OWP | OFFICE OF WATER PREDICTION

Developing Model and Parameter Regionalization Capabilities for the Next Generation Water Resources Modeling Framework (NextGen)

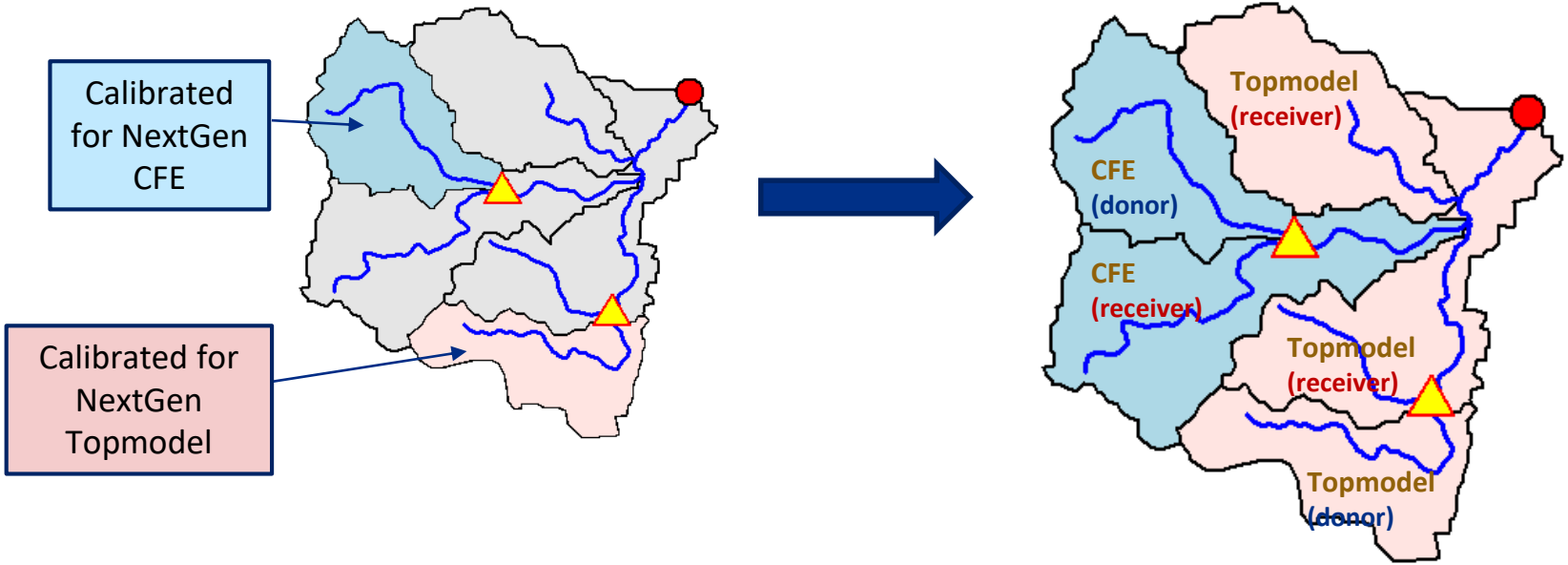
Yuqiong Liu, Xia Feng, Nels Frazier, Jason Ducker, Matthew Williamson, J. Michael Johnson,*Trey Flowers, Brian Cosgrove, and Fred L.Ogden*

Special Acknowledgements: Robert Bartel, Austin Relay, Ahmad Jan, and Keith Jennings

NextGen Regionalization:

