

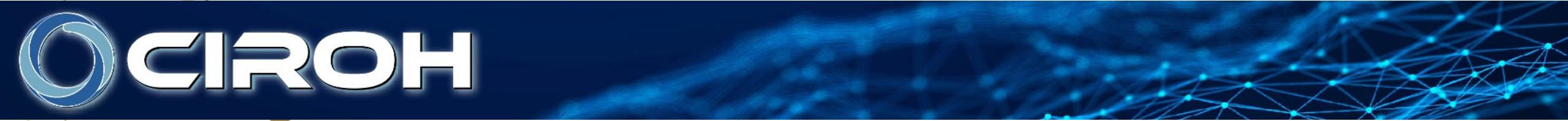
# AWRA 2024 GEOSPATIAL WATER TECHNOLOGY CONFERENCE

MARCH 25-27 | Orlando, FL

Data to Decisions: Managing and Modeling  
Water Challenges



#AWRA2024 | [www.awra.org](http://www.awra.org)



# ACCELERATING COMMUNITY CONTRIBUTION TO THE NEXT GENERATION WATER RESOURCE MODELING CAPABILITY

\*OFFICIAL TITLE: BUILDING AND SHARING A CONTINUOUS RESEARCH DATA STREAM USING THE  
NEXTGEN WATER RESOURCES MODELING FRAMEWORK WITH NEXTGEN IN A BOX

## CIROH Cyberinfrastructure Team

*AWI:* James Halgren, Arpita Patel, Sepehr Karimi, Ben Lee,  
Josh Cunningham, Trupesh Patel, Chad Perry, Karnesh Jain

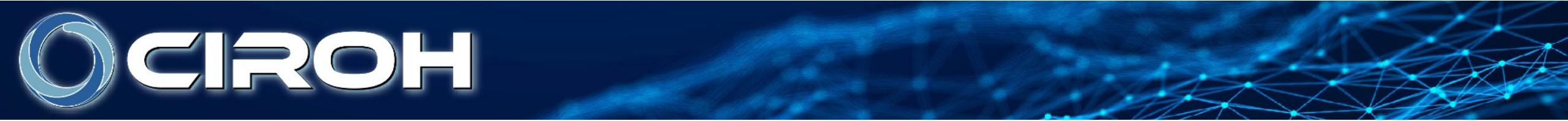
*Lynker:* Zach Wills, Jordan Laser, Nels Frazier

+ Many other CIROH Members and Partners

26 March 2024  
AWRA GWTC Meeting  
Orlando, Florida

Tuesday – Session 16: Big Data Applications in Water Resources - I





National Oceanic and  
Atmospheric Administration  
U.S. Department of Commerce

[Home](#) / [News & Features](#)

Search NOAA sites



# University of Alabama to lead NOAA institute to advance water and flood prediction

New institute will receive up to \$360 million over five years

Focus areas: Research Topics: water

Share:

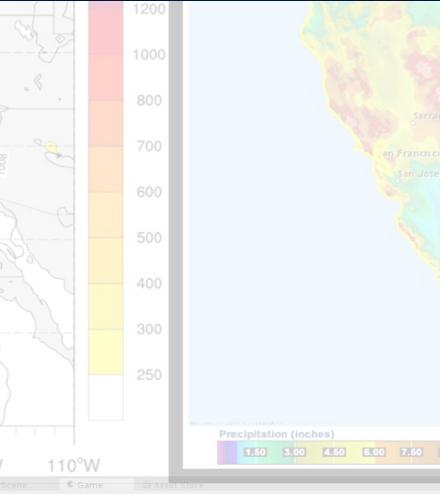


# CIROH is Unique for a NOAA Cooperative Institute

1. Dedicated to Water
2. Operations-guided Research
3. Community of Practice
4. R2O

**OWP NWC** NATIONAL  
WATER  
CENTER  
OFFICE OF  
WATER  
PREDICTION



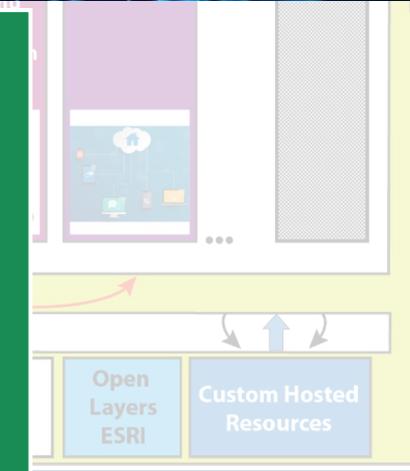


## RT1: Water Prediction Systems and Workflows

- Prediction system testbed
- Model and forecast evaluation
- Inputs, forcings, data assimilation
- Remote sensing and sensor technologies

