Drought Prediction (CA)

Angela Chang, Nan Li, David Sherman
W210, Section 4
February 6th, 2024

Table of contents

01

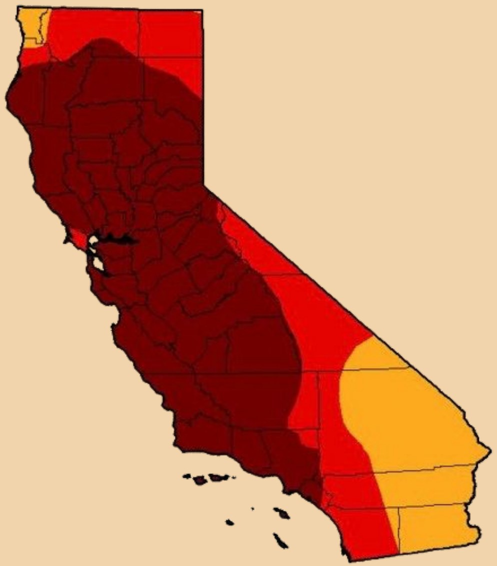
Main Issue

Main Issue

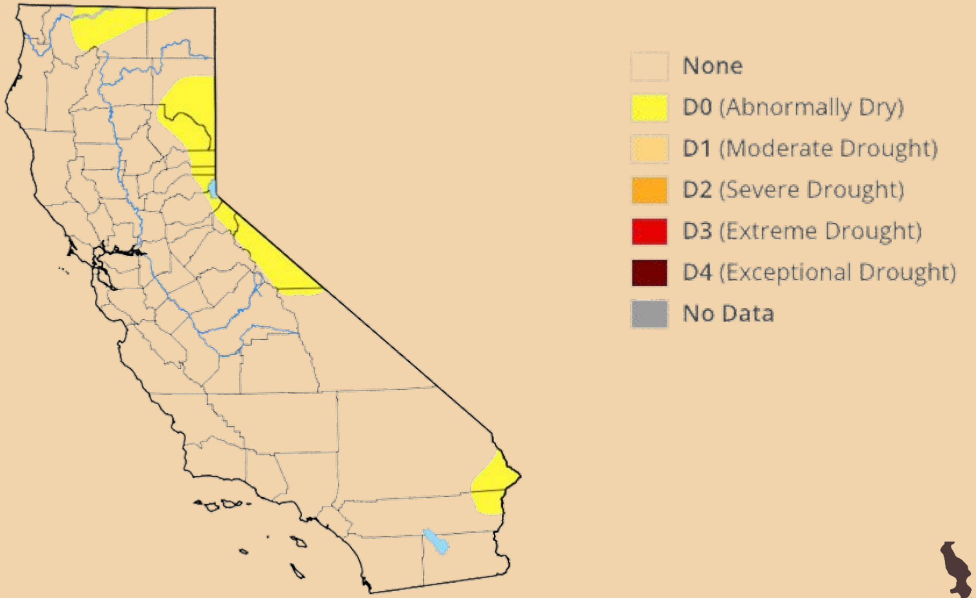
- Drought Intensity
  - Severity in California is evident based on changes in temperature, precipitation, wildfires, etc.
- Impact of Droughts
  - Crop failure, water scarcity, ecosystem degradation, and wildfires all impact agriculture, water resources, and public health
- Need for Predictive Models
  - Drought intensity predictions support implementing preventative measures, like water conservation, agricultural planning, and disaster preparedness

U.S. Drought Monitor (USDM) – California

July 2014



February 2024



https://droughtmonitor.unl.edu/CurrentMap/StateDroughtMonitor.aspx?CA

U.S. Drought Monitor – California (2000–2024)