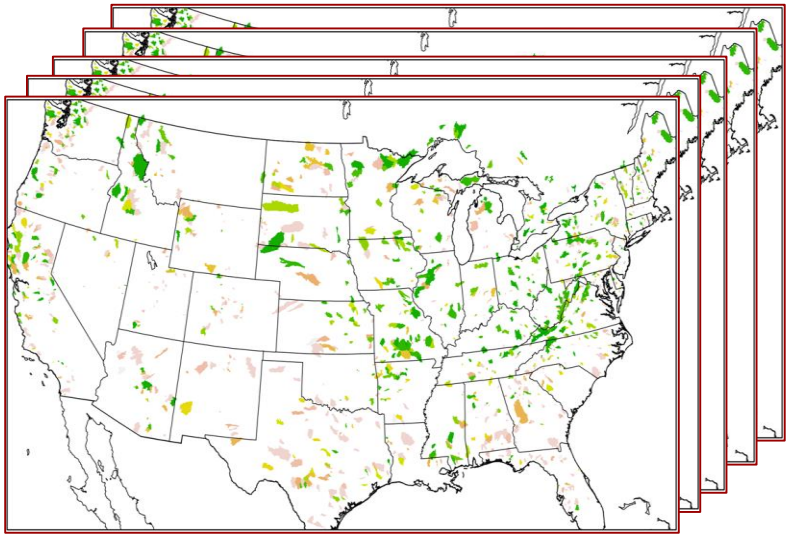
illustration with a 5-catchment toy basin

Goal: Given the calibrated catchments, identify the best donor of model/parameters for each uncalibrated catchment (receiver)

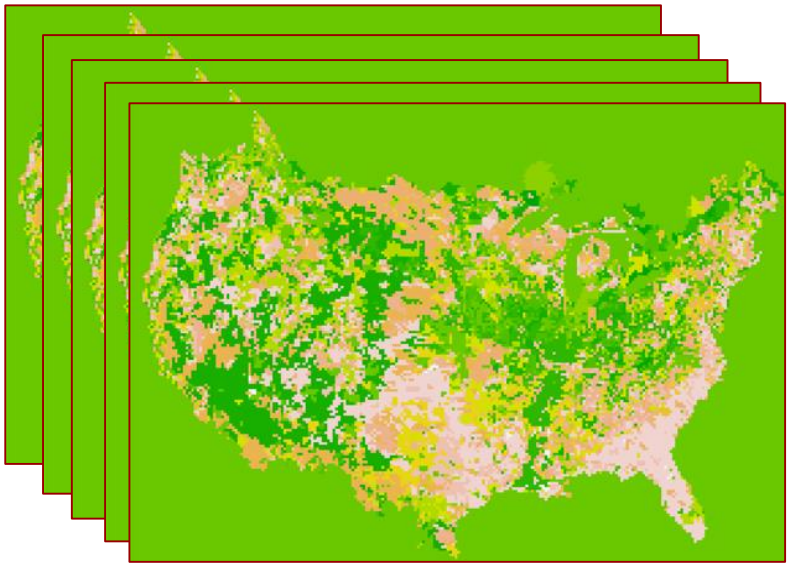


NextGen Regionalization: scale to CONUS

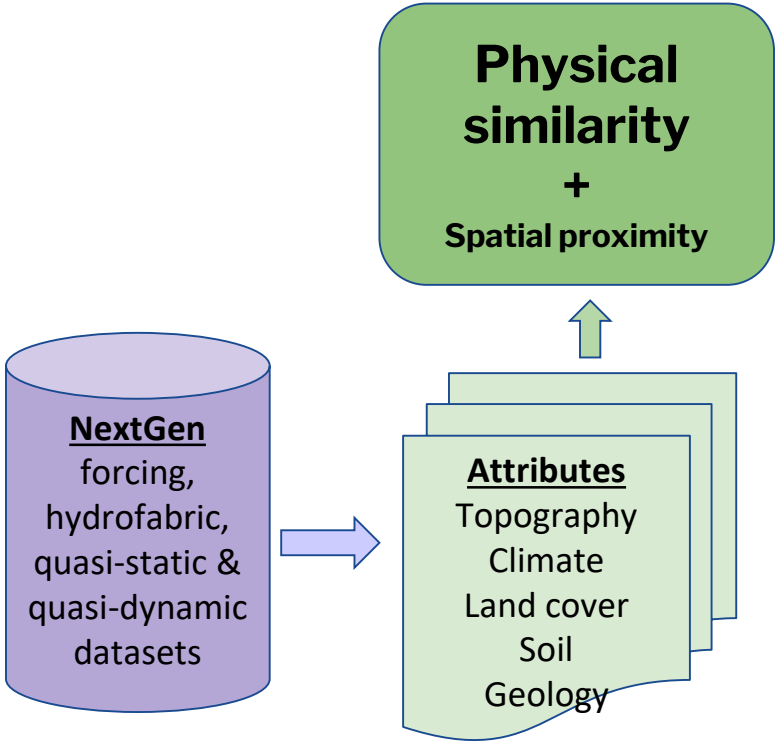
Calibrated NextGen model & parameters
(at calibration/donor basins)