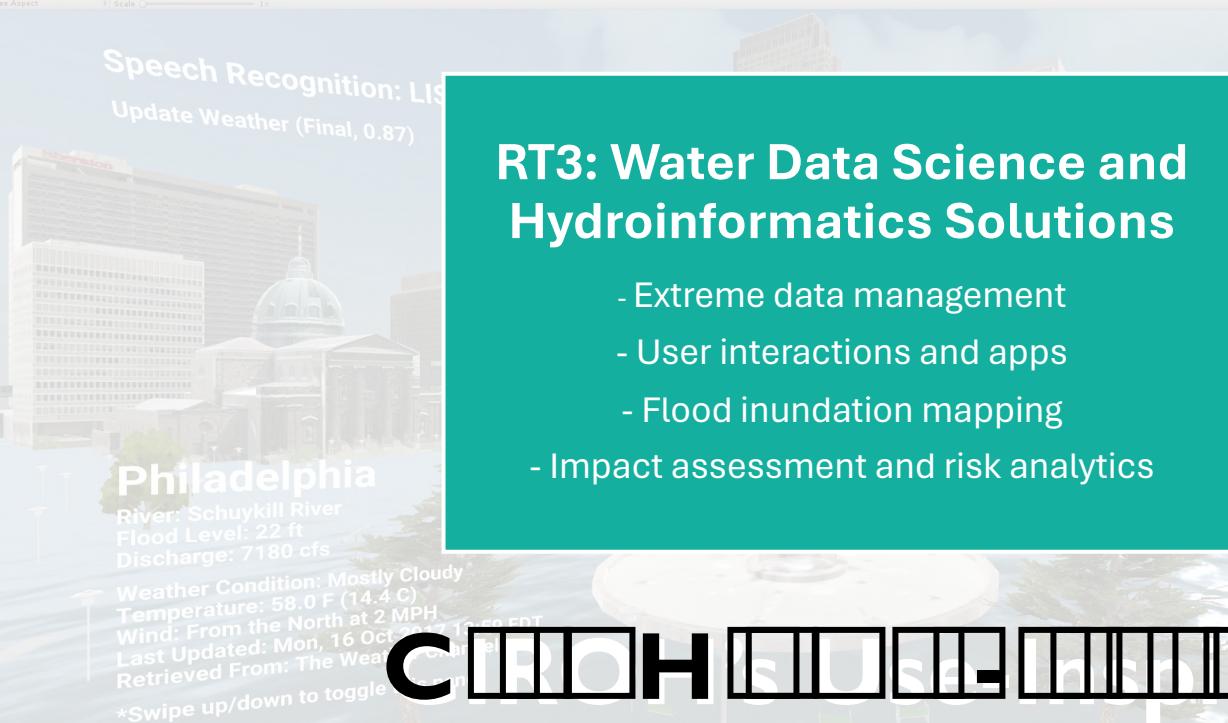
## RT2: Community Water Modeling

- Research cyberinfrastructure
- Hydrologic processes and numerics
- Large-scale hydrologic modeling
- Artificial intelligence and machine learning



## RT3: Water Data Science and Hydroinformatics Solutions

- Extreme data management
- User interactions and apps
- Flood inundation mapping
- Impact assessment and risk analytics



## RT4: Forecast Design and Community Resilience

- Risk perceptions and behaviors
- Impact-based decision science
- Improving forecasts for underserved
- Building community resilience



