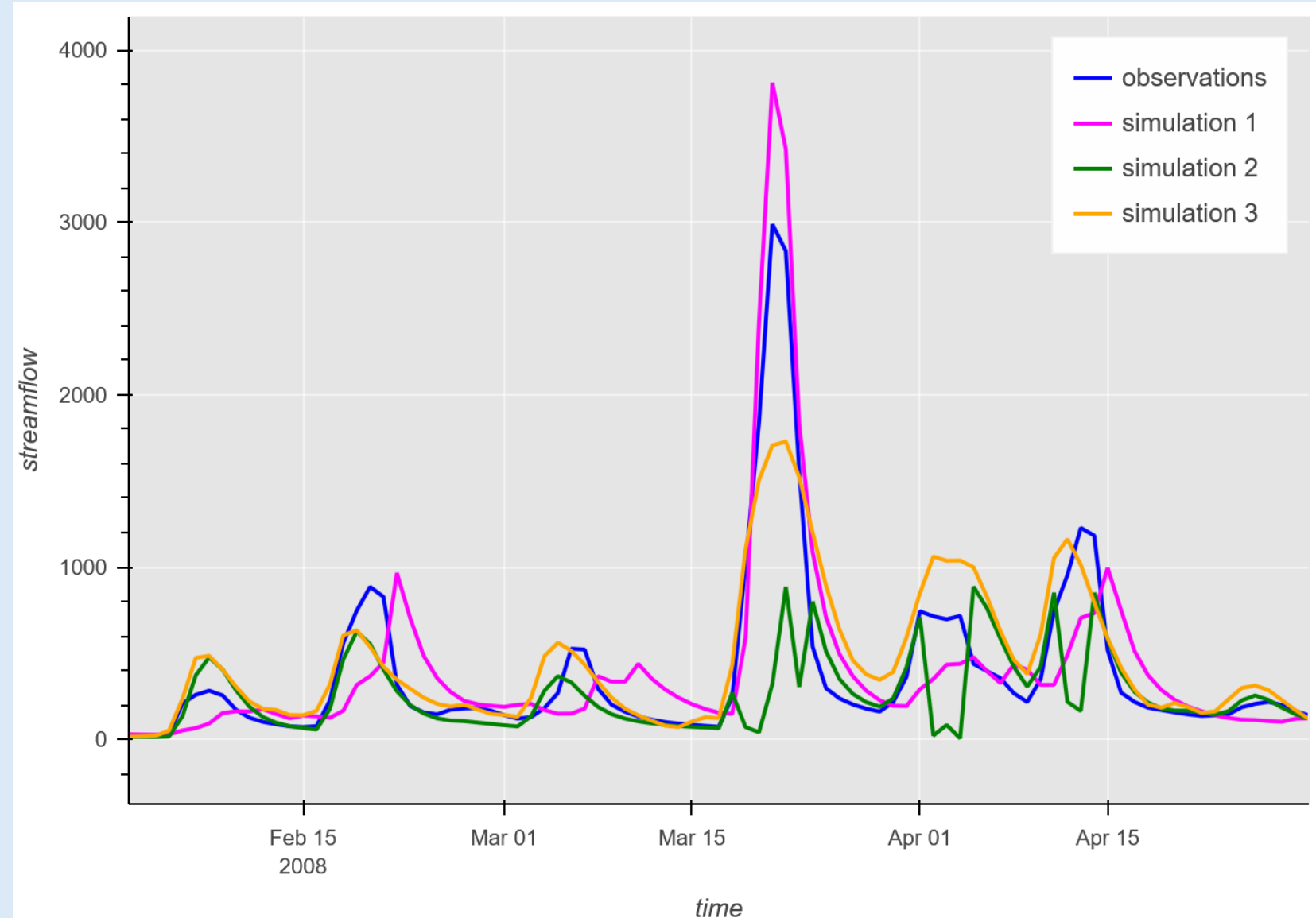
- Introduction
 - Project background and objectives
 - TEEHR-Cloud
 - Exploring the Dashboards
 - Data Availability
 - Retrospective Simulations
 - Forecast Analysis
 - Interactive analysis
 - Querying the data warehouse
 - Incorporating local data
- CIROH Slack channel #teehr
- Workshop materials:
<https://github.com/RTIInternational/teehr-devcon26-workshop>

Project Background

- RTI was engaged through CIROH to research, design, and develop a novel, **flexible and scalable hydrologic evaluation system** to support research and testing of potential forecast system advancements...the evaluation system would be a key factor in CIROH's ability to execute research that leads to **operational improvements** in the quality and use of hydrologic forecasts for public benefit and decision-making.
- The work presented herein is the result of two CIROH projects:
 - Integrated Evaluation System **Prototype** for Testing Research and Operational Advancements.
 - **Expansion** of the CIROH Evaluation System to Accelerate Research to Operations.

