Git-hub Calib

ciroh.ua.edu

CIROH Developers Conference [Community NextGen Workshop] Calibration in the NGIAB Ecosystem

Sifan A. Koriche, Josh Cunningham, Md Shahabul Alam, Xia Feng and James Halgren

Alabama Water Institute, The University of Alabama

May 29, 2025

Cooperative Institute for Research to Operations in Hydrology

Sifan A. Koriche

- Research [Hydrologic] Scientist
- The University of Alabama

Josh Cunningham

- Computer Scientist
- The University of Alabama

Md Shahabul Alam

- Research Scientist
- The University of Alabama

Xia Feng

- Physical Scientist
- The University of Alabama,
- NOAA Affiliate

CIROH
Cooperative Institute for Research to Operations in Hydrology

Git-hub Calib

THE UNIVERSITY OF ALABAMA | Alabama Water Institute

NWC | National Water Center

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

OWP | Office of Water Prediction

THE UNIVERSITY OF ALABAMA Alabama Water Institute — Research What Matters Most.

Lynker

Acknowledgment

This workshop is supported by the Cooperative Institute for Research to Operations in Hydrology (CIROH) with funding from the NOAA Cooperative Institute Program. The tools, workflows, and concepts presented in this workshop build upon the significant efforts and contributions of several key groups and individuals. We gratefully acknowledge:

- The development of the `NextGen In A Box (NGIAB) ecosystem` is a central initiative of `CIROH-UA`, with key development contributions from the `AWI Science and Technology Team`.
- The core `NGEN-Cal tool`, fundamental to the calibration processes showcased, was developed by the `NOAA Office of Water Prediction (NOAA-OWP)` in collaboration with `Lynker`.
- The `NGEN-Cal calibration workflow`, initially developed by `Xia Feng`, was subsequently refactored and integrated to operate within the NGIAB ecosystem by `Josh Cunningham`, with crucial domain science contributions from `Sifan A. Koriche, Md Shahabul Alam, and James Halgren`.
- This `GitHub repository`, which serves as the central hub for all workshop materials, tools, detailed instructions, pre-conference and other relevant information, was prepared and is maintained by `Sifan A. Koriche and Josh Cunningham`.

The statements, findings, conclusions, and recommendations presented in this workshop and its associated materials are those of the author(s) and do not necessarily reflect the official opinions of `NOAA`, `CIROH`, or the other acknowledged contributing organizations and individuals.

Community NextGen UA Team


Arpita Patel


James Halgren


Benjamin Lee

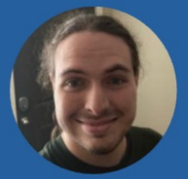

Josh Cunningham


Trupesh Patel


Shahab Alam


Sifan Koriche


Jonathan Frame


Chad Perry


Xia Feng


Manjila Singh


Nia Minor


Savalan Neisary


Quinn Lee


Hari Jajula

CIROH
Cooperative Institute for Research to Operations in Hydrology

CIROH
Cooperative Institute for Research to Operations in Hydrology

Community NextGen Partner Team


Nels Frazier(Lynker)


Zach Wills(Lynker)


Keith Jennings (UVM)


Jordan Laser(Lynker)


Mike Johnson(Lynker)


Josh Sturtevant (Lynker)


Matthew Denno (RTI)


Sam Lamont (RTI)

Giovanni Romero (Aquaveo)


Dan Ames (BYU)

CIROH
Cooperative Institute for Research to Operations in Hydrology

Community NextGen Partner Team


David Tarboton (USU)


Furqan Baig (USU)


Pabitra Dash (USU)


Ayman Nassar (USU)


Anthony Castronova (CUAHSI)


Irene Garousi-Nejad (CUAHSI)

CIROH
Cooperative Institute for Research to Operations in Hydrology

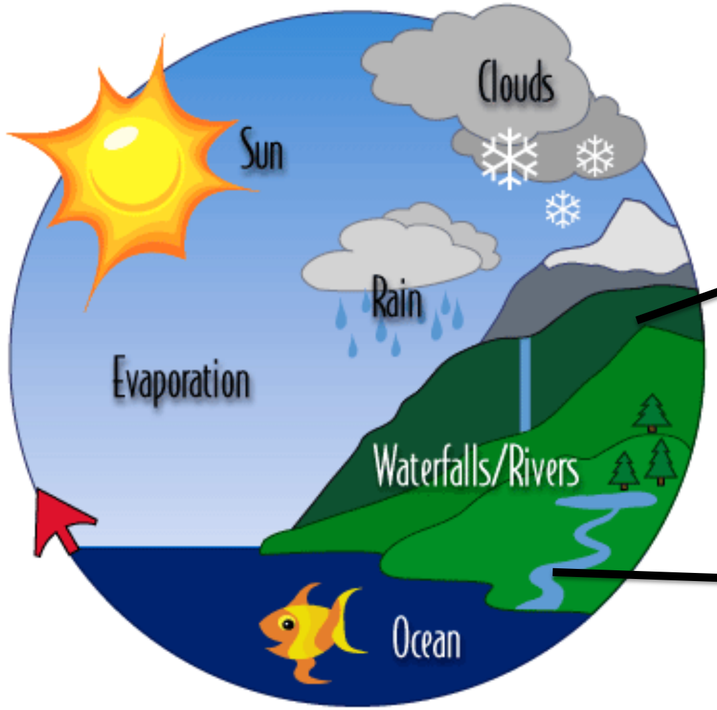
Earth system processes

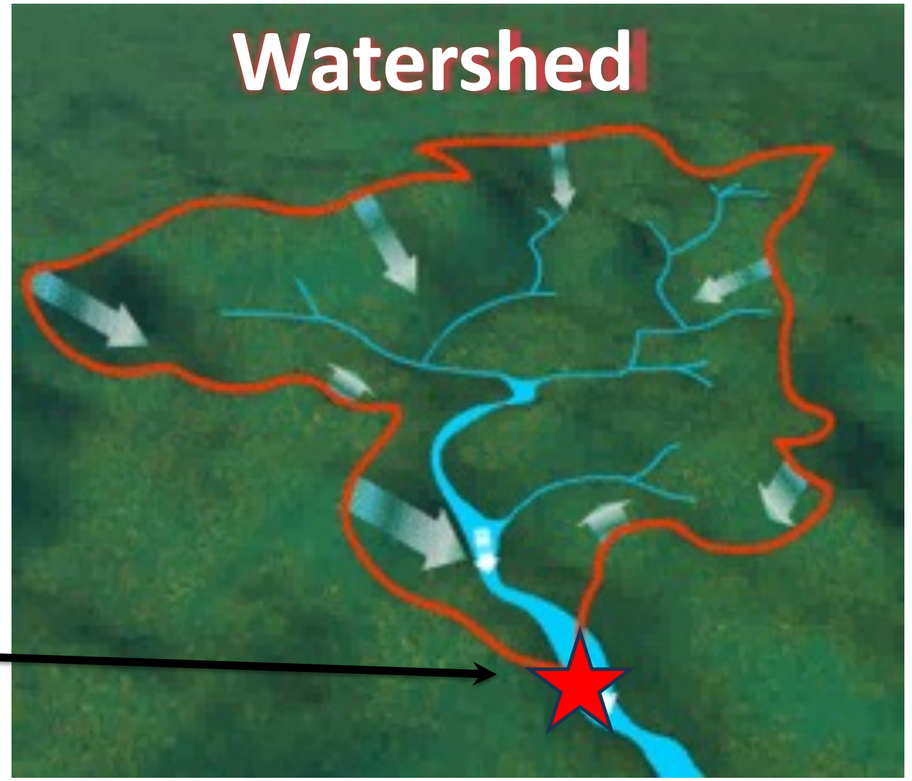
Watershed hydrology

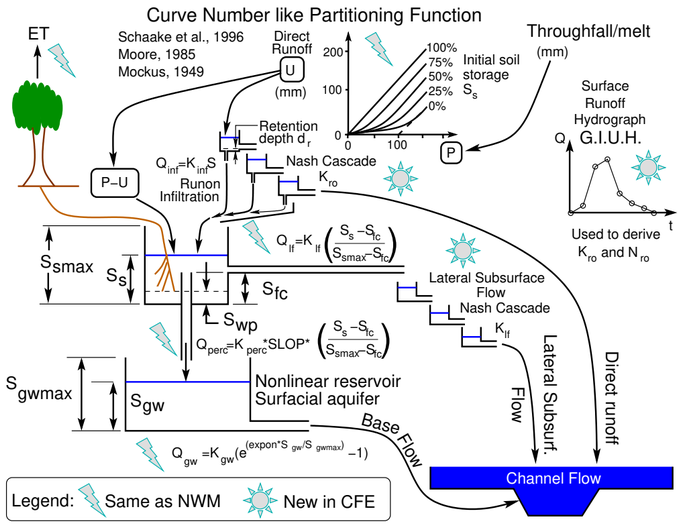
Streamflow

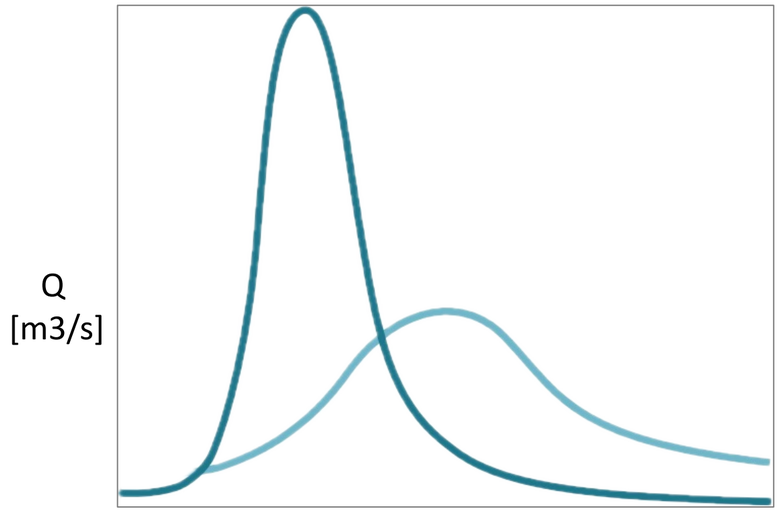
CIROH
Cooperative Institute for Research to Operations in Hydrology

Water Cycle

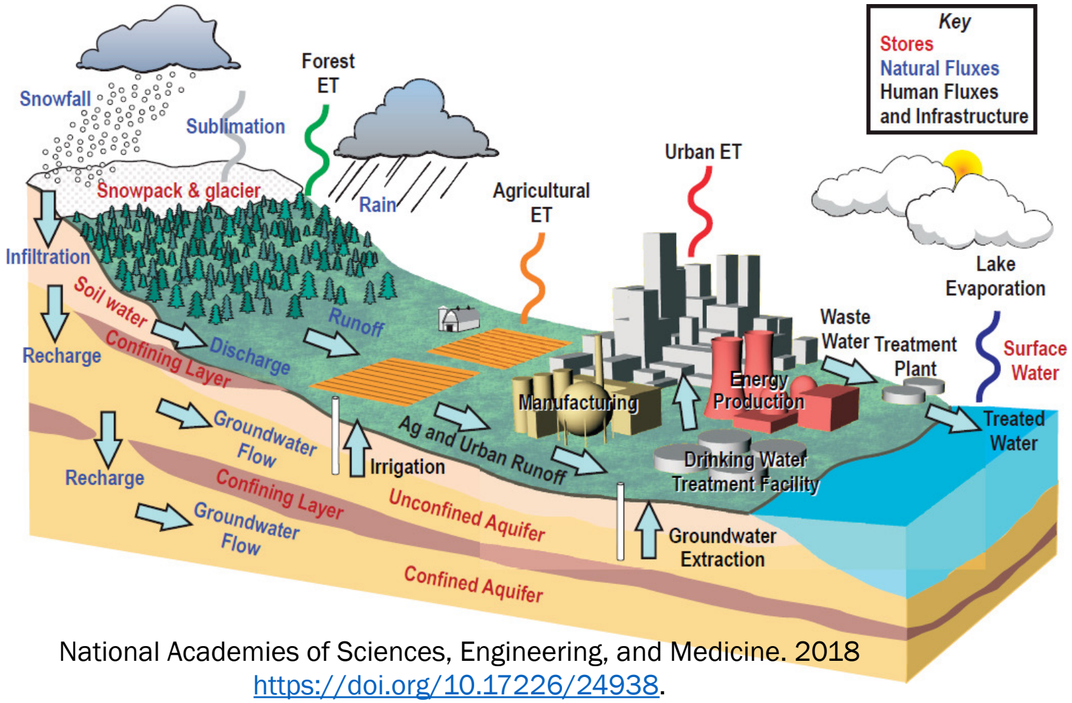








Hydrological processes



National Academies of Sciences, Engineering, and Medicine. 2018 https://doi.org/10.17226/24938.

Simplification of the Natural System

- Conceptual [...]
- Physical [...]
- ML/AI [...]
- Hybrid [...]

parameterization

Hydrological Models

- State/Prognostic Variables → Define system dynamics
- Parameters → Quantities characterizing the system.