## Flood Prod

- Flash F
- Flood W
- Flood W
- Flood A

Wireless Emerg

18

Wednesday

Emergency  
Flash Flood  
EST. Avoid f  
NWS

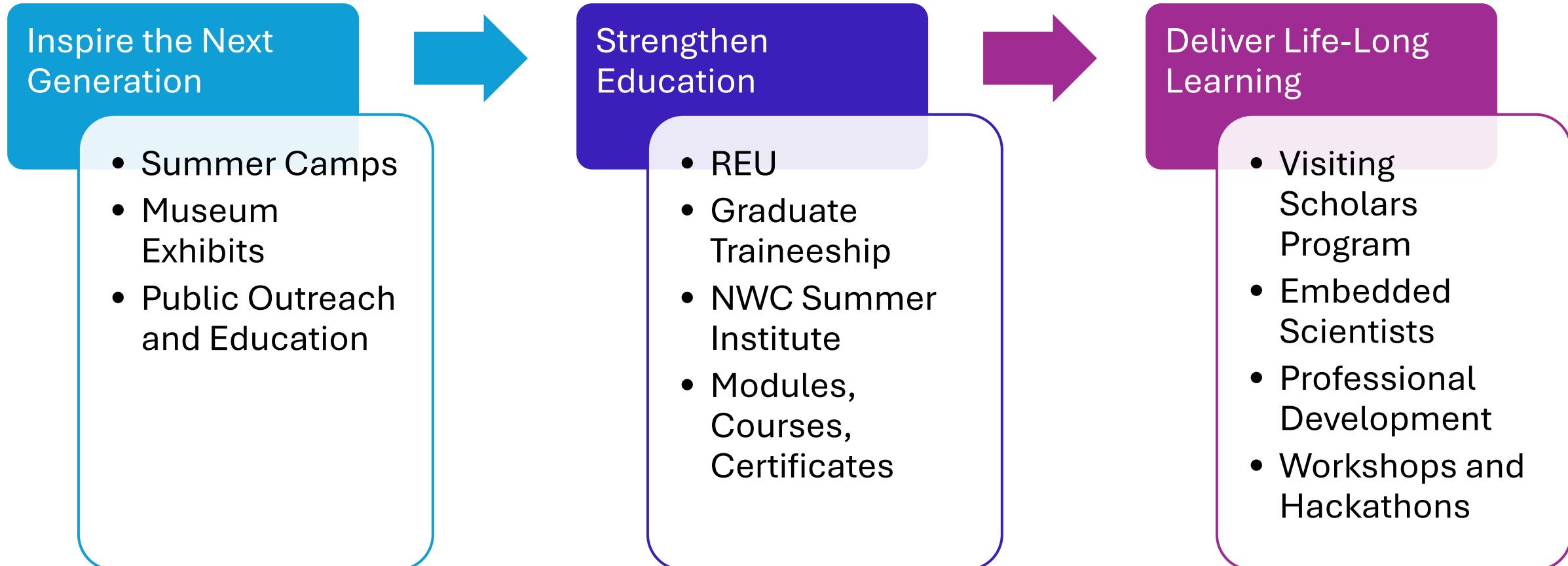
# **CIROH's Current Research Portfolio**

- 93 Projects
  - ~\$89M
  - >30 institutions
  - ~half of U.S. States  
+ Canada
  - 200+ faculty and  
research staff
  - 115+ students





# CIROH



>2000



2023



G 

 C  H



“ ”