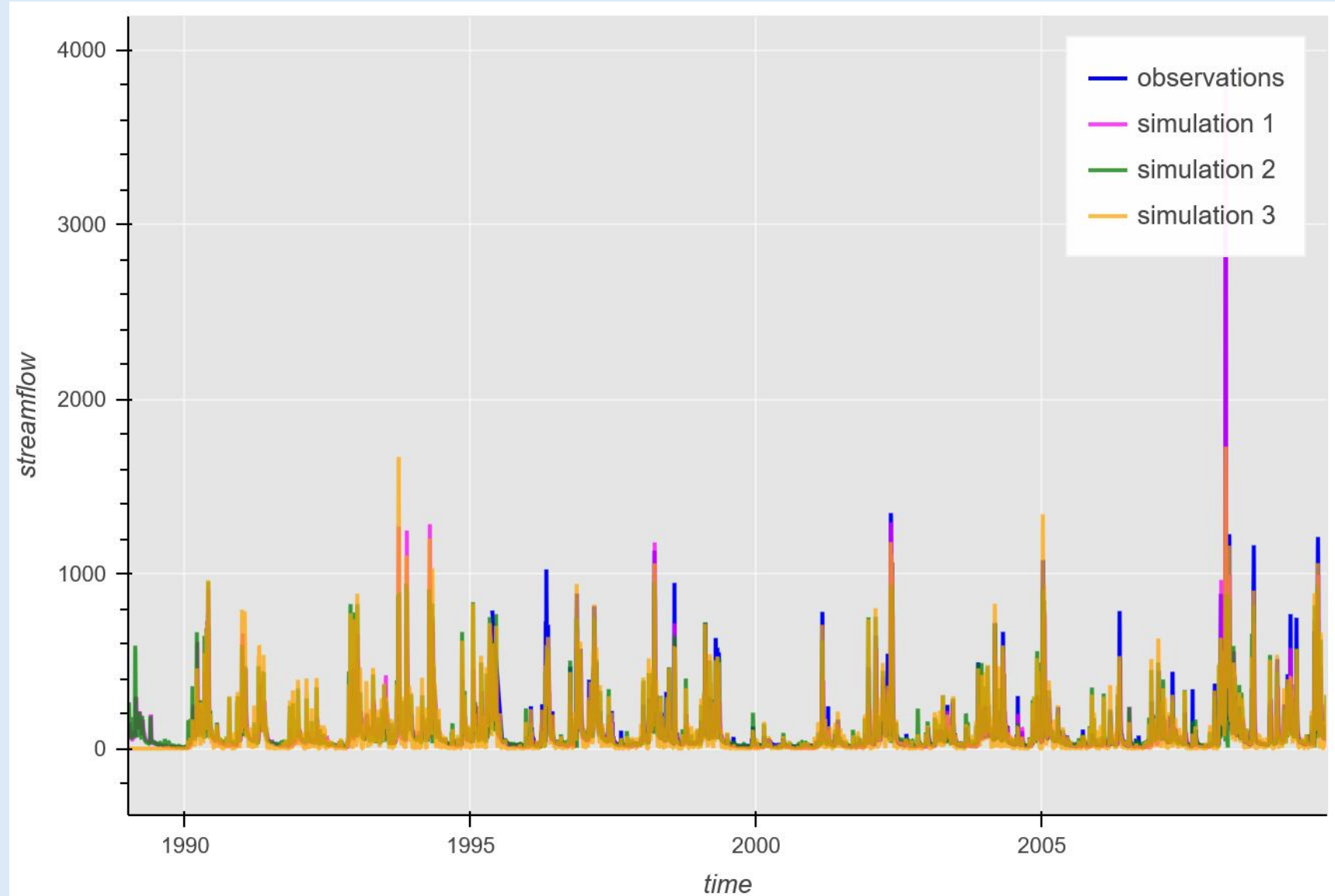
Project Background

Which hydrologic model is better?

