



Sandia National Laboratories

# Machine Learning and Reservoir Computing for Watershed-Scale Drought Prediction

Nicole D. Jackson, Ph.D.

Climate Security Center

American Geophysical Union Fall Meeting 2024

09 December 2024



Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

SAND2024-16489C



# Acknowledgements



The EPIC PLANN (Enhanced Prediction in Chaotic Systems with Physics-learned Autonomous Neural Networks) team



- William Chapman<sup>†</sup>
- Josh Mott<sup>‡</sup>
- J. Darby Smith
- Corinne Teeter

‡



- Louis Scuderi
- Zach Strasberg<sup>‡</sup>

<sup>†</sup> postdoc, <sup>‡</sup> Ph.D. student



Sandia Earth Science Research Foundation

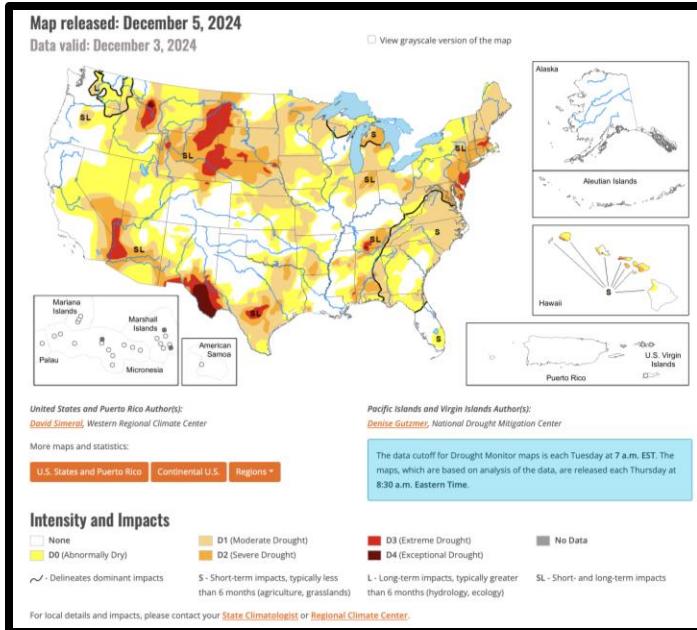
- Program Manager: Stephanie Kuzio

# The US Drought Monitor provides weekly drought estimates



## U.S. Drought Monitor

Last week's report



Source: [U.S. Drought Monitor](#)

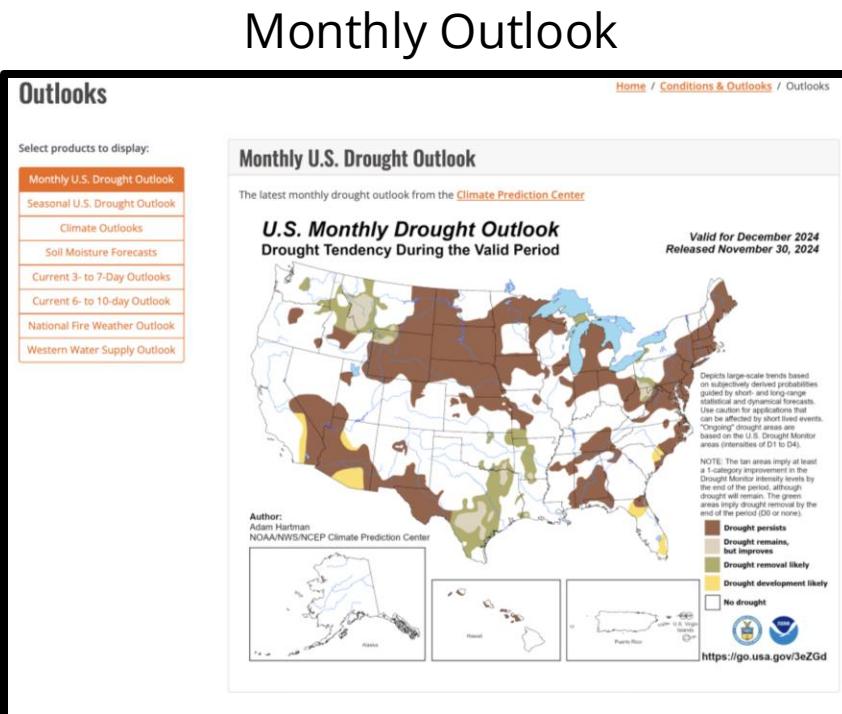
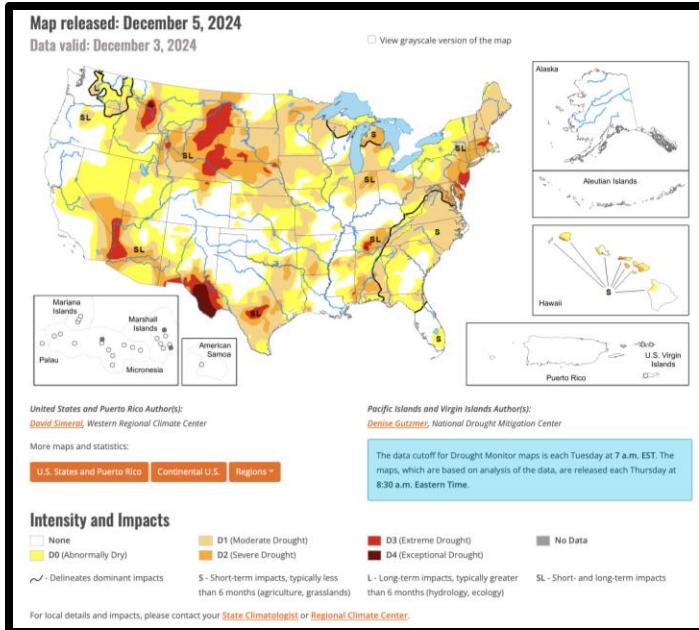


# The US Drought Monitor provides weekly drought estimates as well as monthly forecasts



## U.S. Drought Monitor

Last week's report



Source: [U.S. Drought Monitor](#)