**Monthly R2O Webinar**



**Quarterly Newsletter**

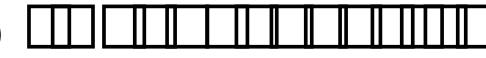


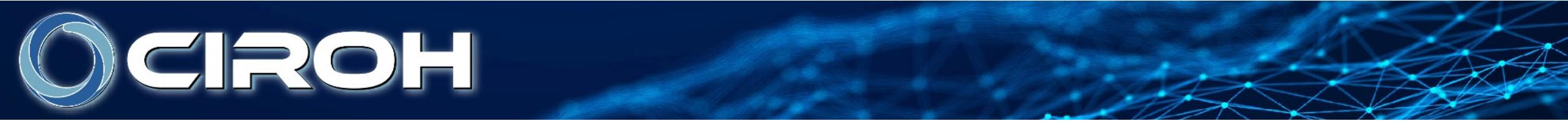
**Social Media**



**Web Site**



E  (@ 



# What has the Cyberinfrastructure team done?

NWM Data Access Acceleration

NextGen in a Box

Model Development Interface

Baseline Research Data Stream

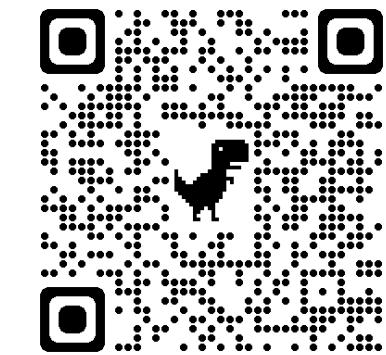
Community NextGen



# What are the options for getting NWM data?

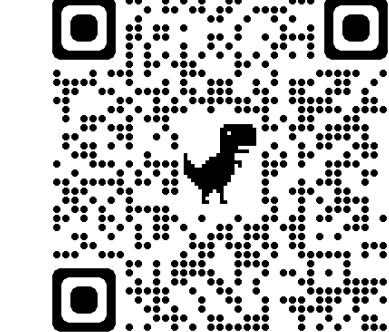
1TB + per day operational output, 100+ TB for complete archive of “retrospective” simulation

`pip install nwmur1`



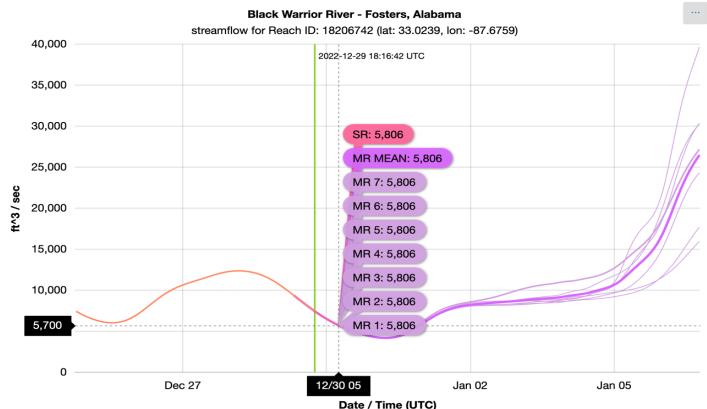
- [nomads.ncep.noaa.gov](http://nomads.ncep.noaa.gov)  
(operational output)

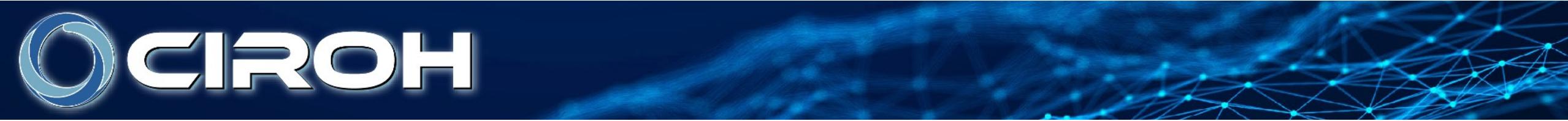
- Google Cloud forecast archive



- AWS-hosted retrospective data

- [water.noaa.gov/map](http://water.noaa.gov/map)

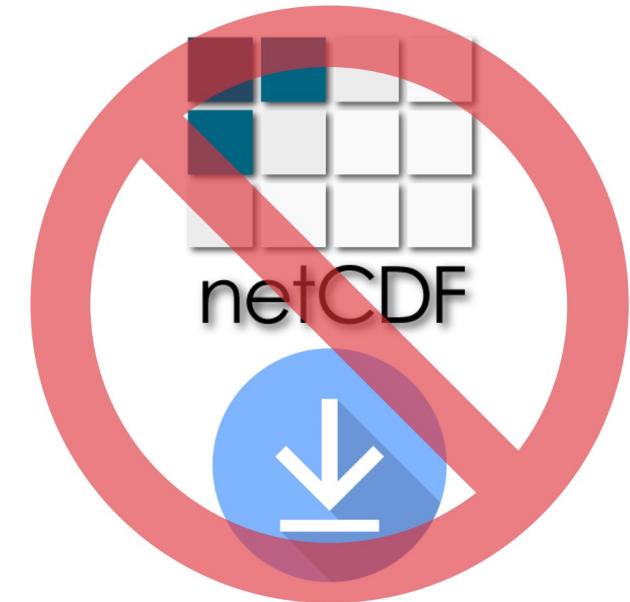




- Faster and lower bandwidth – [dataaccess.ciroh.org](http://dataaccess.ciroh.org)