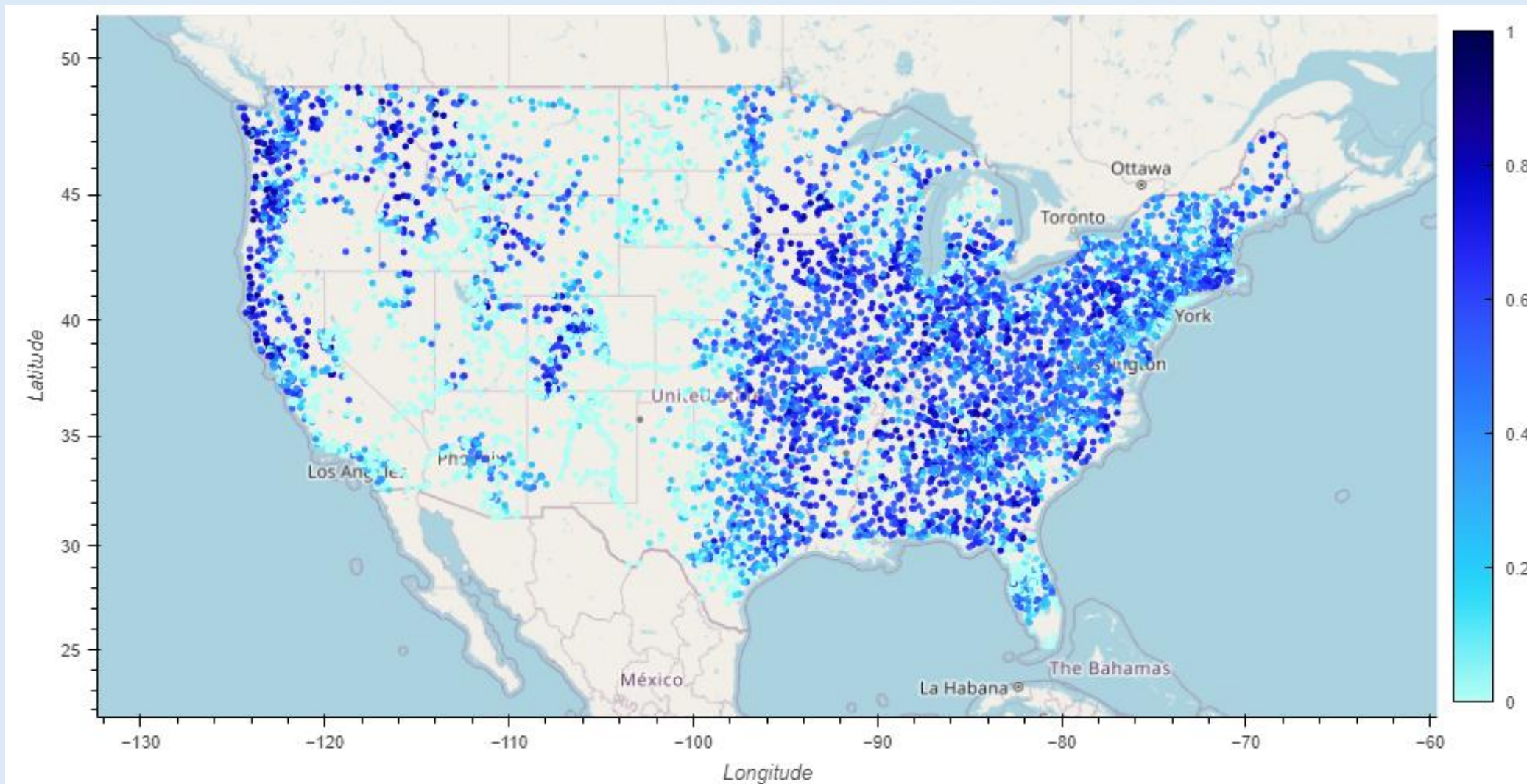


Project Background

Which hydrologic model is better?



Project Background



Project Background

You may need to consider:

- Magnitude
- Timing
- Volume
- Single event or long period of record
- High flow periods
- Low flow periods
- Seasons
- By Geographic Region

This will require more tools:

- Error statistics
- Skill scores
- Categorical metrics
- Hydrologic signatures
- Uncertainty quantification
- Graphical methods
- Some new emergent method...

And ultimately you want to understand:

- What are the drivers?
- Climate
- Time-period
- Regulation
- Basin characteristics
- Forcings
- Model selection
- And so on...

Project Background

These are the challenges TEEHR is designed to address!

- Data access and storage
- Data integrity and consistency
- Tools for metrics and analysis
- Scalability and compute resources
- Ways to share methods, data, and results with the community