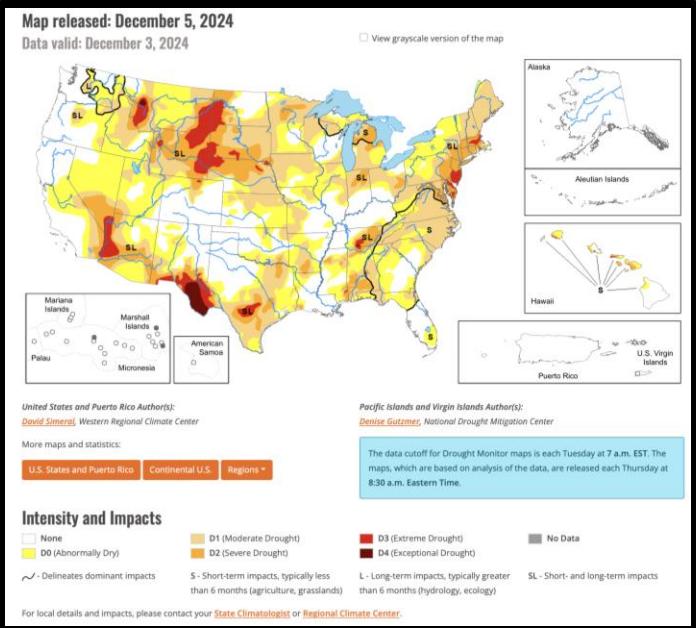


# The US Drought Monitor provides weekly drought estimates as well as monthly and seasonal forecasts

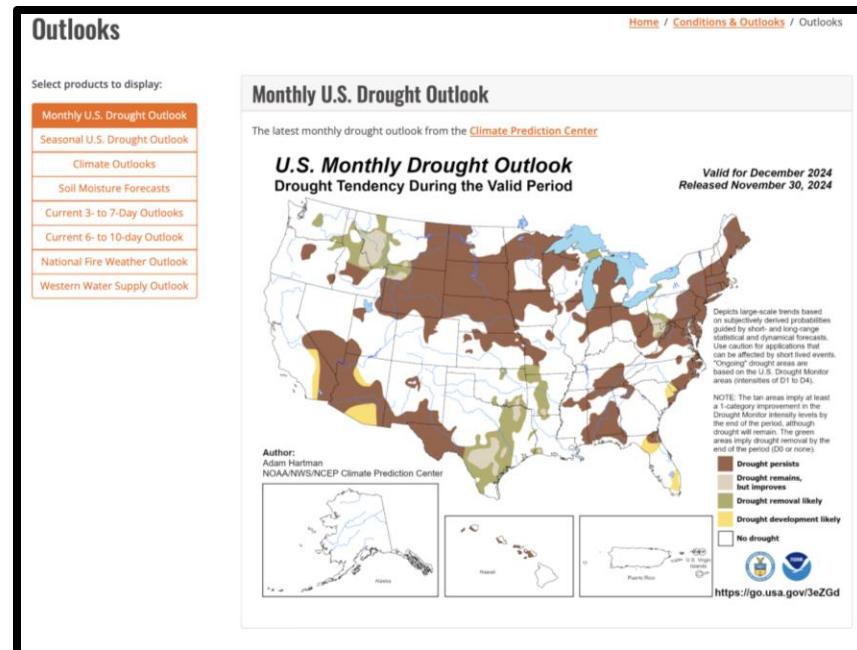


## U.S. Drought Monitor

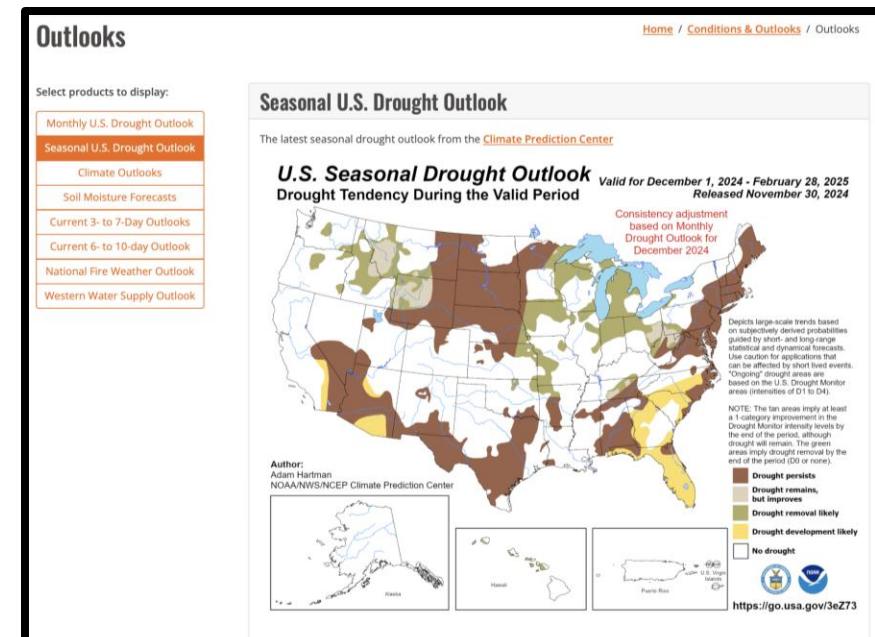
Last week's report



## Monthly Outlook



## Seasonal Outlook



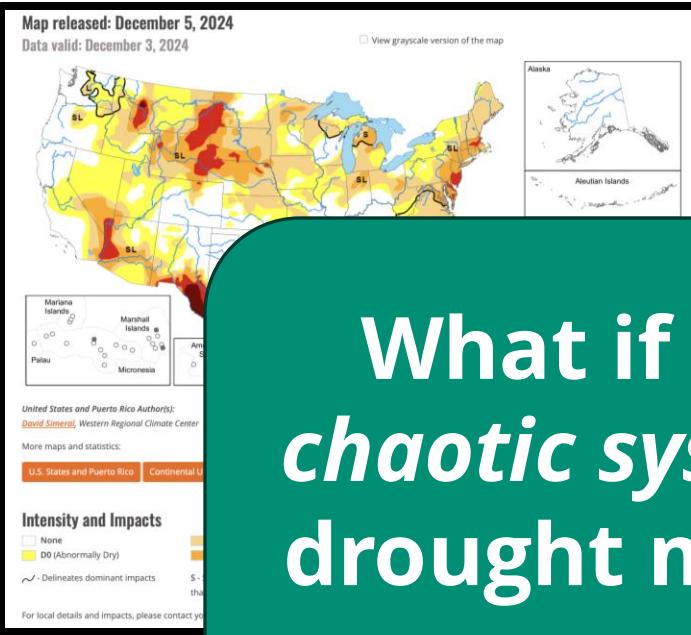
Source: [U.S. Drought Monitor](#)

The US Drought Monitor provides weekly drought estimates as well as monthly and seasonal forecasts

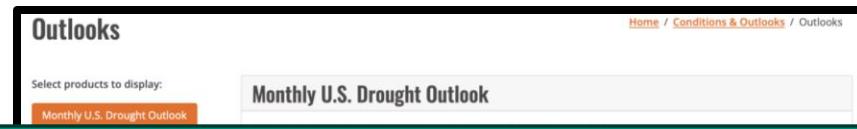


## U.S. Drought Monitor

Last week's report



### Monthly Outlook



### Seasonal Outlook



What if we could use *machine learning* and *chaotic system modeling* to accurately forecast drought much farther (e.g., 1-2 years) into the future?



Source: [U.S. Drought Monitor](https://go.usa.gov/3eZ73)

Multiple machine learning—based algorithms have been used to forecast meteorological drought



# A Deep Learning Based Approach for Long-Term Drought Prediction

Norbert A Agana  
Department of Electrical and  
Computer Engineering  
North Carolina A&T State University  
Greensboro, NC  
Email: naagana@aggies.ncat.edu

at 58 weather-related disasters recorded within a year as a result of droughts and other related events. The ability to design models that can make predictions concerning a significant progress in drought assessment. However, the number of drought variables makes accurate drought management a challenging scientific problem. Efforts conducted in recent years have proposed drove drought forecasting [7][8][9][10][11]. In their proposed model, based on a Bayesian approach, characterized hydrological conditions across the Texas River Basin. The authors used standardized  $z$  (SSI) for their analysis. In [12], a wavelet-programming (WLPG) model was used for drought forecasting (with 3, 6, and 12-months the state of Texas). They demonstrated that the WLPG model is more effective to keep an accurate of drought in lead times more than an autoregressive integrated moving average (ARIMA) or a multistep neural network (RMSNN) as multistep neural network (MSNN) were also

*Index Terms*—Deep Belief Network, unsupervised pre-training

Source: [Agana and Homaifar \(2017\)](#)

- Multilayer Perceptron (MLP)
- Support Vector Regression (SVR)

The flowchart is titled 'Robust Meteorological Drought Prediction Using Antecedent SST Fluctuations and Machine Learning'. It starts with 'Antecedent SST Fluctuations' on the left, which points to 'Machine Learning' (ML) and 'Random Forest'. 'Machine Learning' then branches into 'Support Vector Machine' (SVM) and 'Extreme Learning Machine' (ELM). 'Support Vector Machine' leads to 'Support Vector Regression' (SVR) and 'Random Forest'. 'Extreme Learning Machine' leads to 'Support Vector Machine' (SVM) and 'Extreme Learning Machine' (ELM). 'Support Vector Regression' and 'Random Forest' both lead to 'Drought Prediction'. 'Support Vector Machine' (SVM) and 'Extreme Learning Machine' (ELM) both lead to 'Drought Prediction'. 'Drought Prediction' leads to 'Drought Mitigation' and 'Water Resource Management'.

Source: [Li et al. \(2020\)](#)

- SVR
- Random Forest (RF)
- Extreme Learning Machine (ELM)

IEEE Access

Source: Mokhtar et al. (2021)

- RF
- Extreme Gradient Boost (XGB)
- Convolutional neural network (CNN)
- Long-term short memory (LSTM)

Environmental Modelling and Software 149 (2022) 105327

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: [www.elsevier.com/locate/envsoft](http://www.elsevier.com/locate/envsoft)

## A review of machine learning methods for drought hazard monitoring and forecasting: Current research trends, challenges, and future research directions

Foyez Ahmed Prodhan<sup>a,b,\*</sup>, Jiahua Zhang<sup>a,b,c</sup>, Shaikh Shamim Hasan<sup>a</sup>,  
Til Prasad Pangali Shama<sup>a,b</sup>, Hasiba Pervin Mohana<sup>a</sup>

<sup>a</sup> Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100094, China

<sup>b</sup> University of Chinese Academy of Sciences, Beijing, 100049, China

<sup>c</sup> Department of Agricultural Extension and Rural Development, Bangladesh Shaikh Mehmud Aliwan Agricultural University, Gazipur, 1700, Bangladesh

<sup>d</sup> China Agriculture University, Beijing, 100083, China

### ARTICLE INFO

**Keywords:**  
Machine learning  
Deep learning  
Forecasting  
Drought  
Big data

### ABSTRACT

Machine learning is a dynamic field with wide-ranging applications, including drought modeling and forecasting. Drought is a complex, devastating natural disaster for which it is challenging to develop effective prevention measures. Machine learning forecasting is a promising approach to predict droughts and estimate their potential applications in developing effective and efficient drought forecasting models. We observed that MLMs have achieved significant advances in the robustness, effectiveness, and accuracy of the algorithms for drought modeling. This review summarizes the recent advances in the field of drought modeling and provides a new conception of different model evaluation metrics. Further challenges of MLMs, such as inadequate training data sets, noise, outliers, and observation bias for spatial data sets, are explored. Finally, our review covers the application of machine learning in drought forecasting and provides drought mitigation strategy guidance for policymakers.

Source: Prodhan et al. (2022)

## Meteorological Drought:

- Artificial neural network (ANN)
- SVR
- K-nearest neighbor
- Adaptive Neuro Fuzzy (ANF)
- SVM-Copula
- Wavelet Boosting-Support Vector Regression (WBS-SVM)

# Multiple machine learning-based algorithms have been used to forecast meteorological drought

## A Deep Learning Based Approach for Long-Term Drought Prediction

Noeber A Agana  
Department of Electrical and Computer Engineering  
North Carolina A&T State University  
Greensboro, NC  
Email: naagana@aggies.ncat.edu

Abdullah Homaifar  
Department of Electrical and Computer Engineering  
North Carolina A&T State University  
Greensboro, NC  
Email: homaifar@ncat.edu

**Abstract**—Drought is a natural disaster that comes with high uncertainty. It is a complex phenomenon that must be monitored as hydrological drought. Identifying past droughts and predicting future ones is very vital in limiting their effects. However, the random and nonlinear nature of drought variables makes accurate drought prediction a challenging scientific problem. Neural networks have shown great promise over the last two decades in modeling nonlinear time series, but the issue of nonlinear optimization of the hidden layers or the number of layers are required for highly complex phenomena. This research looks into the drought prediction problem using deep learning algorithms to propose a Deep Belief Network (DBN) using a Restricted Boltzmann Machine (RBMs) for drought prediction using lagged values of Standardized Streamflow Index (SSI) as inputs. The proposed model is applied to predict different time scales of hydrological droughts in the Gunnison River Basin in the Upper Colorado River Basin. The study compares the efficiency of the proposed model to that of traditional approaches such as Multilayer Perceptron (MLP), Support Vector Regression (SVR) for predicting the different time scale drought conditions. The proposed model shows an edge in performance over the other models in terms of Root Mean Square Error and Mean Absolute Error as metrics.

**Index Terms**—Deep Belief Network, unsupervised pretraining, hydrological drought, streamflow index, Gunnison River Basin.

