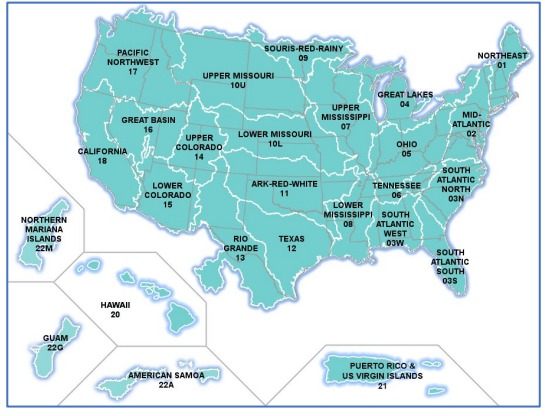
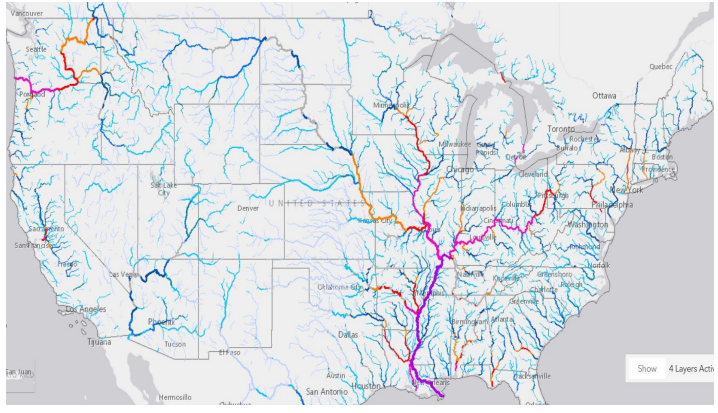
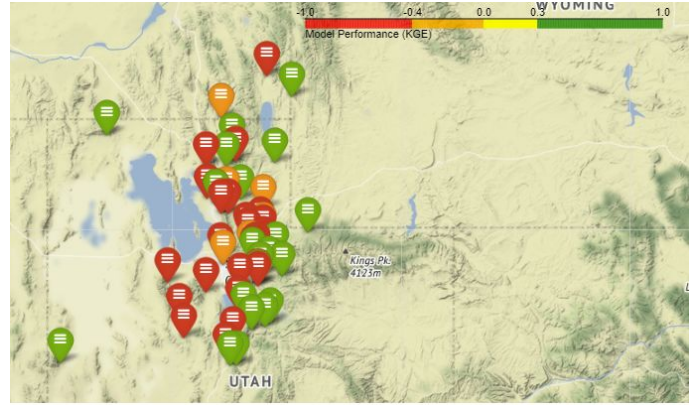
Investigating the Roles of Basin Attributes and Hydrologic Signatures in Predicting Extreme Hydrological Events using the National Water Model







Md. Shahabul Alam[1], Sifan A. Koriche[1], Ahmad J. Khattak[2], Ryan Johnson[3], Xia Feng[1], Fred L. Ogden[4], Mukesh Kumar[1], Josh Cunningham[1], Savalan Neisary[1], James Halgren[1], Arpita Patel[1], Trupesh Patel[1], Steven Burian[1]

[1]The University of Alabama
[2]Lynker
[3]The University of Utah
[4]NOAA OWP

AGU24
Washington, D.C. | 9–13 December 2024
WHAT'S NEXT FOR SCIENCE

December 10, 2024

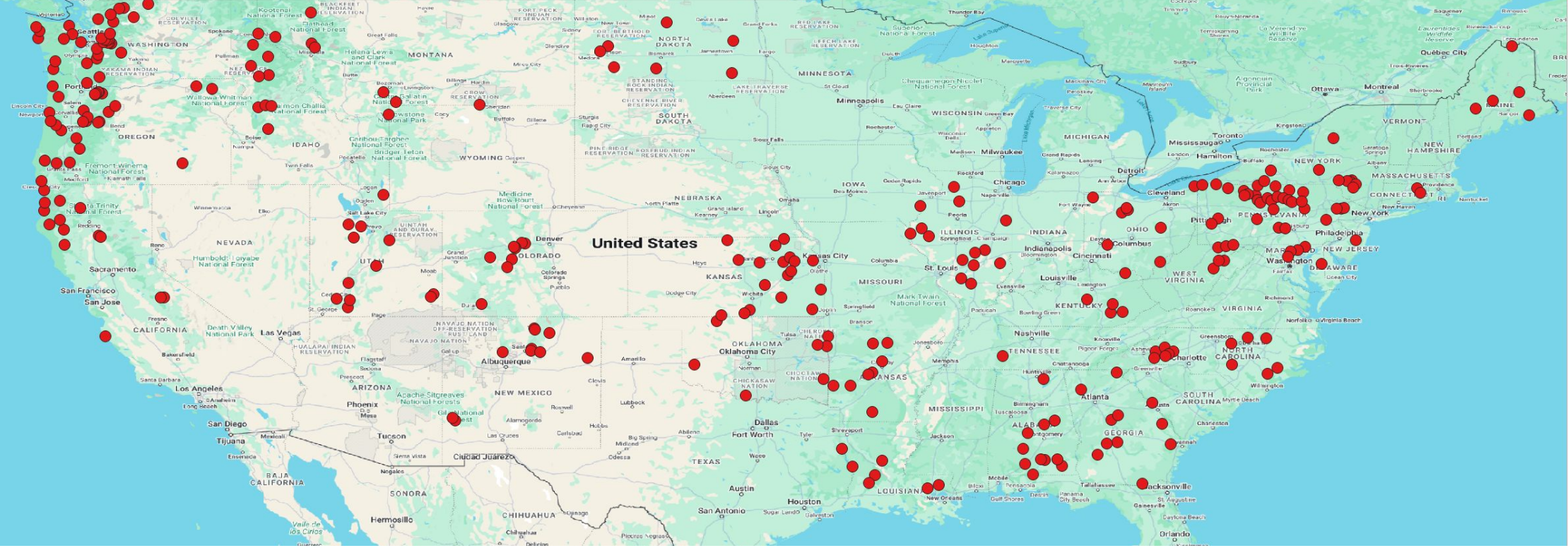
NOAA
OWP | OFFICE OF WATER PREDICTION
THE UNIVERSITY OF ALABAMA
THE UNIVERSITY OF ALABAMA Alabama Water Institute Research What Matters Most.
CIROH

