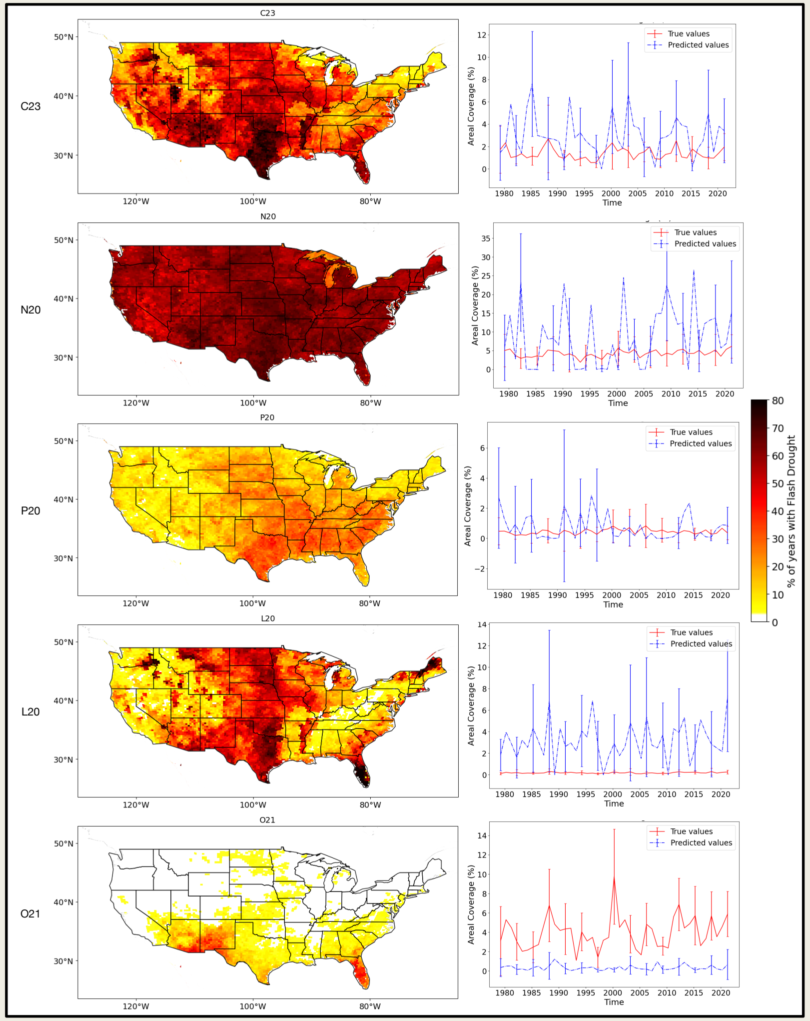
EVALUATION OF FLASH DROUGHT IDENTIFICATION WITH MACHINE LEARNING TECHNIQUES, PART 3: GLOBAL PERSPECTIVES

Stuart Edris[1], Jeffrey Basara[2] , Jordan Christian[1], Jason Furtado[1], Amy McGovern[1,3], Xiaoming Xiao[4]

[1]School of Meteorology, University of Oklahoma Norman, Oklahoma, USA; [2]Department of Environmental, Earth, and Atmospheric Sciences, University of Massachusetts Lowell, Lowell, Massachusetts, USA;
[3]School of Computer Science, University of Oklahoma, Norman, Oklahoma, USA; [4]Department of Microbiology and Plant Biology, University of Oklahoma, Norman, Oklahoma, USA