Source: [Agana and Homaifar \(2017\)](#)

- Multilayer Perceptron (MLP)
- Support Vector Regression (SVR)

## Water Resources Research

### RESEARCH ARTICLE

#### Robust Meteorological Drought Prediction Using Antecedent SST Fluctuations and Machine Learning

Jun Li<sup>1,2</sup>, Zhaoli Wang<sup>2</sup>, Xushu Wu<sup>1,2</sup> , Cheng-Yu Xu<sup>1</sup> , Shenglan Guo<sup>4</sup>, Xiaodong Chen<sup>1,2</sup>, and Zhenxing Zhang<sup>2</sup>

<sup>1</sup>School of Civil Engineering and Transportation, State Key Laboratory of Subtropical Building Science, South China University of Technology, Guangzhou, China; <sup>2</sup>Guangdong Engineering Technology Research Center of Safety and Governance for Water Conservation Project, Guangzhou, China; <sup>3</sup>Department of Geosciences, University of Oslo, Oslo, Norway; <sup>4</sup>Center of Water Resources and Hydropower Engineering Sciences, Wuhan University, Wuhan, China; <sup>5</sup>Pearl River Research Institute, University of Illinois, Champaign, IL, USA

**Abstract** While reliable drought prediction is fundamental for drought mitigation and water resources management, it is still a challenge to develop robust drought prediction models due to complex factors that influence droughts. Machine learning algorithms have been widely used as the fundamental predictor to develop drought prediction models. However, traditional models usually extract SST signals from one or several specific zones within a given time span, which limits full use of SST signals for drought prediction. Here, we introduce a new meteorological drought prediction approach by using the antecedent SST fluctuation pattern (ASPF) and machine learning techniques (e.g., support vector regression, random forest, and extreme learning machine) to predict meteorological droughts. Three models (i.e., ASFS-SVR, ASFS-ELM, and ASFS-RF) are developed for ensemble, probability, and deterministic drought predictions. The Colorado, Danube, Orange, and Pearl River basins with frequent droughts over different continents are selected as the cases, where a standardized precipitation evapotranspiration index (SPEI) are predicted in the 1<sup>st</sup>–1<sup>st</sup> resolution with 1- and 3-month lead times. Results show that the ASFS model can effectively predict the time evolution of drought events with probability skills, outperforming the ASFS-SVR and ASFS-ELM models. Our study has potential to provide a reliable tool for drought prediction, which further supports the development of drought early warning systems.

### 1. Introduction

Source: [Li et al. \(2020\)](#)

- SVR
- Random Forest (RF)
- Extreme Learning Machine (ELM)

**Gap: forecasting for periods beyond 3-6 months**



AGU  
ADVANCING  
EARTH AND  
SPACE SCIENCE

IEEE Access  
A Multidisciplinary Journal

Received March 26, 2021; accepted April 7, 2021; date of publication April 20, 2021; date of current version May 6, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3104840

## Estimation of SPEI Meteorological Drought Using Machine Learning Algorithms

ALI MOKHTAR<sup>1,2</sup>, MOHAMMADNAJIB JALALI<sup>2</sup>, HONGMING HU<sup>1,4</sup>, NADHIR AL-ANSARI<sup>1,5</sup>,

SAAD SH. SAMMEN<sup>3</sup>, YEBODA GYASI-AGYEI<sup>1,11</sup>, AND JESÚS RODRIGO-COMINO<sup>1,12,13</sup>

<sup>1</sup>Sohbet Key Laboratory of Soil and Water Erosion and Dryland Farming on the Loess Plateau, Institute of Soil and Water Conservation, Chinese Academy of Agricultural Sciences, Yangling, Shaanxi 723105, China

<sup>2</sup>Department of Agricultural Engineering, Faculty of Agriculture, Cairo University, Giza 12613, Egypt

<sup>3</sup>School of Geographical Sciences, East China Normal University, Shanghai 200062, China

<sup>4</sup>College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310088, China

<sup>5</sup>College of Water Resources and Hydropower Engineering, China University of Geosciences, Wuhan, Wuhan, China

<sup>6</sup>College of Environmental and Resource Sciences, Yangzhou University, Yangzhou 225002, China

<sup>7</sup>Geography Department, Faculty of Arts and Humanities, Tanta University, Tanta 31527, Syria

<sup>8</sup>Department of Geology, Faculty of Science, Tanta University, Tanta 31527, Syria

<sup>9</sup>Department of Geology, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

<sup>10</sup>Department of Geology, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

<sup>11</sup>Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China

<sup>12</sup>College of Geodesy and Geomatics Engineering, China University of Geosciences, Beijing 100083, China

<sup>13</sup>China Agricultural University, Beijing 100083, China

IEEE

Access

10.1109/ACCESS.2021.3104840

10

Environmental Modelling and Software 149 (2022) 105327

Contents lists available at ScienceDirect

Environmental Modelling and Software  
journal homepage: [www.elsevier.com/locate/emosoft](http://www.elsevier.com/locate/emosoft)



A review of machine learning methods for drought hazard monitoring and forecasting: Current research trends, challenges, and future research directions

Foyez Ahmed Prodhan , Jiahua Zhang , Shaikh Shamim Hasan , Til Prasad Pangali Sharma , Hasiq Pervin Mohanna

<sup>1</sup>Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China

<sup>2</sup>College of Geodesy and Geomatics Engineering, China University of Geosciences, Beijing 100083, China

<sup>3</sup>Department of Geodesy and Geomatics Engineering, China University of Geosciences, Beijing 100083, China

<sup>4</sup>China Agricultural University, Beijing 100083, China

### ARTICLE INFO

Keywords:  
Machine learning  
Drought  
Forecasting  
Droughts  
Big data

**ABSTRACT** Machine learning is a dynamic field with wide-ranging applications, including drought modeling and forecasting. Drought is a complex, devastating natural disaster for which it is challenging to develop effective prediction models. Therefore, our review focuses on basic information about machine learning methods (MLMs) and their potential in drought modeling and forecasting. We also introduce the basic concepts of machine learning. We discuss the advantages achieved in terms of the robustness, effectiveness, and accuracy of the algorithms for drought modeling in recent years. The performance comparison of MLMs with other models provides a comprehensive comparison of the performance of different models. The effects of different parameters on the data sets, noise, outliers, and observation bias for spatial data sets, are explored. Finally, our review conveys indepth understanding to researchers on machine learning applications in forecasting and modeling and provides drought mitigation strategy guidance for policymakers.

Source: [Prodhan et al. \(2022\)](#)

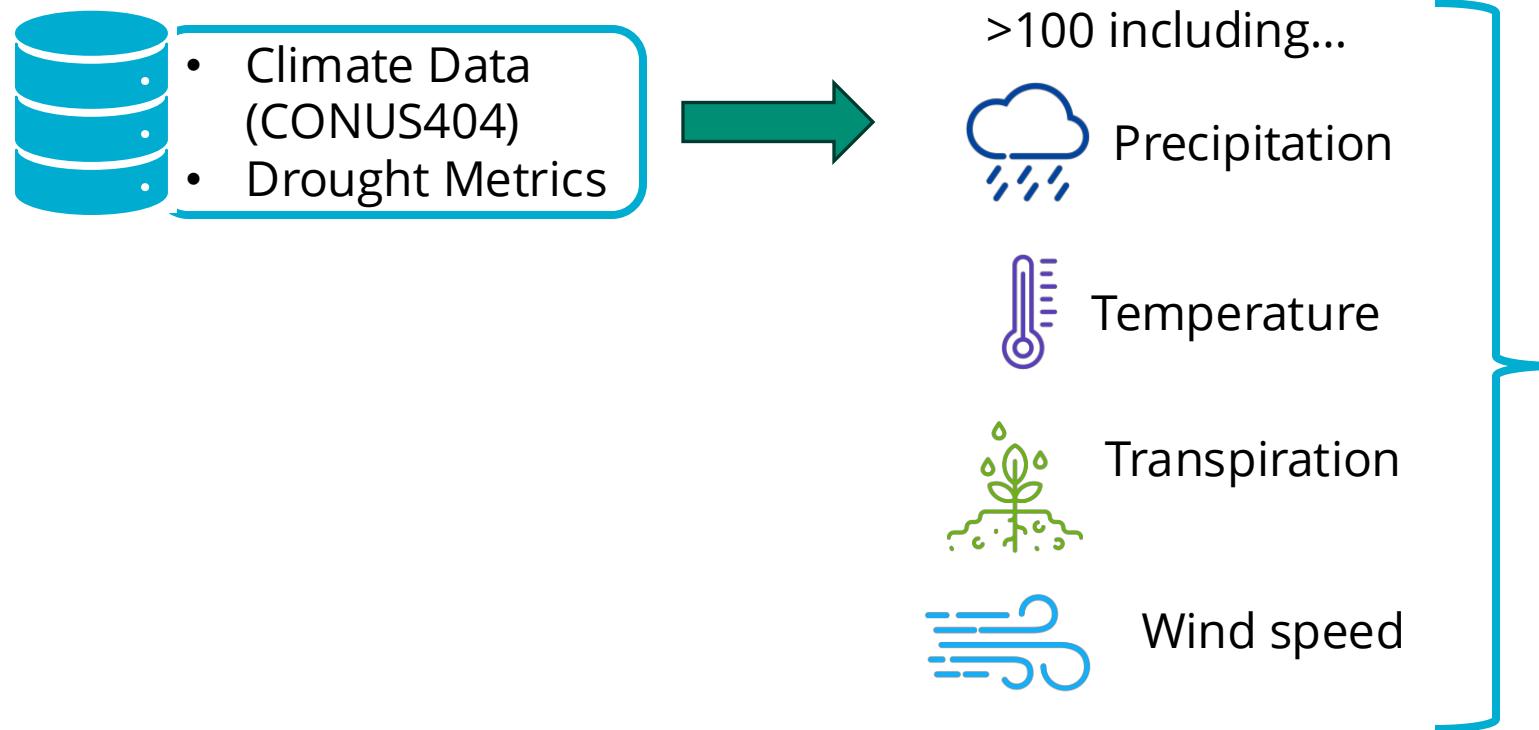
## Meteorological Drought:

- Artificial neural network (ANN)
- SVR
- K-nearest neighbor
- Adaptive Neuro Fuzzy (ANF)
- SVM-Copula
- Wavelet Boosting-Support Vector Regression (WBS-SVM)

Source: [Mokhtar et al. \(2021\)](#)

- RF
- Extreme Gradient Boost (XGB)
- Convolutional neural network (CNN)
- Long-term short memory (LSTM)