A Tale of Two TEEHRs

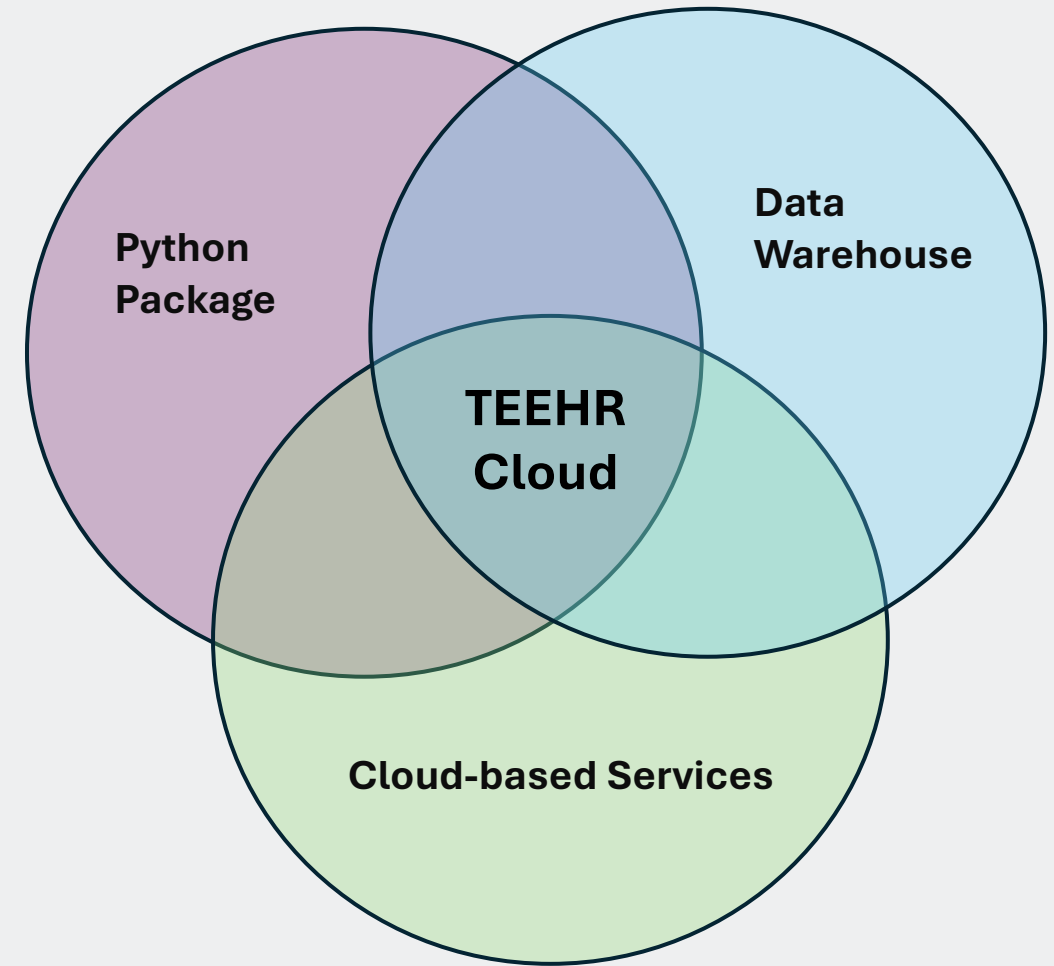
- **Tools for Exploratory Evaluation in Hydrologic Research (TEEHR)** is an opinionated Python framework for fetching, loading, storing, and processing large amounts hydrologic simulation data for the purpose of exploring and evaluating the datasets to assess their skill and performance.
 - Aimed at researchers, developers and scientists who need flexible, customizable tools to explore and assess hydrologic simulation data.
- **TEEHR Evaluation System (TEEHR Cloud)** is a cloud-based Evaluation Platform that leverages TEEHR to support standardized evaluations and provide “evaluation-ready” datasets and dashboards (and analytics capabilities).
 - Aimed more at agencies, decision-makers, operational forecasters, and modelers who need accessible, automated evaluation tools without deep technical customization.

A Tale of Two TEEHRs

- **Tools for Exploratory Evaluation in Hydrologic Research (TEEHR)** is an opinionated Python framework for fetching, loading, storing, and processing large amounts hydrologic simulation data for the purpose of exploring and evaluating the datasets to assess their skill and performance.
 - Aimed at researchers, developers and scientists who need flexible, customizable tools to explore and assess hydrologic simulation data.
- **TEEHR Evaluation System (TEEHR Cloud)** is a cloud-based Evaluation Platform that leverages TEEHR to support standardized evaluations and provide “evaluation-ready” datasets and dashboards (and analytics capabilities).
 - Aimed more at agencies, decision-makers, operational forecasters, and modelers who need accessible, automated evaluation tools without deep technical customization.

TEEHR Cloud

- Facilitate the evaluation of model and forecasting improvements
- Scalar time series evaluation
 - Continuous simulation (retrospective)
 - Deterministic forecasting
 - Ensemble forecasting
- Growing list of additional users/projects.