Comparing Data usage: Zarr vs Native NetCDF

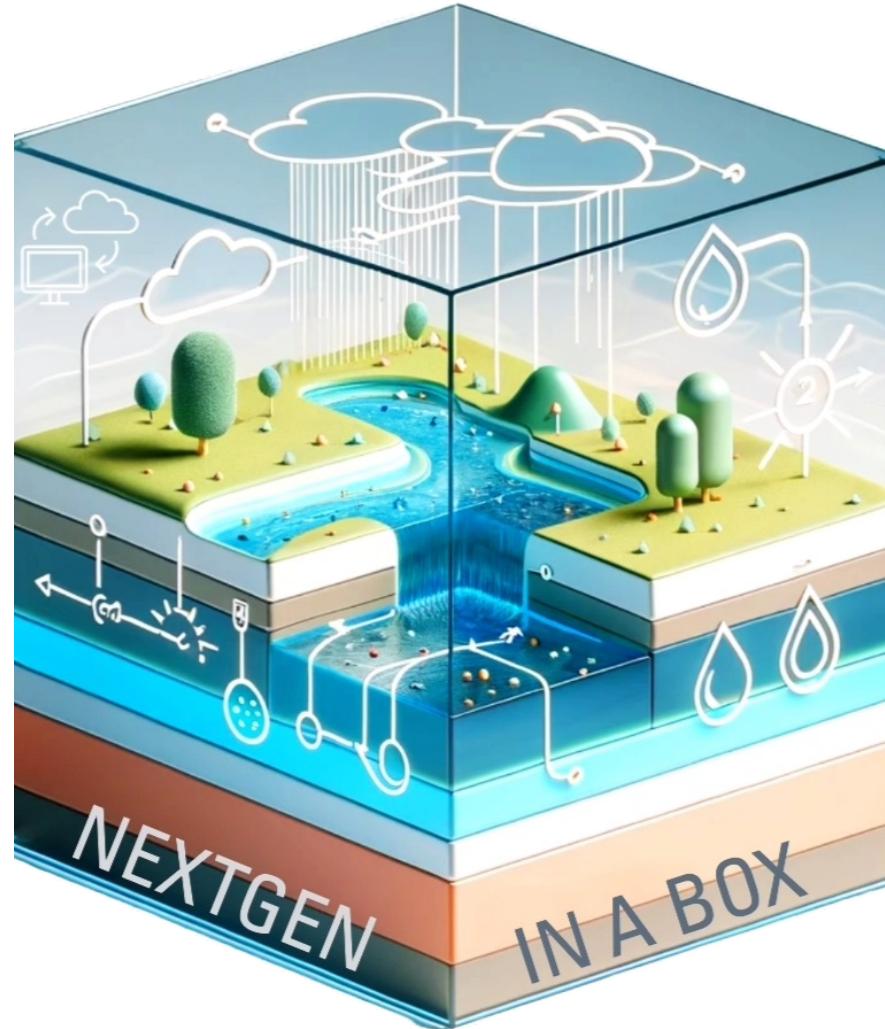
	Zarr 1 Feature ID	Zarr 12k Feature IDs	Native NetCDF
<b>Retrospective</b> <b>1 year</b>	39.2 GB	78.7 GB	395.3 GB
<b>Operational</b> <b>Medium Range</b> <b>240 hours</b>	318.7 MB	617.3 MB	3336 MB
<b>Operational</b> <b>Short Range</b> <b>18 hours</b>	24.9 MB	48.4 MB	248 MB