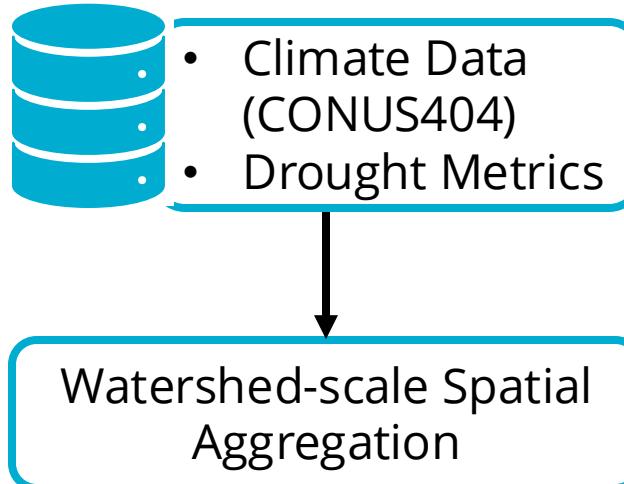
# Modeling supported by the high spatial resolution (4 km) NCAR/USGS CONUS404 dataset for period 1979-2021



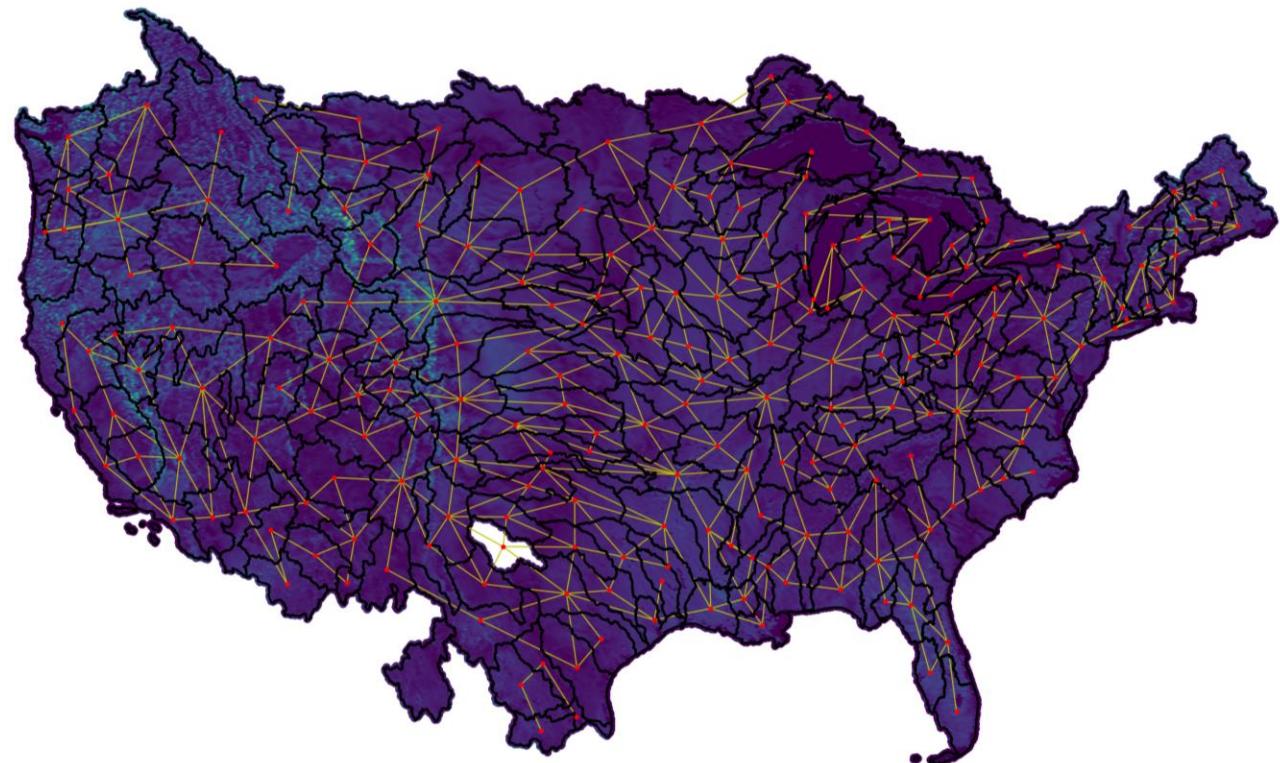
Calculated **meteorological drought** metrics:

- Standardized Precipitation Index (SPI)
- Standardized Precipitation Evapotranspiration Index (SPEI)

# Data is statistically reduced to the weekly timescale at the HUC4 and HUC8 spatial scales



Sample HUC8 aggregation for precipitation



Connections across  
watersheds

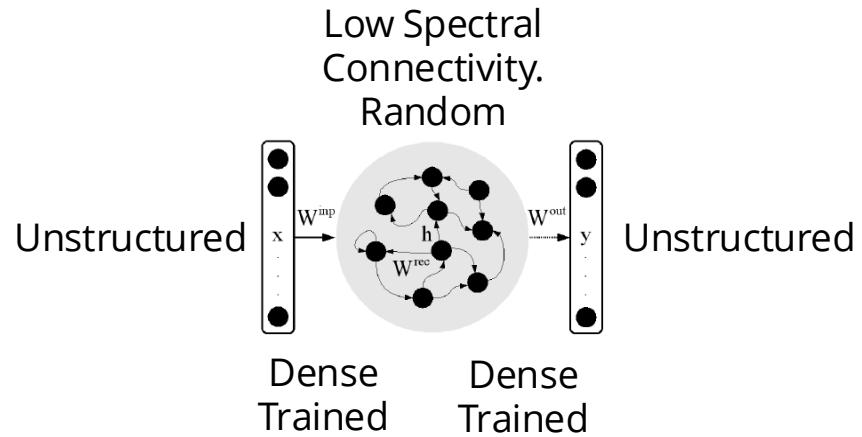
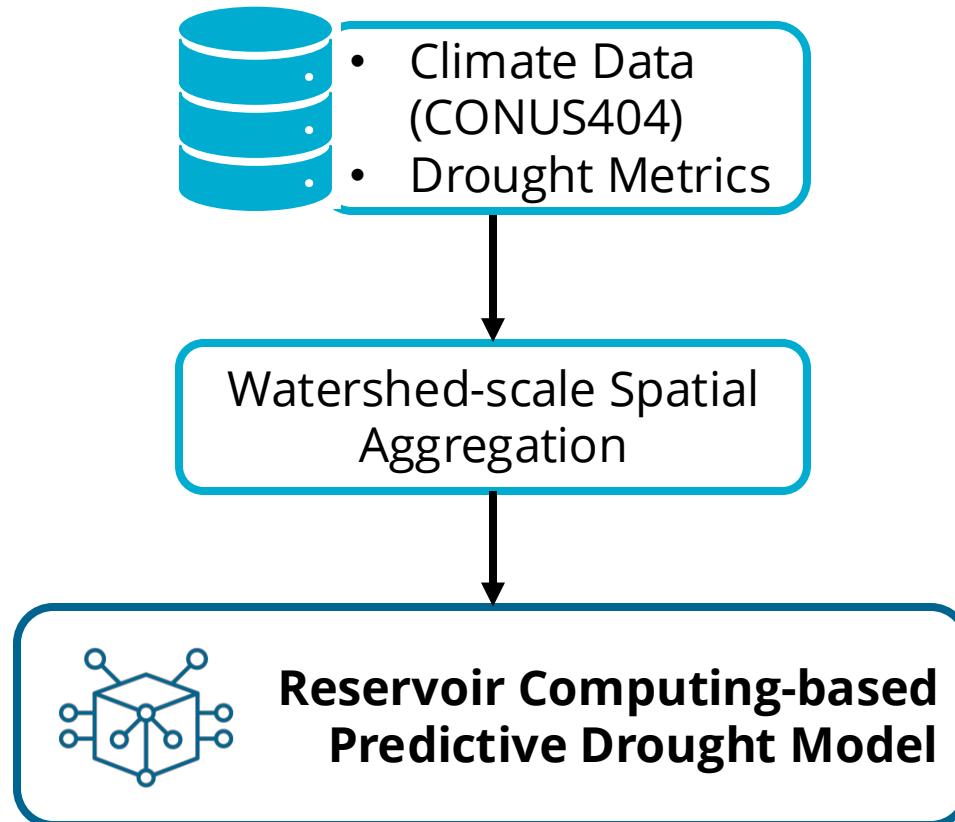