Background

- **The National Water Model (NWM)** effectively predicts catastrophic flooding, with performance varying across basin scales (Viterbo et al., 2020)
- Predicting rare and extreme **flood** as well as **drought events** remains a challenge
- **Traditional** hydrological model evaluation reveals system behavior but struggles with capturing **impactful extreme events** like **floods and droughts**
- Understanding the relative role of **basin attributes and hydrologic signatures** is key to improving hydrologic models' capabilities
- **ML methods**- such as **Random Forests (RF)** algorithm has shown promising performance in identifying and analyzing the relative importance

Objectives

- **Identify when (season)** the NWM predicts extreme events better
- **Identify where (region)** the NWM predicts extreme events better
- Determine which extreme events (**floods or droughts**) the NWM predicts more accurately
- **Investigate what (features)** causes the model to perform better
- **Analyze how the features** impact the model performance

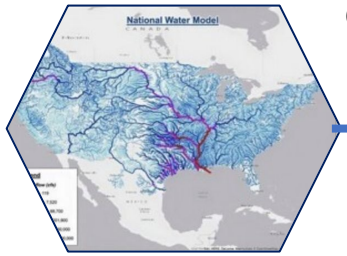


Study sites

- ~300 CAMELS basins across CONUS
- Collocated with the NWM reaches
- CAMELS data comes with basin attributes and hydrologic signatures

Methodology

NWM



NHDPlus
Cross-walk
Hourly to daily streamflow
1980-2020

NWIS

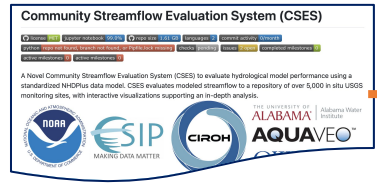
USGS

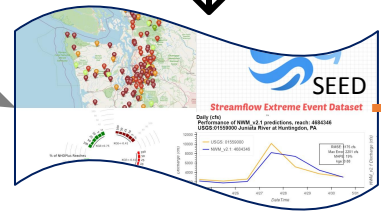
dataRetrieval
15-min to daily streamflow

streamflow
.h5

aws

S3





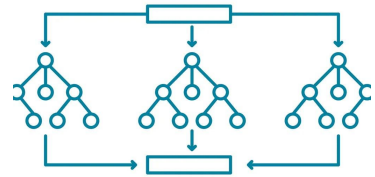
Select flood and drought events

KGE

CAMELS Dataset (Addor et al., 2017)

Soil
Topography
Geology

Climate
Hydrologic Signatures
Vegetation



Random Forests

Feature Importance

SHAP Analysis

*SHAP: SHapley Additive exPlanations

Flood: f(Annual Maximum Flow, T)
Drought: f(Annual Minimum Flow, T)