Calibration

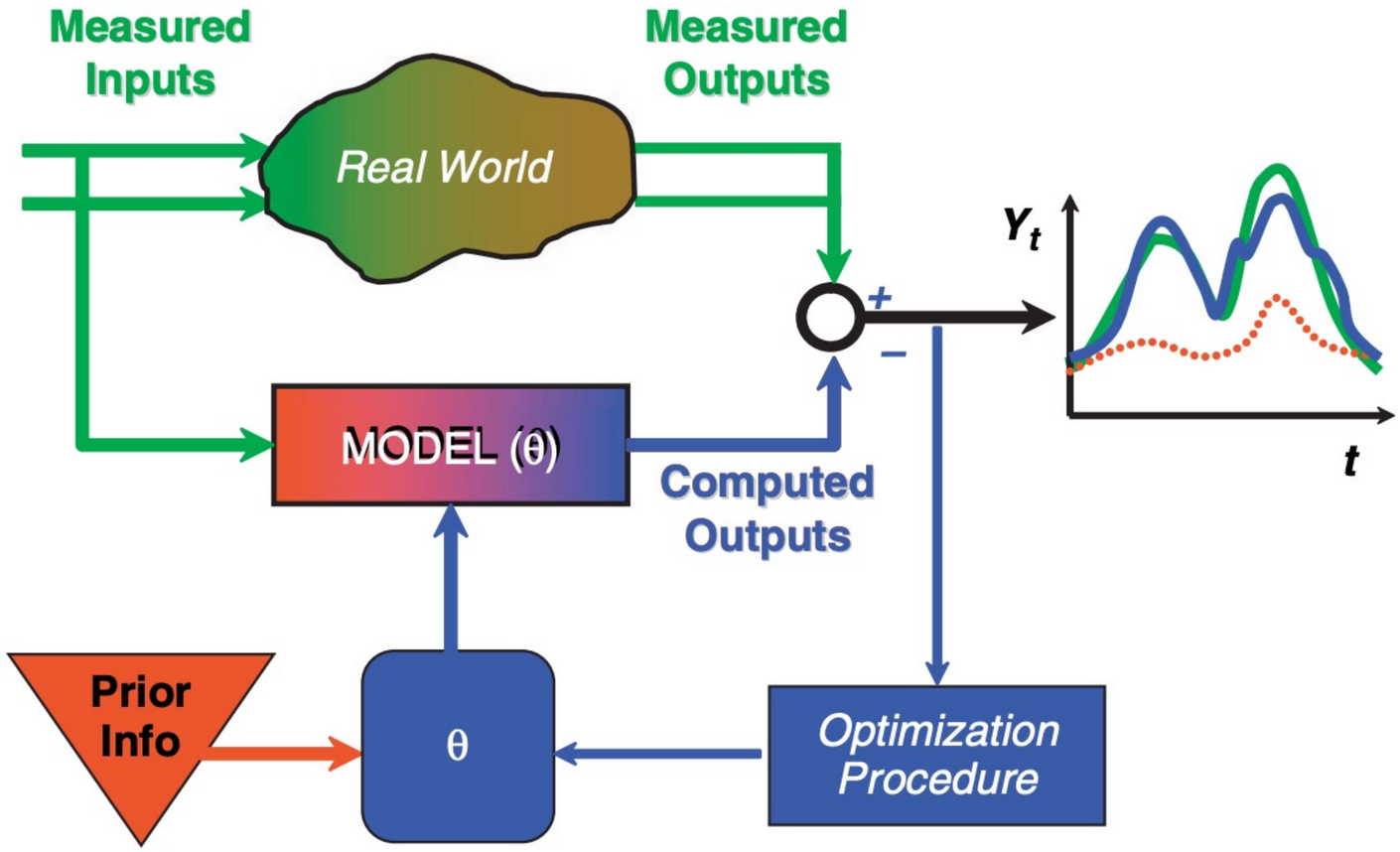
Parameter Types:

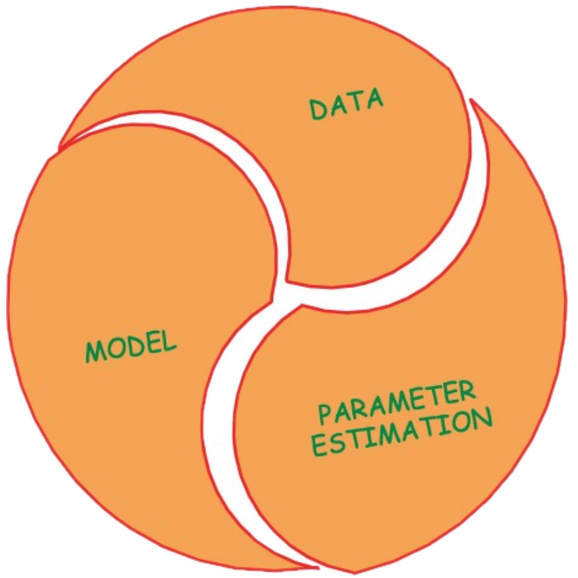
- Physical Parameters → Directly measurable, independent of basin response. E.g., Watershed area, impervious area, local permeability, vegetated area fraction, areal percentage of water bodies.
- Process Parameters → Quantities characterizing the system. E.g., Effective soil moisture storage depth, effective lateral interflow, drainage rates for lumped storages, mean hydraulic conductivity, surface runoff coefficient.

CIROH
Cooperative Institute for Research to Operations in Hydrology

Calibration

Calibrating a hydrologic model is the process of modifying parameters within acceptable ranges to obtain simulated results that replicate known conditions.





The concept of model calibration as an optimization problem (Moradkhani and Sorooshian, 2008)

CIROH
Cooperative Institute for Research to Operations in Hydrology

Calibration

Methods | Types | Algorithms

Manual Calibration Method	Trial-and-error, guided visual tuning (expert)	• single-objective optimization • deterministic	Utilizes expert knowledge; intuitive but slow, subjective, and often impractical for complex models
Automatic Calibration Methods	Shuffled Complex Evolution of the University of Arizona (SCE-UA)	• single-objective optimization • adapted for multi-objective optimization	• Employs the Concept of "Complexes" and "Shuffling" • Combines Deterministic [local search - Simplex] and Probabilistic [global search - randomized] Approaches
	Genetic Algorithm (GA)	Globally optimal or near-optimal parameter sets	• Heuristic optimization algorithms based on the evolutionary process of natural selection and genetics (Zhang et al., 2023; Seibert, 2000).
	Dynamically Dimensioned Search (DDS)	• single-objective optimization. • near-optimal parameter sets	• stochastically perturbs a dynamically changing number of parameters at each step to explore the solution space. | Not Guaranteed Global Optimum
	Particle Swarm Optimization (PSO)	• Typically, Single-Objective (but can be adapted) • Global Optimization	• population-based stochastic optimization technique • uses a swarm of particles that explore the search space [each particle's movement influenced by its own best-found position and the best-found position of the entire swarm]
	Generalized Likelihood Uncertainty Estimation (GLUE), Gauss-Newton-type methods, Artificial Bee Colony (ABC), Invasive Weed (IWA), Harmony Search (HS), Artificial Neural Networks, Differential Evolution (DE), … etc		

Calibration

DDS

PSO

Dynamically Dimensioned Search:

- The algorithm **searches globally at the start of the search and becomes a more local search** as the number of iterations approaches the maximum allowable number of function evaluations.

STEP 1. Define DDS inputs:
- neighborhood perturbation size parameter, r (0.2 is default)
- maximum # of function evaluations, m
- vectors of lower, $\boldsymbol{x}^{min}$, and upper, $\boldsymbol{x}^{max}$, bounds for all D decision variables
- initial solution, $\boldsymbol{x}^0 = [x_1, \ldots, x_D]$

STEP 2. Set counter to 1, i = 1, and evaluate objective function F at initial solution, $F(\boldsymbol{x}^0)$:
- $F_{best} = F(\boldsymbol{x}^0)$, and $\boldsymbol{x}^{best} = \boldsymbol{x}^0$

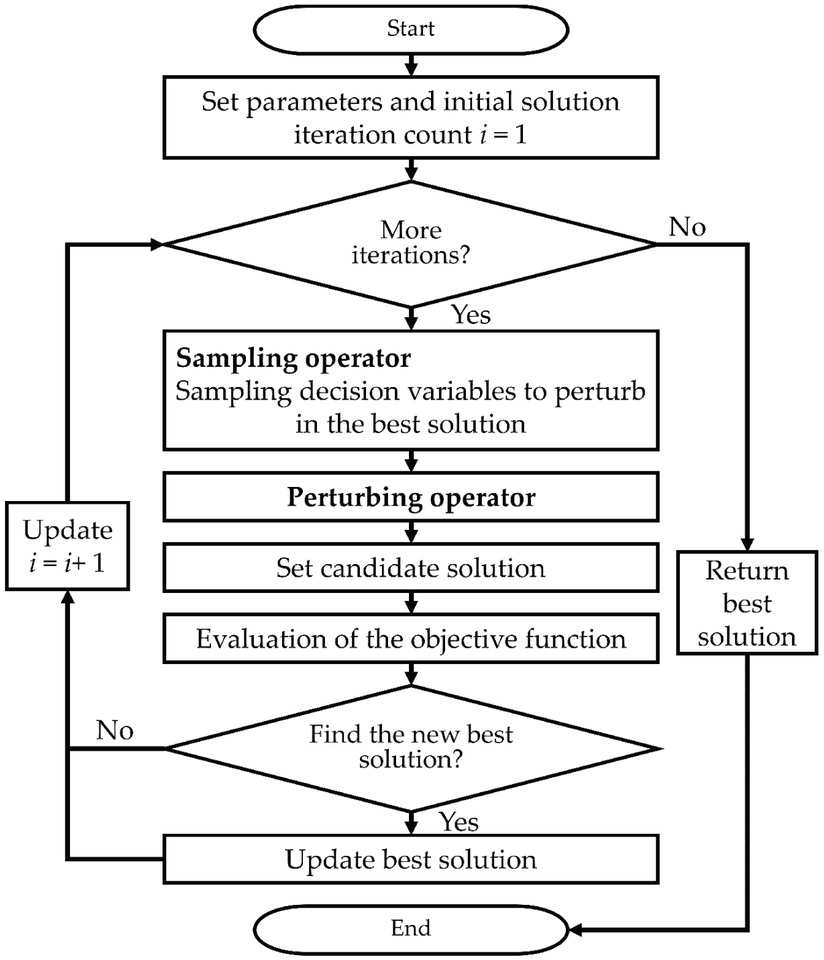
STEP 3. Randomly select J of the D decision variables for inclusion in neighborhood, $\{N\}$:
- calculate probability each decision variable is included in $\{N\}$ as a function of the current iteration count: $P(i) = 1 - \ln(i)/\ln(m)$
- FOR d = 1, … D decision variables, add d to $\{N\}$ with probability P
- IF $\{N\}$ empty, select one random d for $\{N\}$

STEP 4. FOR j = 1, …, J decision variables in $\{N\}$, perturb x_j^{best} using a standard normal random variable, N(0,1), reflecting at decision variable bounds if necessary:
- $x_j^{new} = x_j^{best} + \sigma_j N(0,1)$, where $\sigma_j = r(x_j^{max} - x_j^{min})$
- IF $x_j^{new} < x_j^{min}$, reflect perturbation:
  - $x_j^{new} = x_j^{min} + (x_j^{min} - x_j^{new})$
  - IF $x_j^{new} > x_j^{max}$, set $x_j^{new} = x_j^{min}$
- IF $x_j^{new} > x_j^{max}$, reflect perturbation:
  - $x_j^{new} = x_j^{max} - (x_j^{new} - x_j^{max})$
  - IF $x_j^{new} < x_j^{min}$, set $x_j^{new} = x_j^{max}$

STEP 5. Evaluate $F(\boldsymbol{x}^{new})$ and update current best solution if necessary:
- IF $F(\boldsymbol{x}^{new}) \le F_{best}$, update new best solution:
  - $F_{best} = F(\boldsymbol{x}^{new})$ and $\boldsymbol{x}^{best} = \boldsymbol{x}^{new}$

STEP 6. Update iteration count, i = i+1, and check stopping criterion:
- IF i = m, STOP, print output (e.g. F_{best} & $\boldsymbol{x}^{best}$)
- ELSE go to STEP 3

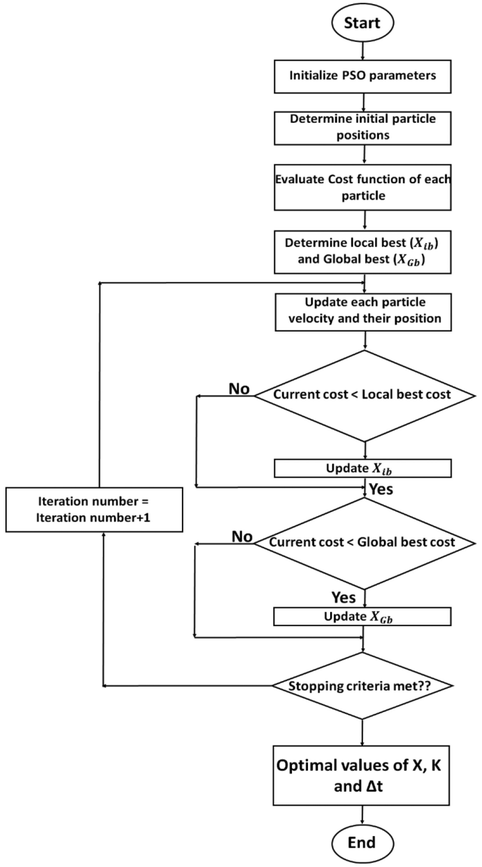
Tolson and Shoemaker et al., 2007



Jin et al., 2022

Particle swarm optimization (PSO):

- uses a swarm of particles that explore the search space, with **each particle's movement influenced by its own best-found position and the best-found position of the entire swarm** (or its neighborhood).