Enable standardized continental-scale hydrologic model evaluations



TEEHR

Tools for Exploratory Evaluation in Hydrologic Research

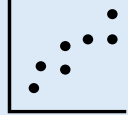
Fetching & Loading



Data Validation & Storage



Analytics

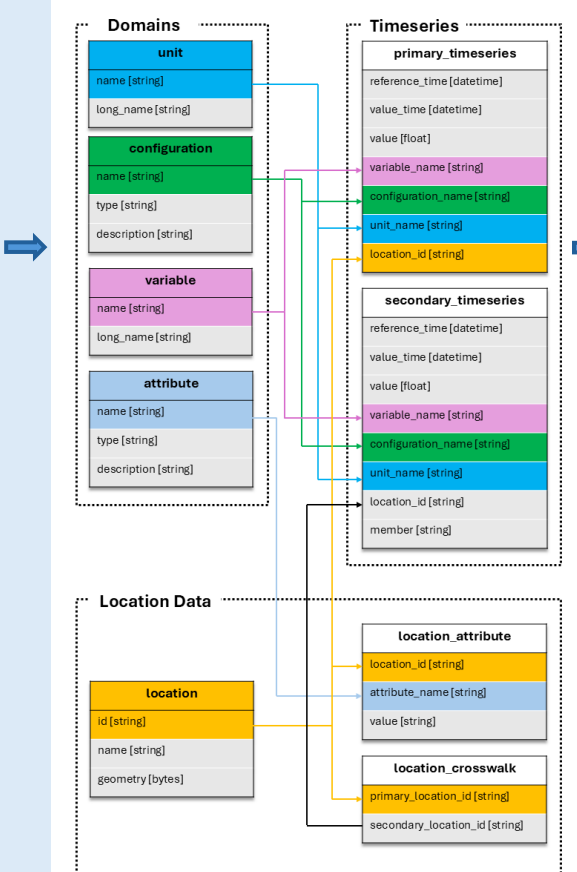


Results



Data Sources and Formats

NextGen

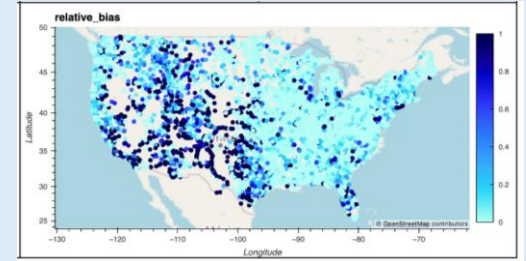


TEEHR Dataset

- dataset
 - attributes
 - configurations
 - joined_timeseries
 - location_attributes
 - location_crosswalks
 - locations
 - primary_timeseries
 - secondary_timeseries
 - units
 - variables

- Joining timeseries
- Built-in metrics
- Event detection
- Baseflow separation
- Uncertainty
- Location attributes
- Grouping
- Filtering
- Scalable

	configuration_name	primary_location_id	nash_sutcliffe_efficiency	klng_gupta_efficiency	relative_bias	root_mean_standard_deviation_ratio
0	nwm30_retrospective	usgs-01038000	0.686707	0.709544	-0.058372	0.559726
1	nwm30_retrospective	usgs-01401000	0.170434	0.185527	-0.133946	0.910805
2	nwm30_retrospective	usgs-01514850	0.656467	0.747603	0.170950	0.586117
3	nwm30_retrospective	usgs-01643500	0.193687	0.434383	-0.009543	0.897949
4	nwm30_retrospective	usgs-02326526	-10.448203	-0.273122	-0.907320	3.383519



The TEEHR Data Warehouse

“Iceberg is a high-performance format for huge analytic tables.”

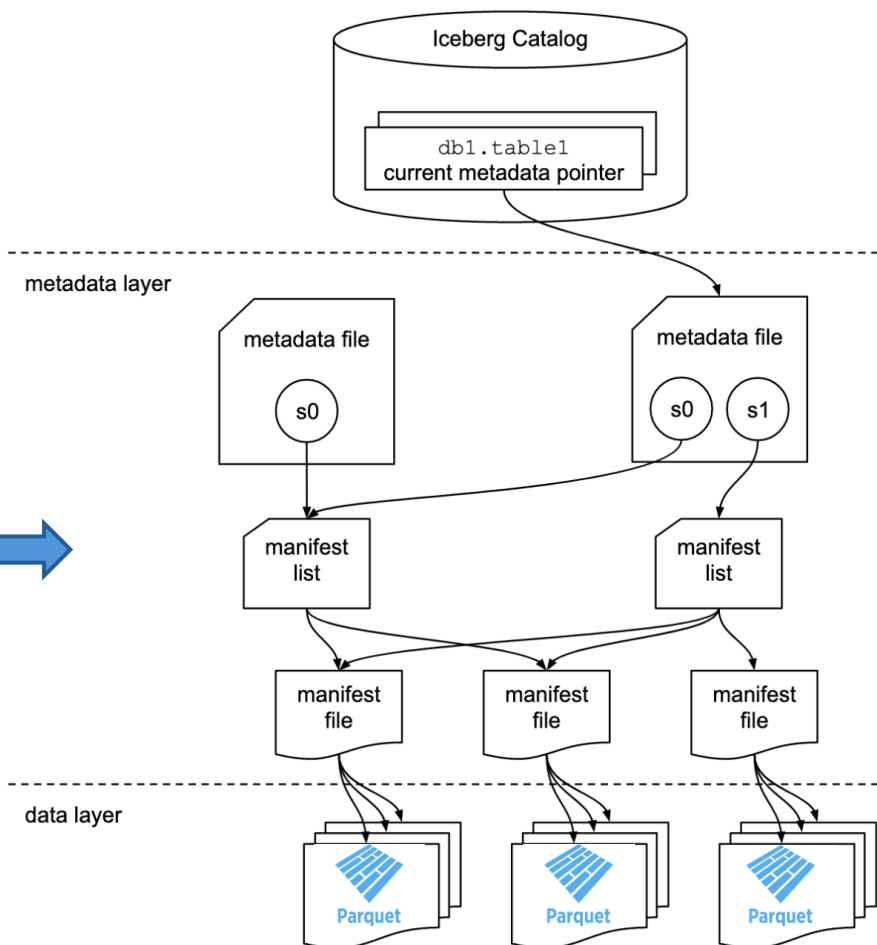
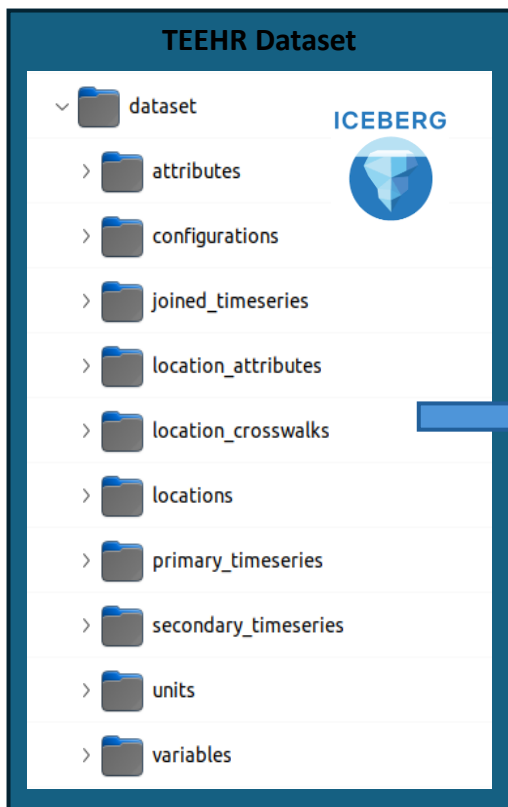