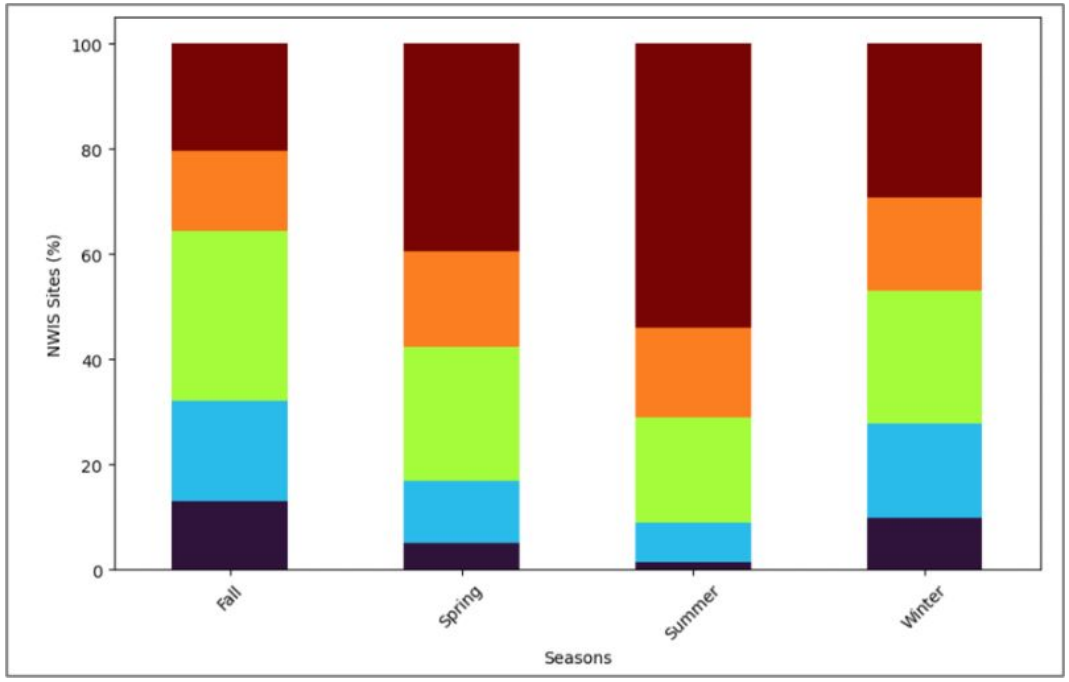
CIROH DocuHub

CSES GitHub

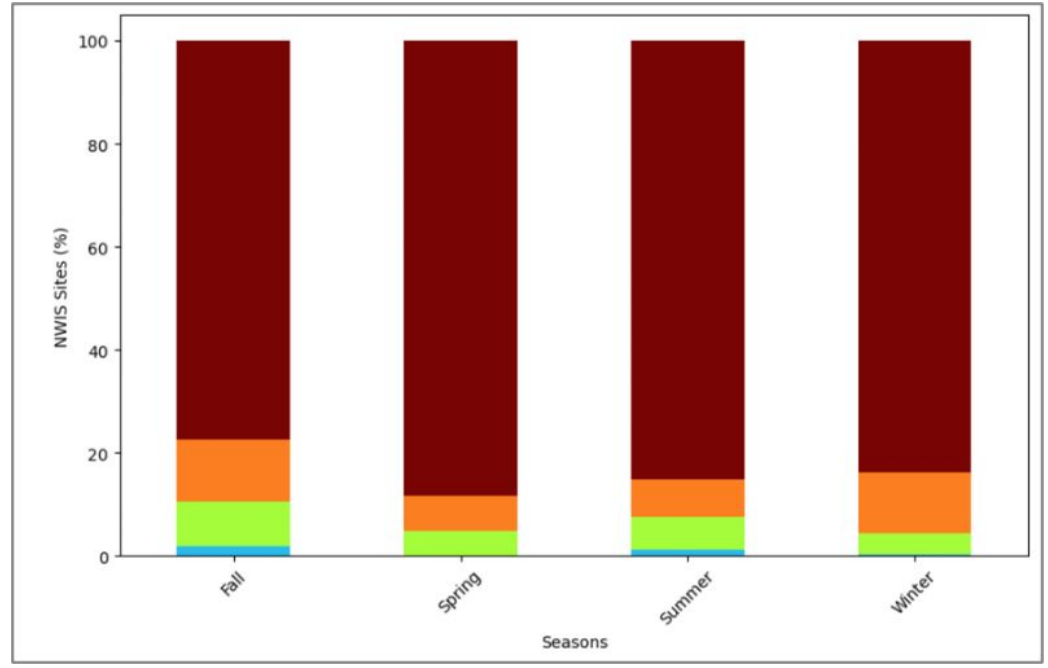
SEED-CSES GitHub

Results: Seasonality affecting NWM performance
(When does the model perform better?)