Watershed centroid

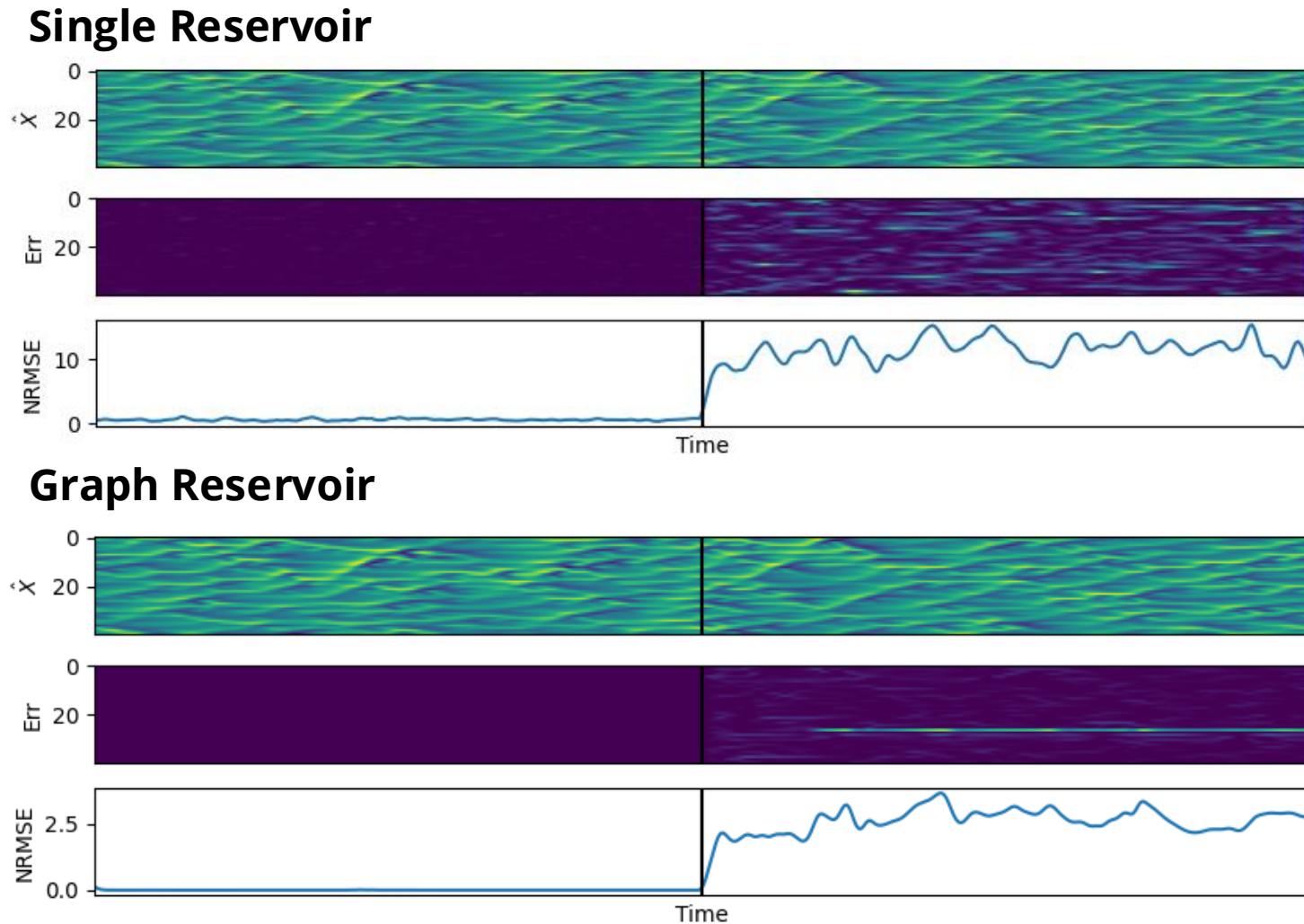
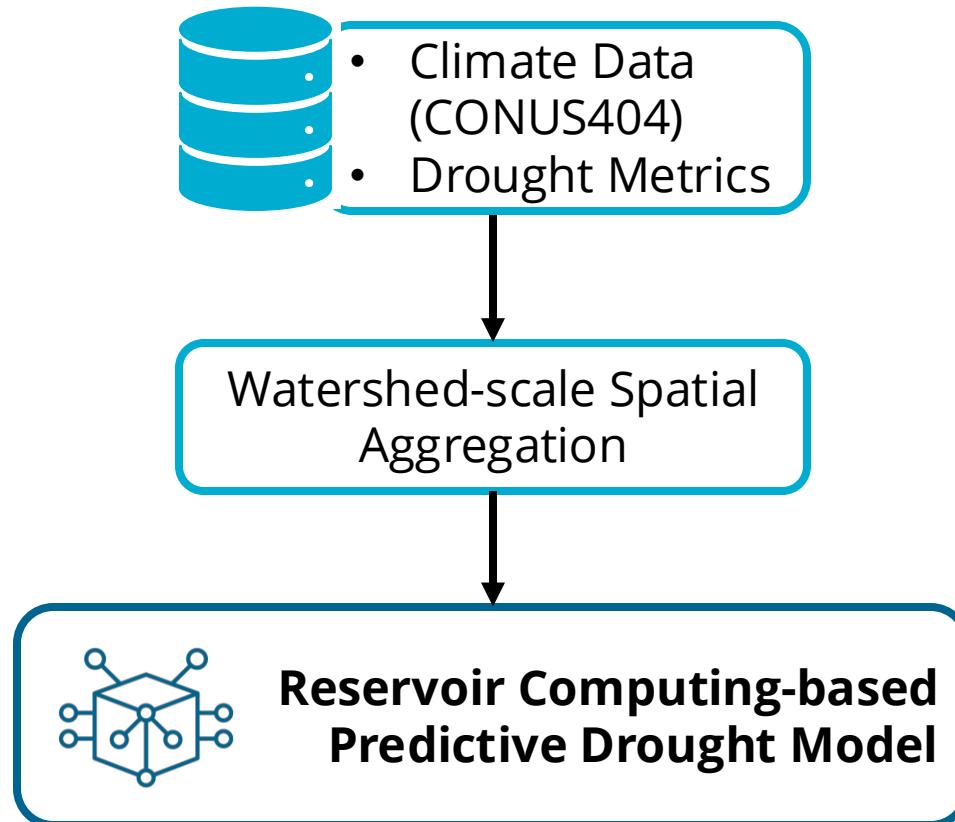


Precipitation

# Reservoir networks use spectral normalization and random connectivity to autonomously mimic chaotic system

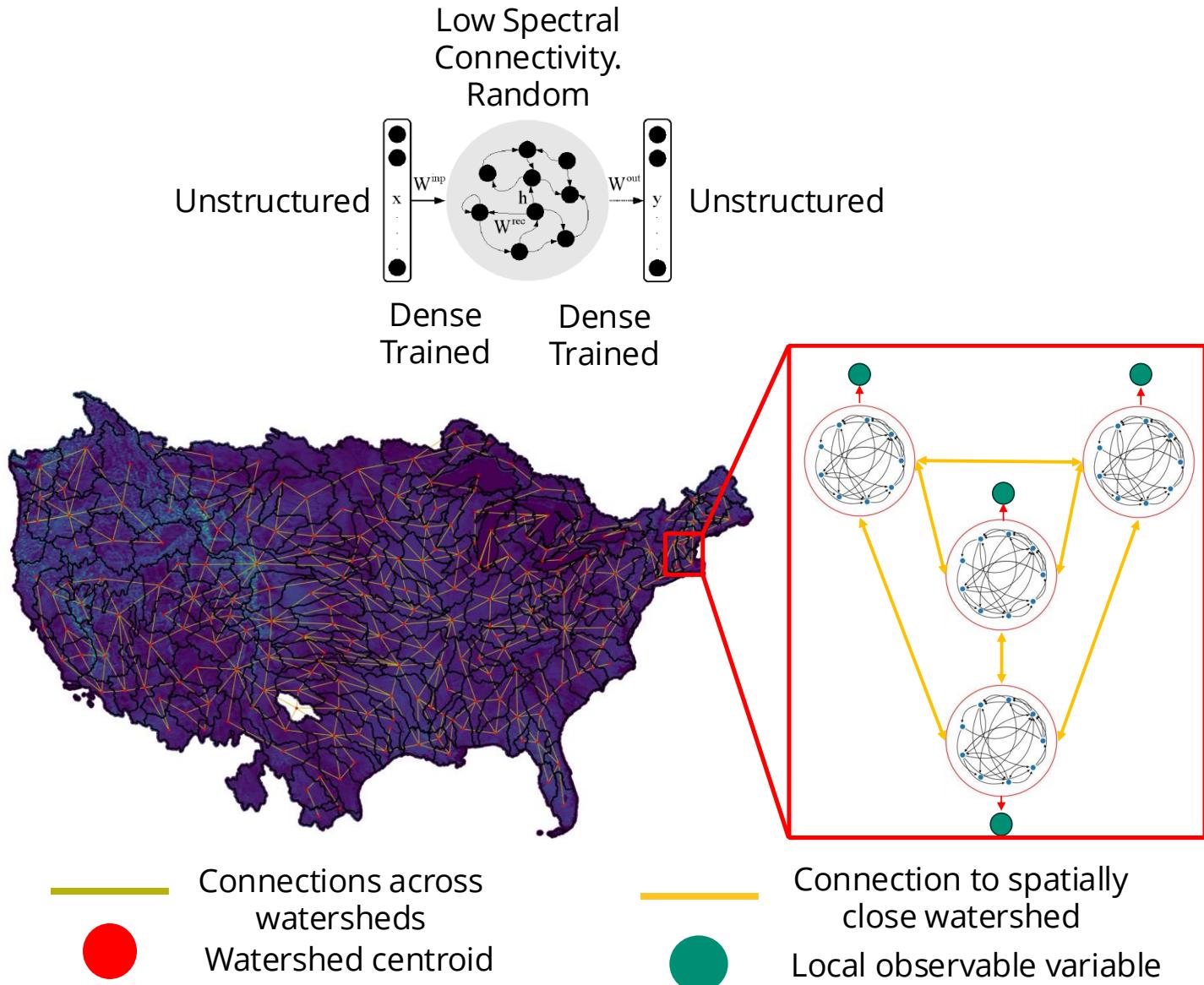
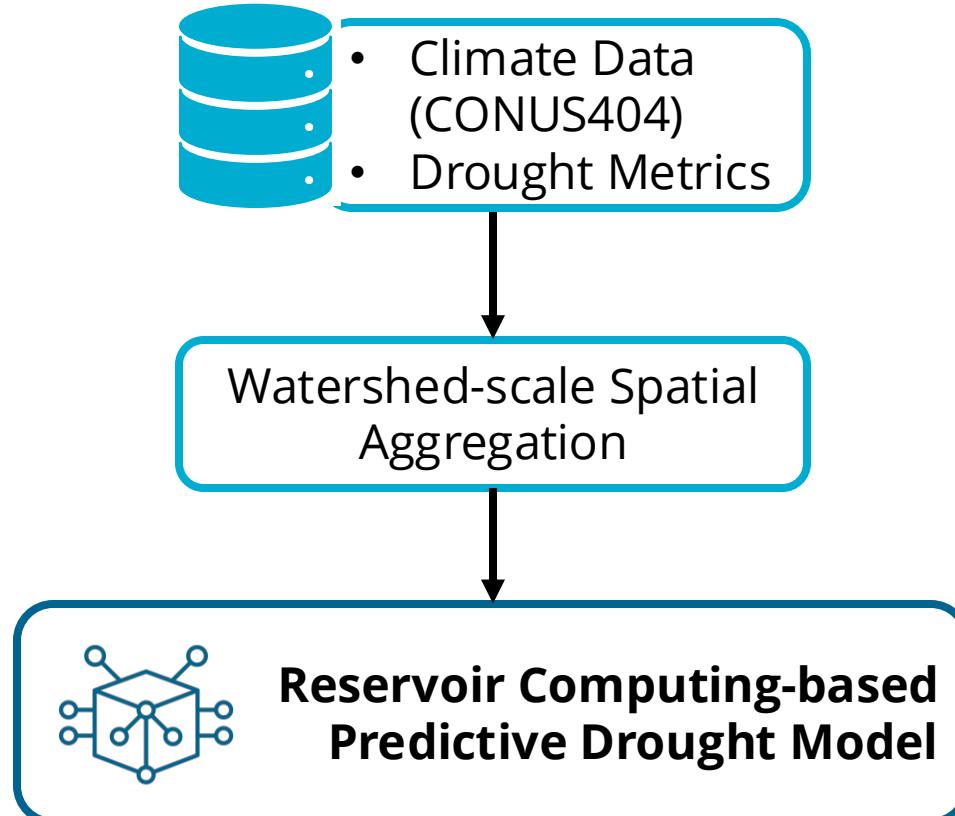