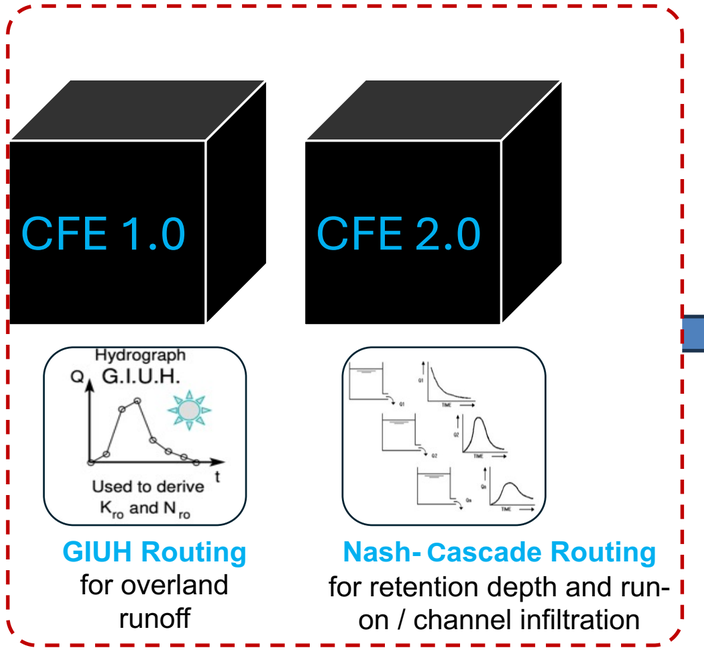


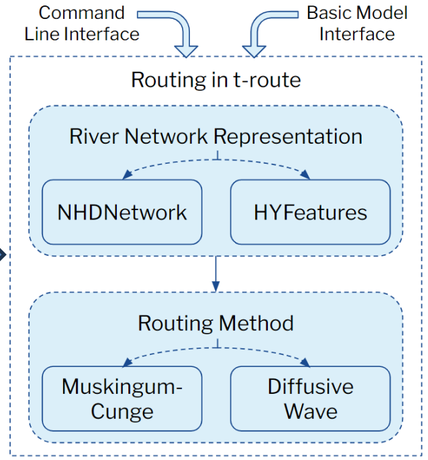
Norouzi and Bazargan, 2020

CIROH
Cooperative Institute for Research to Operations in Hydrology

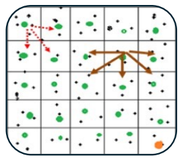
Conceptual Model [NextGen Framework]

https://github.com/NOAA-OWP/ngen
https://github.com/NOAA-OWP/ngen-cal
https://github.com/NOAA-OWP/cfe
https://github.com/NOAA-OWP/t-route
https://github.com/NOAA-OWP/topmodel
https://github.com/NOAA-OWP/evapotranspiration
https://github.com/NOAA-OWP/SoilFreezeThaw
https://github.com/NOAA-OWP/noah-owp-modular
https://github.com/NOAA-OWP/sac-sma

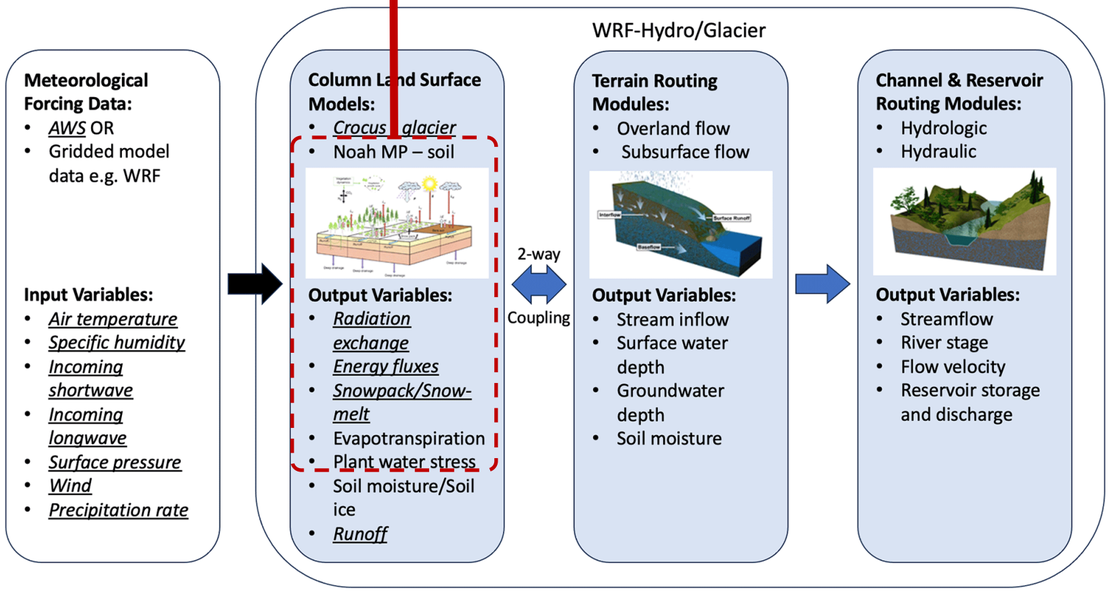




WRF-Hydro [NWM]



Grid-based Routing
to for surface and subsurface runoff

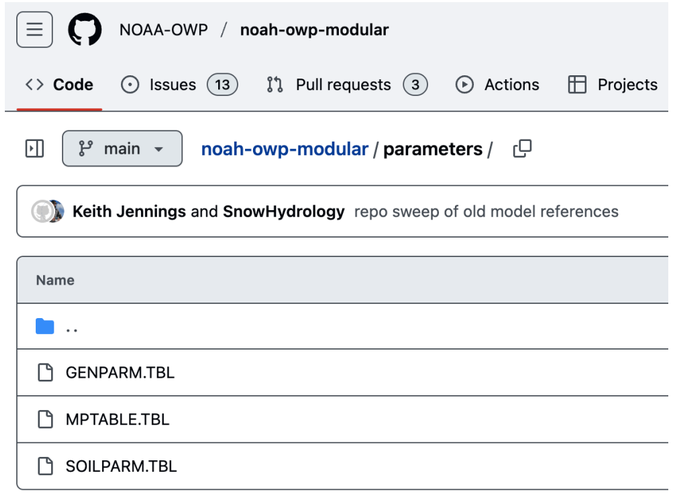


CIROH
Cooperative Institute for Research to Operations in Hydrology

CFE and Noah-OWP parameters

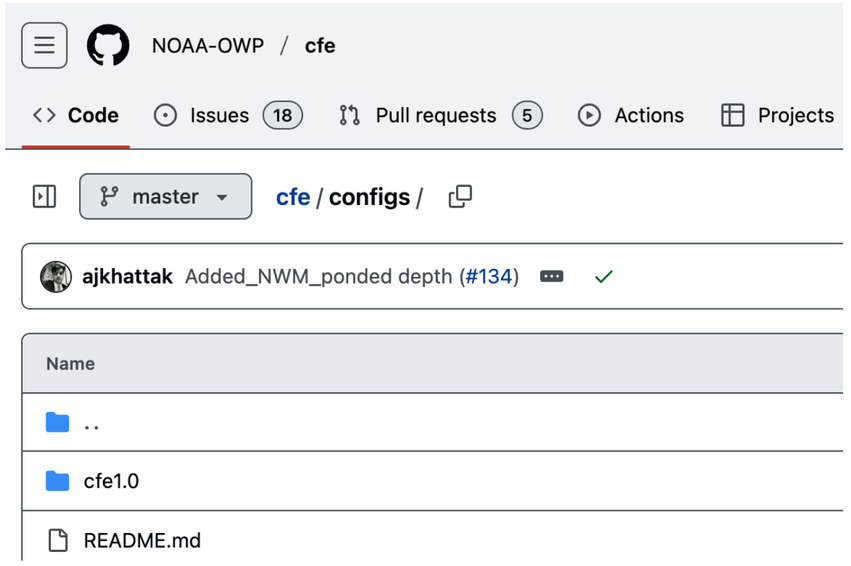
```yaml
NoahOWP: &id002
  - name: RSURF_EXP
    min: 1.0
    max: 6.0
    init: 5.0
  - name: CWP
    min: 0.09
    max: 0.36
    init: 0.18
  - name: MP
    min: 3.6
    max: 12.6
    init: 9.0
  - name: VCMX25
    min: 24.0
    max: 112.0
    init: 52.2
  - name: MFSNO
    min: 0.5
    max: 4.0
    init: 2.0
  - name: RSURF_SNOW
    min: 0.136
    max: 100.0
    init: 50.0
  - name: SCAMAX
    min: 0.7
    max: 1.0
    init: 0.9
```

Noah-owp-Modular parameters




```yaml
CFE: &id001
  - name: b
    min: 2.0
    max: 15.0
    init: 4.05
  - name: satpsi
    min: 0.03
    max: 0.955
    init: 0.355
  - name: satdk
    min: 1.0e-07
    max: 0.000726
    init: 3.38e-06
  - name: maxsmc
    min: 0.16
    max: 0.59
    init: 0.439
  - name: refkdt
    min: 0.1
    max: 4.0
    init: 1.0
  - name: slope
    min: 0.0
    max: 1.0
    init: 0.1
  - name: max_gw_storage
    min: 0.01
    max: 0.25
    init: 0.05
  - name: expon
    min: 1.0
    max: 8.0
    init: 3.0
  - name: Cgw
    min: 1.8e-06
    max: 0.0018
    init: 1.8e-05
  - name: Klf
    min: 0.0
    max: 1.0
    init: 0.01
  - name: Kn
    min: 0.0
    max: 1.0
    init: 0.03
NoahOWP: &id002
  - name: RSURF_EXP
    min: 1.0
    max: 6.0
    init: 5.0
  - name: CWP
    min: 0.09
```

CFE parameters



CIROH
Cooperative Institute for Research to Operations in Hydrology

Ngen-Cal:

Configuration Options

General

- contains model agnostic configuration of the software.

```yaml
#An exampmle ngen-calibration configuration file
general:
  # Strategy configuration
  strategy:
      # Type of strategey, currently supported is estimation
      type: estimation
      #defaults to dds
      algorithm: "dds"
      # To adjuts the neighborhood size parameter of the dds algorithm, uncomment the following two lines
      #parameters:
      #  neighborhood: 0.5
      # To use PSO optmization, select the pso algorithm and configure its parameters as follows
      #algorithm: "pso"
      #parameters:
      #    pool: 4 #number of processors to use (by default, uses 1)
      #    particles: 8 #number of particles to use (by default, uses 4)
      #    options: #the PSO parameters (defaults to c1: 0.5, c2: 0.3, w:0.9)
      #        c1: 0.1
      #        c2: 0.1
      #        w: 0.42

  #iteration control
  # In general, the start iteration should probably always be 0
  # if `restart: true`, this is overridden by the detected existing iterations saved
  # otherwise, this can have some odd consequences, and will likely be removed from the
  # configuration support in the near future
  start_iteration: 0
  # The total number of search iterations to run
  # Note that if using `restart: true`, you shouldn't change the number of total iterations
  # though it may be tempting to use this to extend a previous run, some algorithms are sentstive
  # to the total number of iterations, and changing it unexpectedly across restarts may lead
  # to undesirable results
  iterations: 100
```

Parameter references

```yaml
#Describe the model parameters you want to use, valid for independent and uniform
#ngen strategies
cfe_params: &cfe_params
  -
      name: maxsmc
      min: 0.2
      max: 1.0
      init: 0.439
  -
      name: satdk
      min: 0.0
      max: 0.000726
      init: 3.38e-06
  -
      name: slope
      min: 0.0
      max: 1.0
      init: 0.01
```

Model definition [Model specific configuration]

```yaml
#Model specifc configuration
model:
    # Which model to execute for the search optimization
    # Currently only support `ngen` and `none` (for testing purposes)
    type: ngen
    # A binary in $PATH or a qualified path to the binary to run
    binary: "ngen"

    # Required path to ngen realization config (with calibration info included)
    realization: ../cfe-cal-2/realization_config.json
    # Required path to catchment hydrofabirc file
    catchments: ../cfe-cal-2/hydrofabric/catchment_data.geojson
    # Required path to nexus hydrofabric file
    nexus: ../cfe-cal-2/hydrofabric/nexus_data.geojson
    # Required path to hydrofabric crosswalk file
    crosswalk: ../cfe-cal-2/hydrofabric/crosswalk.json
    #ngen calibration strategies include
    #uniform: Each catchment shares the same parameter space, evaluates at one observable nexus
    #independet: Each catchment upstream of observable nexus gets its own permuated parameter space, evaluta
    #explicit: only calibrates basins in the realization_config with a "calibration" definition and an obse
    strategy: independent
    params:
        CFE: *cfe_params

    eval_params:
      # This is the range of the hydrograph dates to run the objective function over
      # To evaluate the entire period, you can comment these lines out
      #evaluation_start: '2015-12-15 12:00:00'
      #evaluation_stop: '2015-12-30 23:00:00'
      # choices are "kling_gupta", "nnse", "custom", "single_peak", "volume"
      objective: "kling_gupta"
      # one can also provide a module path to any function that takes
      # obs, sim array-like arguments and produces a single value float
      # for example, nnse above could be called this way
      #objective: "ngen_cal.objectives.normalized_nash_sutcliffe"
      # Can choose to minimize the objective function or maximixe it (only when using the DDS algorithm)
      # choices are 'min', 'max'.
      # An an explicit floating point value can be supplied instead, and the
      # optmization will attempt to converge on that value
      # Default: min
      #target: 0.0
```

https://github.com/NOAA-OWP/ngen-cal/wiki

CIROH
Cooperative Institute for Research to Operations in Hydrology

CIROH DevCon 2025 Workshop: Hydrological Model Calibration in the NGIAB Ecosystem

Workshop CIROH DevCon 2025 NextGen Framework Docker Ready

"How do I calibrate, regionalize and share the calibrated parameters in a NextGen Framework ecosystem"

- 📅 **Date**: Thursday, May 29, 2025
- ⏰ **Time**: 1:30 PM – 3:00 PM (US Eastern Time)
- 📍 **Location**: UVM Dudley H. Davis Center – Level 4, 590 Main Street, Burlington, VT 05405
- 🏛️ **Organized by**: Cooperative Institute for Research to Operations in Hydrology - CIROH

Git-hub Calib

CIROH
Cooperative Institute for Research to Operations in Hydrology

Git-hub Calib

skoriche / NGIAB-Calibration-DevCon25

Code | Issues | Pull requests | Actions | Projects

NGIAB-Calibration-DevCon25 Public

main | 2 Branches | 0 Tags

skoriche Update Acknowledgment

- Data/Provo_River
- Information
- Slides
- tools
- .gitignore
- .gitmodules
- LICENSE
- README.md
- dev_install.sh

NGIAB-Calibration-DevCon25 / Data / Provo_River /

README.md

Data Preparation Command

The following command was used to create the folder structure:

```
uvx ngiab-prep -i gage-10154200 --start 2007-10-01 --end 2013-09-30 -sfr --source nwm
```

NGIAB-Calibration-DevCon25 / Information /

- figures
- ReadME-Data-Preparation.md
- ReadME-NGIAB-Installation-and-Setup.md

NGIAB-Calibration-DevCon25 / tools /

- NGIAB-CloudInfra @ 4157009
- NGIAB_data_preprocess @ a6dd733
- ngen-cal @ f450a9d
- ngiab-cal @ 19d5ded

dev_install.sh

CIROH
Cooperative Institute for Research to Operations in Hydrology

Wiki page

NGIAB Calibration Workshop

Quick Start

- 🚀 Quickstart Guide

Overview

- Home
- What is NextGen?