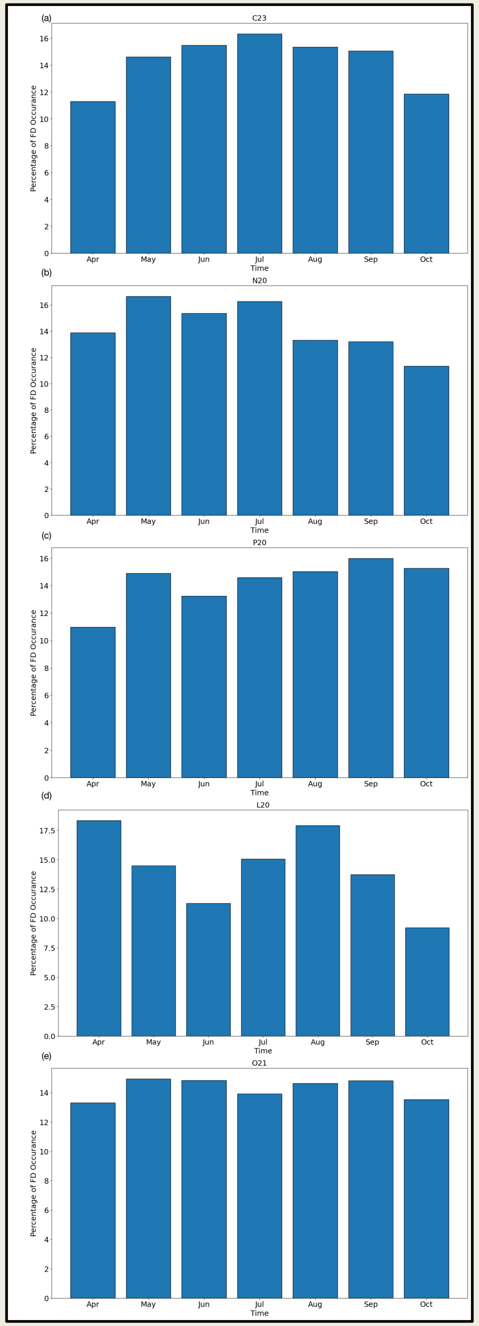
Background

- Advances in flash drought (FD) studies have delivered numerous results
  - *Understanding in FD hotpots and seasonality and the reasons for them*
  - *Numerous different methods to identify and quantify FD*
  - *Parts 1 and 2 used machine learning (ML) to represent FD events*



