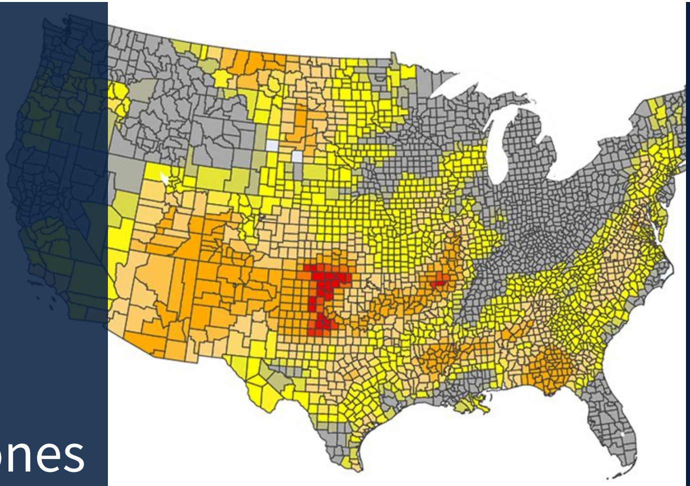
DroughtED

A dataset and methodology for drought forecasting spanning multiple climate zones



Christoph Minixhofer
Calum McMeekin
Mark Swan
Dr. Pavlos Andreadis

THE UNIVERSITY of EDINBURGH

PROBLEM

- The frequency and duration of droughts are being exacerbated by climate change
- Due to this, drought forecasting is increasingly important

THE UNIVERSITY of EDINBURGH

PRIOR WORK

- shows deep learning is promising
    - covers single climate region
    - uses distinct models for regions

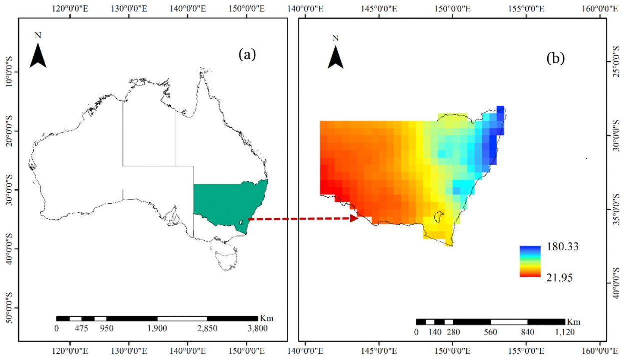


Image: example drought forecasting study area in Australia (Diskhit et al., 2021)

THE UNIVERSITY *of* EDINBURGH

DROUGHT FORECASTING ACROSS CLIMATE REGIONS

- Prior work focuses on forecasting drought for solitary regions
- *A more diverse dataset could lead to generalisation across regions*

THE UNIVERSITY *of* EDINBURGH

TARGET VALUE: UNITED STATES DROUGHT MONITOR (USDM)

- Expert labels (5 drought categories)
- Measures agricultural + meteorological drought
- As categories are ordinal we convert to numerical values
- Evaluation: *Macro F1* and *MAE/RMSE*

THE UNIVERSITY of EDINBURGH

DroughtED

- Globally available input features
- Time-invariant features
- Seasonal reference data
- *Currently covers continental US*
- *Can be expanded to other regions*

THE UNIVERSITY *of* EDINBURGH

NASA POWER PROJECT

- Globally available
- Wind speed, surface pressure, temperature, humidity, precipitation (21 values) + previous drought values
- 180 days of data leading up to prediction

$$\text{current:} \;\; \vec{x}_1, ..., \vec{x}_{180}$$

SEASONAL REFERENCE DATA

- Include past values offset by 1 year

- *Previous meteorological data + drought values in the same season can help indicate if current values are normal or abnormal*

$$\text{current: } \vec{x}_{c,1}, \ldots, \vec{x}_{c,180}$$

$$\text{past: } \vec{x}_{p,1}, \ldots, \vec{x}_{p,180}$$

HARMONIZED WORLD SOIL DATABASE

- time-invariant (indirectly identifies location)
- Elevation, Slope, Aspect, Land Type, Soil Quality (29 Values)
- Enables model to generalise across large areas