# Graph-based reservoirs reduce normalized root mean square error during prediction period versus a single reservoir

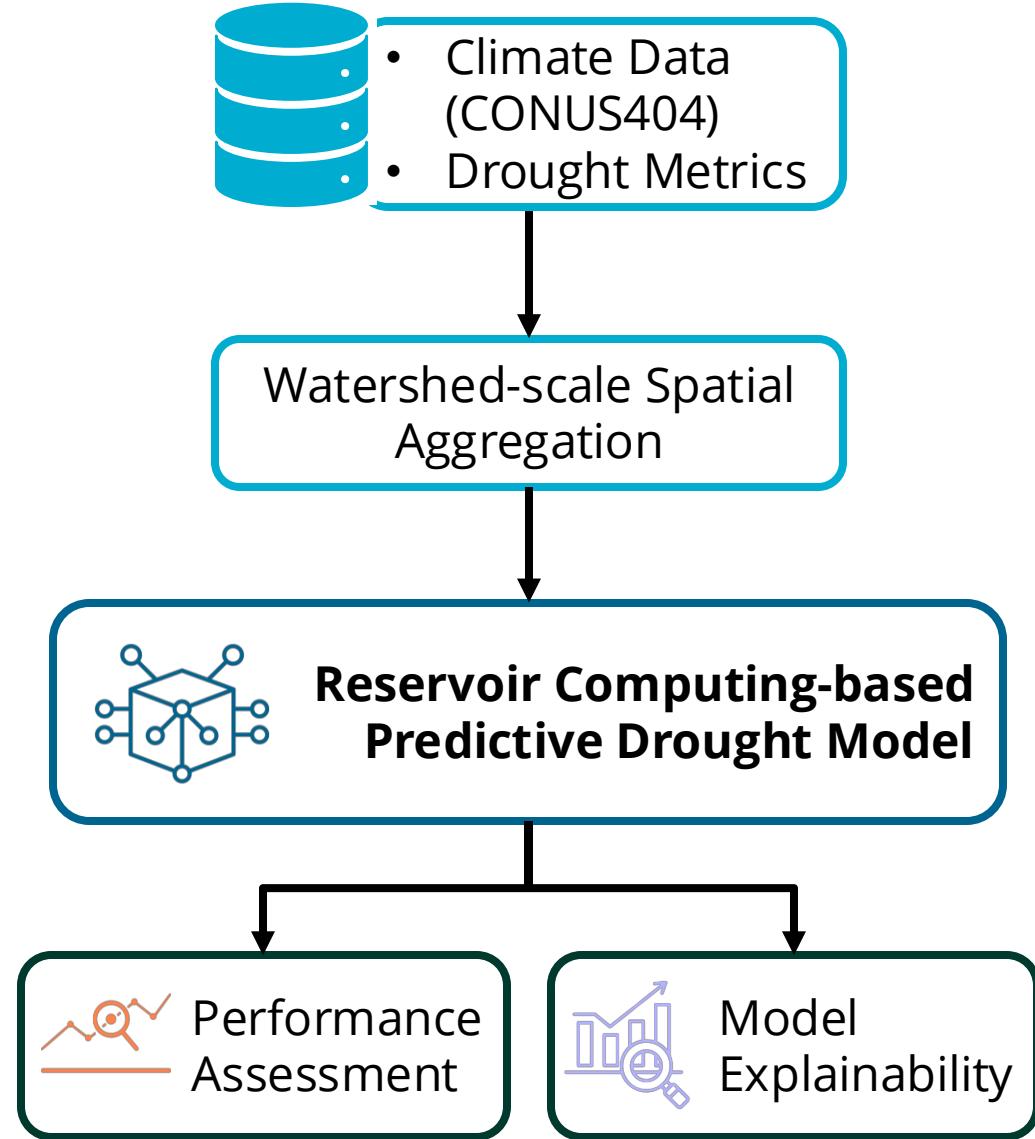




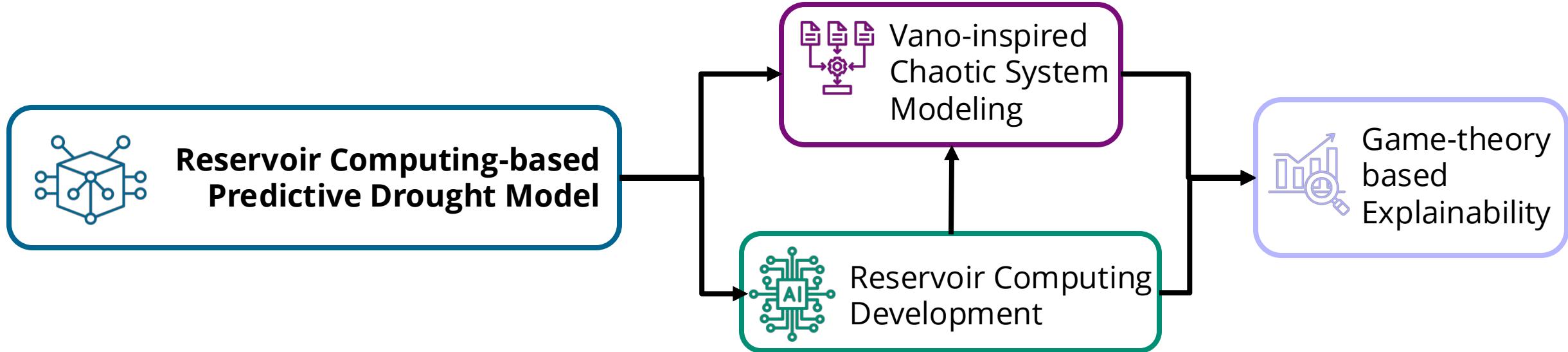
# Graph reservoir nodes are generated for each watershed while maintaining the input layer's spatial structure



# The EPIC PLANN project is using a standard framework to develop its ML-based predictive model for drought



# Reservoir model development inspired by chaotic systems modeling and game theory

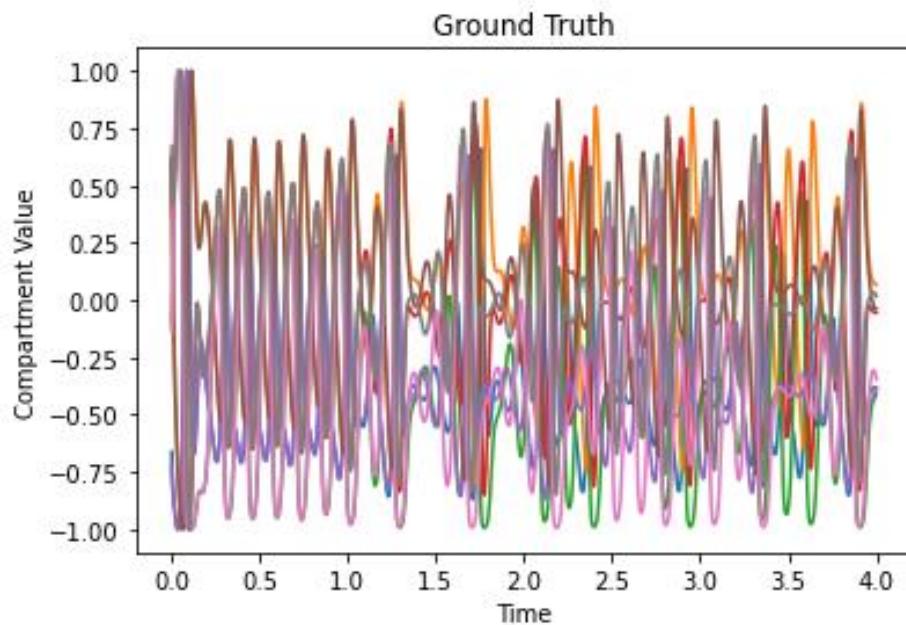


# A graph-based reservoir network is trained to predict two sets of Vano models side by side



## Input layer

$X_1(t_0), \dots, X_1(t_{N_k})$   
 $X_2(t_0), \dots, X_2(t_{N_k})$   
 $X_3(t_0), \dots, X_3(t_{N_k})$   
 $X_4(t_0), \dots, X_4(t_{N_k})$   
 $Y_1(t_0), \dots, Y_1(t_{N_k})$   
 $Y_2(t_0), \dots, Y_2(t_{N_k})$   
 $Y_3(t_0), \dots, Y_3(t_{N_k})$   
 $Y_4(t_0), \dots, Y_4(t_{N_k})$

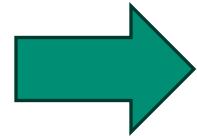


A graph-based reservoir network is trained to predict two sets of Vano models side by side