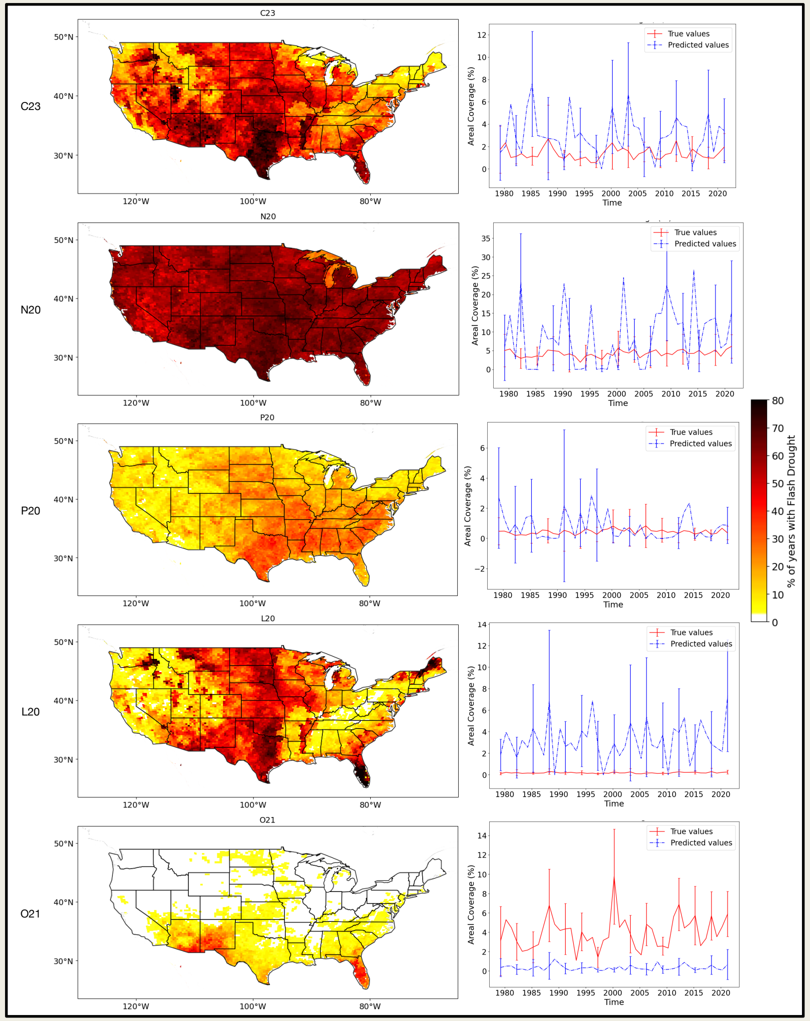
Background

- Advances in flash drought (FD) studies have delivered numerous results
  - *Understanding in FD hotpots and seasonality and the reasons for them*
  - *Numerous different methods to identify and quantify FD*
  - *Parts 1 and 2 used machine learning (ML) to represent FD events*
  - *Parts 1 and 2 also highlighted some of the differences in different FD identification methods*



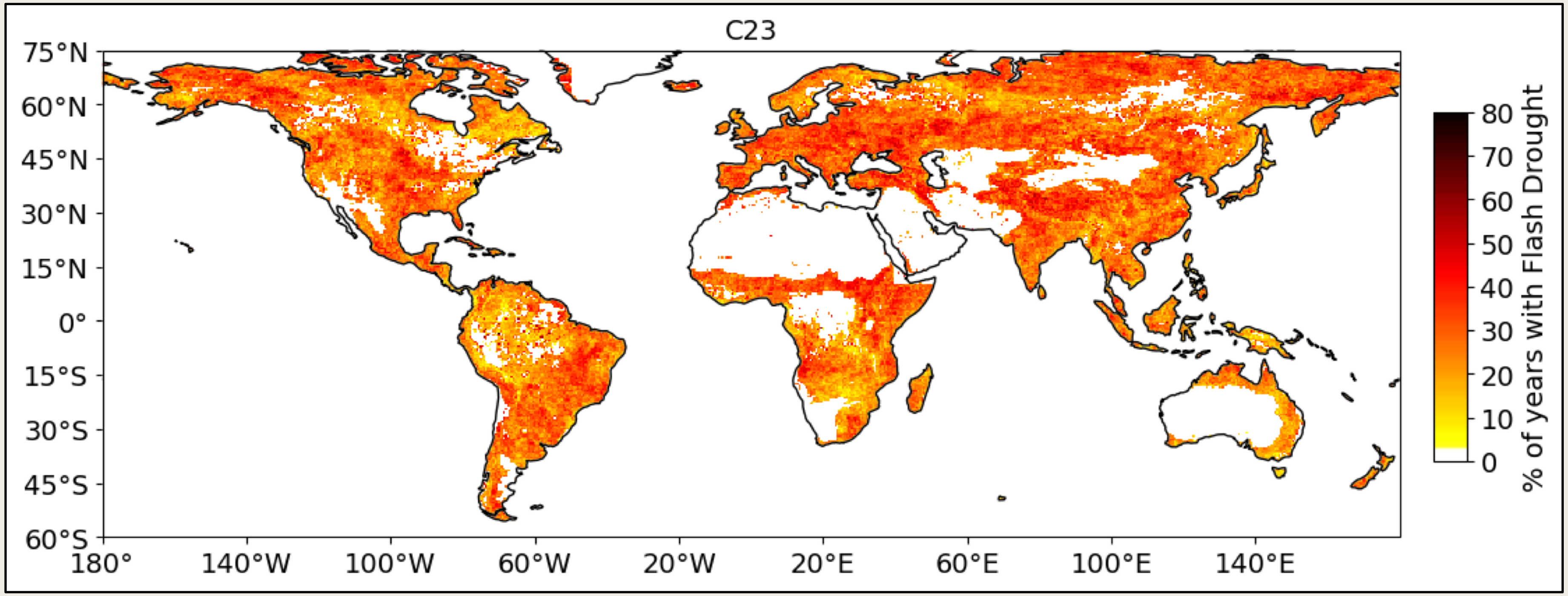
Background

- Advances in flash drought (FD) studies have delivered numerous results
  - *Understanding in FD hotpots and seasonality and the reasons for them*
  - *Numerous different methods to identify and quantify FD*
  - *Parts 1 and 2 used machine learning (ML) to represent FD events*
  - *Parts 1 and 2 also highlighted some of the differences in different FD identification methods*
- However, many studies focus on a specific region, with only a handful expanding to global analyses