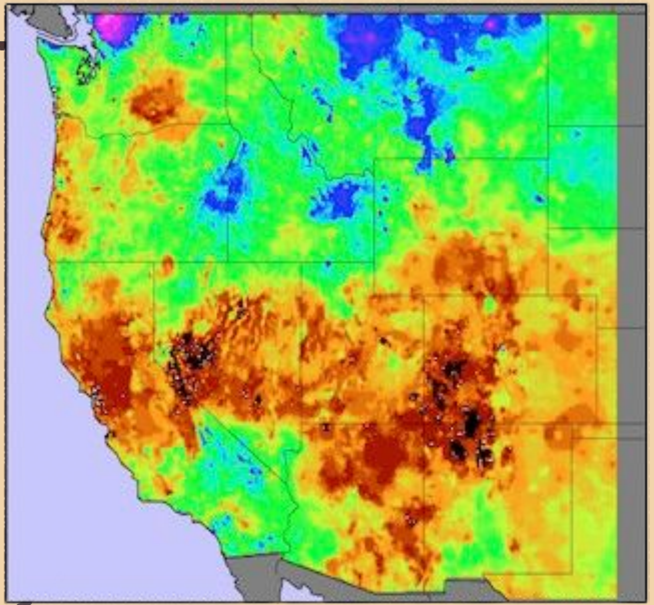
- Table below represents the average proportion of land that could be classified as None – D4 throughout the 21st century
    - The longest drought period in CA history took place between 2012–2016, and **85%** of the state was in at least a "Moderate Drought" state (D1)

Time Period	None	D0-D4	D1-D4	D2-D4	D3-D4	D4
2000-2004	51.85%	48.15%	25.86%	12.64%	4.22%	0.00%
2005-2009	40.43%	59.57%	45.26%	23.20%	5.42%	0.00%
2010-2011	82.97%	17.03%	5.81%	1.57%	0.00%	0.00%
2012-2016	**4.58%**	**95.42%**	**85.50%**	**67.77%**	**40.08%**	**21.20%**
2017-2019	55.33%	44.67%	22.28%	7.61%	1.42%	0.14%
2020s	26.11%	73.89%	65.00%	50.25%	27.50%	8.07%

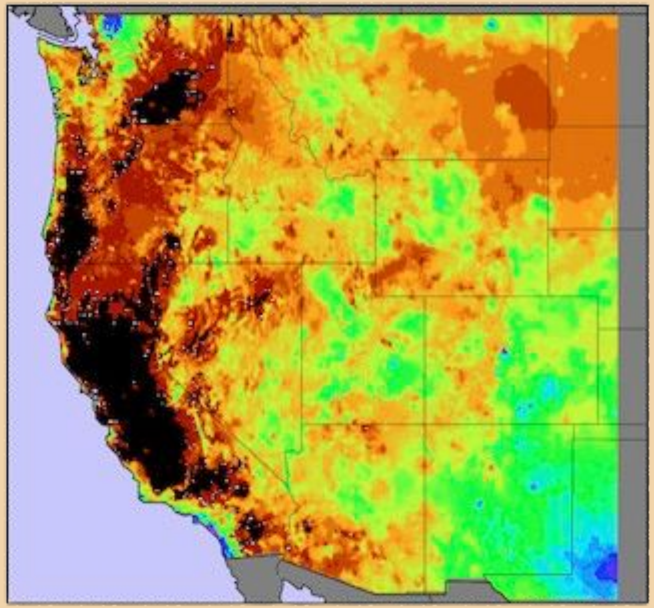
https://droughtmonitor.unl.edu/CurrentMap/StateDroughtMonitor.aspx?CA

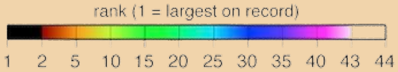
Evaporative Demand Maps

2020



2021





- Evaporative demand = the potential loss of water from the surface as driven by atmospheric factors (temperature, wind speed, humidity and cloud cover)
- Periods of high evaporative demand are connected to droughts and increased fire danger

https://www.drought.gov/drought-status-updates/california-nevada-drought-status-update-2023-10-19