image source: <https://iceberg.apache.org/spec/#goals>

Key Iceberg Features:

- Optimized for huge data (Petabytes)
- Separates storage from compute
- Time travel
- Schema evolution
- Hidden partitioning
- REST catalog

Current Warehouse Contents:

- “Evaluation Ready Data”
- Historical simulations
 - NWM v3.0 hourly 40 yr retrospective
 - Corresponding USGS hourly streamflow
 - MAP processed to the corresponding USGS basins
- Real-time forecasts
 - NWM Short and medium range
 - NRDS CFE short and medium range
 - NRDS LSTM short and medium range
 - NWS RFC Forecasts
 - Corresponding USGS gage data
- ~50 location attributes (slope, % impervious, ...)



Amazon S3



Apache Iceberg™

The Cloud-based Services

Data Exploration



kubernetes

REST Service



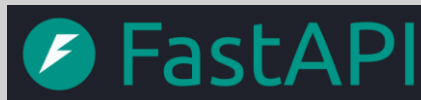
Automated Processing



PREFECT



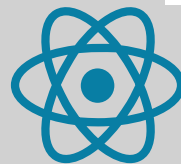
Dashboards



trino



Panel



React

Workflow/Evaluation Manager



Flows

7 Flows

Dashboard

Runs

Flows

Deployments

Work Pools

Blocks

Variables

Automations

Event Feed

Concurrency

<input type="checkbox"/>	Name	Last run	Next run	Deployments	Activity
<input type="checkbox"/>	ingest-datastream-forecasts <small>Created 2026/02/22 09:39:37 PM</small>	<input checked="" type="checkbox"/> ingest-datastream-forecasts	<input type="checkbox"/> enthusiastic-swan	4 Deployments	⋮
<input type="checkbox"/>	ingest-nwm-streamflow-forecasts <small>Created 2026/02/22 09:39:31 PM</small>	<input type="checkbox"/> ingest-nwm-streamflow-forecasts	<input type="checkbox"/> quixotic-barnacle	2 Deployments	⋮
<input type="checkbox"/>	ingest-nwps-rfc-forecasts <small>Created 2026/02/22 09:39:38 PM</small>	<input type="checkbox"/> ingest-nwps-rfc-forecasts	<input type="checkbox"/> ingest-nwps-rfc-forecasts	1 Deployment	⋮
<input type="checkbox"/>	ingest-usgs-streamflow-obs <small>Created 2026/02/22 09:39:26 PM</small>	<input type="checkbox"/> ingest-usgs-streamflow-obs	<input type="checkbox"/> ingenious-ferret	1 Deployment	⋮
<input type="checkbox"/>	routine-table-maintenance <small>Created 2026/02/22 09:39:39 PM</small>		<input type="checkbox"/> delightful-woodlouse	1 Deployment	⋮
<input type="checkbox"/>	update-forecast-metrics-table <small>Created 2026/02/22 09:39:39 PM</small>	<input type="checkbox"/> update-forecast-metrics-table	<input type="checkbox"/> defiant-jaybird	1 Deployment	⋮
<input type="checkbox"/>	update-joined-forecast-table <small>Created 2026/02/22 09:39:39 PM</small>	<input checked="" type="checkbox"/> update-joined-forecast-table	<input type="checkbox"/> unnatural-shrimp	1 Deployment	⋮

Items per page

TEEHR Dashboards

TEEHR Dashboard Hub

Tools for Exploratory Evaluation in Hydrologic Research. Choose from our collection of specialized dashboards for hydrologic data analysis.

- Application-specific evaluation dashboards

- Retrospective model comparisons
- Deterministic flood forecasting
- Ensemble water supply forecasting
- Forecast Informed Reservoir Operations (FIRO)
- Flash Flood Forecasts
- Low Flow Forecasts
- Reservoir Operations
- etc.

Available Dashboards

Data Availability

Explore contents of the data warehouse spatially and temporally.

Features:

- Data summaries by model configuration
- Data summaries by location
- Completeness analysis of timeseries data

Open Dashboard

Retrospective Simulations

Analyze and compare historical simulation data with observed values. Explore metrics, timeseries, and spatial patterns.

Features:

- Interactive maps with simulation metrics
- Time series visualization and comparison
- Statistical performance metrics
- Multi-configuration analysis

Open Dashboard

Forecast Analysis

Real-time forecast analysis and validation tools for operational and research data streams.