Workshop Resources

- Workshop Materials
- Pre-Workshop Checklist

Instructions

- Step-by-Step Instructions

Technical Documentation

- Tools and Modules
- Workflow Process
- Directory Structure

Developer Resources

- Development Setup

Support

ReadMe

📚 Workshop Resources

🚀 Get Started

- **Quickstart Guide** - Start calibrating in 5 minutes
- Step-by-Step Instructions - Detailed tutorial
- Workshop Materials - Slides and datasets

🔧 Understanding the Tools

- What is NextGen? - Framework overview
- Tools and Modules - Component explanations
- Workflow Process - How everything connects

🔬 Advanced Topics

- Development Setup - Modify and extend tools
- Directory Structure - File organization reference

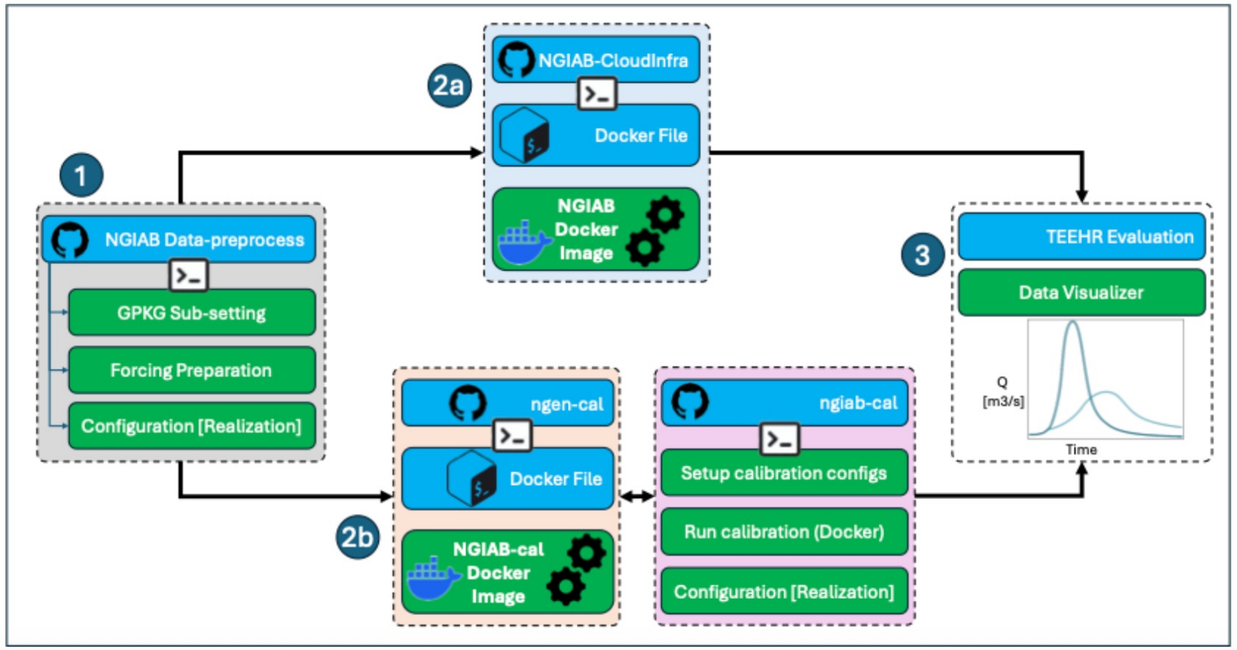
📞 Support

- Troubleshooting - Common issues and solutions
- Contact Information - Get help and support
- GitHub Issues - Report bugs or ask questions

What You'll Learn

- ✅ **Model Calibration Fundamentals** - Theory and practice of improving hydrological model accuracy
- ✅ **Hands-on NextGen Experience** - Complete calibration workflows using real tools and data
- ✅ **Parameter Management** - Share and collaborate on calibrated parameters effectively
- ✅ **Best Practices** - Learn from experts and discuss real-world challenges

Workshop Workflow



CIROH
Cooperative Institute for Research to Operations in Hydrology

James Halgren

NextGen In A Box Calibration

Overview and Demo

File: **calibrate.sh**

```
#!/bin/bash
gages=("10132500" "10132000" "10131000" "10130500" "10129500")
for gage in "${gages[@]}"; do
    echo $gage
    uvx --from ngiab_data_preprocess cli -i gage-"$gage" -sfr --start 2007-10-01 --end 2013-09-30
    uvx ngiab-cal /home/josh/swe_calibration/gage-"$gage" -g "$gage" --run -i 200
done
```

It’s that simple*

* As long as our default calibration parameters work well for your basin
  And you have docker and astral-uv installed
  And you have run the preprocessor setup

CIROH
Cooperative Institute for Research to Operations in Hydrology

Architecture

- Installed ngen-cal inside nextgen in a box
- Uses a modified branch of ngen-cal [CIROH-UA/ngen-cal@ngiab_cal]
- Created a small python cli tool to generate a default calibration configuration
- ngiab and the datastream use the same folder structure, other tools can too

```
FROM awiciroh/ciroh-ngen-image AS base      Josh Cunningham, 2 weeks ago
RUN dnf install -y git gcc-c++ make cmake python3-devel python3-pip
WORKDIR /calibration
RUN chmod -R 777 /calibration/
COPY ngen-cal/requirements.txt .
RUN uv venv
RUN uv pip install -r requirements.txt
COPY ngen-cal /calibration/ngen-cal
RUN chmod -R 777 /calibration/ngen-cal
RUN uv pip install -e ngen-cal/python/runCalibValid/ngen_cal
RUN uv pip install -e ngen-cal/python/runCalibValid/ngen_conf
RUN uv pip install numpy==1.26.0 netCDF4 geopandas==1.* xarray colorama rich

COPY mpi-ngen /dmod/bin/mpi-ngen

RUN echo "/calibration/.venv/bin/python /calibration/ngen-cal/python/runCalibValid/calibr
RUN echo "/calibration/.venv/bin/python /calibration/ngen-cal/python/runCalibValid/valida
RUN echo "/calibration/calibrate.sh && /calibration/validate.sh" >> run.sh
RUN chmod +x run.sh calibrate.sh validate.sh

ENV VIRTUAL_ENV=/ngen/.venv/

# This is to stop matplotlib complaining
RUN mkdir -p /.config/
RUN mkdir -p /.cache/
RUN chmod -R 777 /.config/
RUN chmod -R 777 /.cache/
ENV PS1="ngiab-cal\[\033[01;32m\]@demo\[\033[00m\]:\[\033[01;35m\]\w\[\033[00m\]$ "
ENTRYPOINT [ "/bin/bash" ]
```

```
folder_layout/
├── calibration
│   ├── crosswalk.json
│   ├── ngen_cal_conf.yaml
│   ├── obs_hourly_discharge.csv
│   ├── Output
│   │   ├── Calibration_Run
│   │   └── Validation_Run
│   ├── realization.json
│   ├── realization.json_original
│   └── troute.yaml
├── config
│   ├── cat_config
│   │   ├── CFE
│   │   │   ├── cat-2863621.ini
│   │   │   └── cat-2863622.ini
│   │   └── NOAH-OWP-M
│   │       ├── cat-2863621.input
│   │       └── cat-2863622.input
│   ├── gage-10154200_subset.gpkg
│   ├── realization.json
│   └── troute.yaml
├── forcings
│   └── forcings.nc
├── metadata
├── outputs
│   ├── ngen
│   │   ├── cat-2863621.csv
│   │   ├── cat-2863622.csv
│   │   └── nex-2863847_output.csv
│   └── troute
│       └── troute_output_201008010000.nc
└── teehr
```

CIROH
Cooperative Institute for Research to Operations in Hydrology

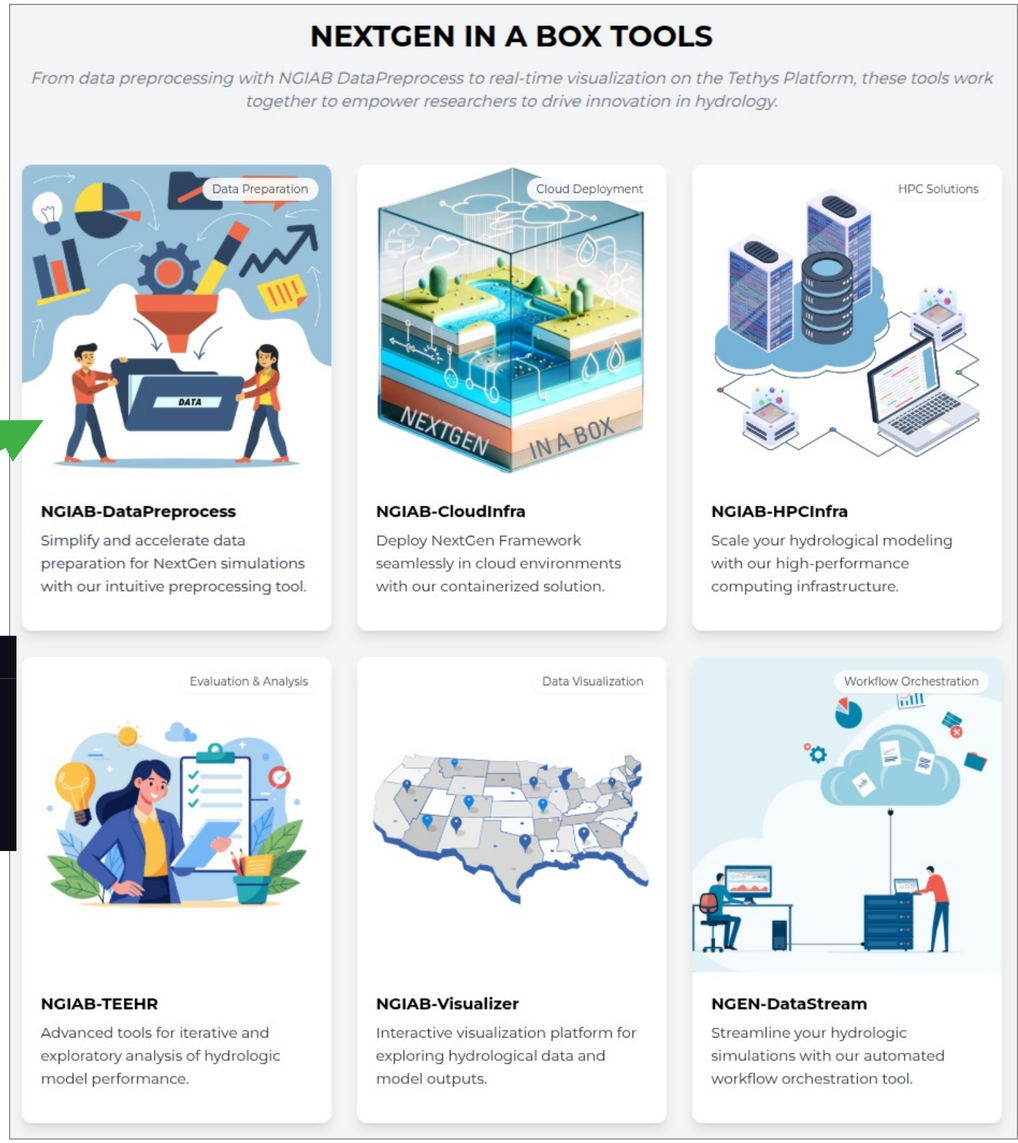
Architecture

NGIAB ecosystem / run workflow / requirements

- All of these leverage docker
- Astral UV is a powerful recommended tool but not needed:
  - allows use of these tools with no virtual environments
  - no clashing with other environments or global python packages

File: **calibrate.sh**

```
#!/bin/bash
gages=("10132500" "10132000" "10131000" "10130500" "10129500")
for gage in "${gages[@]}"; do
    echo $gage
    uvx --from ngiab_data_preprocess cli -i gage-"$gage" -sfr --start 2007-10-01 --end 2013-09-30
    uvx ngiab-cal /home/josh/swe_calibration/gage-"$gage" -g "$gage" --run -i 200
done
```



#5: using the preprocessor, select the gage, subset the hydrofabric, calculate mean average area forcings, generate ngen realization and model configuration

#6: using the new python tool, generate calibration configuration for 200 iterations, optionally specifying the gage inside that hydrofabric you want to calibrate on

CIROH
Cooperative Institute for Research to Operations in Hydrology

What is the default calibration?

Expert informed, but still just defaults.

- CFEv1 + Noah-owp-modular (cfe1 uses giuh not nash cascade for surface runoff)
- Dynamically Dimensioned Search (DDS) for 100 iterations
- Uses kge objective function
- Lumped calibration (model parameters are modified globally for all catchments being simulated)
- Single objective (one gage)

```
"kge": metric_functions.KGE,
"nse": metric_functions.NSE,
"rmse": metric_functions.root_mean_squared_error,
"rsr": metric_functions.rmse_std_ratio,
"nnse": metric_functions. Weighted_NSE,
```

- You can configure whatever models you like, I just can't generate model config for them yet
- Number of iterations is easy to configure
- PSO (particle swarm is available but I've not tested it properly in this configuration)
- Objective function configurable
- Unsure about lumped calibration and single objective

Live demo –

what could go wrong?