Wildfires

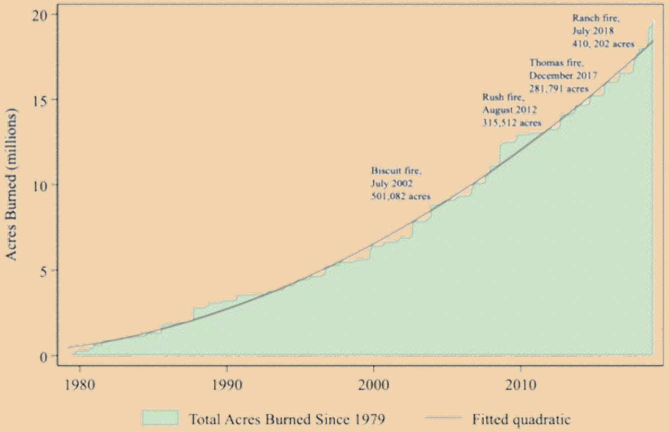


Table 1. Top 10 Counties with Highest Estimated Structure Value Loss for SRA Fires, by Decade.

	1979 - 1988		1989 - 1998		1999 - 2008		2009-2018	
	County	Loss ($BIL)	County	Loss ($BIL)	County	Loss ($BIL)	County	Loss ($BIL)
1	San Bernardino	0.08	Alameda	1.01	San Diego	1.11	Butte	3.52
2	Napa	0.04	Santa Barbara	0.22	Los Angeles	0.35	Sonoma	2.14
3	Nevada	0.04	Shasta	0.14	San Bernardino	0.30	Los Angeles	0.94
4	Los Angeles	0.03	Orange	0.11	Shasta	0.28	Napa	0.48
5	Ventura	0.03	Riverside	0.08	Butte	0.10	Shasta	0.37
6	Monterey	0.02	Ventura	0.06	Santa Barbara	0.08	Ventura	0.36
7	San Luis Obispo	0.01	San Diego	0.03	Orange	0.07	Calaveras	0.22
8	Riverside	0.01	Yuba	0.02	Riverside	0.05	Lake	0.19
9	Orange	0.01	El Dorado	0.02	Santa Cruz	0.05	San Diego	0.15
10	Lake	0.01	Butte	0.02	Trinity	0.04	Mendocino	0.14

https://www.fire.ca.gov/our-impact/statistics

CA Droughts' Scale

- The 2012–2016 drought took $2.7 billion toll on agriculture industry
- The 2021 drought cost the agriculture industry ~8.7K jobs
- Jan 2024: California's Sierra Nevada snowpack — the source of nearly one-third of the state's water supply — was at its lowest level in a decade

02
Significance

Solution Significance

- Improved Preparedness
    - Predictive models for droughts enable early warnings and proactive measures, reducing vulnerability for both communities and ecosystems
    - Providing timely predictions allow stakeholders to implement preventive actions such as water conservation measures, agricultural planning, and disaster preparedness efforts
- Resource Allocation
    - Can optimally allocate resources, particularly water supplies and relief efforts, by identifying regions at high risk of drought intensity
    - Directing resources to most vulnerable areas to drought will facilitate efficient resource management
- Environmental Impact
    - Early intervention can mitigate the ecological impacts of droughts, preserving biodiversity and ecosystem services

03
Datasets

Datasets

- Main Dataset
  - A Kaggle dataset offered by the NASA Power Project and authors of the U.S. Drought Monitor
  - Time series weather data at county level (2000–2020): wind speed, temperature, humidity, precipitation, *etc.*
  - Time series drought score at county level (2000–2020)
  - Soil property data at county level

- Data Preprocessing
  - Averaging daily weather data and integrating it with weekly drought scores
  - Cleaning data and standardizing features to ensure consistency and compatibility across datasets

- Supplemental Dataset

Datasets

- 1,096 per county weekly data points for each of 58 California counties represented, or a total of 63,568 data points