Final regionalized NextGen model & parameters
(across the entire model domain)



Overarching approach to NextGen regionalization



How to compute and assess physical similarities between receivers & donors and then properly pair them?

Regionalization Algorithms!

Regionalization algorithms: current capability

- Classical clustering algorithm
- Easy to implement
- sensitive to outliers

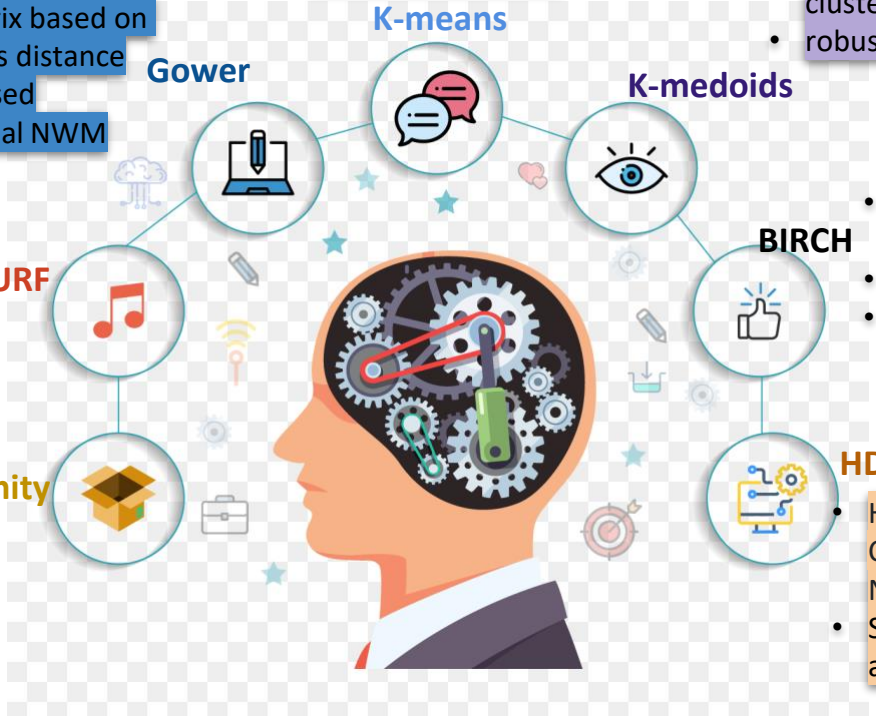
K-means

- Variant of K-means clustering
- robust to noise and outliers

K-medoids

- Dissimilarity matrix based on weighted Gower's distance
- not clustering based
- used in operational NWM

Gower



URF

- Unsupervised random forest (machine learning)
- Robust & accurate for complex problems
- Key output: dissimilarity matrix

BIRCH

- Balanced Iterative Reducing & Clustering using Hierarchy
- Suitable for large datasets
- Does not scale well to high dimensional data

Proximity

- easy baseline
- pure spatial proximity
- no physical similarity

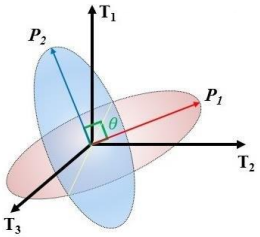
HDBSCAN

- Hierarchical Density-Based Spatial Clustering of Applications with Noise
- Suitable for datasets with outliers and clusters with varying densities

Implementation of NextGen regionalization algorithms

All algorithms are python-based

- machine learning library scikit-learn
- Built-in parallelization
- Modular & flexible framework to support expansion