Compute Resource	Cloud				Local			
	Zarr Parallel	NC Parallel	Zarr Serial	NC Serial	Zarr Parallel	NC Parallel	Zarr Serial	NC Serial
Retrospective 1 year	12 m 30 s	36 m 18 s	3 h 17 m	4 h 55 m	37 m	2h 22 m	6 h 19 m	19 h 26 m
Operational Medium Range 240 hours	18.2 s	23 s	2 m 23 s	2 m 51 s	28 s	46 s	3 m 7 s	4 m 30 s
Operational Short Range 18 hours	4 s	7 s	10 s	11.9 s	5.5 s	9.4 s	13.8 s	20.6 s

# NextGen In A Box (NGIAB)

- "NextGen In a Box" (NGIAB) is a **containerized version of the NextGen** National Water Resources Modeling Framework.
- The project's key achievement is the development and distribution of NGIAB, which makes the **NextGen framework more accessible to the research community**.
- NGIAB enhances research infrastructure and promotes the integration of the NextGen Framework.
- NGIAB is a **community version of the National Water Model (NWM)**. CIROH, in collaboration with Lynker, is establishing the processes, protocols, and research infrastructure that bridges the academic and research community with NOAA's National Water Center's NextGen National Water Resources Modeling Framework (NextGen Framework).
- It is a community version, allowing the NextGen framework to run on local machines for watersheds of any size.

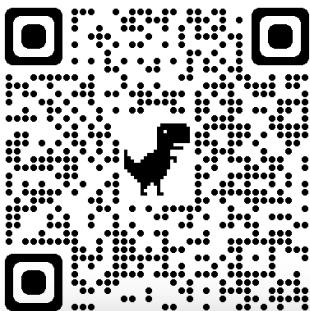


1. Clone
2. Run
3. Repeat



# NextGen Model preprocessor

From selecting a basin to a baseline  
NextGen-based hydrologic  
simulation in a few minutes.



1. Clone
2. Run
3. Repeat

DON'T RUN CFE IN SNOWY PLACES WITHOUT A SNOW MODEL

it will return garbage and you will deserve it

NextGen-In-A-Box Data Preparation Tool (run a NextGen-based simulation in just a few clicks!)

Interactive Map

**1**

Selected Basins  
wb-410618 Select outlet basin from latest hydrofabric:  
<https://noaa-owp.github.io/hydrofabric/>  
tributary watersheds determined from topology

**2**

Catchments Successfully Subset!  
Subset

**3**

Start Time: 01 / 01 / 2010, 00 : 00  
End Time: 02 / 01 / 2010, 00 : 00

**4**

Create Forcing From Zarrs Warning: This will take a while  
Retrospective v3.0 Range: Feb 1979 – Jan 2023

**5**

Create Realization  
Generates CFE basic simulation; PET and Snow coming soon!

Realization.json Configuration Generated!

**6**

\$ git clone git@github.com:CIROH-UA/NGIAB-CloudInfra.git  
./NGIAB-CloudInfra/guide.sh

Point your NGIAB instance directly at the data package from the data preparation tool

Latest observed value: 22 kcfs  
0:30 PM CST 15-Feb-2024  
Flood Stage is 129 ft

Black Warrior River at Oliver Lock and Dam  
NWSU: TODA1, Reach ID: 18229923

Model Guidance  
Not Official Forecast

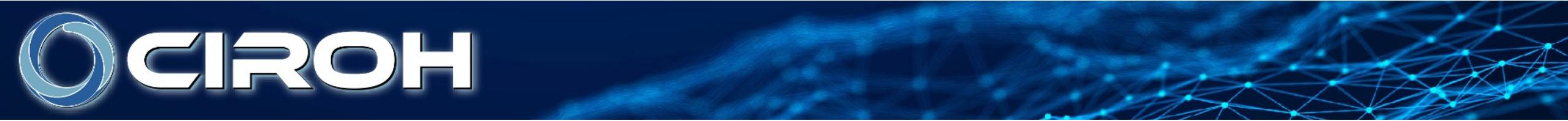
Do hydrology!