Flood predictions



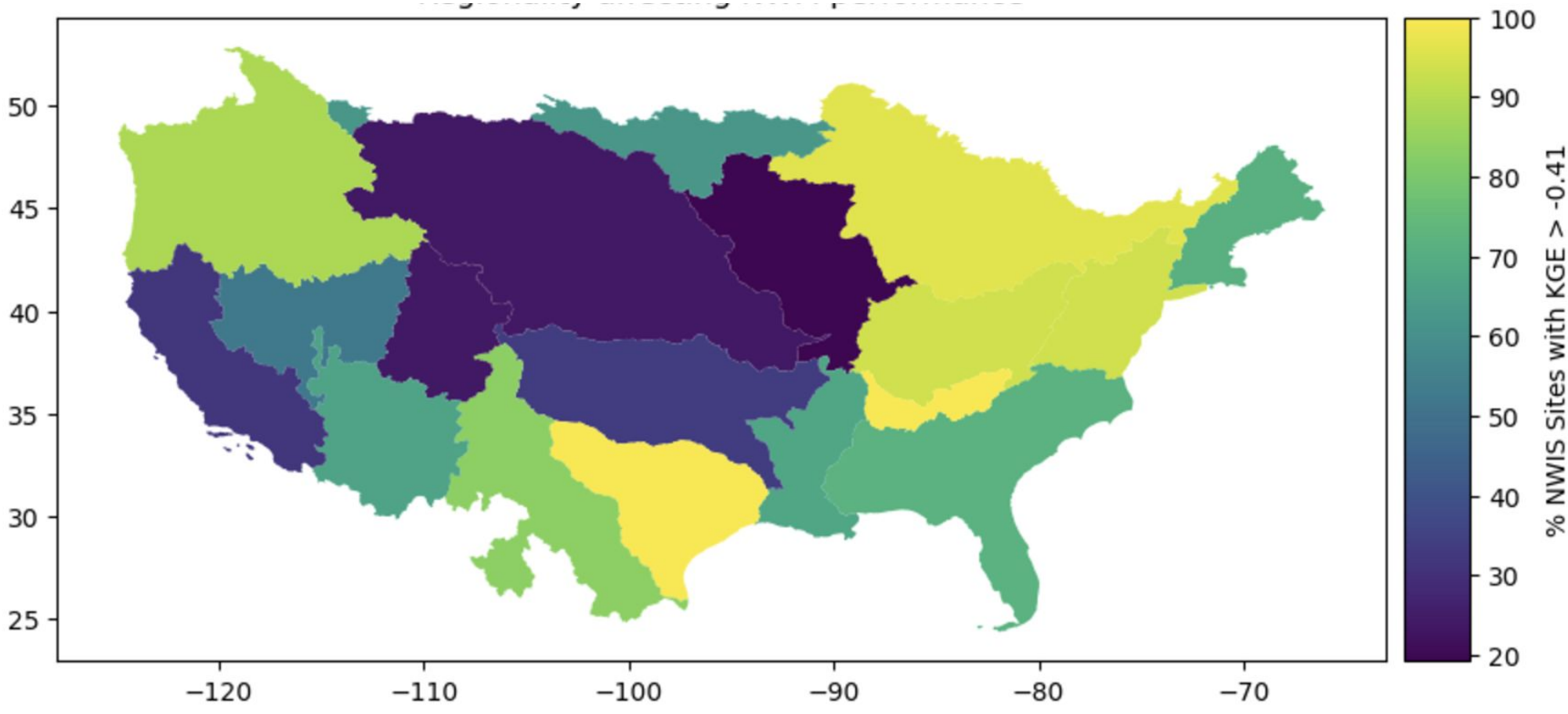
Drought predictions



Winter: DJF
Spring: MAM
Summer: JJA
Fall: SON

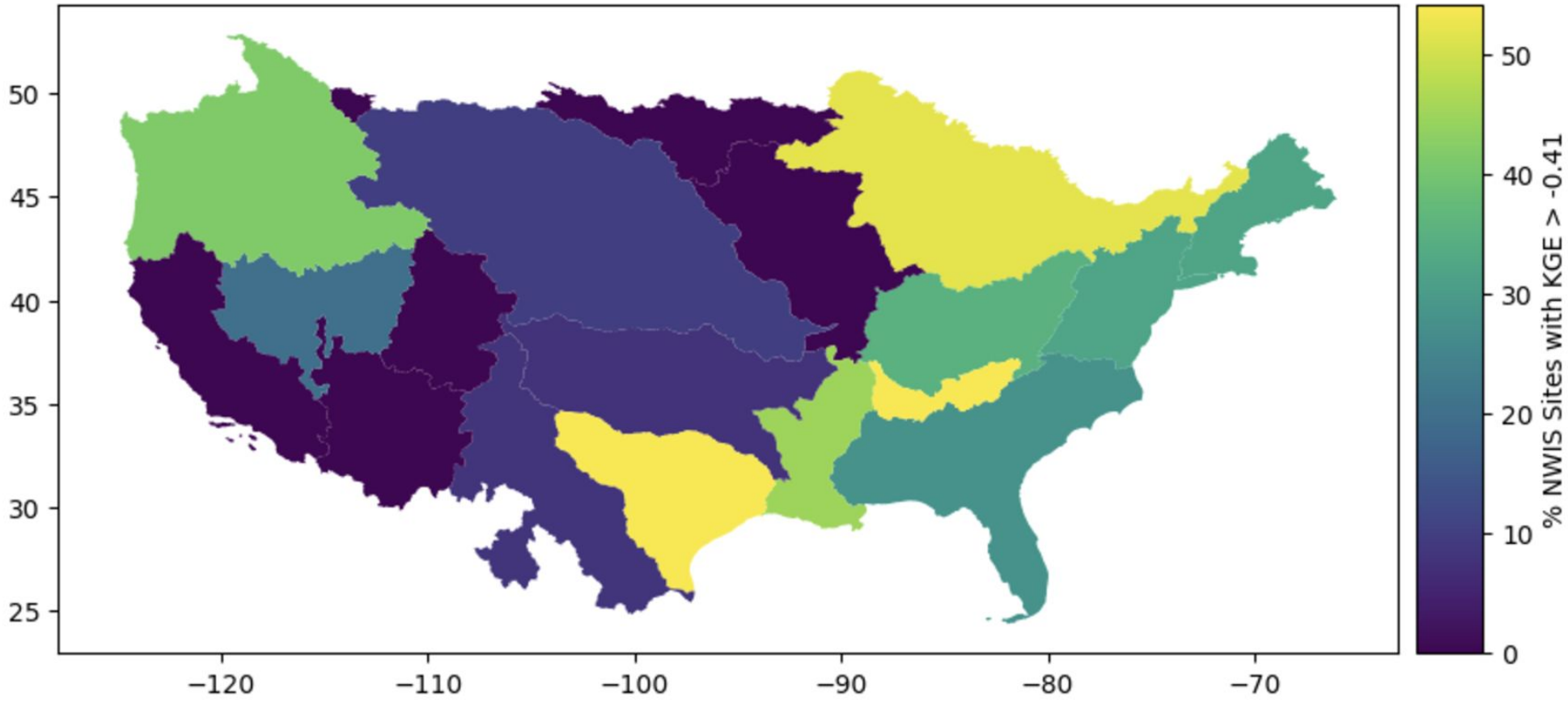
KGE Category
- Very good
- Good
- Fair
- Poor
- Unacceptable

Results: Regionality affecting NWM performance
(Where does the model perform better?)