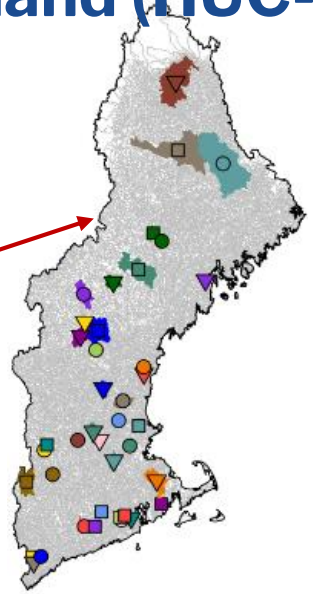
Inputs (catchment attributes) processing:

- Principal component analysis
- Remove interdependence
- Reduce data dimension

Built-in iterative search procedure:

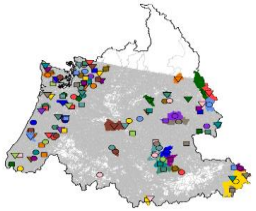
- To handle data gaps
- To identify donors with optimal combination of physical similarity & spatial proximity

Study Region: New England (HUC-01)

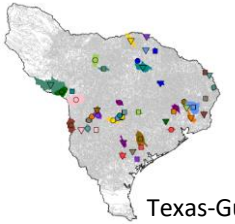


- Total: **20256** NextGen catchments
- **40** calibration basins (colored symbols in map)
- **2293** calibrated catchments (colored areas in map)
- **17974** uncalibrated catchments

Side note: similar analyses were also conducted for two other regions: Pacific Northwest and Texas-Gulf but will not be discussed here



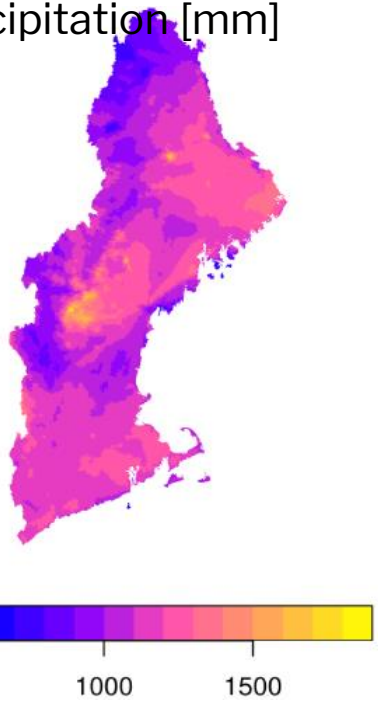
Pacific Northwest (HUC-17)



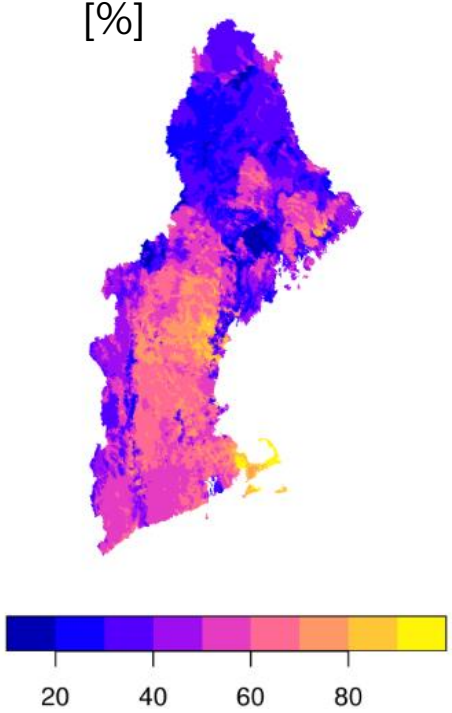
Texas-Gulf (HUC-12)

Spatial distribution of attributes (examples)