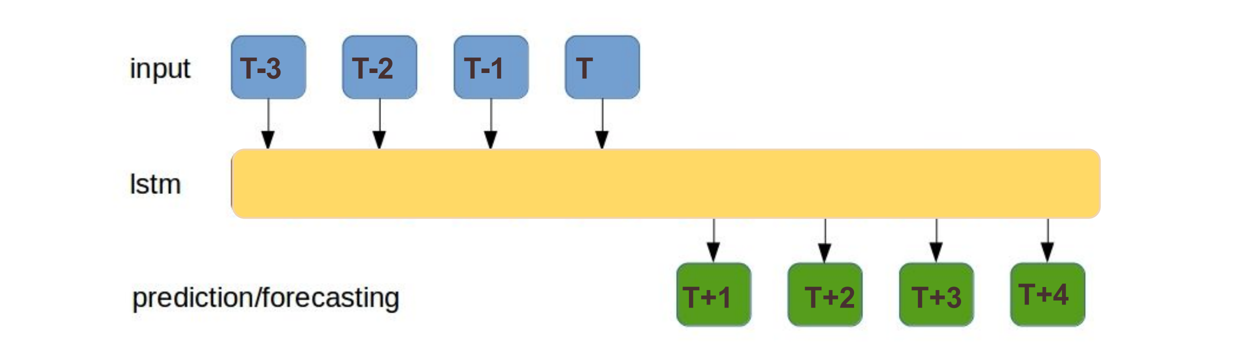
- Train–validation–test split:
    - Training (2000 – 2009)
    - Validation (2010 – 2011)
    - Test (2012 – 2020)

04

Approach

Approach

- Modeling Techniques
    - Machine learning/deep learning algorithms such as LSTM and CNN for drought intensity forecasting
    - Time series forecasting methods, such as SARIMA, ARIMA, etc



Approach

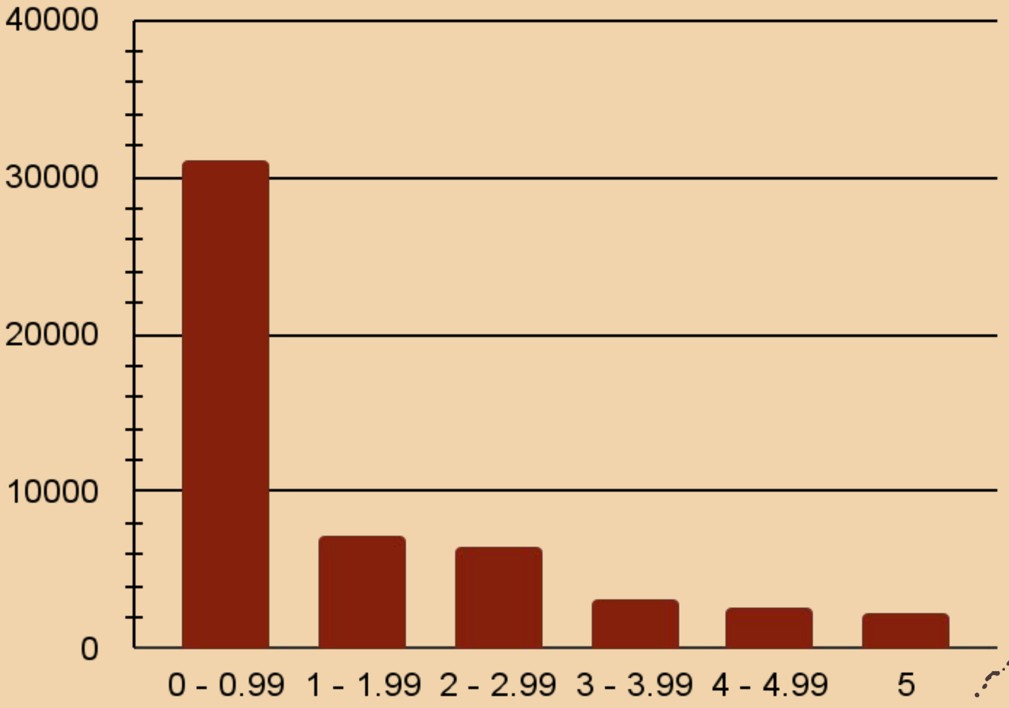
- Feature Selection
  - Using relevant meteorological variables (temperature, humidity, wind speed, etc.) based on domain knowledge and statistical analysis to capture the dynamics of drought
- Model Evaluation
  - Evaluating model performance using metrics such as RMSE and MAE to assess prediction accuracy and reliability