Flood predictions

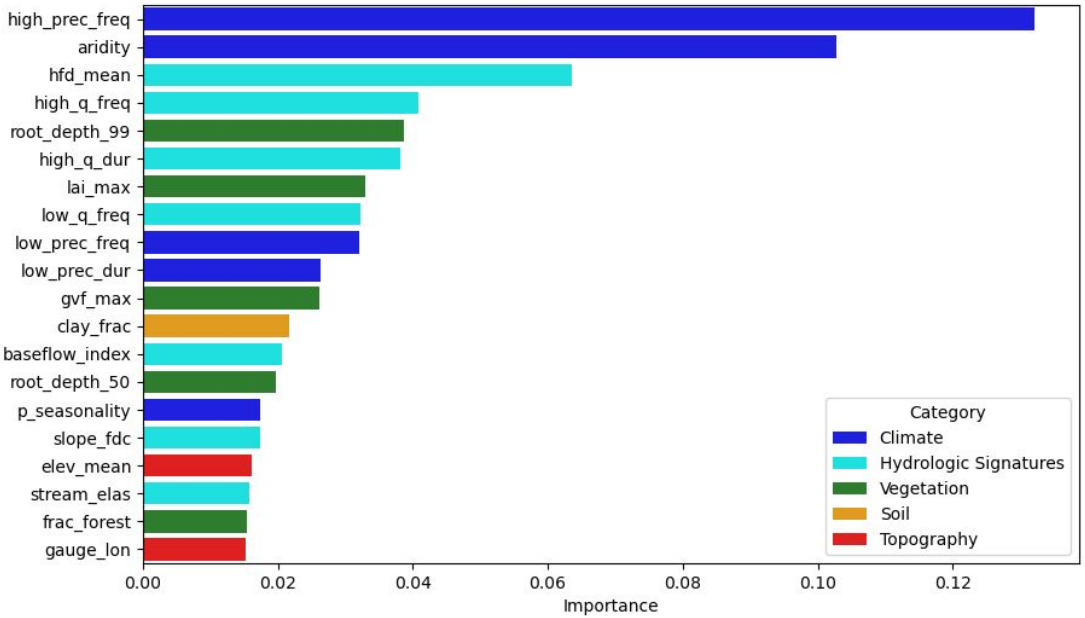
Results: Regionality affecting NWM performance
(Where does the model perform better?)



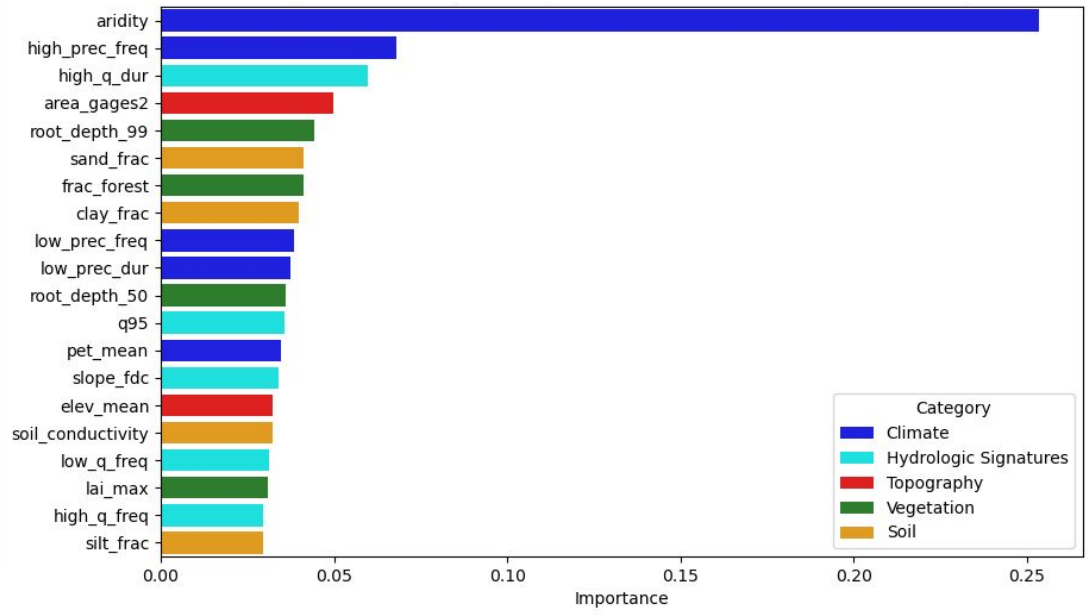
Drought predictions

Results: Important features affecting NWM performance
(What causes the model to perform better?)

2-Year Flood Events



100-Year Flood Events



Top 5 features include:

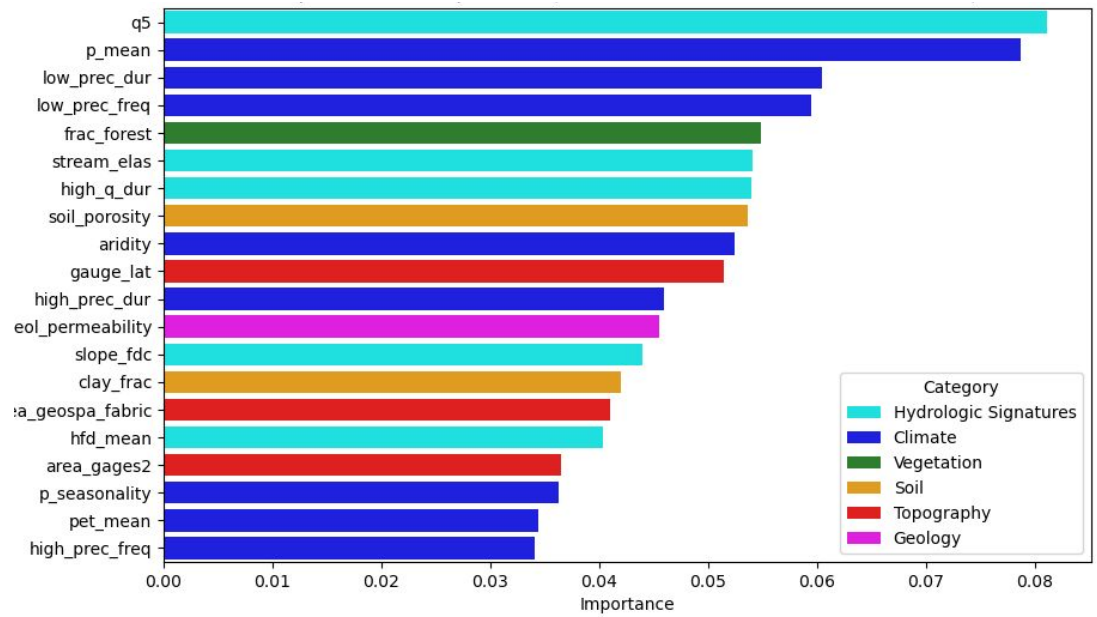
- Climate
- Hydrologic signatures
- Vegetation