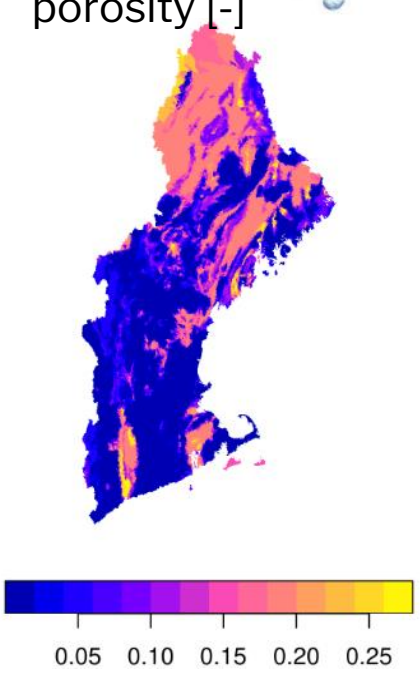
Mean annual precipitation [mm]



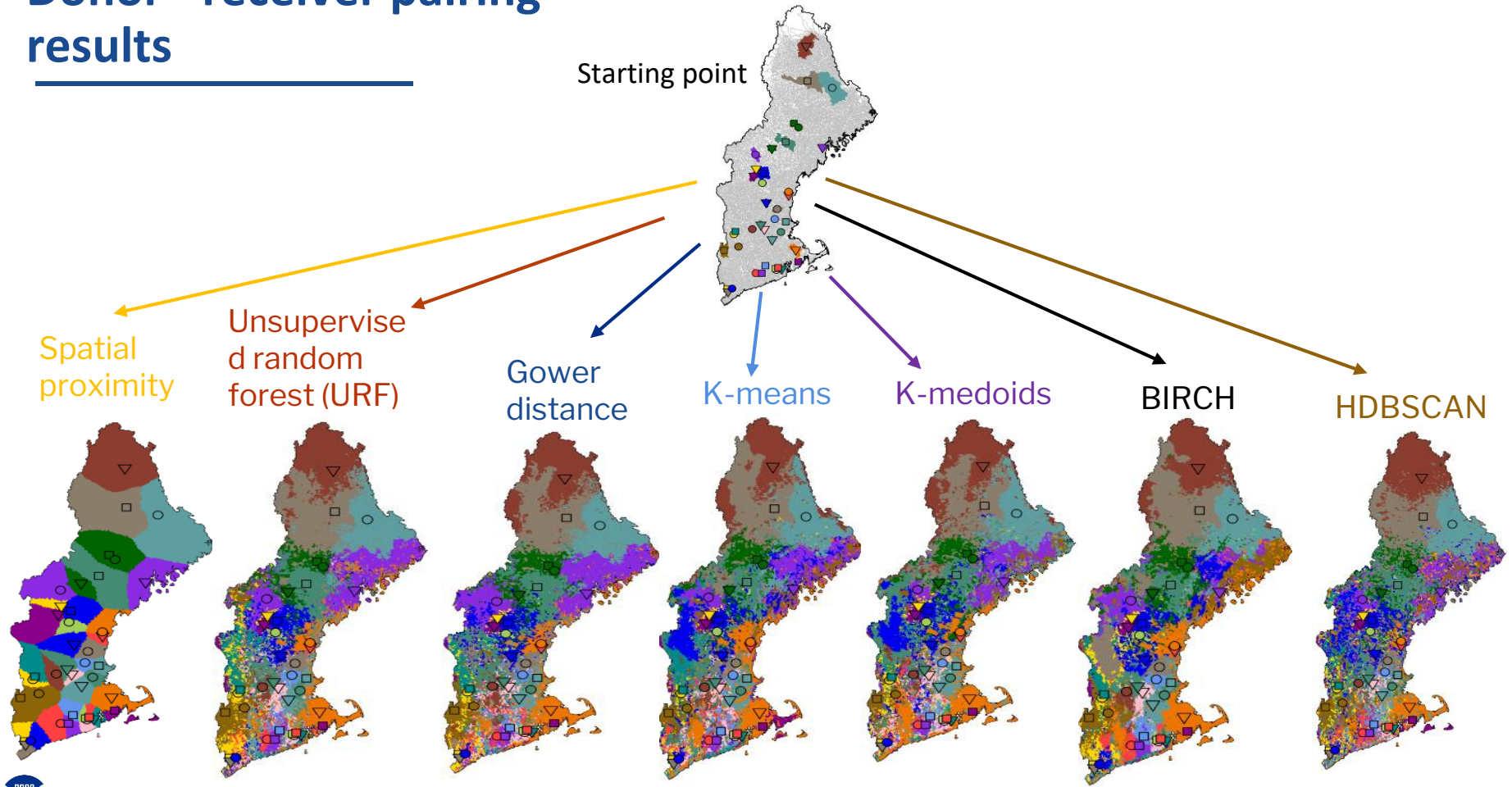
sand fraction in soil [%]