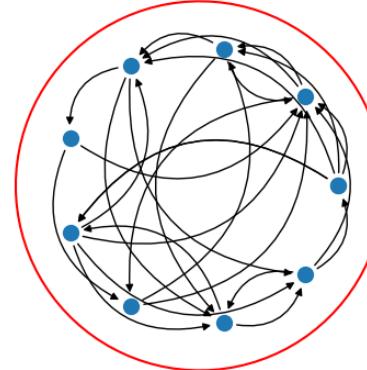


### Input layer

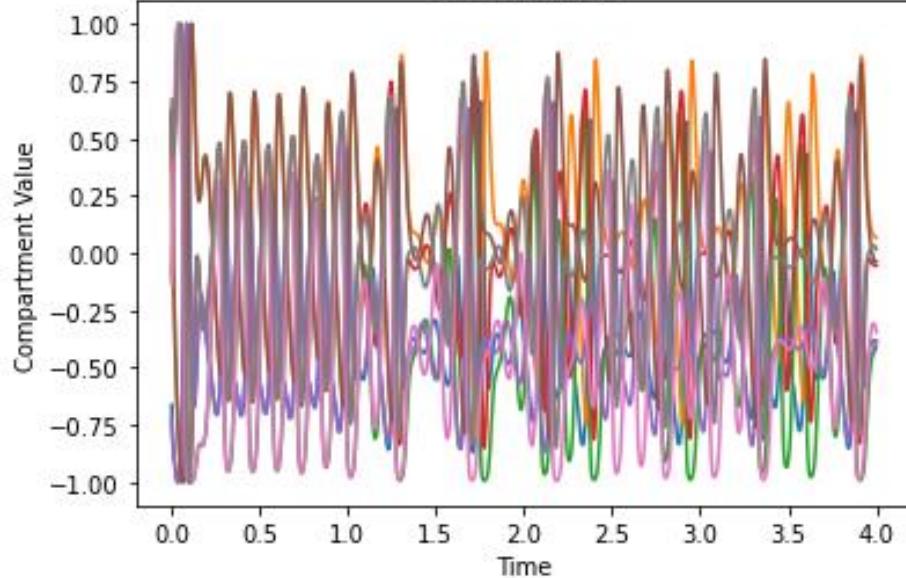
$X_1(t_0), \dots, X_1(t_{N_k})$   
 $X_2(t_0), \dots, X_2(t_{N_k})$   
 $X_3(t_0), \dots, X_3(t_{N_k})$   
 $X_4(t_0), \dots, X_4(t_{N_k})$   
 $Y_1(t_0), \dots, Y_1(t_{N_k})$   
 $Y_2(t_0), \dots, Y_2(t_{N_k})$   
 $Y_3(t_0), \dots, Y_3(t_{N_k})$   
 $Y_4(t_0), \dots, Y_4(t_{N_k})$



### Trained Reservoirs



Ground Truth

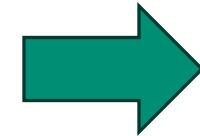


# A graph-based reservoir network is trained to predict two sets of Vano models side by side

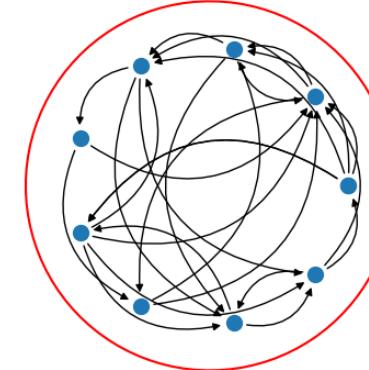


## Input layer

$$\begin{aligned} X_1(t_0), \dots, X_1(t_{N_k}) \\ X_2(t_0), \dots, X_2(t_{N_k}) \\ X_3(t_0), \dots, X_3(t_{N_k}) \\ X_4(t_0), \dots, X_4(t_{N_k}) \\ Y_1(t_0), \dots, Y_1(t_{N_k}) \\ Y_2(t_0), \dots, Y_2(t_{N_k}) \\ Y_3(t_0), \dots, Y_3(t_{N_k}) \\ Y_4(t_0), \dots, Y_4(t_{N_k}) \end{aligned}$$



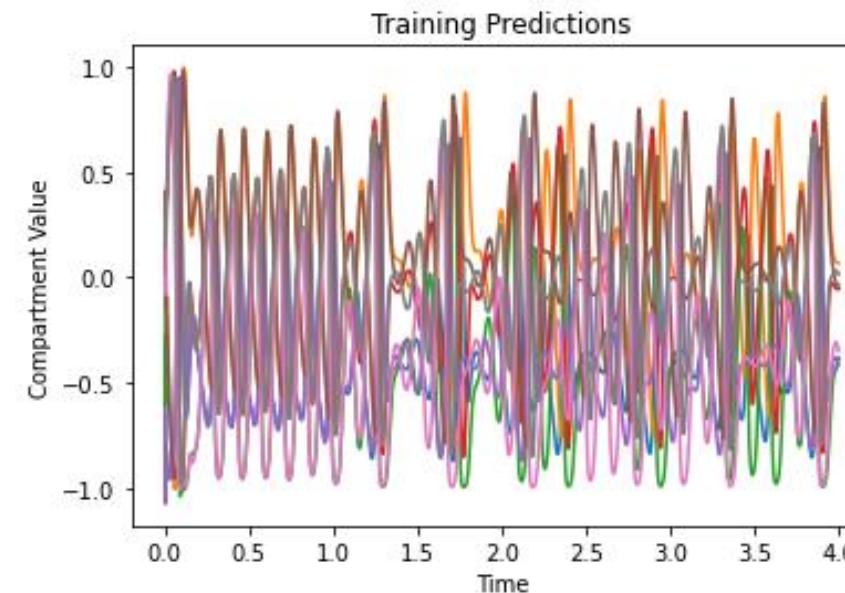
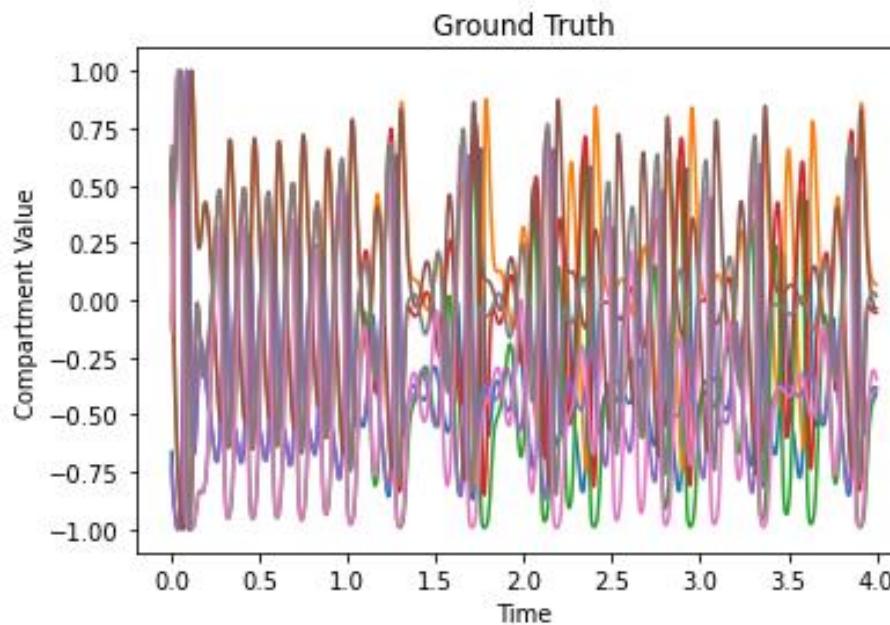
## Trained Reservoirs



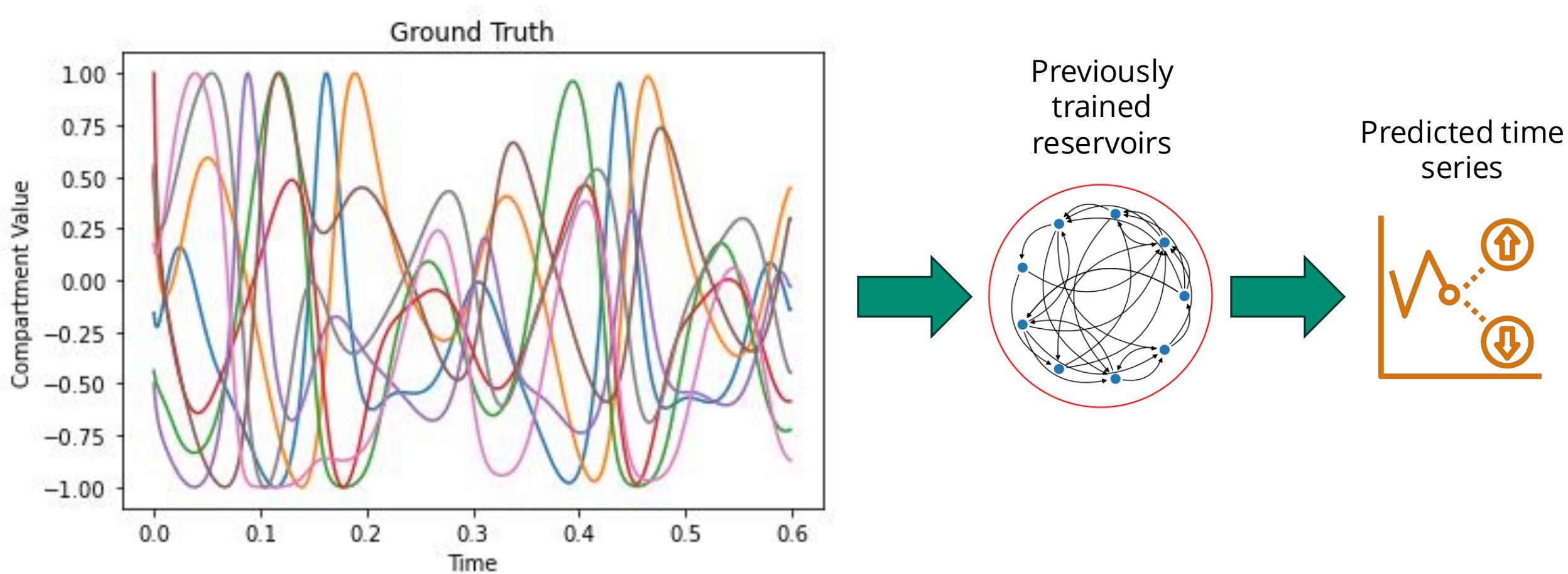
## Output layer

$$\begin{aligned} X_1(t_{N_k+1}), \dots, X_1(t_{N_p}) \\ X_2(t_{N_k+1}), \dots, X_2(t_{N_p}) \\ X_3(t_{N_k+1}), \dots, X_3(t_{N_p}) \\ X_4(t_{N_k+1}), \dots, X_4(t_{N_p}) \\ Y_1(t_{N_k+1}), \dots, Y_1(t_{N_p}) \\ Y_2(t_{N_k+1}), \dots, Y_2(t_{N_p}) \\ Y_3(t_{N_k+1}), \dots, Y_3(t_{N_p}) \\ Y_4(t_{N_k+1}), \dots, Y_4(t_{N_p}) \\ g(t_{N_k+1}), \dots, g(t_{N_p}) \end{aligned}$$