05

EDA

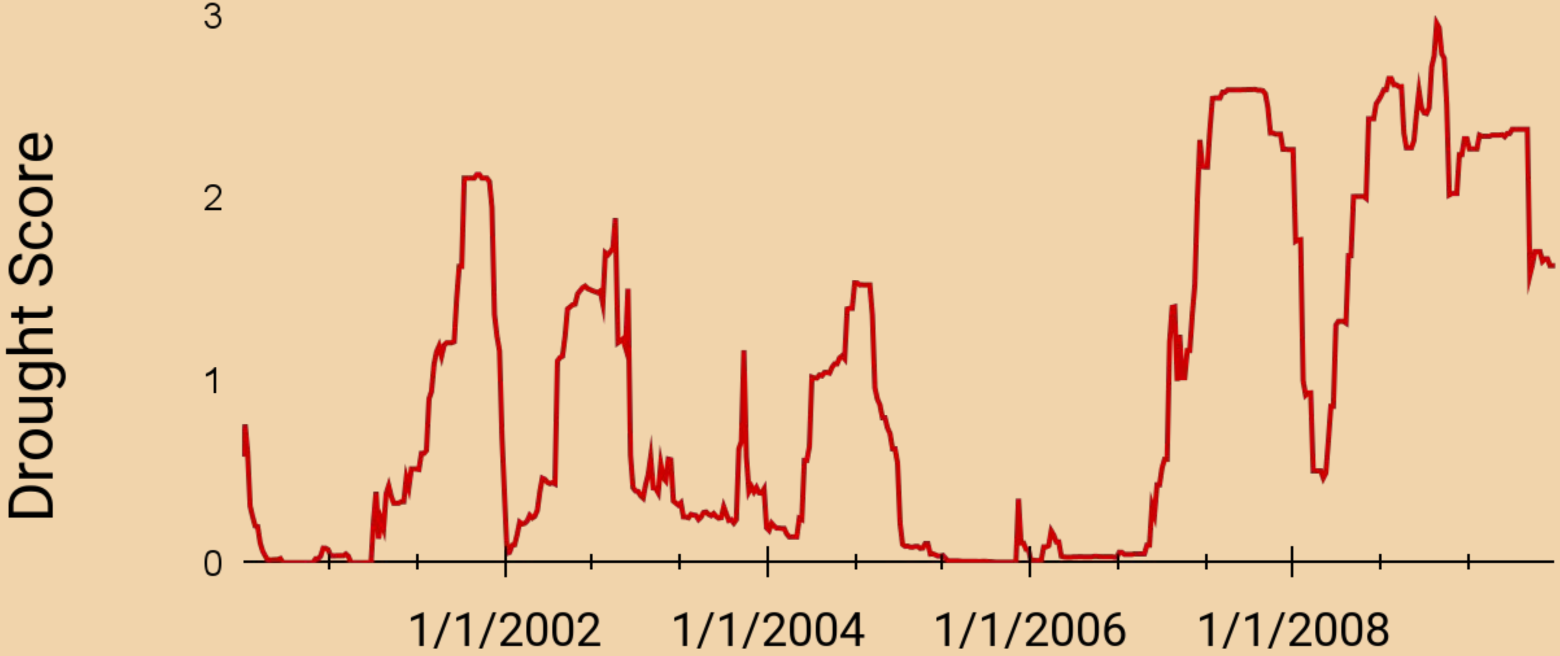
Drought Scores, 2009 – 2020: Distribution

0	No drought
1	D0: Abnormally dry
2	D1: Moderate drought
3	D2: Severe drought
4	D3: Extreme drought
5	D4: Exceptional Drought