Features:

- Live forecast data
- Forecast vs observation comparison
- Lead time analysis

Open Dashboard

Retrospective Model Comparisons

Coming Soon

Evaluate multiple historical simulations compared to a baseline, considering sampling uncertainty.

Planned Features:

- Probability of improvement of alternative models over a baseline.
- Interactive maps, heatmaps and plots
- Aggregated and detailed summaries
- Filter by physical, geographical, temporal or hydrologic attributes to see trends

Coming Soon

Deterministic Forecast Evaluation for Flood Protection

Coming Soon

Evaluate deterministic forecasts for flood protection applications.

Planned Features:

- Event-based evaluation
- Evaluate event peak and timing
- Categorical metrics
- Variable flood thresholds
- Interactive maps and plots
- Warning and lead time assessments
- Sampling uncertainty estimated

Coming Soon

Ensemble Forecast Evaluation for Water Supply

Coming Soon

Evaluate ensemble forecasts for water supply applications.

Planned Features:

- Continuous and seasonal evaluation
- Ensemble metrics
- Evaluate low flows, low flow duration
- Variable low flow thresholds
- Interactive maps and plots
- Filter by physical, geographical, temporal or hydrologic attributes to see trends
- Sampling uncertainty estimated

Coming Soon

Post Flood Event Evaluations

Coming Soon

Evaluate ensemble forecasts for water supply applications.

Deterministic Forecast Evaluation for Dam Safety and Hydropower

Coming Soon

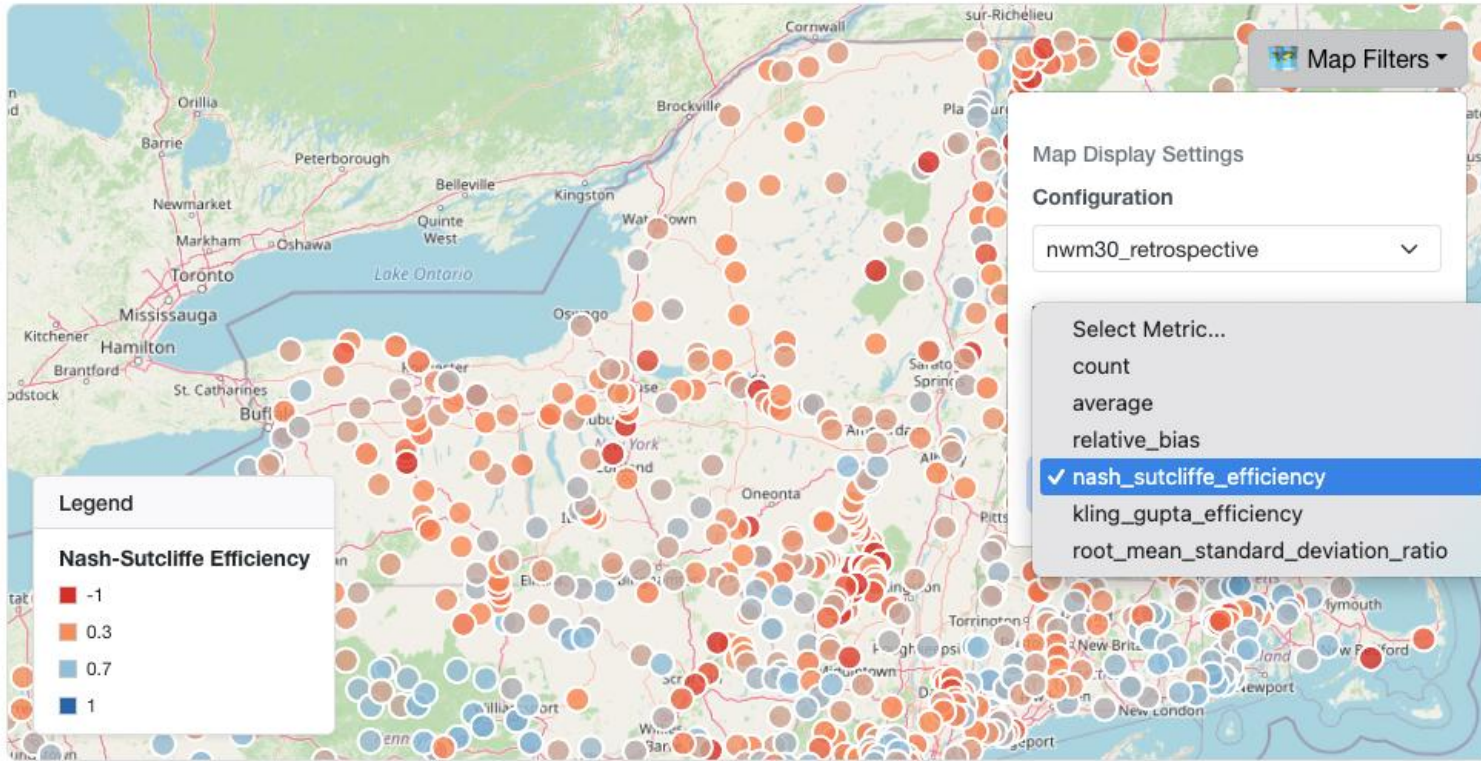
Real-time forecast analysis for dam safety and hydropower applications.