Top 5 features include:

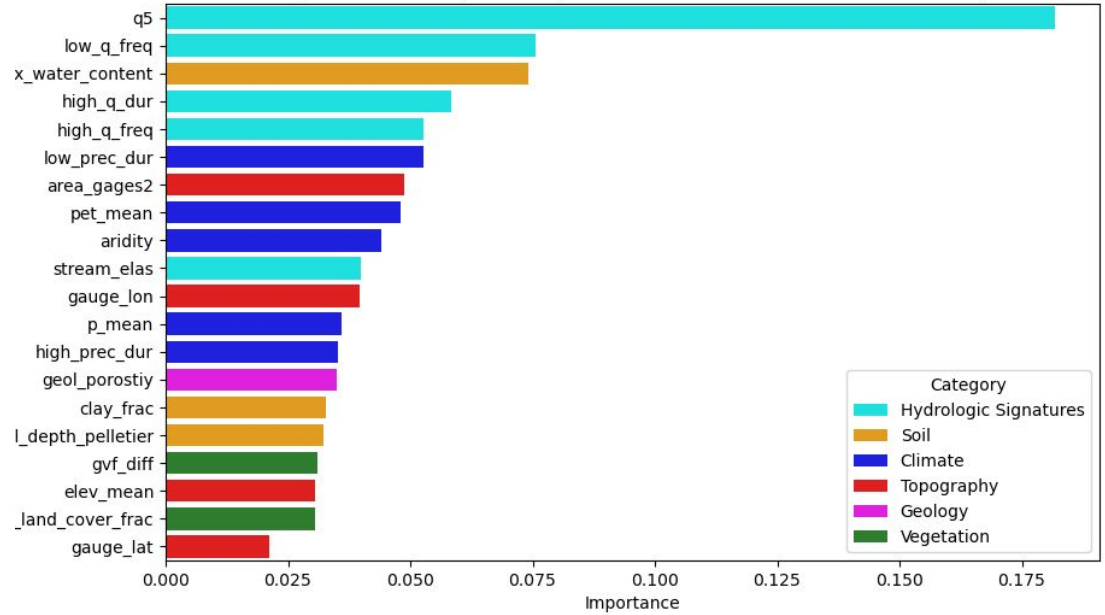
- Climate
- Hydrologic signatures
- Topography
- Vegetation

Results: Important features affecting NWM performance
(What causes the model to perform better?)

2-Year Drought Events



100-Year Drought Events



Top 5 features include:

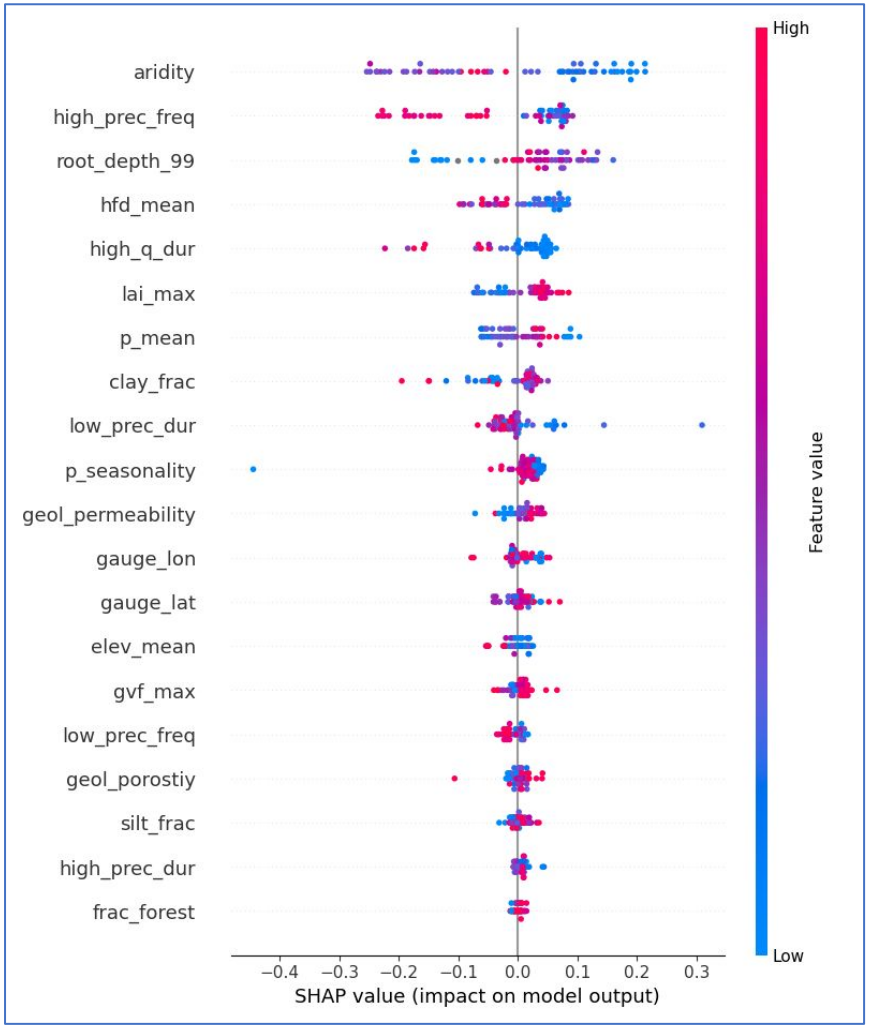
- Climate
- Hydrologic signatures
- Vegetation