US DEPARTMENT OF COMMERCE  
NOAA  
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Site Time (CST)  
TODA1 (plotting HTGR) "Gage 0" Datum (N/A): 0'

Zoom 1d 2d 7d 14d All

Category Stage



# What is the Cyberinfrastructure team doing?

NWM Data Access Acceleration

NextGen in a Box

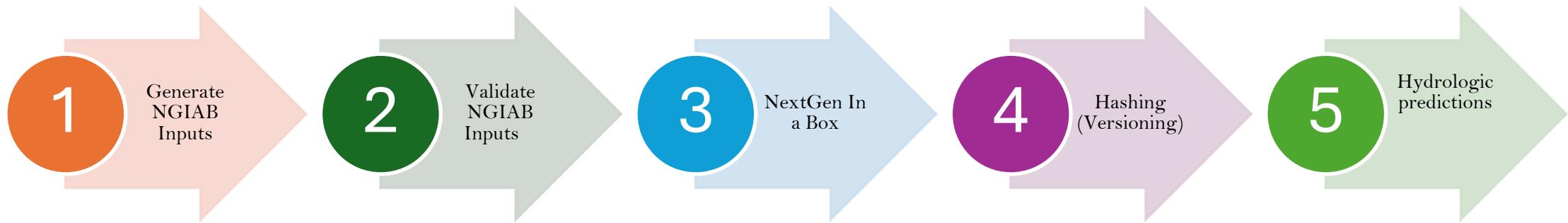
Model Development Interface

Baseline Research Data Stream

Community NextGen



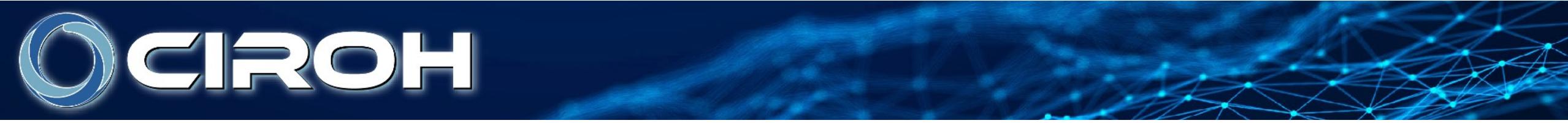
# NextGen Data Stream



Ngen-datastream refers to the software chain that builds valid ngen-run input packages, executes NEXTGEN with NGIAB, and versions the entire run.

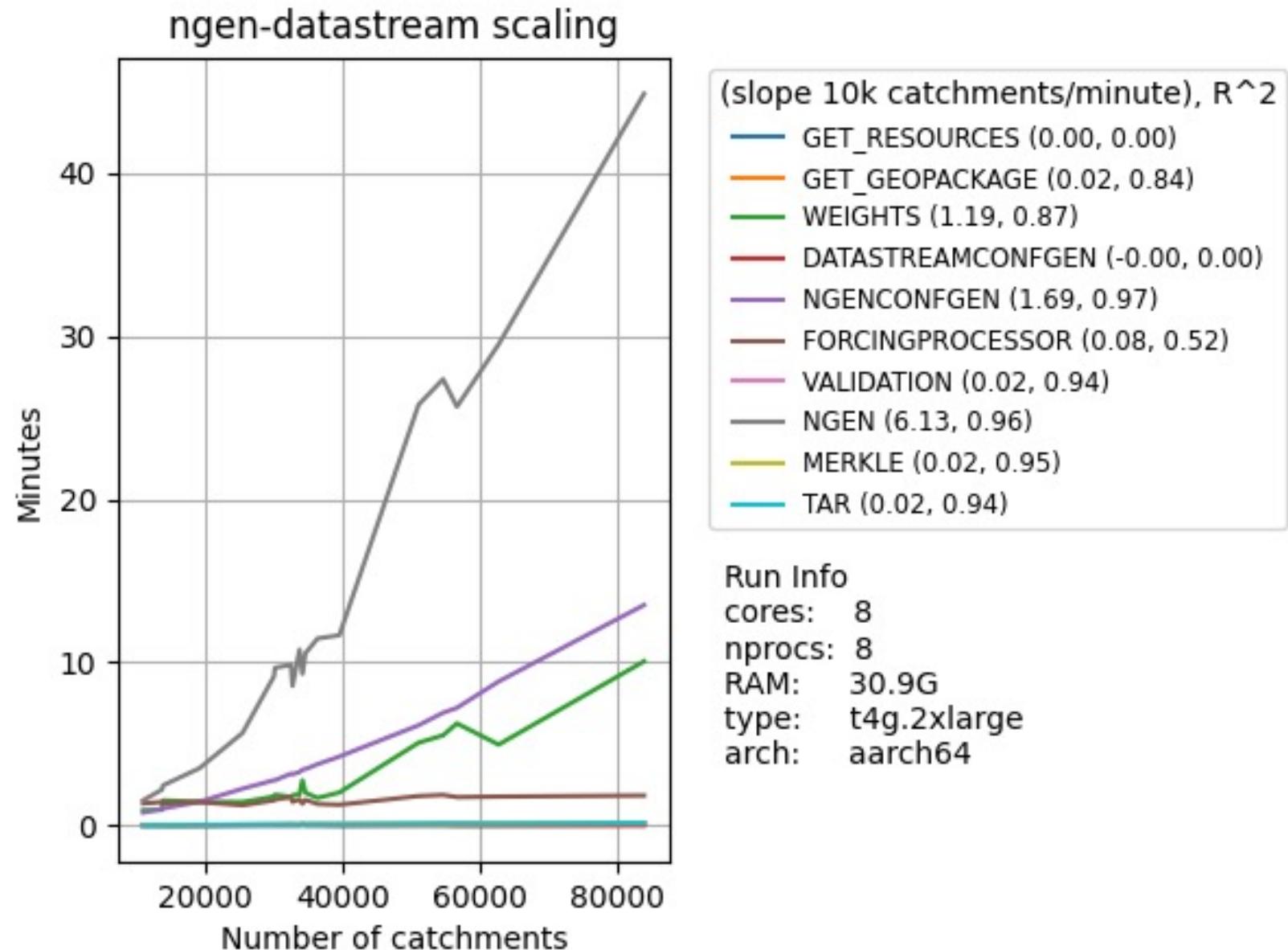
<https://github.com/CIROH-UA/ngen-datastream>