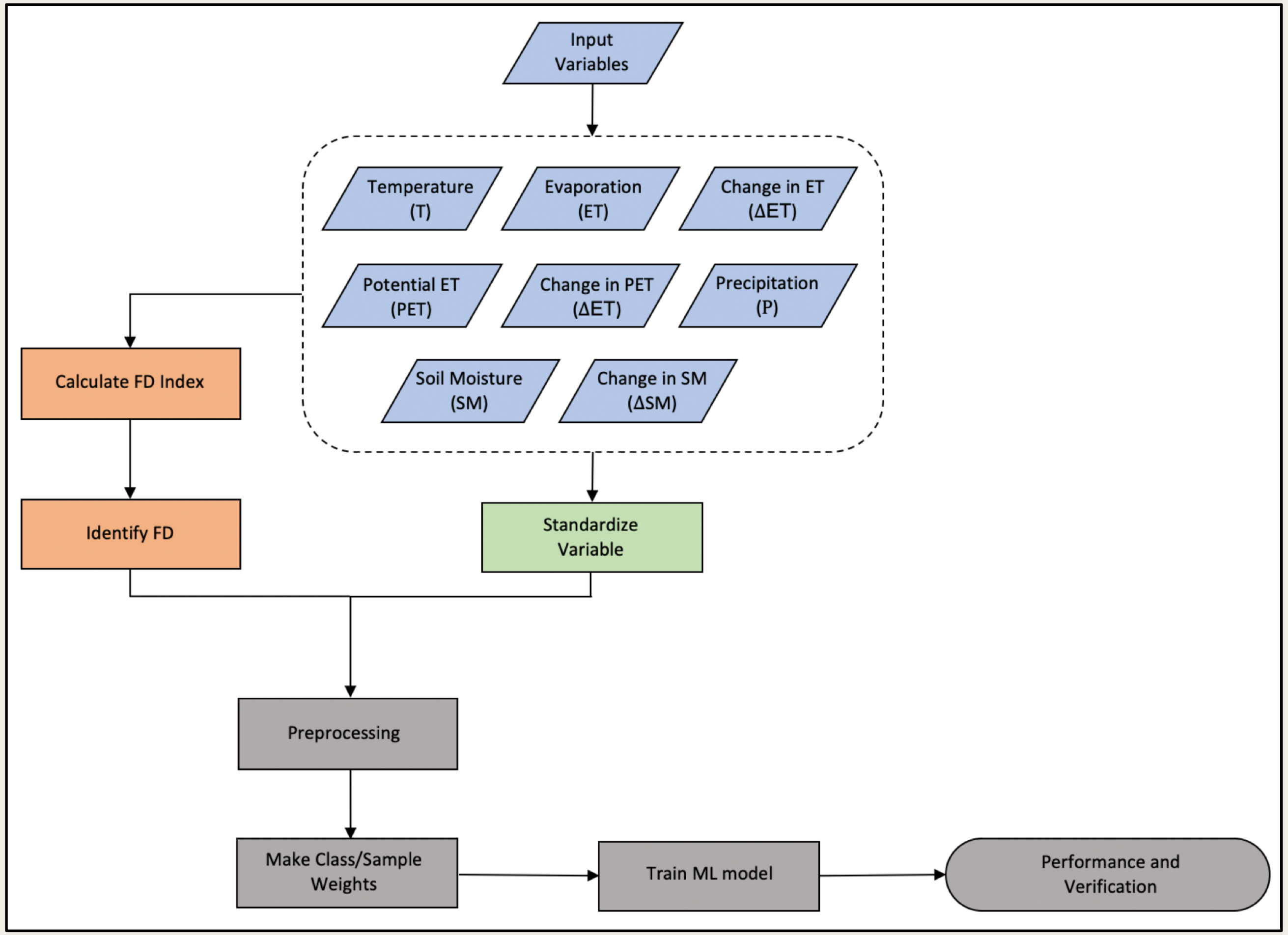
Research Goals

- Expand on parts 1 and 2 and investigate the ability of ML algorithms to generalize and learn global patterns
  - *Can it capture trends in FD in different locations? Different seasonalities? Different hotspots?*
  - *This also allows the opportunity to look at FD events in locations that are not often talked about*