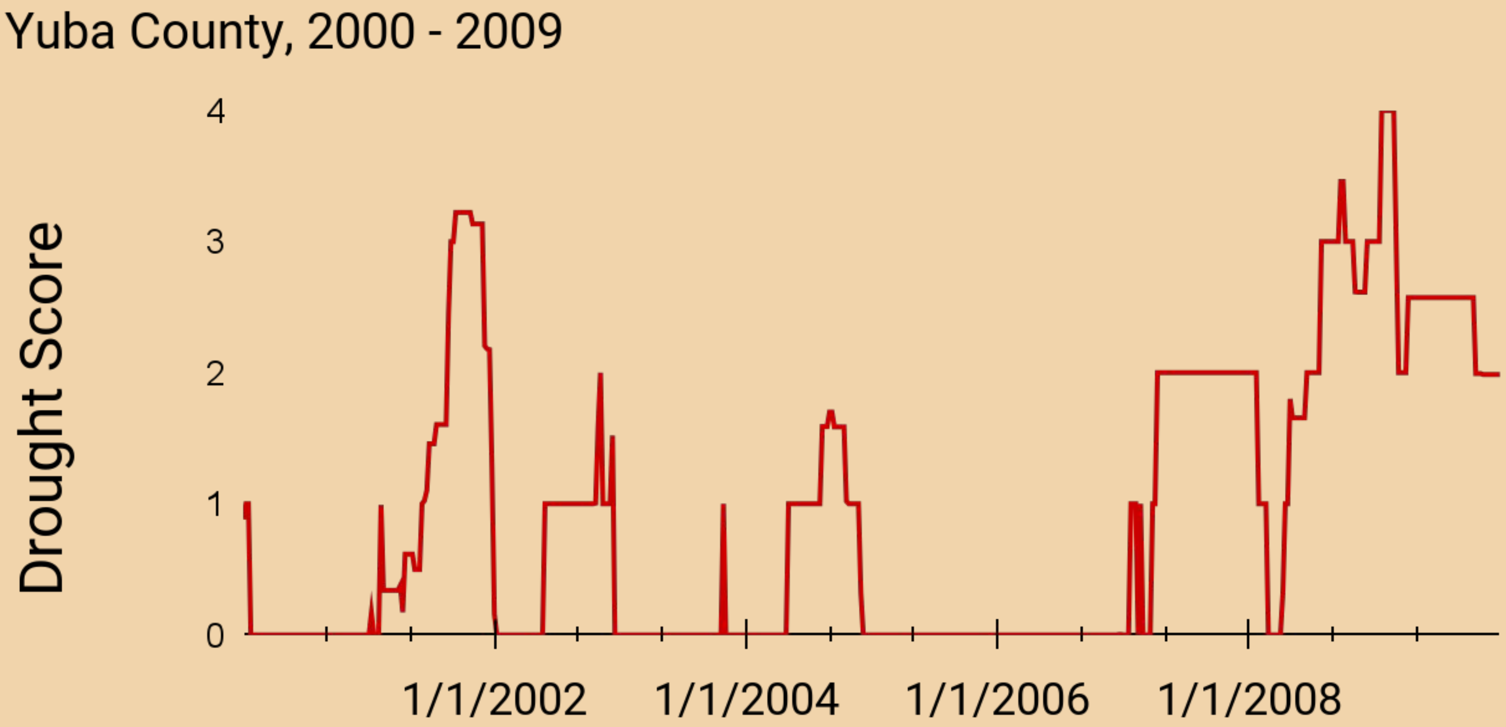


Drought Score: Time Series

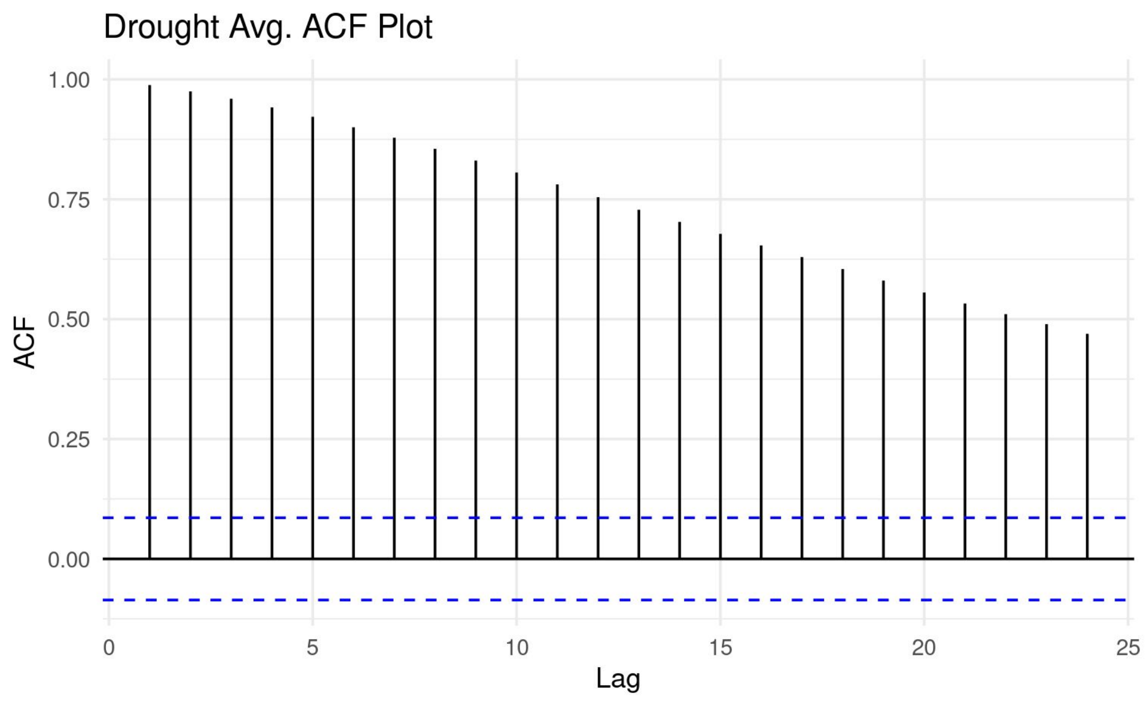
California (Average), 2000 - 2009



Drought Score: Time Series



Drought Scores: ACF

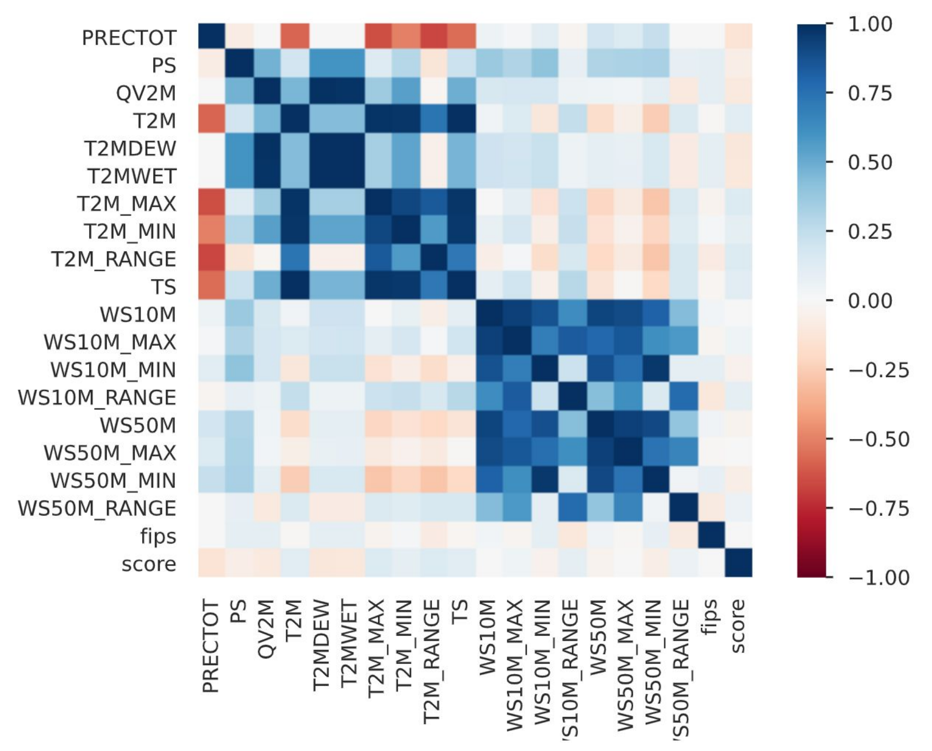


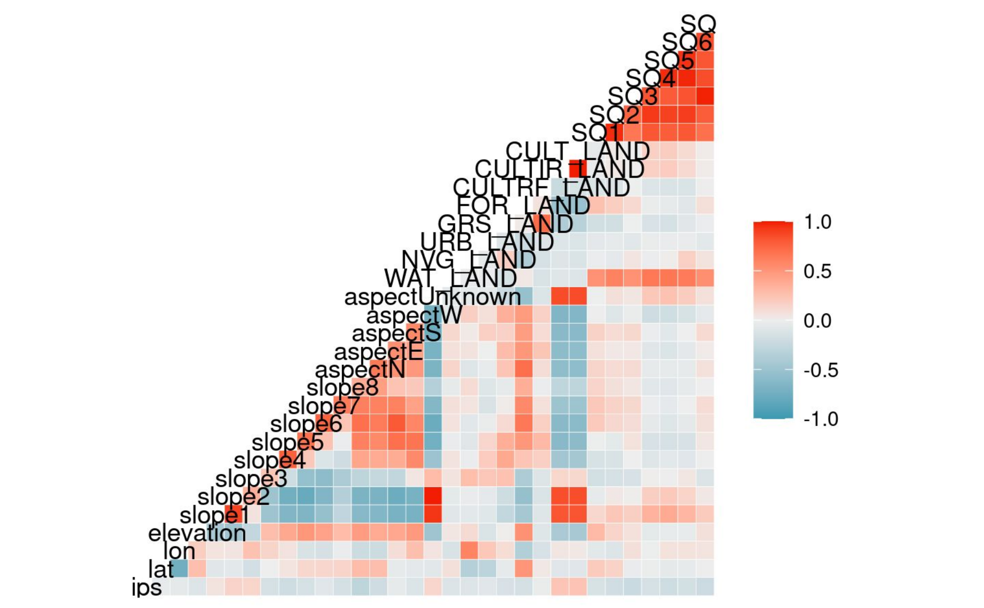
Drought Score vs. Other Variables

Variable	Corr	Variable	Corr
lon (longitude)	0.201745425	T2MDEW (Dewpoint)	-0.157316374
T2M_RANGE (Temperature range)	0.201348937	T2MWET (Wet bulb)	-0.156117648
T2M_MAX (Maximum temp)	0.178775957	QV2M (Humidity)	-0.144131260
T2M (Temperature)	0.155216999	PRECTOT (Precipitation)	-0.134216136

*2M indicates measurement at 2 meters

Explanatory Variable Correlation





06
Next Steps

Next Steps

- Continue EDA, particularly time series analysis on the time series variables and feature engineering
- Make changes to project based on today’s feedback
- Connect with a SME
- Start a baseline model

Thanks!

Do you have any questions?

am_chang@berkeley.edu
nankli@berkeley.edu
davidsherman@berkeley.edu

References

- Drought Prediction: A Comprehensive Review of Different Drought Prediction Models and Adopted Technologies
- Frontiers | DroughtCast: A Machine Learning Forecast of the United States Drought Monitor (frontiersin.org)
- Forecasting drought using neural network approaches with transformed time series data - PMC (nih.gov)
- https://www.scientificamerican.com/article/drought-takes-2-7-billion-toll-on-california-agriculture/
- https://www.mercurynews.com/2024/01/02/sierra-nevada-snowpack-at-lowest-level-in-10-years-what-it-means-for-californias-water-supply/
- https://droughtmonitor.unl.edu/CurrentMap/StateDroughtMonitor.aspx?CA
- https://www.drought.gov/drought-status-updates/california-nevada-drought-status-update-2023-10-19