Top 5 features include:

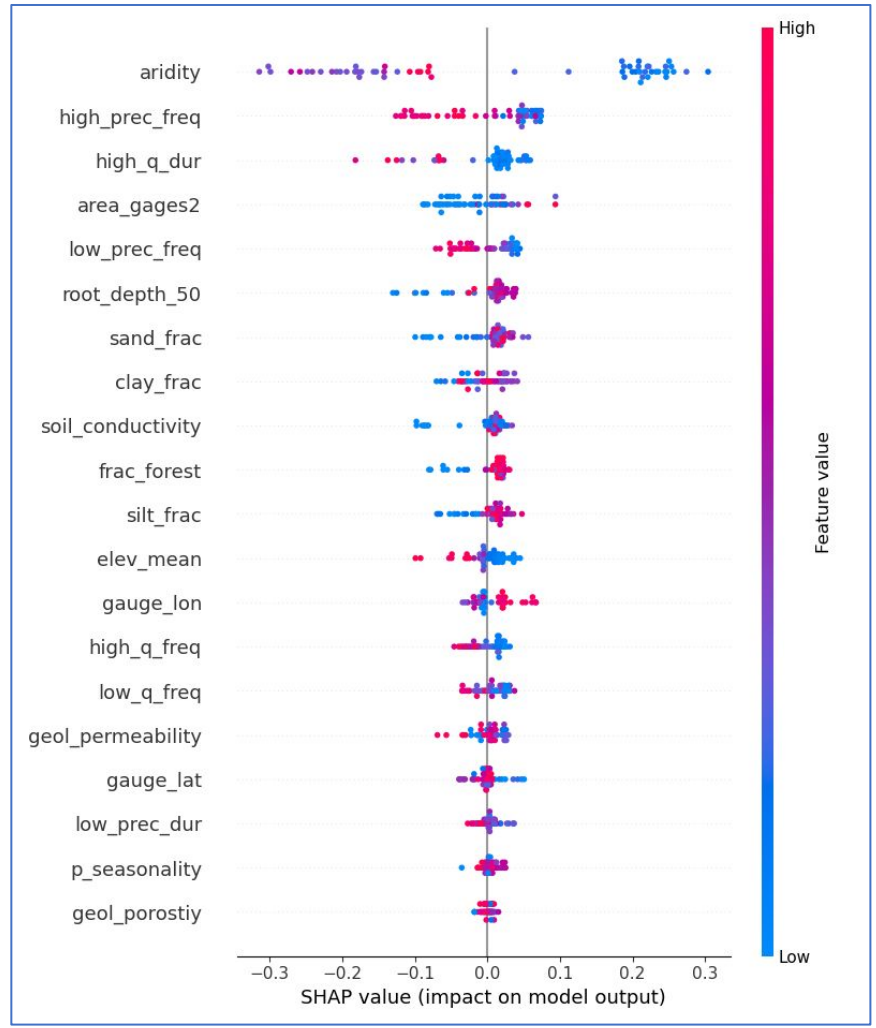
- Hydrologic signatures
- Soil

Results: SHAP analysis to interpret feature importance

(Why does the model perform better?)