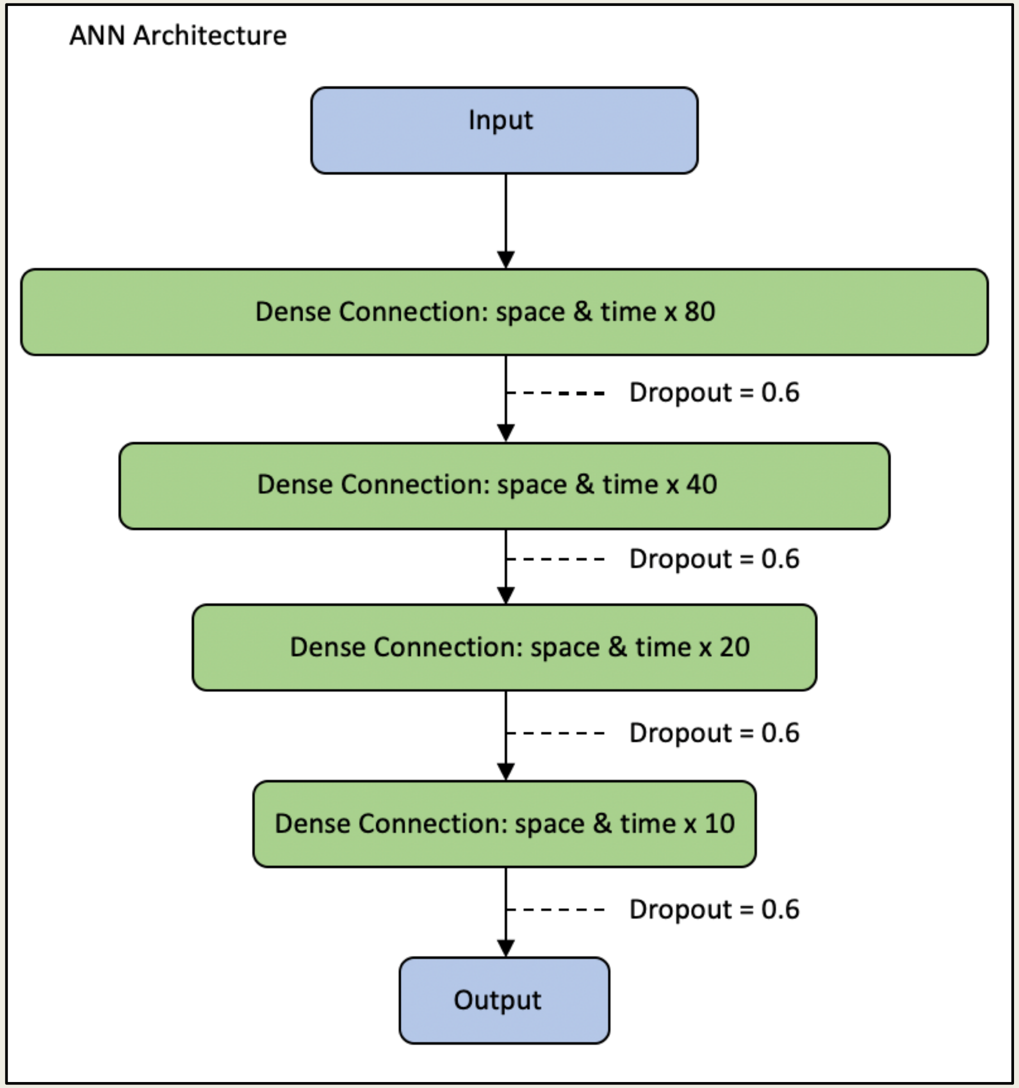
Dataset

- Global data was taken from the ERA5 reanalysis
  - *1979 – 2021 year range to match the previous parts*
  - *Data averaged to the pentad timescale*
  - *Change variables are the 30 day mean in pentad differences*
- Growing season defined as:
  - *Northern Hemisphere: April – October*
  - *Southern Hemisphere: September – March*
- Presented results focus on the FD identification method outlined in Christian et al. 2023 (C23)