$$\vec{Location} = \vec{HWSD} \oplus \vec{LatLon}$$

THE UNIVERSITY *of* EDINBURGH

Input Variables

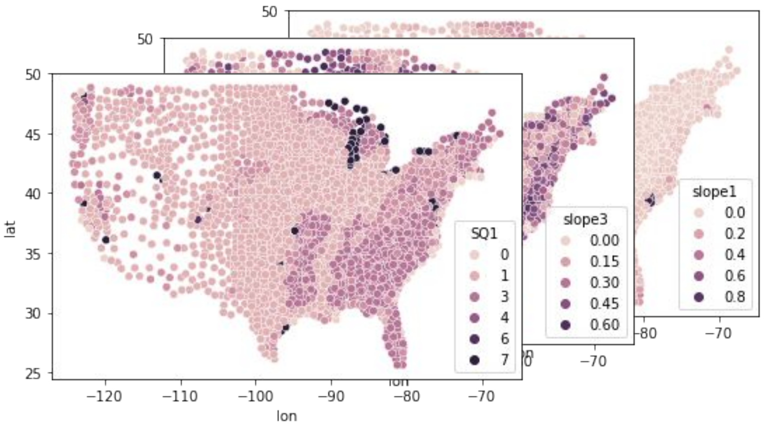
current: $\vec{x_{c,1}}, ..., \vec{x_{c,180}}$

past: $\vec{x_{p,1}}, ..., \vec{x_{p,180}}$

&

$\vec{Location} = \vec{HWSD} \oplus \vec{LatLon}$

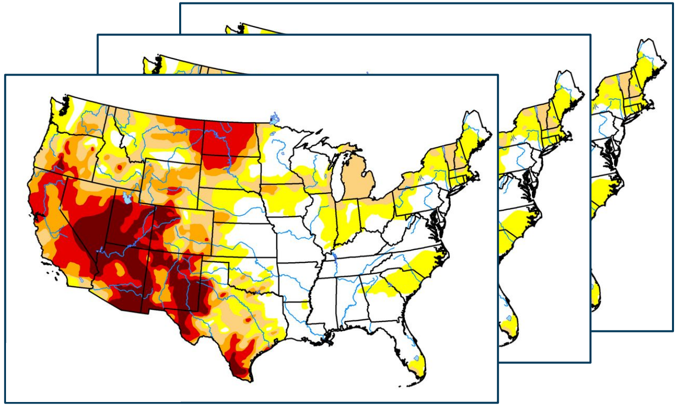
Latitude



Longitude

Δt

1 to 6 weeks

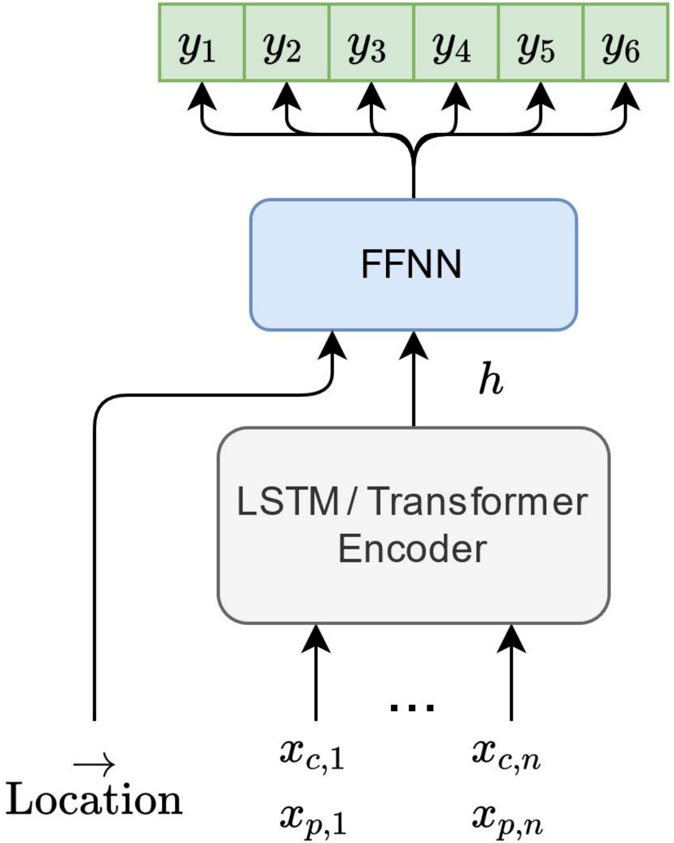


Target Variables

Time (daily, 2000-2020)

MODELS

- Use DroughtED to predict 6 future values (weekly)



EXPERIMENT RESULTS

- Comparing Model Performance on Local vs National Training Data

Training Data	Evaluation Data	Week 1 (%)
Iowa	Iowa	88.4
Montana	Montana	53.1
Oklahoma	Oklahoma	70.9
All	Iowa	**90.1**
	Montana	**55.8**
	Oklahoma	**75.8**

CONCLUSION

- Baseline models performed better on multi-regional data
- Baseline models performed favourably to SOTA

FUTURE WORK

- Expand to regions beyond the US, test further models
- kaggle.com/cdminix/us-drought-meteorological-data
- github.com/minixc/droughted_scripts