Working on integration with hydrologic evaluation and FIM output and evaluation.



Where we are  
right now...

Coming soon:  
*<URL with  
datastream  
data>*





# What comes next for CIROH Cyberinfrastructure?

NWM Data Access Acceleration

NextGen in a Box

Model Development Interface

Baseline Research Data Stream

Community NextGen



# Cunningham's Law: Don't like it? Go fix it!

- Democratized participation in the improving and updating of the modeling of the nation's water.
- By Next Year DevCon: Run the baseline, Make a change, Submit the update, See the change in the Research Data Stream.



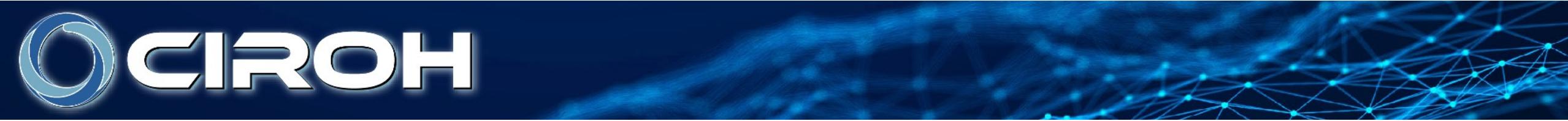
**CIROH wants YOU  
to Improve the  
National Water Model**

In version 4.0, slated for release in 2026, the U.S. National Water Model (NWM) will be transitioning to the NextGen Water Resources Modeling Framework. Engage with CIROH – the Cooperative Institute for Research to Operations in Hydrology – to make a difference in water forecasting for your basin.



Come use our CIROH community versions of the NextGen tools to get started:

- NextGen in a Box (NGIAB): NextGen prepackaged execution environment.
- NGIAB Data Preparation Tool: Easy map-based tool to create CFE-based NextGen realizations with all necessary inputs.
- NextGen Datastream: Benchmark dataset for demonstrating model improvements. (Coming Soon!)



# Questions?

## CIROH Jupyter Hub

<https://ciroh.awi.2i2c.cloud/>



Welcome to the Cooperative Institute  
for Research to Operations in  
Hydrology **2i2c JupyterHub**.

This is a pilot service running on open source  
infrastructure. See [the 2i2c Pilot documentation](#) for  
usage and deployment information.



## CIROH DocuHub

<https://docs.ciroh.org/>