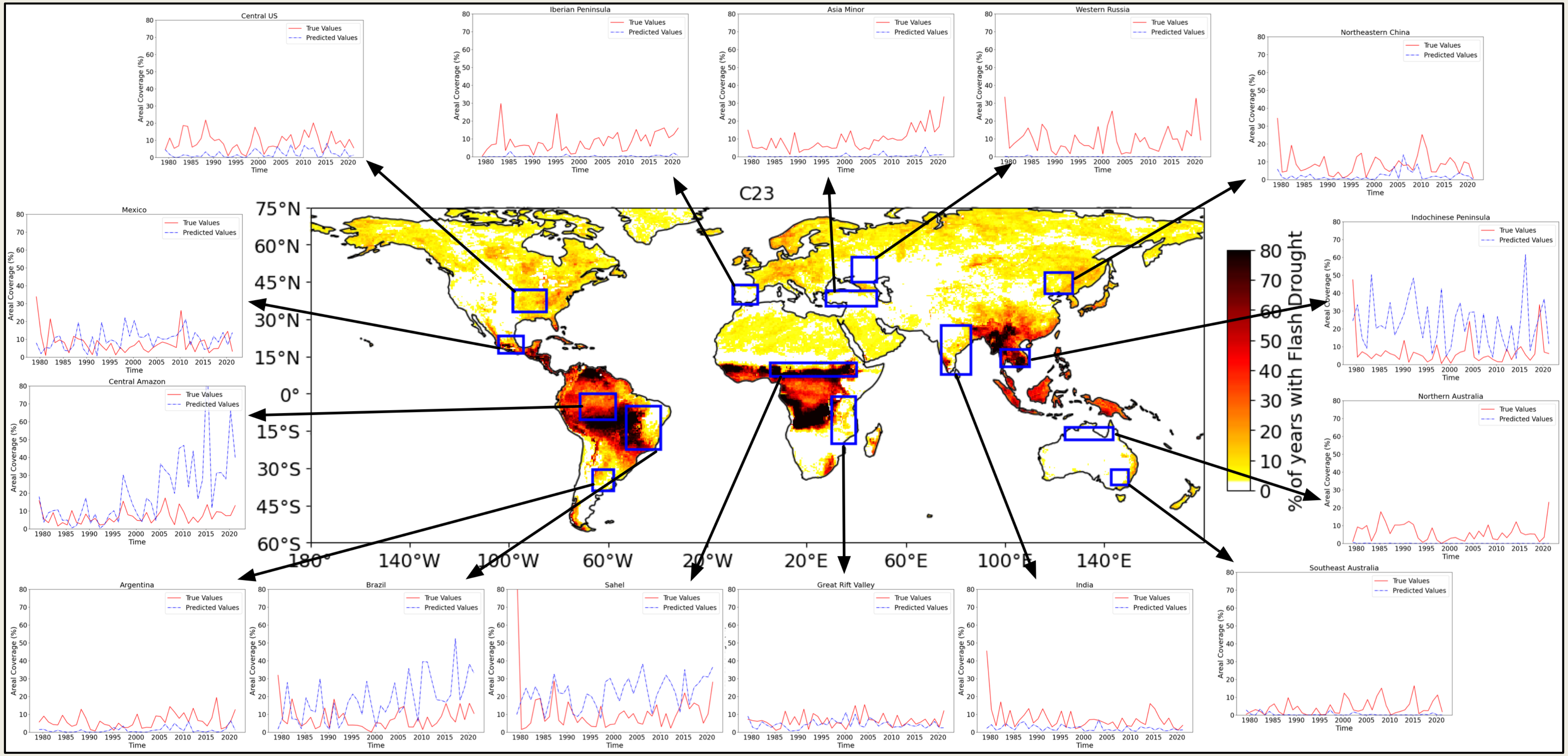


Neural Network Design

- Current results trained on a densely connected artificial neural network (ANN)
  - *Traditional NNs connect each variable/node to in a layer to all the nodes in the next layer*
- Used K-Fold cross validation:
  - *Data was split into 42 folds (1 fold = 1 growing season)*
  - *40 folds were used for training, 1 for validation, 1 for testing*
- Each (land) grid point and pentad was treated as an example
  - *ANN takes in any set of input variables and gives a probability of FD*