- uvx --from ngiab_data_preprocess cli -i gage-10154200 --start 2010-08-01 --end 2013-08-01 -sfr --source aorc
- uvx ngiab-cal /home/josh/demo/gage-10154200 -i 3

Backup (things went wrong)

```
09:51|josh@local:~$ uvx ngiab-cal /mnt/raid0/cal_testing/cal_test/ -f -i 3
INFO - Validating input files
Select a gage to calibrate from ['10153800', '10153500', '10154200']: 10154200
INFO - Downloading USGS data for 10154200 between 2010-08-01 00:00:00 and 2013-08-01 00:00:00
INFO - Writing ngiab-cal configuration
WARNING - This is still experimental, run the following command to start calibration:
WARNING - docker run -it -v "/mnt/raid0/cal_testing/cal_test:/ngen/ngen/data" --user $(id -u):$(id -g) joshcu/ngiab-cal:demo
```

```
09:52|josh@local:~$ docker run -it -v "/mnt/raid0/cal_testing/cal_test:/ngen/ngen/data" --user $(id -u):$(id -g) joshcu/ngiab
-cal:demo
ngiab-cal@demo:/calibration$ ls
calibrate.sh  ngen-cal  requirements.txt  run.sh  validate.sh
ngiab-cal@demo:/calibration$ cat calibrate.sh
/calibration/.venv/bin/python /calibration/ngen-cal/python/runCalibValid/calibration.py /ngen/ngen/data/calibration/ngen_cal_
conf.yaml
ngiab-cal@demo:/calibration$ cat validate.sh
/calibration/.venv/bin/python /calibration/ngen-cal/python/runCalibValid/validation.py /ngen/ngen/data/calibration/Output/Val
idation_Run/ngen_cal_conf.yaml
ngiab-cal@demo:/calibration$ cat run.sh
/calibration/calibrate.sh && /calibration/validate.sh
```

CIROH
Cooperative Institute for Research to Operations in Hydrology

Backup (things went wrong)

```
ngiab-cal@demo:/calibration$ ./calibrate.sh
14:54:29,998 INFO: generated new fontManager
2025-05-07 14:54:31 - INFO - Reading configuration from /ngen/ngen/data/calibration/ngen_cal_conf.yaml
2025-05-07 14:54:31 - INFO - Starting calibration process
2025-05-07 14:54:31 - INFO - {'type': 'ngen', 'binary': '/dmod/bin/mpi-ngen', 'realization': '/ngen/ngen/data/calibration/realization.json', 'catchments': '/ngen/ngen/data/config/cal_test_subset.gpkg', 'nexus': '/ngen/ngen/data/config/cal_test_subset.gpkg', 'crosswalk': '/ngen/ngen/data/calibration/crosswalk.json', 'obsflow': '/ngen/ngen/data/calibration/obs_hourly_discharge.csv', 'strategy': 'uniform', 'params': {'CFE': [{'name': 'b', 'min': 2.0, 'max': 15.0, 'init': 4.05}, {'name': 'satpsi', 'min': 0.03, 'max': 0.955, 'init': 0.355}, {'name': 'satdk', 'min': 1e-07, 'max': 0.000726, 'init': 3.38e-06}, {'name': 'maxsmc', 'min': 0.16, 'max': 0.59, 'init': 0.439}, {'name': 'refkdt', 'min': 0.1, 'max': 4.0, 'init': 1.0}, {'name': 'slope', 'min': 0.0, 'max': 1.0, 'init': 0.1}, {'name': 'max_gw_storage', 'min': 0.01, 'max': 0.25, 'init': 0.05}, {'name': 'expon', 'min': 1.0, 'max': 8.0, 'init': 3.0}, {'name': 'Cgw', 'min': 1.8e-06, 'max': 0.0018, 'init': 1.8e-05}, {'name': 'Klf', 'min': 0.0, 'max': 1.0, 'init': 0.01}, {'name': 'Kn', 'min': 0.0, 'max': 1.0, 'init': 0.03}], 'NoahOWP': [{'name': 'RSURF_EXP', 'min': 1.0, 'max': 6.0, 'init': 5.0}, {'name': 'CWP', 'min': 0.09, 'max': 0.36, 'init': 0.18}, {'name': 'MP', 'min': 3.6, 'max': 12.6, 'init': 9.0}, {'name': 'VCMX25', 'min': 24.0, 'max': 112.0, 'init': 52.2}, {'name': 'MFSNO', 'min': 0.5, 'max': 4.0, 'init': 2.0}, {'name': 'RSURF_SNOW', 'min': 0.136, 'max': 100.0, 'init': 50.0}, {'name': 'SCAMAX', 'min': 0.7, 'max': 1.0, 'init': 0.9}]}, 'eval_params': {'objective': 'kge', 'evaluation_start': '2011-08-01 00:00:00', 'evaluation_stop': '2012-07-31 00:00:00', 'valid_start_time': '2010-08-01 00:00:00', 'valid_end_time': '2013-08-01 00:00:00', 'valid_eval_start_time': '2012-07-31 00:00:00', 'valid_eval_end_time': '2013-08-01 00:00:00', 'full_eval_start_time': '2010-08-01 00:00:00', 'full_eval_end_time': '2013-08-01 00:00:00', 'save_output_iteration': 0, 'save_plot_iteration': 0, 'save_plot_iter_freq': 3, 'basinID': 10154200, 'threshold': None, 'site_name': 'USGS 10154200: ', 'user': ''}, 'workdir': PosixPath('/ngen/ngen/data/calibration/Output/Calibration_Run/ngen_hommc10g_worker')}
2025-05-07 14:54:32 - INFO - Using DDS algorithm, starting at iteration 0
2025-05-07 14:54:32 - INFO - Starting calibration loop with strategy: uniform
2025-05-07 14:54:32 - WARNING - Simulation output troute_output_201008010000.nc not found. Current working directory is /ngen/ngen/data/calibration
2025-05-07 14:54:32 - WARNING - This is expected on the first iteration.
2025-05-07 14:54:32 - WARNING - Simulation output troute_output_201008010000.nc not found. Current working directory is /ngen/ngen/data/calibration
2025-05-07 14:54:32 - WARNING - This is expected on the first iteration.
2025-05-07 14:54:32 - INFO - Running /dmod/bin/mpi-ngen /ngen/ngen/data/config/cal_test_subset.gpkg "all" /ngen/ngen/data/config/cal_test_subset.gpkg "all" realization.json to produce initial simulation
──────────────────────────────────────── ngen simulation live output ────────────────────────────────────────
NGen top-level timings:
        NGen::init: 4.32685
        NGen::simulation: 38.2611
        NGen::routing: 29.8939
──────────────────────────────────────── tail -n 4 /ngen/ngen/data/calibration/Output/Calibration_Run/ngen_hommc10g_worker/ngen.log ────────────────────────────────────────
2025-05-07 14:55:46 - INFO - Current score 1.702826551445217
Best score 1.702826551445217
2025-05-07 14:55:46 - INFO - Best parameters at iteration 0
2025-05-07 14:55:47 - INFO - Plotting Streamflow Time Series
2025-05-07 14:55:47 - INFO - Plotting Scatterplot of Streamflow between Observation and Other Runs
2025-05-07 14:55:49 - INFO - Plotting FDC of Observation and Other Runs
2025-05-07 14:55:50 - INFO - Plotting Streamflow Time Series with Precipitation
2025-05-07 14:55:51 - INFO - Plotting Scatterplot between Objective Funtion and Iteration
2025-05-07 14:55:51 - INFO - Plotting Scatterplot between Variables and Iteration
2025-05-07 14:55:53 - INFO - Plotting Scatterplot between Objective Function and Metric
2025-05-07 14:55:55 - INFO - Plotting Scatterplot between Variables and Iteration
2025-05-07 14:55:58 - INFO - Iteration 1/3 Time elapsed: 0:01:25.923350, estimated remaining: 0:04:17.770050
──────────────────────────────────────── ngen simulation live output ────────────────────────────────────────
Updating layer: surface layer
Running timestep 17700
Updating layer: surface layer
Running timestep 17800
Updating layer: surface layer
Running timestep 15100
Updating layer: surface layer
Running timestep 17700
Updating layer: surface layer
Running timestep 17700
──────────────────────────────────────── tail -n 10 /ngen/ngen/data/calibration/Output/Calibration_Run/ngen_hommc10g_worker/ngen.log ────────────────────────────────────────
```

```
ngiab-cal@demo:/calibration$ ./run.sh
22:46:13,957 INFO: generated new fontManager
2025-05-06 22:46:14 - INFO - Reading configuration from /ngen/ngen/data/calibration/ngen_cal_conf.yaml
2025-05-06 22:46:14 - INFO - Starting calibration process
2025-05-06 22:46:14 - INFO - {'type': 'ngen', 'binary': '/dmod/bin/mpi-ngen', 'realization': '/ngen/ngen/data/calibration/realization.json', 'catchment
on/obs_hourly_discharge.csv', 'strategy': 'uniform', 'params': {'CFE': [{'name': 'b', 'min': 2.0, 'max': 15.0, 'init': 4.05}, {'name': 'satpsi', 'min'
t': 1.0}, {'name': 'slope', 'min': 0.0, 'max': 1.0, 'init': 0.1}, {'name': 'max_gw_storage', 'min': 0.01, 'max': 0.25, 'init': 0.05}, {'name': 'expon'
 'NoahOWP': [{'name': 'RSURF_EXP', 'min': 1.0, 'max': 6.0, 'init': 5.0}, {'name': 'CWP', 'min': 0.09, 'max': 0.36, 'init': 0.18}, {'name': 'MP', 'min'
}, {'name': 'SCAMAX', 'min': 0.7, 'max': 1.0, 'init': 0.9}]}, 'eval_params': {'objective': 'kge', 'evaluation_start': '2003-12-01 00:00:00', 'evaluatio
ll_eval_start_time': '2002-12-01 00:00:00', 'full_eval_end_time': '2022-01-02 00:00:00', 'save_output_iteration': 0, 'save_plot_iteration': 0, 'save_pl
2025-05-06 22:46:16 - INFO - Using DDS algorithm, starting at iteration 0
2025-05-06 22:46:16 - INFO - Starting calibration loop with strategy: uniform
2025-05-06 22:46:16 - WARNING - Simulation output troute_output_200212010000.nc not found. Current working directory is /ngen/ngen/data/calibration
2025-05-06 22:46:16 - WARNING - This is expected on the first iteration.
2025-05-06 22:46:16 - WARNING - Simulation output troute_output_200212010000.nc not found. Current working directory is /ngen/ngen/data/calibration
2025-05-06 22:46:16 - WARNING - This is expected on the first iteration.
2025-05-06 22:46:16 - INFO - Running /dmod/bin/mpi-ngen /ngen/ngen/data/config/gage-06719505_subset.gpkg "all" /ngen/ngen/data/config/gage-06719505_sub
──────────────────────────────────────── ngen simulation live output ────────────────────────────────────────
NGen top-level timings:
        NGen::init: 3.08741
        NGen::simulation: 1390.25
        NGen::routing: 110.22
──────────────────────────────────────── tail -n 4 /ngen/ngen/data/calibration/Output/Calibration_Run/ngen_pu1m3qjg_worker/ngen.log ────────────────────────────────────────
2025-05-06 23:11:25 - INFO - Current score 1.4488313186304642
Best score 1.4488313186304642
2025-05-06 23:11:25 - INFO - Best parameters at iteration 0
2025-05-06 23:11:28 - INFO - Plotting Streamflow Time Series
2025-05-06 23:11:29 - INFO - Plotting Scatterplot of Streamflow between Observation and Other Runs
2025-05-06 23:11:31 - INFO - Plotting FDC of Observation and Other Runs
2025-05-06 23:11:32 - INFO - Plotting Streamflow Time Series with Precipitation
2025-05-06 23:11:32 - INFO - Plotting Scatterplot between Objective Funtion and Iteration
2025-05-06 23:11:32 - INFO - Plotting Scatterplot between Variables and Iteration
2025-05-06 23:11:32 - INFO - Plotting Scatterplot between Objective Function and Metric
2025-05-06 23:11:33 - INFO - Plotting Scatterplot between Variables and Iteration
2025-05-06 23:11:34 - INFO - Iteration 1/3 Time elapsed: 0:25:17.841281, estimated remaining: 1:15:53.523843
──────────────────────────────────────── ngen simulation live output ────────────────────────────────────────
NGen top-level timings:
        NGen::init: 3.80876
        NGen::simulation: 1521.12
        NGen::routing: 99.9203
──────────────────────────────────────── tail -n 4 /ngen/ngen/data/calibration/Output/Calibration_Run/ngen_pu1m3qjg_worker/ngen.log ────────────────────────────────────────
2025-05-06 23:38:47 - INFO - Iteration 2/3 Time elapsed: 0:52:31.558247, estimated remaining: 0:52:31.558248
──────────────────────────────────────── ngen simulation live output ────────────────────────────────────────
NGen top-level timings:
        NGen::init: 5.11636
        NGen::simulation: 1552.94
        NGen::routing: 100.097
──────────────────────────────────────── tail -n 4 /ngen/ngen/data/calibration/Output/Calibration_Run/ngen_pu1m3qjg_worker/ngen.log ────────────────────────────────────────
2025-05-07 00:06:34 - INFO - Iteration 3/3 Time elapsed: 1:20:18.736845, estimated remaining: 0:26:46.245615
──────────────────────────────────────── ngen simulation live output ────────────────────────────────────────
NGen top-level timings:
        NGen::init: 4.46685
        NGen::simulation: 1544.42
        NGen::routing: 101.207
──────────────────────────────────────── tail -n 4 /ngen/ngen/data/calibration/Output/Calibration_Run/ngen_pu1m3qjg_worker/ngen.log ────────────────────────────────────────
2025-05-07 00:34:14 - INFO - Plotting Streamflow Time Series
2025-05-07 00:34:14 - INFO - Plotting Scatterplot of Streamflow between Observation and Other Runs
2025-05-07 00:34:17 - INFO - Plotting FDC of Observation and Other Runs
2025-05-07 00:34:17 - INFO - Plotting Streamflow Time Series with Precipitation
2025-05-07 00:34:18 - INFO - Plotting Scatterplot between Objective Funtion and Iteration
2025-05-07 00:34:18 - INFO - Plotting Scatterplot between Variables and Iteration
2025-05-07 00:34:18 - INFO - Plotting Scatterplot between Objective Function and Metric
2025-05-07 00:34:19 - INFO - Plotting Scatterplot between Variables and Iteration
2025-05-07 00:34:20 - INFO - Calibration process completed
2025-05-07 00:34:22 - INFO - Reading configuration from /ngen/ngen/data/calibration/Output/Validation_Run/ngen_cal_conf.yaml
2025-05-07 00:34:22 - INFO - Starting Control Run
2025-05-07 00:34:22 - INFO - {'type': 'ngen', 'binary': '/dmod/bin/mpi-ngen', 'realization': PosixPath('/ngen/ngen/data/calibration/Output/Validation_Run
n', 'obsflow': '/ngen/ngen/data/calibration/obs_hourly_discharge.csv', 'strategy': 'uniform', 'params': {'CFE': [{'name': 'b', 'min': 2.0, 'max': 15.0
': 'refkdt', 'min': 0.1, 'max': 4.0, 'init': 1.0}, {'name': 'slope', 'min': 0.0, 'max': 1.0, 'init': 0.1}, {'name': 'max_gw_storage', 'min': 0.01, 'max
, 'min': 0.0, 'max': 1.0, 'init': 0.03}], 'NoahOWP': [{'name': 'RSURF_EXP', 'min': 1.0, 'max': 6.0, 'init': 5.0}, {'name': 'CWP', 'min': 0.09, 'max': 0
 'min': 0.136, 'max': 100.0, 'init': 50.0}, {'name': 'SCAMAX', 'min': 0.7, 'max': 1.0, 'init': 0.9}]}, 'eval_params': {'objective': 'kge', 'evaluation
val_end_time': '2022-01-02 00:00:00', 'full_eval_start_time': '2002-12-01 00:00:00', 'full_eval_end_time': '2022-01-02 00:00:00', 'save_output_iteratio
_Run/ngen_0vb2rtc9_worker')}
2025-05-07 00:34:24 - INFO - Executing validation with best parameters
2025-05-07 00:34:24 - INFO - --- Start valid_control ---
──────────────────────────────────────── ngen simulation live output ────────────────────────────────────────
NGen top-level timings:
        NGen::init: 4.09035
        NGen::simulation: 1648.26
        NGen::routing: 125.66
──────────────────────────────────────── tail -n 4 /ngen/ngen/data/calibration/Output/Validation_Run/ngen_0vb2rtc9_worker/ngen.log ────────────────────────────────────────
2025-05-07 02:32:35 - INFO - Validation process completed
2025-05-07 02:32:35 - INFO - Starting Validation Run
2025-05-07 02:32:35 - INFO - {'type': 'ngen', 'binary': '/dmod/bin/mpi-ngen', 'realization': '/ngen/ngen/data/calibration/Output/Validation_Run/realiza
en/ngen/data/calibration/obs_hourly_discharge.csv', 'strategy': 'uniform', 'params': {'CFE': [{'name': 'b', 'min': 2.0, 'max': 15.0, 'init': 4.05}, {'
 0.1, 'max': 4.0, 'init': 1.0}, {'name': 'slope', 'min': 0.0, 'max': 1.0, 'init': 0.1}, {'name': 'max_gw_storage', 'min': 0.01, 'max': 0.25, 'init': 0
: 1.0, 'init': 0.03}], 'NoahOWP': [{'name': 'RSURF_EXP', 'min': 1.0, 'max': 6.0, 'init': 5.0}, {'name': 'CWP', 'min': 0.09, 'max': 0.36, 'init': 0.18}
': 100.0, 'init': 50.0}, {'name': 'SCAMAX', 'min': 0.7, 'max': 1.0, 'init': 0.9}]}, 'eval_params': {'objective': 'kge', 'evaluation_start': '2003-12-01
2-01-02 00:00:00', 'full_eval_start_time': '2002-12-01 00:00:00', 'full_eval_end_time': '2022-01-02 00:00:00', 'save_output_iteration': 0, 'save_plot_t
worker')}
2025-05-07 02:32:37 - INFO - Executing validation with best parameters
2025-05-07 02:32:37 - INFO - --- Start valid_best ---
──────────────────────────────────────── ngen simulation live output ────────────────────────────────────────
NGen top-level timings:
        NGen::init: 4.3512
        NGen::simulation: 1850.82
        NGen::routing: 129.003
──────────────────────────────────────── tail -n 4 /ngen/ngen/data/calibration/Output/Validation_Run/ngen_kanqy852_worker/ngen.log ────────────────────────────────────────
2025-05-07 11:41:49 - INFO - Plotting Streamflow Time Series
2025-05-07 11:41:49 - INFO - Plotting FDC of Observation and Other Runs
2025-05-07 11:41:51 - INFO - Plotting Streamflow Time Series with Precipitation
2025-05-07 11:41:51 - INFO - Plotting Barplot of Metrics
2025-05-07 11:41:52 - INFO - Validation process completed
ngiab-cal@demo:/calibration$
```