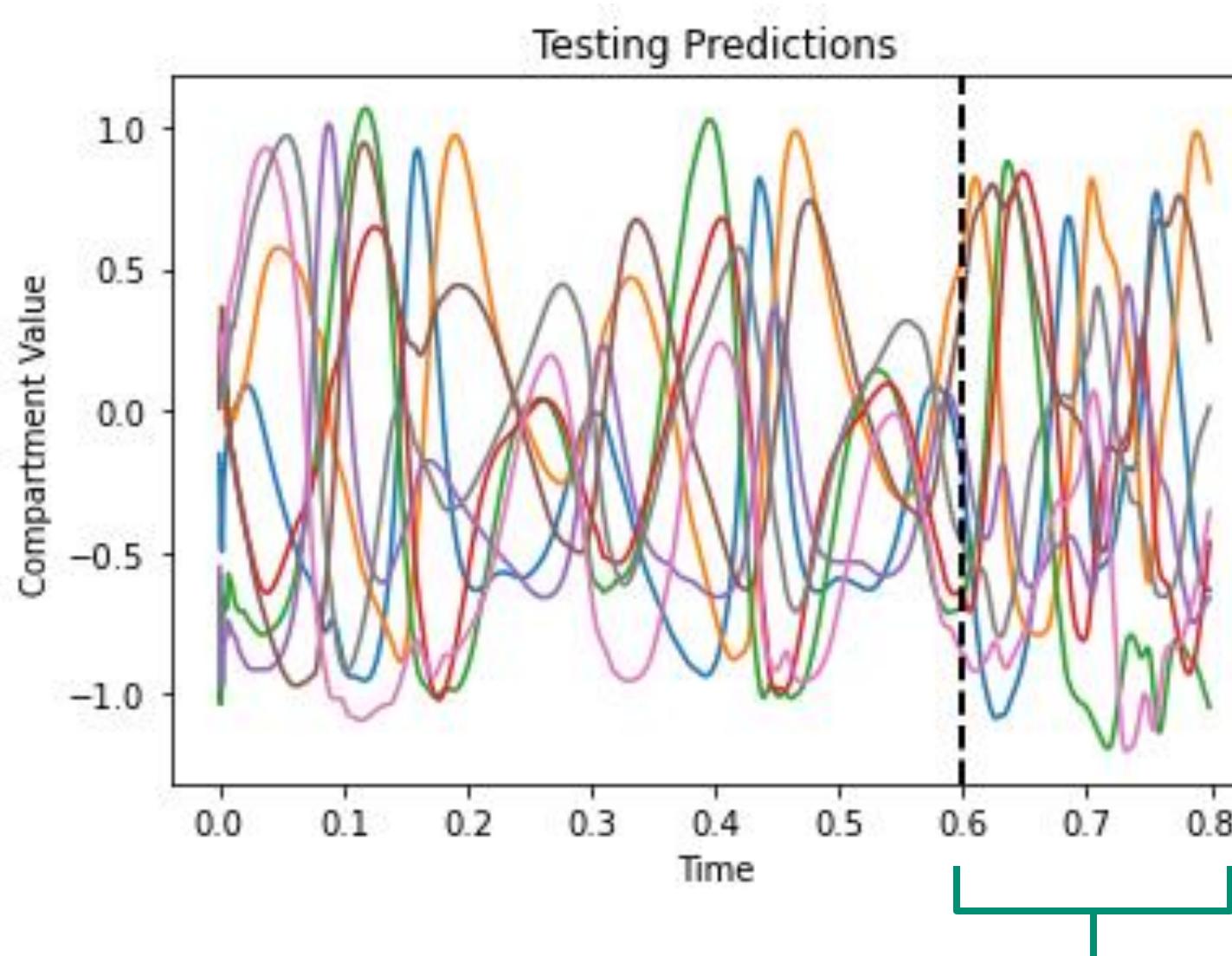
Mystery function



# A new time series using perturbed initial conditions is generated to support prediction with trained reservoirs



# Predicted time series using perturbed data and trained reservoirs with rollout



# Shapley Additive Explanation Values (SHAP) are used to determine influence of inputs on changes to outputs



## Input layer

$$X_1(t_0), \dots, X_1(t_{N_k})$$

$$X_2(t_0), \dots, X_2(t_{N_k})$$

$$X_3(t_0), \dots, X_3(t_{N_k})$$

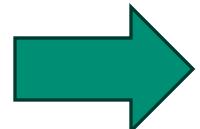
$$X_4(t_0), \dots, X_4(t_{N_k})$$

$$Y_1(t_0), \dots, Y_1(t_{N_k})$$

$$Y_2(t_0), \dots, Y_2(t_{N_k})$$

$$Y_3(t_0), \dots, Y_3(t_{N_k})$$

$$Y_4(t_0), \dots, Y_4(t_{N_k})$$



## Output layer

$$X_1(t_{N_k+1}), \dots, X_1(t_{N_p})$$

$$X_2(t_{N_k+1}), \dots, X_2(t_{N_p})$$

$$X_3(t_{N_k+1}), \dots, X_3(t_{N_p})$$

$$X_4(t_{N_k+1}), \dots, X_4(t_{N_p})$$

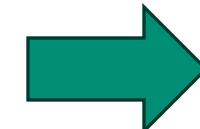
$$Y_1(t_{N_k+1}), \dots, Y_1(t_{N_p})$$

$$Y_2(t_{N_k+1}), \dots, Y_2(t_{N_p})$$

$$Y_3(t_{N_k+1}), \dots, Y_3(t_{N_p})$$

$$Y_4(t_{N_k+1}), \dots, Y_4(t_{N_p})$$

$$g(t_{N_k+1}), \dots, g(t_{N_p})$$



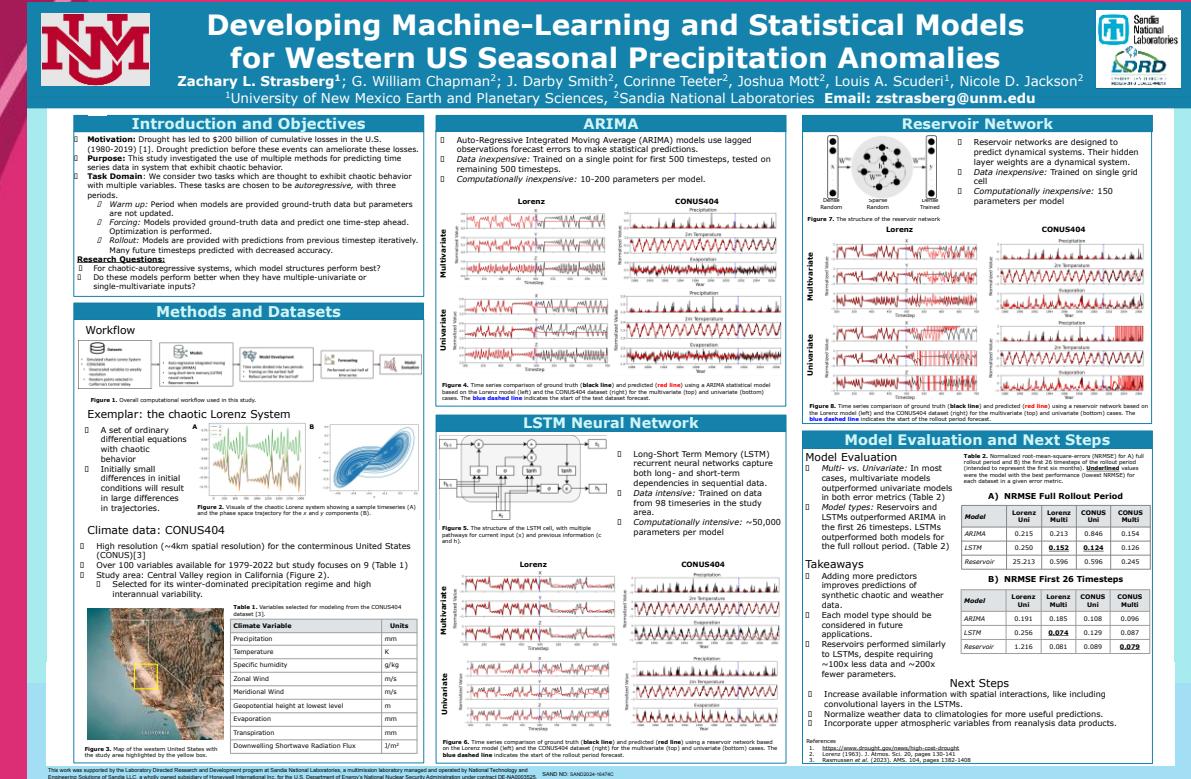
## SHAP Values

0.02147524
0.02505524
0.02095524
0.01256381
0.00879333
0.00837357
0.00999976
0.00241643

# A51L-1838: Developing Machine-Learning and Statistical Models for Western US Seasonal Precipitation Anomalies

- Presenter: Zach Strasberg
- Friday, 13 December 2024
- 08:30 – 12:20
- Poster Halls B-C

# AGU24



In conclusion, we have developed reservoir computing-based predictive time series models for chaotic, dynamic systems



- Graph reservoirs reduce error compared to single reservoirs
- Reservoir-based predictive model developed using the Vano system with chaotic behavior
- SHAP identified 3 key variables driving Vano model performance

---

## Future Work

- Increase variables and explainability techniques for the Vano model
- Consideration of multiple drought metrics (SPEI, SPI, etc.)
- Increase spatial extent of modeled watersheds