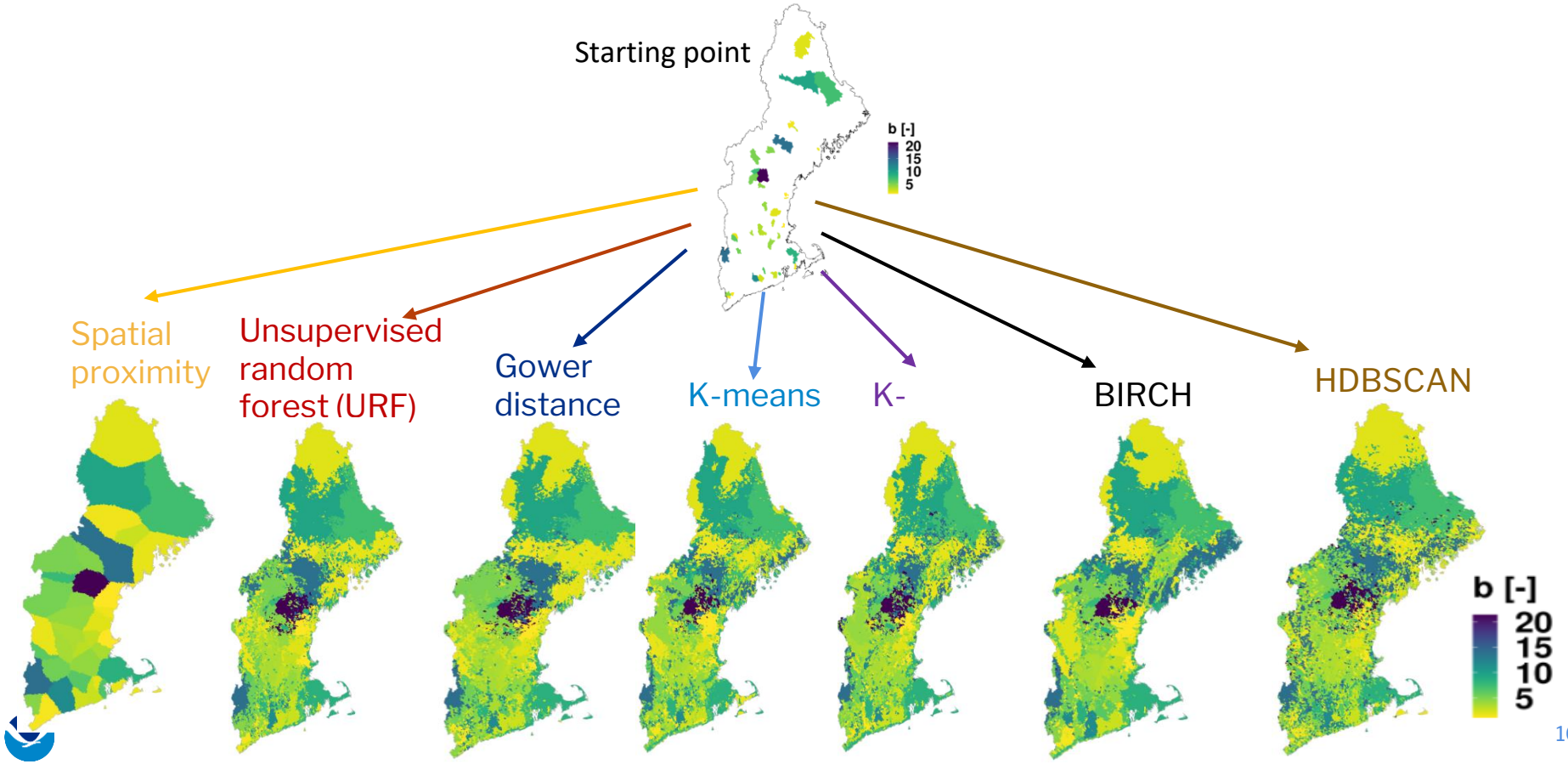
Subsurface porosity [-]