Cooperative to Operation

Parameter output

```
{
    "name": "bmi_fortran",
    "params": {
        "name": "bmi_fortran",
        "model_type_name": "NoahOWP",
        "library_file": "/dmod/shared_libs/libsurfacebmi.so",
        "forcing_file": "",
        "init_config": "./config/cat_config/NOAH-OWP-M/{{id}}.input",
        "allow_exceed_end_time": true,
        "main_output_variable": "QINSUR",
        "variables_names_map": {
            "PRCPNONC": "precip_rate",
            "Q2": "SPFH_2maboveground",
            "SFCTMP": "TMP_2maboveground",
            "UU": "UGRD_10maboveground",
            "VV": "VGRD_10maboveground",
            "LWDN": "DLWRF_surface",
            "SOLDN": "DSWRF_surface",
            "SFCPRS": "PRES_surface"
        },
        "uses_forcing_file": false,
        "model_params": {
            "RSURF_EXP": 1.1446477876537466,
            "CWP": 0.228607017822218,
            "MP": 5.15798212196691,
            "VCMX25": 44.43670283180776,
            "MFSNO": 1.8189116998241324,
            "RSURF_SNOW": 8.529872610767484,
            "SCAMAX": 0.9960120721149563
        }
    }
},
```

Using --run

```
gage-10154200/
├── calibration
│   ├── crosswalk.json
│   ├── ngen_cal_conf.yaml
│   ├── obs_hourly_discharge.csv
│   ├── Output
│   ├── realization.json
│   ├── realization.json_original
│   └── troute.yaml
├── config
│   ├── calibrated_params.json
│   ├── cat_config
│   ├── gage-10154200_subset.gpkg
│   ├── realization.json
│   ├── realization._old
│   └── troute.yaml
├── forcings
│   ├── forcings.nc
│   └── raw_gridded_data.nc
├── metadata
│   └── num_partitions
├── outputs
│   ├── ngen
│   └── troute
└── partitions_30.json
```

can be found here too ->

```
gage-10154200/calibration/Output/
├── Calibration_Run
│   ├── 10154200_Calib_Run_Complete
│   └── ngen_fnug56vr_worker
│       ├── 10154200_last_iteration.csv
│       ├── 10154200_metrics_iteration.csv
│       ├── 10154200_objective_log.txt
│       ├── 10154200_output_best_iteration.csv
│       ├── 10154200_output_last_iteration.csv
│       ├── 10154200_params_iteration.csv
│       ├── ngen.log
│       ├── Output_Calib
│       ├── Output_Iteration
│       │   └── 10154200_output_iteration_0000.csv
│       ├── parameter_df_state_nex-2863632.parquet
│       ├── partitions_30.json
│       ├── Plot_Iteration
│       │   ├── 10154200_fdc_iteration.png
│       │   ├── 10154200_hydrograph_iteration.png
│       │   ├── 10154200_metric_iteration.png
│       │   ├── 10154200_metric_objfun.png
│       │   ├── 10154200_objfun_iteration.png
│       │   ├── 10154200_param_iteration.png
│       │   ├── 10154200_scatterplot_streamflow_iteration.png
│       │   └── 10154200_streamflow_precip_iteration.png
│       ├── realization.json
│       └── troute_output_201008010000.nc_last
└── Validation_Run
    ├── 10154200_metrics_valid_best.csv
    ├── 10154200_metrics_valid_control.csv
    ├── 10154200_output_valid_best.csv
    ├── 10154200_output_valid_control.csv
    ├── 10154200_Valid_best_Run_Complete
    ├── 10154200_Valid_control_Run_Complete
    ├── control_realization.json
    ├── control_realization.json_original
    ├── ngen_1dnbxt1c_worker
    │   ├── ngen.log
    │   ├── Output_Valid
    │   ├── partitions_30.json
    │   ├── realization.json
    │   └── troute_output_201008010000.nc_valid_best
    ├── ngen_cal_conf.yaml
    ├── ngen_ls85bx8r_worker
    │   ├── control_realization.json
    │   ├── ngen.log
    │   ├── Output_Valid
    │   ├── partitions_30.json
    │   └── troute_output_201008010000.nc_valid_control
    ├── Plot_Valid
    │   ├── 10154200_barplot_metrics_valid_run.png
    │   ├── 10154200_fdc_valid_run.png
    │   ├── 10154200_hydrograph_valid_run.png
    │   └── 10154200_streamflow_precip_valid_run.png
    ├── realization.json
    └── realization.json_original
```

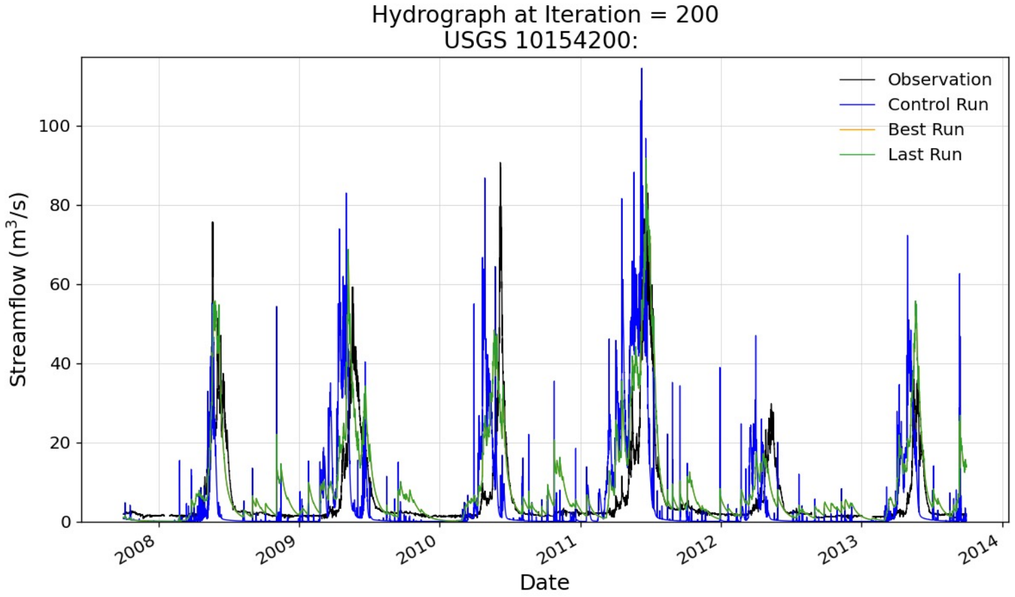
CIROH
Cooperative Institute for Research to Operations in Hydrology

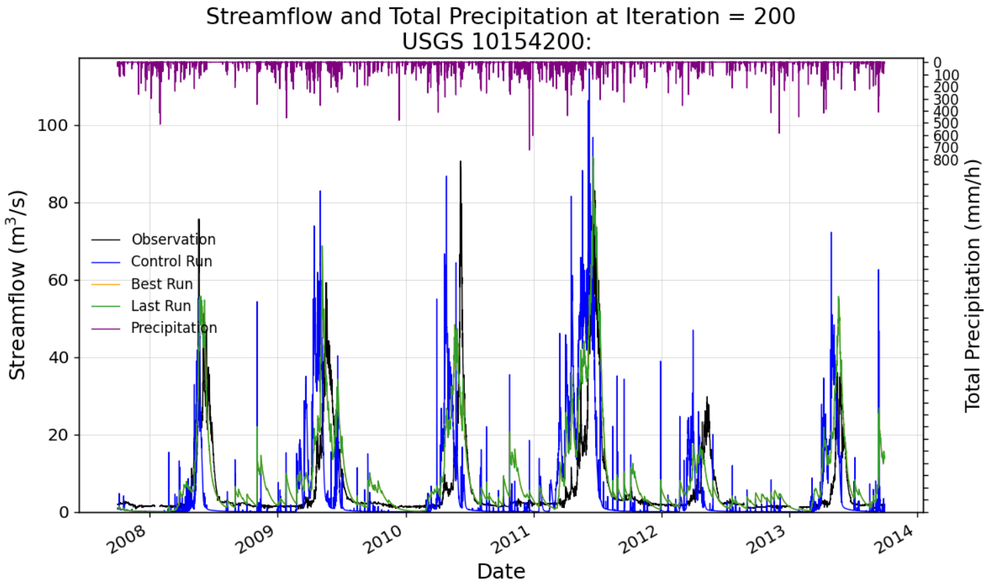
Lots of plots!

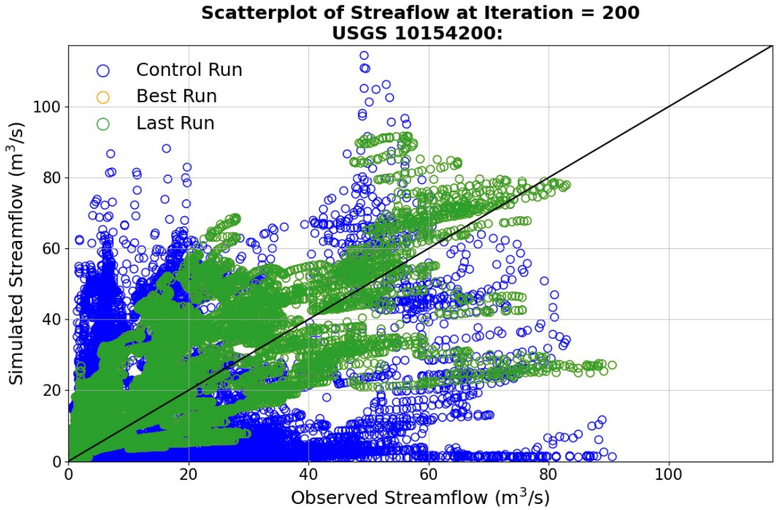
```
gage-10154200/calibration/Output/
├── Calibration_Run
│   ├── 10154200_Calib_Run_Complete
│   └── ngen_fnug56vr_worker
│       ├── 10154200_last_iteration.csv
│       ├── 10154200_metrics_iteration.csv
│       ├── 10154200_objective_log.txt
│       ├── 10154200_output_best_iteration.csv
│       ├── 10154200_output_last_iteration.csv
│       ├── 10154200_params_iteration.csv
│       ├── ngen.log
│       ├── Output_Calib
│       ├── Output_Iteration
│       │   └── 10154200_output_iteration_0000.csv
│       ├── parameter_df_state_nex-2863632.parquet
│       ├── partitions_30.json
│       ├── Plot_Iteration
│       │   ├── 10154200_fdc_iteration.png
│       │   ├── 10154200_hydrograph_iteration.png
│       │   ├── 10154200_metric_iteration.png
│       │   ├── 10154200_metric_objfun.png
│       │   ├── 10154200_objfun_iteration.png
│       │   ├── 10154200_param_iteration.png
│       │   ├── 10154200_scatterplot_streamflow_iteration.png
│       │   └── 10154200_streamflow_precip_iteration.png
│       ├── realization.json
│       └── troute_output_201008010000.nc_last
└── Validation_Run
    ├── 10154200_metrics_valid_best.csv
    ├── 10154200_metrics_valid_control.csv
    ├── 10154200_output_valid_best.csv
    ├── 10154200_output_valid_control.csv
    ├── 10154200_Valid_best_Run_Complete
    ├── 10154200_Valid_control_Run_Complete
    ├── control_realization.json
    ├── control_realization.json_original
    ├── ngen_1dnbxt1c_worker
    │   ├── ngen.log
    │   ├── Output_Valid
    │   ├── partitions_30.json
    │   ├── realization.json
    │   └── troute_output_201008010000.nc_valid_best
    ├── ngen_cal_conf.yaml
    ├── ngen_ls85bx8r_worker
    │   ├── control_realization.json
    │   ├── ngen.log
    │   ├── Output_Valid
    │   ├── partitions_30.json
    │   └── troute_output_201008010000.nc_valid_control
    ├── Plot_Valid
    │   ├── 10154200_barplot_metrics_valid_run.png
    │   ├── 10154200_fdc_valid_run.png
    │   ├── 10154200_hydrograph_valid_run.png
    │   └── 10154200_streamflow_precip_valid_run.png
    ├── realization.json
    └── realization.json_original
```