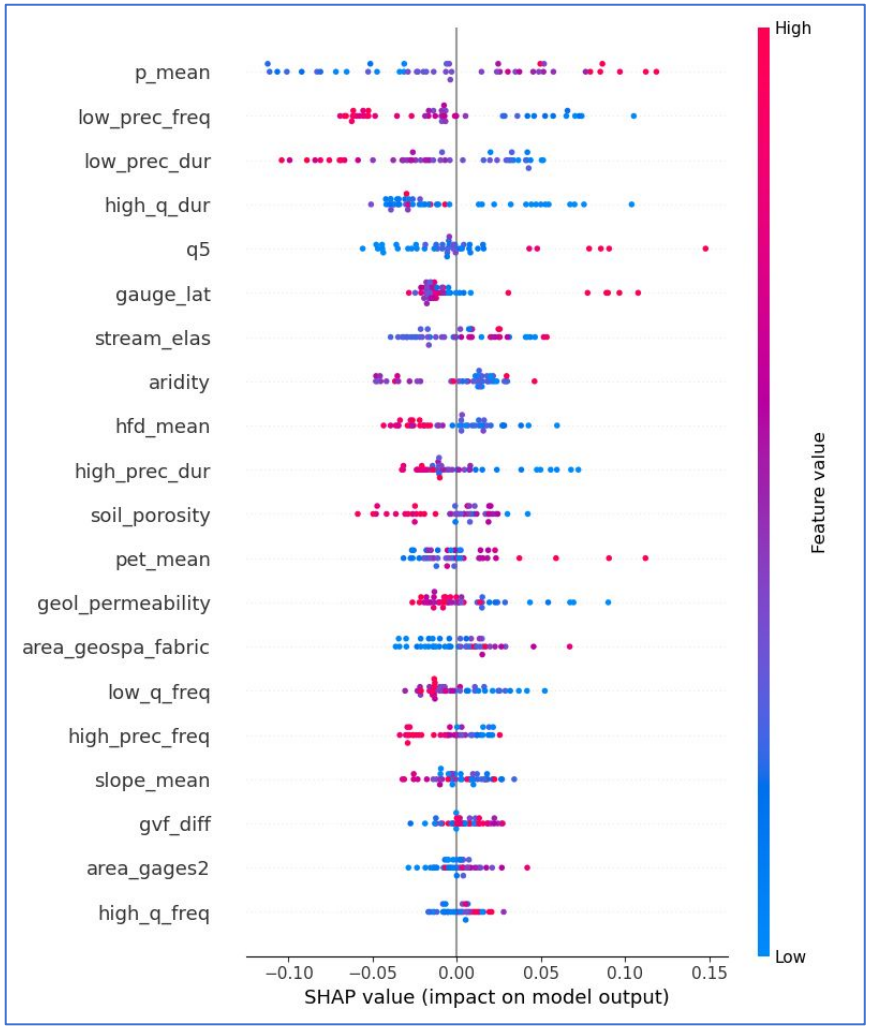
2-Year Flood Events



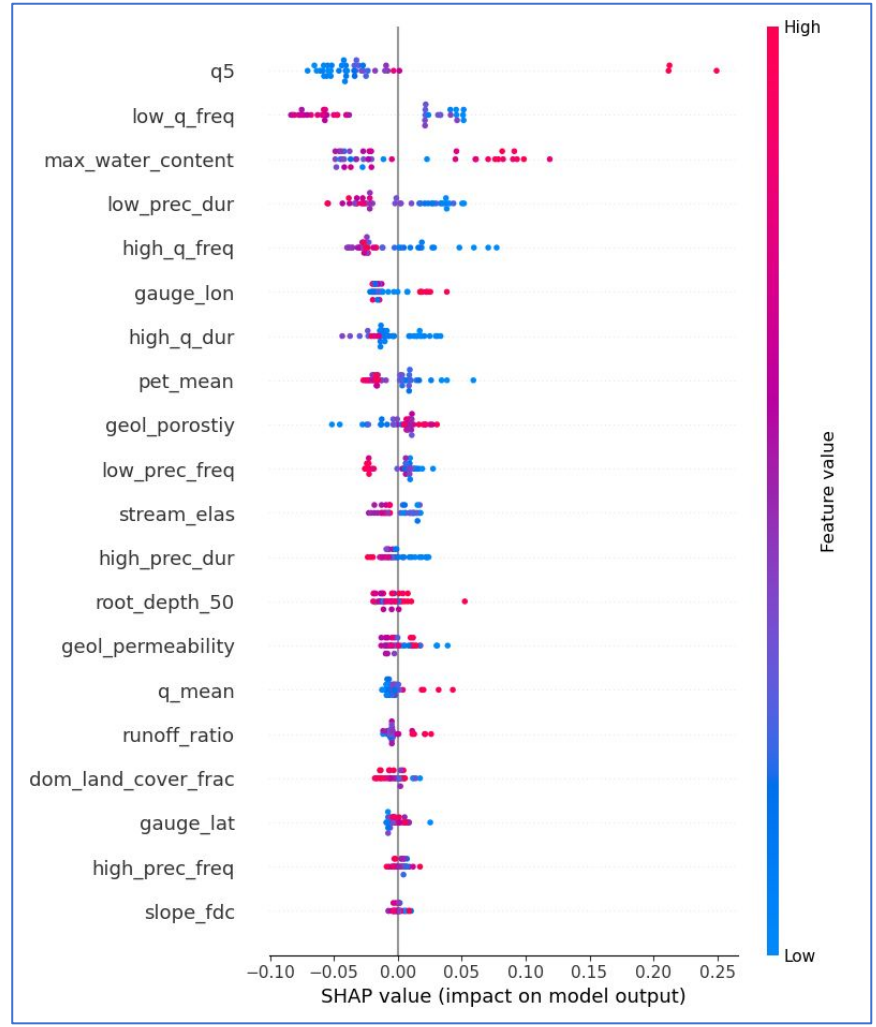
100-Year Flood Events

Results: SHAP analysis to interpret feature importance

(Why does the model perform better?)



2-Year Drought Events



100-Year Drought Events

Conclusions

- The NWM v2.1 performs better in the **Fall season** followed by the **Winter**
- The model performs relatively better in the **east, northeast, and southeast regions**
- The NWM performs better in predicting **flood events than the drought events**
- **Climate** (aridity), **hydrologic signatures** (frequency of high-flow days), and **vegetation** (forest fraction) are among the most important features
- **SHAP analysis** shows how each feature affects the model performance, e.g., **high aridity leads to poor model performance**

Next steps

- Include different methods of **flood and drought estimation** in SEED
- Application of SEED-CSES to other **versions of NWM** as well **NextGen-based models** (e.g., CFE)
- Couple the **forcing and soil moisture** data into the SEED events for process understanding
- Investigate how **regionality and seasonality** play in defining US hydroclimate regimes in NextGen-based hydrologic models using ML methods

THE UNIVERSITY OF ALABAMA

CIROH

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

NOAA