Deterministic Forecast Evaluation for Water Supply

Coming Soon

Real-time forecast analysis for water supply applications.

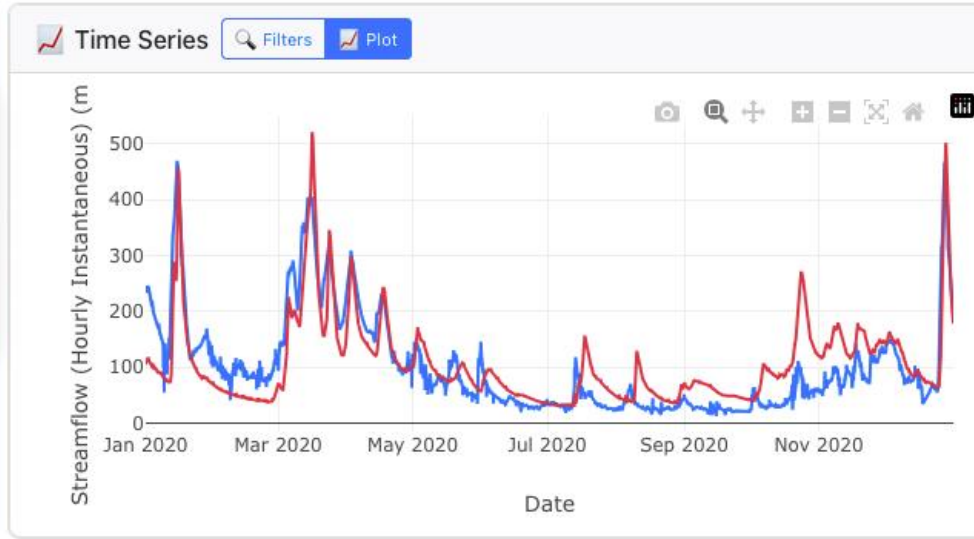
TEEHR Dashboards



Selected Location ✕

BLACK RIVER AT WATERTOWN NY

ID: usgs-04260500
43.9859, -75.9251



Metrics Filters Table

primary_location_id	configuration_name	variable_name	unit_name	count	average	relative_bias	nash_sutcliffe_efficiency	kling_gupta_efficiency	root_mean_stanc
usgs-04260500	nwm30_retrospective	streamflow_hourly_inst	m ³ /s	262629.0000	135.9299	-0.0404	0.4263	0.7198	

TEEHR-Hub Analytics

The screenshot displays a JupyterLab environment with the following components:

- File Explorer:** Shows a directory structure for 'accessing_teehr_warehouse' with files like '01_warehouse_contents.i...', '02_timeseries_plot.ipynb', '03_basic_metrics.ipynb', and '04_timeseries_with_events.ipynb'.
- Code Notebook:** Contains Python code for creating a Spark session, reading data from USGS, filtering for specific configurations, and performing event detection. The code is as follows:

```
[3]: spark = create_spark_session()
ev = teeHR.RemoteReadOnlyEvaluation(spark=spark)

USGS_GAGE_ID = "usgs-01010070"
PRIMARY_CONFIGURATION_NAME = "usgs_observations"
SECONDARY_CONFIGURATION_NAME = "nwm30_retrospective"

events_df = ev.table("sim_joined_timeseries").filter([
  {
    "column": "configuration_name",
    "operator": "=",
    "value": SECONDARY_CONFIGURATION_NAME
  },
  {
    "column": "primary_location_id",
    "operator": "=",
    "value": USGS_GAGE_ID
  }
]).add_calculated_fields([
  tcf.AbovePercentileEventDetection()
]).to_pandas()
```
- Terminal:** Shows the execution of the code.
- Map:** A scatter plot of the United States showing event locations. The x-axis is Longitude (ranging from -130 to -76) and the y-axis is Latitude (ranging from 25 to 50). A color scale on the right indicates event density from 0 to 1.0.
- Bar Chart:** Titled 'Observed Events Over 85th Percentile', showing the number of events per year from 2016 to 2022. The y-axis is 'primary_value' (0 to 160) and the x-axis is 'value_time' (years). The chart shows significant variability, with several years exceeding 100 events.

Scalable Compute Example (378 cores/~\$10/hr)

- Basic metric for the 40-yr retrospective (about 1 min)
- Basic metric with bootstrap uncertainty, 1000 reps (about 5 mins)
- We are still learning but the system is highly scalable and tunable for various workloads.
- Utilized in notebooks and automated workflows

Next Steps

- More NRDS integration
- Dashboard development
- Data processing
- TEEHR-Python feature dev
- Connecting
 - CIROH Slack channel #teehr
 - GitHub repos
 - <https://github.com/RTIInternational/teehr>
 - <https://github.com/RTIInternational/teehr-hub>
 - TEEHR-Python Documentation
 - <https://rtiinternational.github.io/teehr/>
 - Dashboards
 - <https://dashboards.teehr.rtiamanzi.org/>