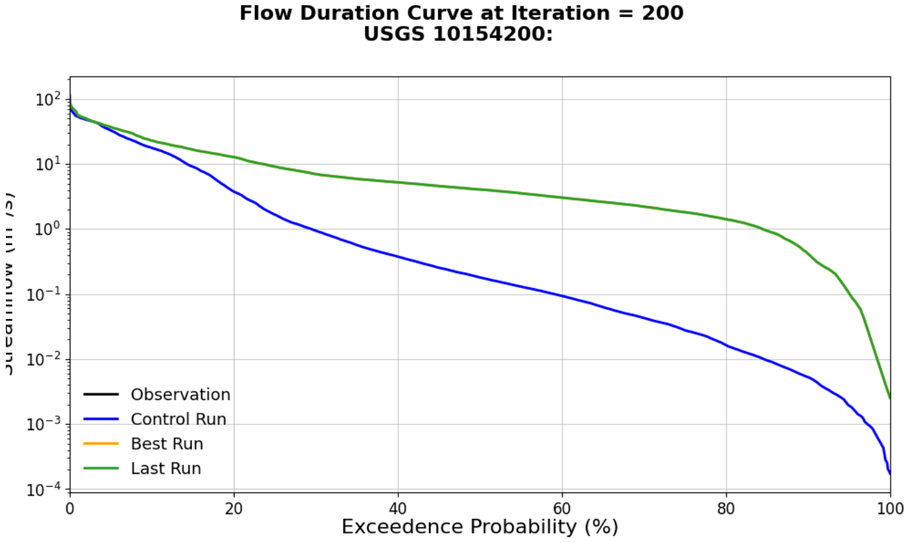
CIROH
Cooperative Institute for Research to Operations in Hydrology

Plots streamflow (from 200 iterations on Provo 10145200)



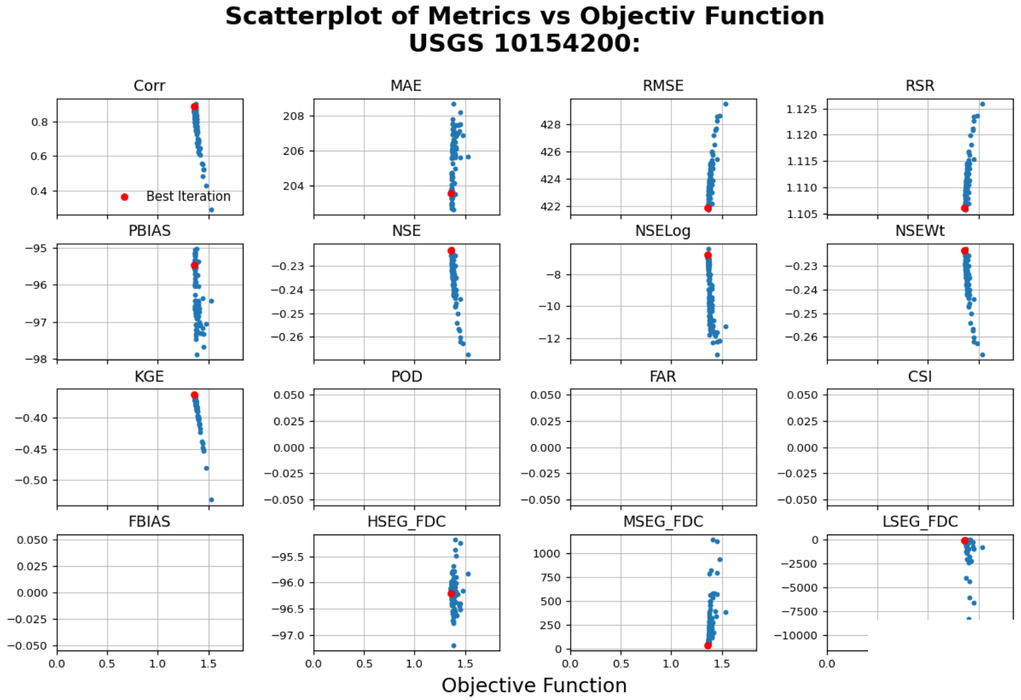


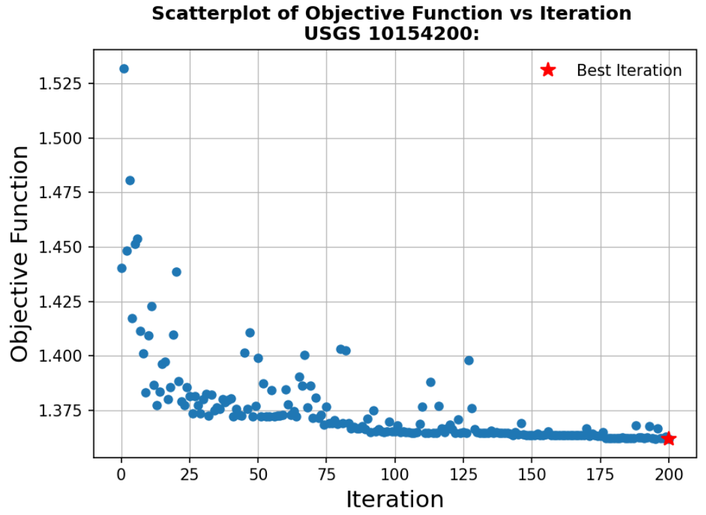


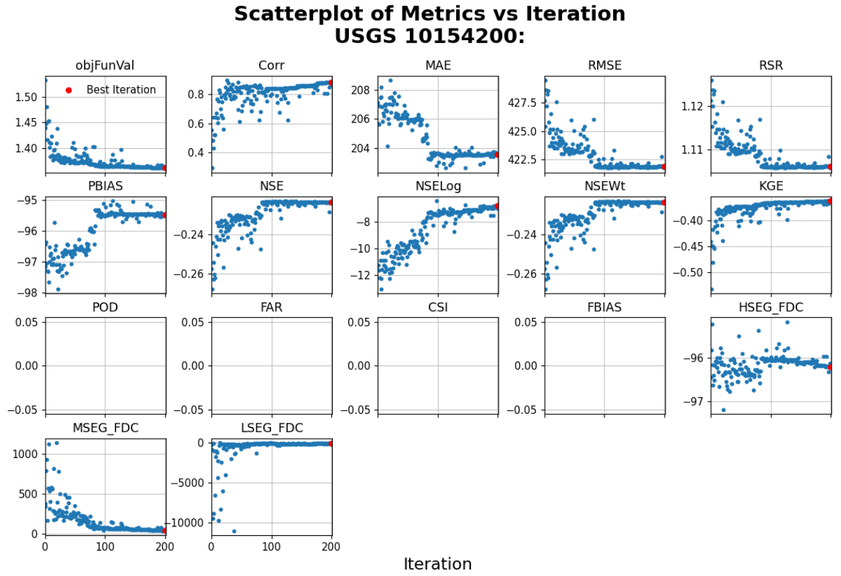


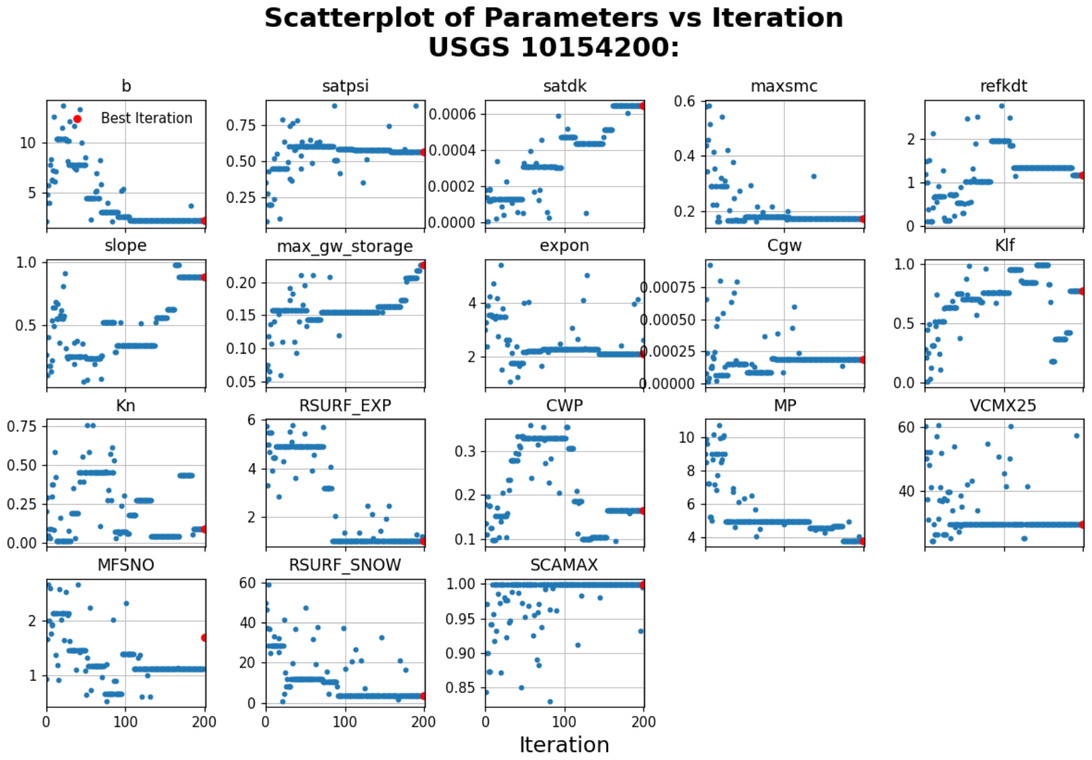
CIROH
Cooperative Institute for Research to Operations in Hydrology

Plots metrics (from 200 iterations on Provo 10145200)





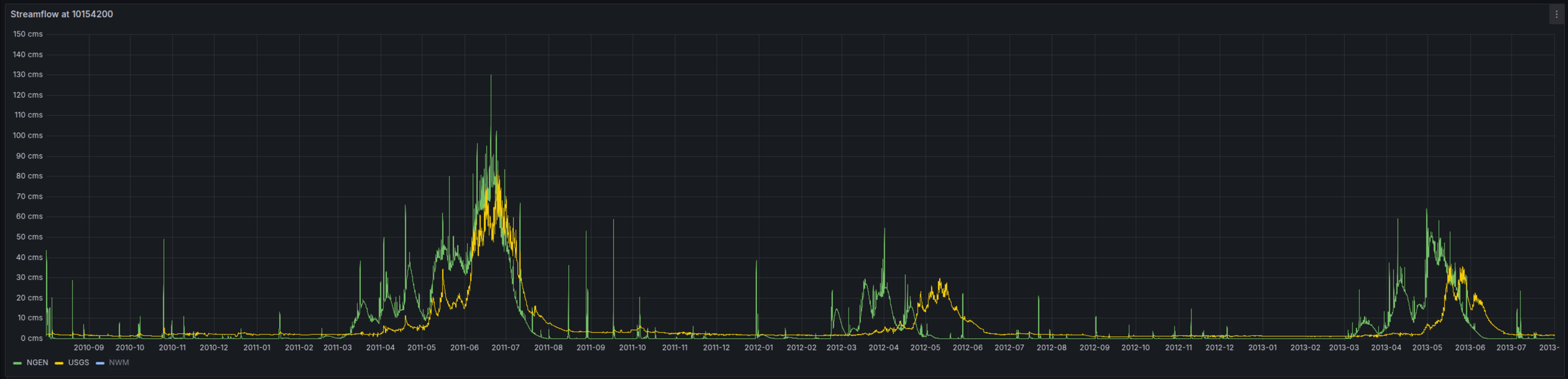


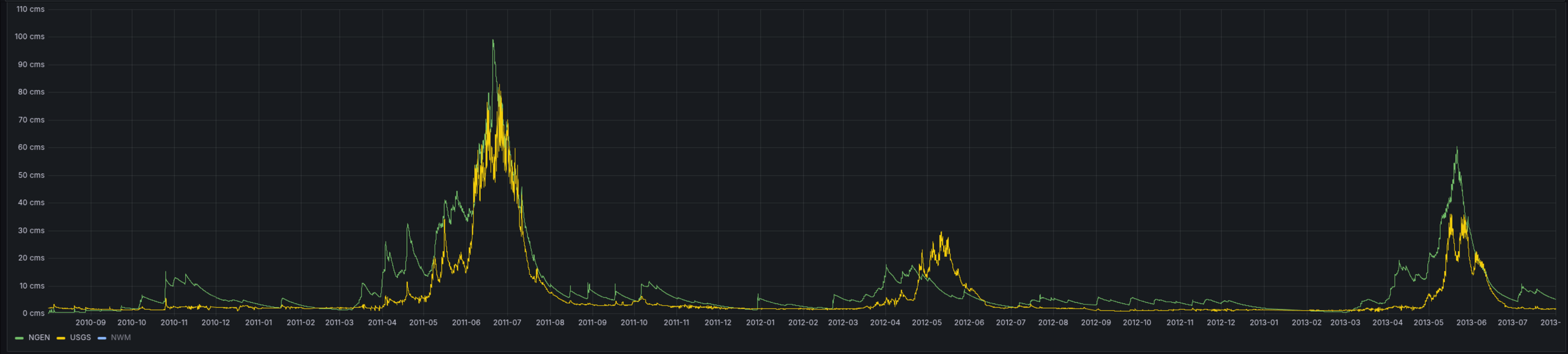


CIROH
Cooperative Institute for Research to Operations in Hydrology

Plots metrics (from 200 iterations on Provo 10145200)