Donor - receiver pairing results



Regionalized parameter example: **b** - soil pore size distribution index

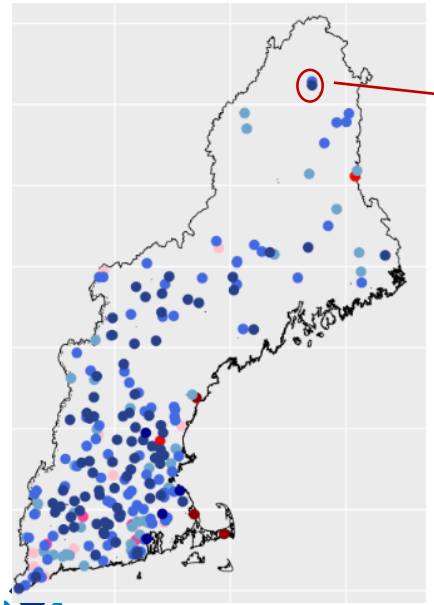


Preliminary assessment with streamflow simulations

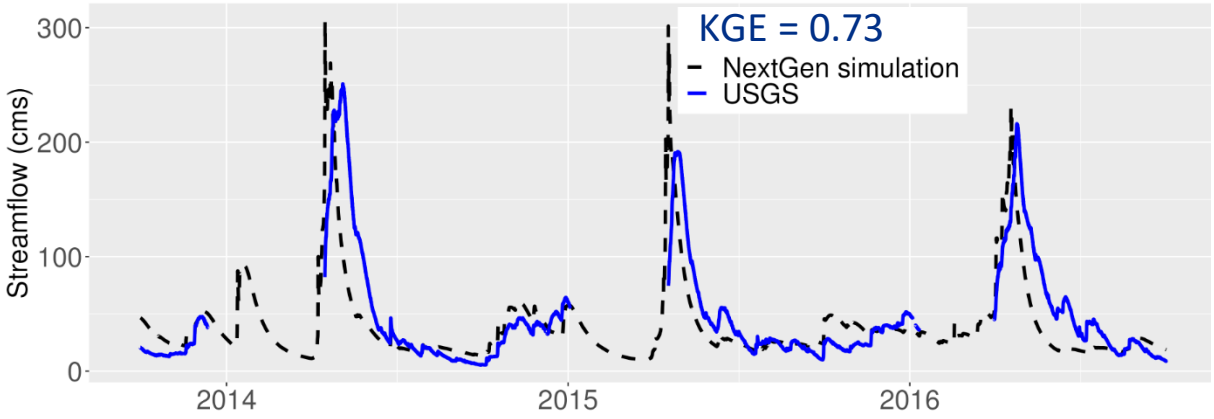
KGE from **independent validation** at 200+ uncalibrated basins

Experimental setup

- ✔ Regionalization algorithm: **Gower**
- ✔ NextGen formulation: mosaic **CFE + Topmodel**
- ✔ Validation period: **WY 2014-2016**



Independent Validation at 01013500
(Fish River near Fort Kent, Maine)



Concluding remarks

- Developed a python-based regionalization functionality with multiple algorithms for NextGen
- Preliminary streamflow assessment indicates successful regionalization of NextGen modules/parameters
- Future work: enhance regionalization functionalities & expand implementation to other regions

Thank You!

Yuqiong Liu

Yuqiong.Liu@noaa.gov

https://water.noaa.gov

NOAA

OWP | OFFICE OF WATER PREDICTION