USGS Gage ID 10154200 | 2010-08-01 01:00:00 to 2013-08-01 00:00:00 | Refresh





KGE NSE PBIAS

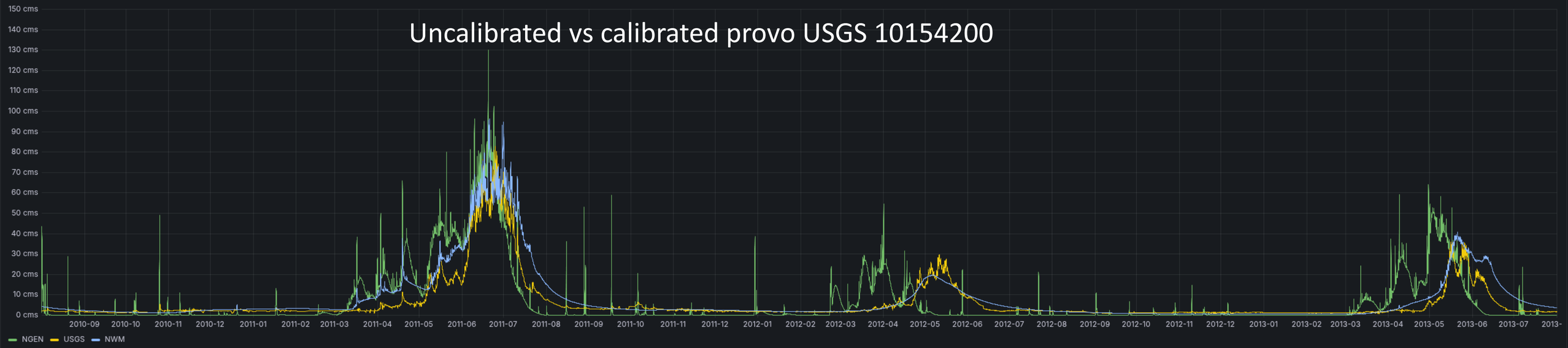
gage_id	source	kge	kge_r	kge_a	kge_b	nse	pbias
10154200	ngen	0.327	0.900	1.24	1.62	0.572	-61.9

USGS Gage ID 10154200

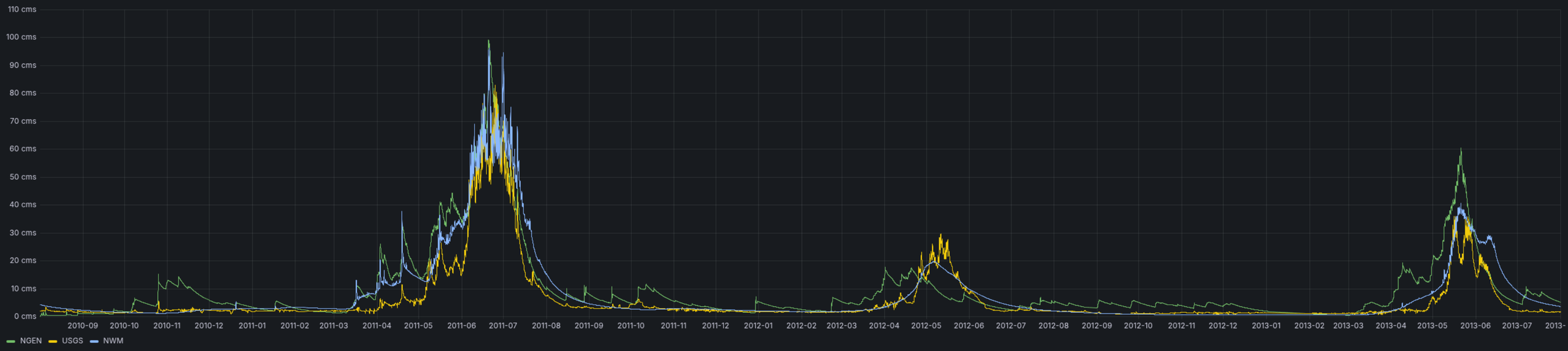
2010-08-01 01:00:00 to 2013-08-01 00:00:00 Refresh

Uncalibrated vs calibrated provo USGS 10154200

Streamflow at 10154200



Streamflow at 10154200

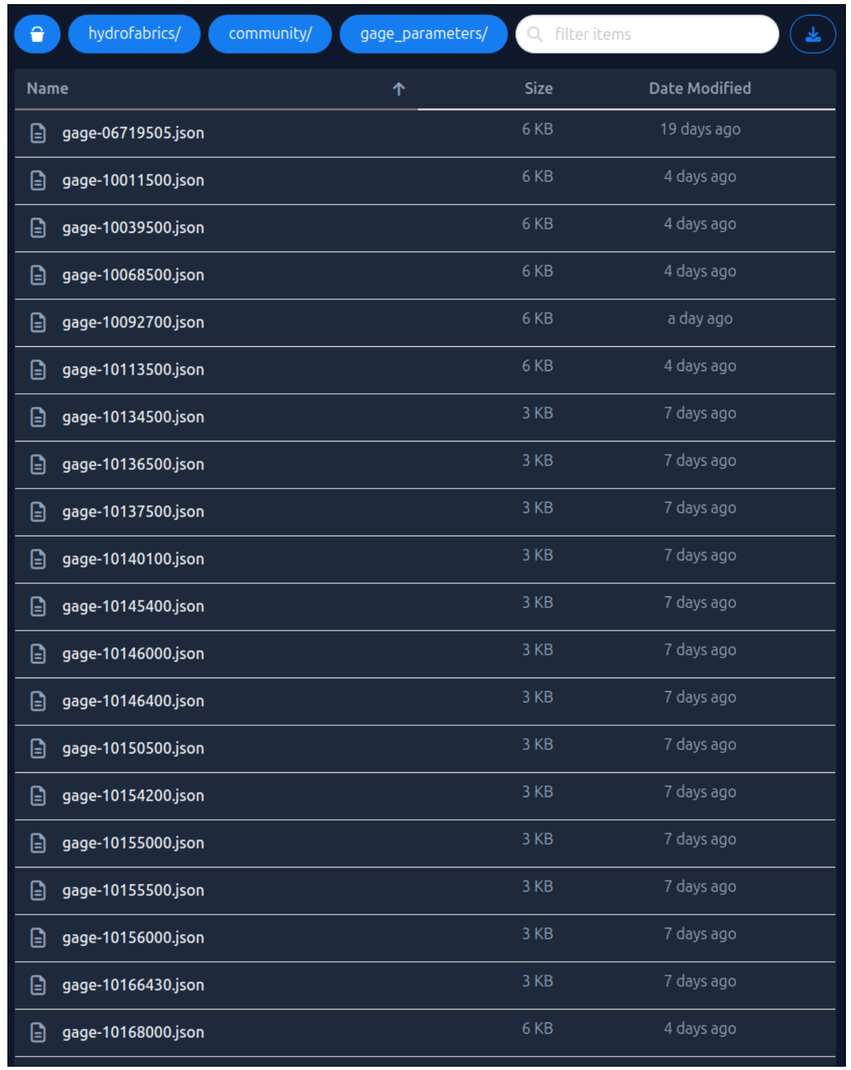


KGE NSE PBIAS

gage_id	source	kge	kge_r	kge_a	kge_b	nse	pbias
10154200	ngen	0.327	0.900	1.24	1.62	0.572	-61.9
10154200	nwm	0.603	0.946	1.21	1.33	0.790	-33.1

Community calibrated parameters - https://communityhydrofabric.s3.us-east-1.amazonaws.com/index.html

- Temporary solution while we work on something better
- After calibrating a gage, the realization is uploaded to this s3 bucket.
- When using the preprocessor, if you specify a gage on this list, it downloads the model parameters from S3
- Similar limitations to the default calibration
  - Lumped parameters
  - Only works for single gages



CIROH
Cooperative Institute for Research to Operations in Hydrology

Teehr + tethys ngiab visualiser

- Run the folder using guide.sh to use them both

TEEHR consolidates data from the USGS and NWM, allowing side-by-side visual comparisons of observed and simulated over the model run intervals. Figure 1 shows the default comparison of the modeled outlet hydrograph and the corresponding time series from the NWM 3.0. While the default configuration produces this view, the TEEHR user documentation provides additional examples of capabilities for customized plotting functions.

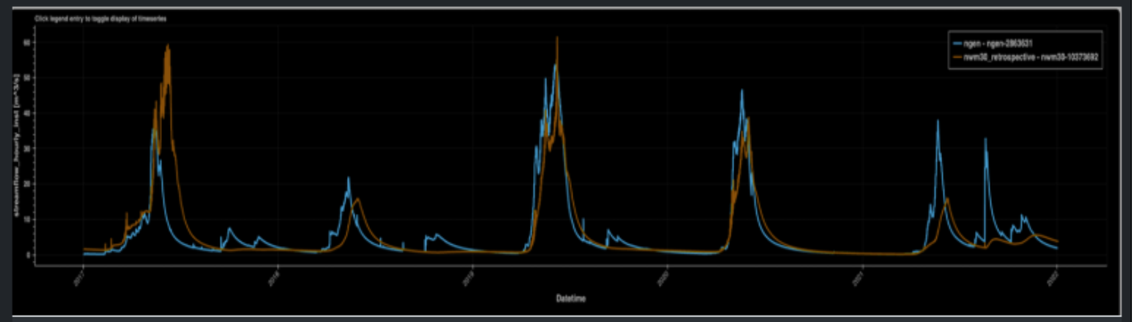


Figure 1: Comparison of the NextGen-based model (labeled “ngen”; blue line) and the NWM time series (labeled “nwm30_retrospective”; orange line) for the same location. The figure is automatically generated by the TEEHR-based analysis that accompanies the `guide.sh` script included with NGIAB and is named `timeseries_plot_streamflow_hourly_inst.html` in the `teehr` folder.

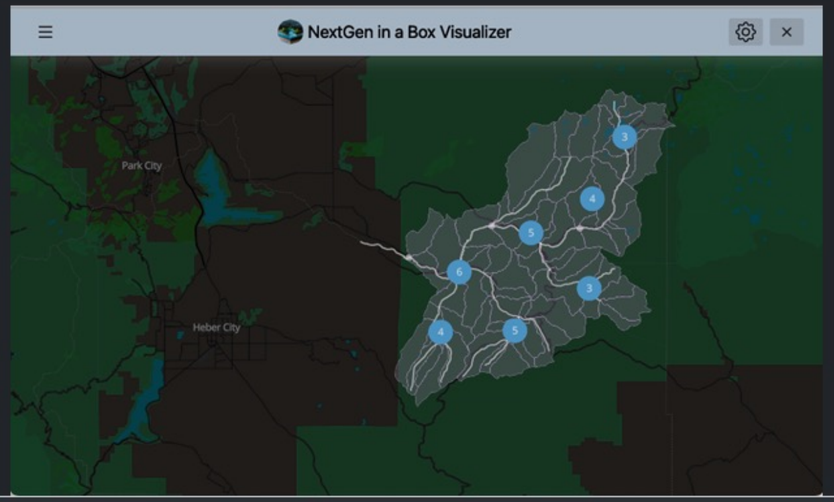


Figure 1: A map showing the geospatial visualization using the Data Visualizer within the Tethys framework for an entire study area (Provo River near Woodland, UT).

Figure 2: A map showing the geospatial visualization using the Data Visualizer within the Tethys framework for a selected outlet nexus point as well as displaying a time series plot between observed (labeled “USGS”; blue line) and simulated (labeled “ngen”; orange line) with the performance metrics (KGE, NSE, and relative bias). The Visualizer can also show the performance of the NWM 3.0 compared to the observed time series.

CIROH
Cooperative Institute for Research to Operations in Hydrology

Performance considerations

- Current configuration runs one simulation at a time, with as many cores as it can
- For small basins t-route doesn’t parallelize as well as model execution
- On very large machines, parallel ranks aren’t always balanced. Using the partitioning scheme that reduces cross-process mpi communication
- For maximum efficiency and “gage per hour” throughput it would be best to run many simulations in parallel using one core each
- Depending on docker setup there can be significant disk I/O overhead due to the -v “/local/folder/:/container/folder/” bind mount. Docker on linux without docker desktop works well by default.

CIROH
Cooperative Institute for Research to Operations in Hydrology

Future work

- Other than small changes and bugfixes leading up to devcon, future updates will be done after divergence from ngen-cal@main has been addressed
- Objectives will be steered by the needs of the datastream and community suggestions / contributions.
- Possible additions (after devcon)
  - Automatic support for more models
  - Tensorboard integration or equivalent for tracking calibration

Thank You!

calibrated gage parameters

github.com/CIROH-UA/NGIAB-CloudInfra

github.com/CIROH-UA/ngiab_cal

github.com/CIROH-UA/ngen-cal/tree/ngiab_cal

github.com/CIROH-UA/NGIAB_data_preprocess/

File: camels.sh

```
1  #!/bin/bash
2  # Download a list of CAMELS gage ids
3  wget https://raw.githubusercontent.com/peckhams/nextgen_basin_repo/5e1317256a9365ae3a24a250358314e1e9ffc339/CAMELS/Data/cam
   els_name.txt ./camels_name.txt
4  output_folder=$(cat ~/.ngiab/preprocessor)
5  while read line
6  do
7      gage=$(echo "$line" | cut -d ';' -f 1)
8      echo $gage
9      # subset the hydrofabric, calculate mean-average area forcings, generate model config files
10     uvx --from ngiab_data_preprocess cli -i gage-"$gage" -sfr --start 2007-10-01 --end 2013-09-30
11     # calibrate gage for 200 iterations
12     uvx ngiab-cal "$output_folder"/gage-"$gage" -g "$gage" --run -i 200
13 done < <(tail -n +2 ./camels_name.txt)
```

CIROH
Cooperative Institute for Research to Operations in Hydrology