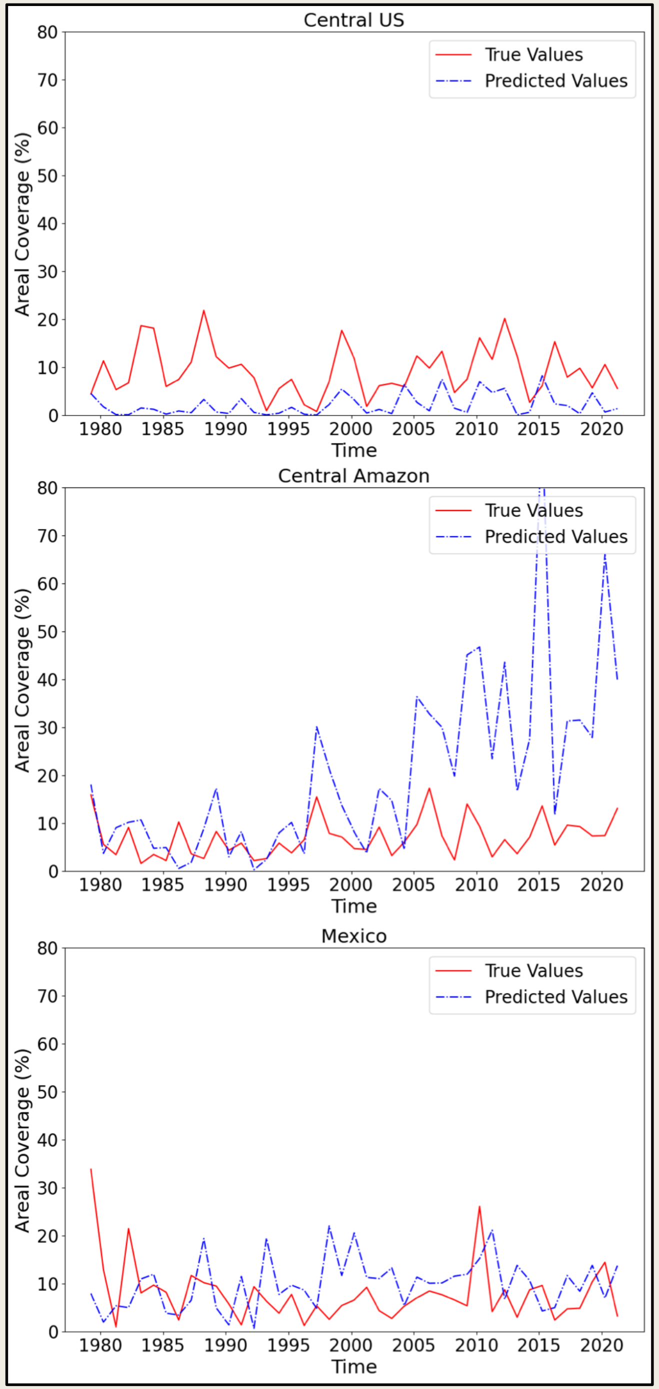


Climatological Predictions

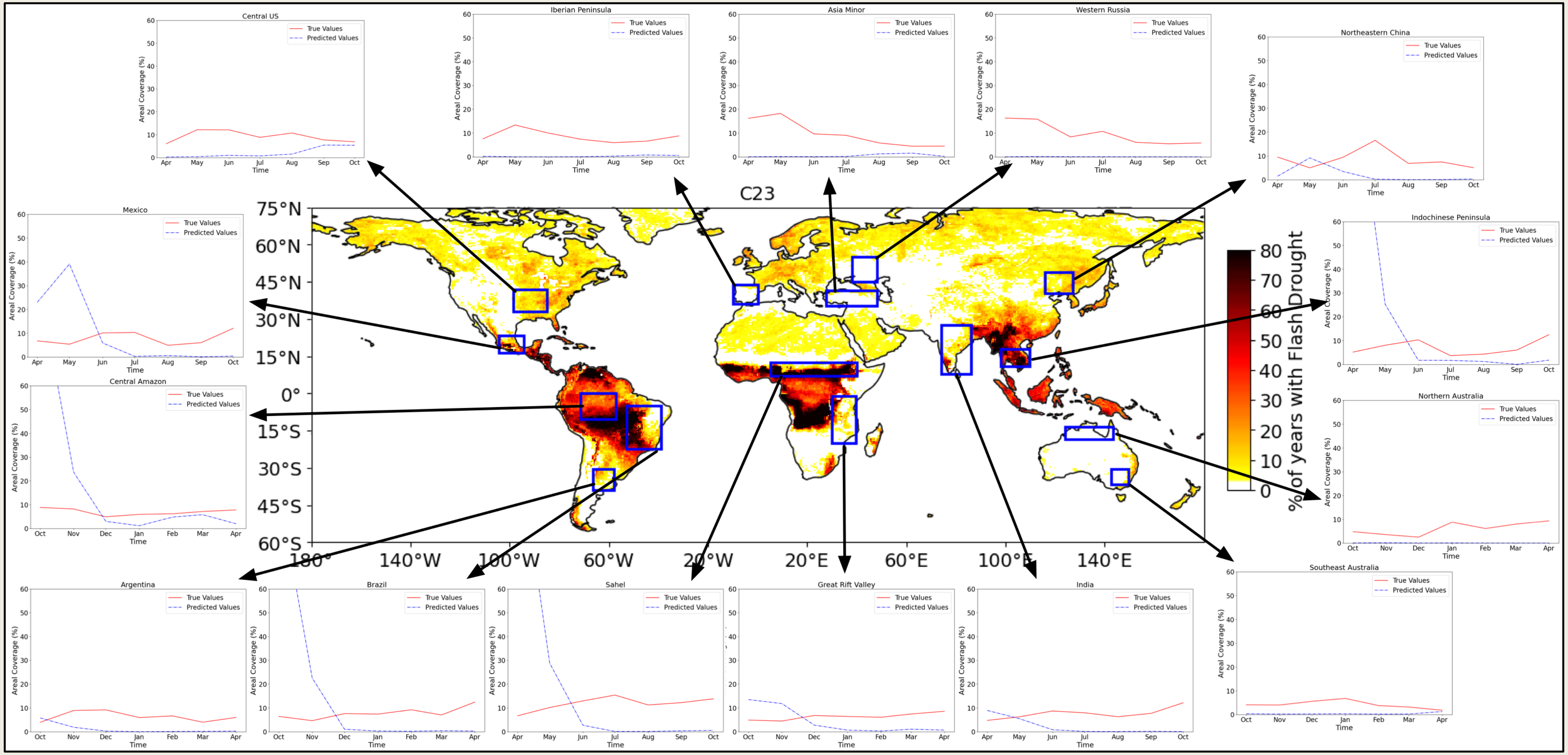


Climatological Predictions

- Three general patterns occur across all regions
  - *ANNs heavily under predict FD coverage (namely happens in higher latitudes)*
  - *ANNs start by performing well, then start to think there is an increasing trend in FD coverage that is not seen in ERA5 (more common for some tropical regions)*
  - *ANNs were able to represent FD coverage reasonably well and learn what year had more widespread events (few transitional zones)*

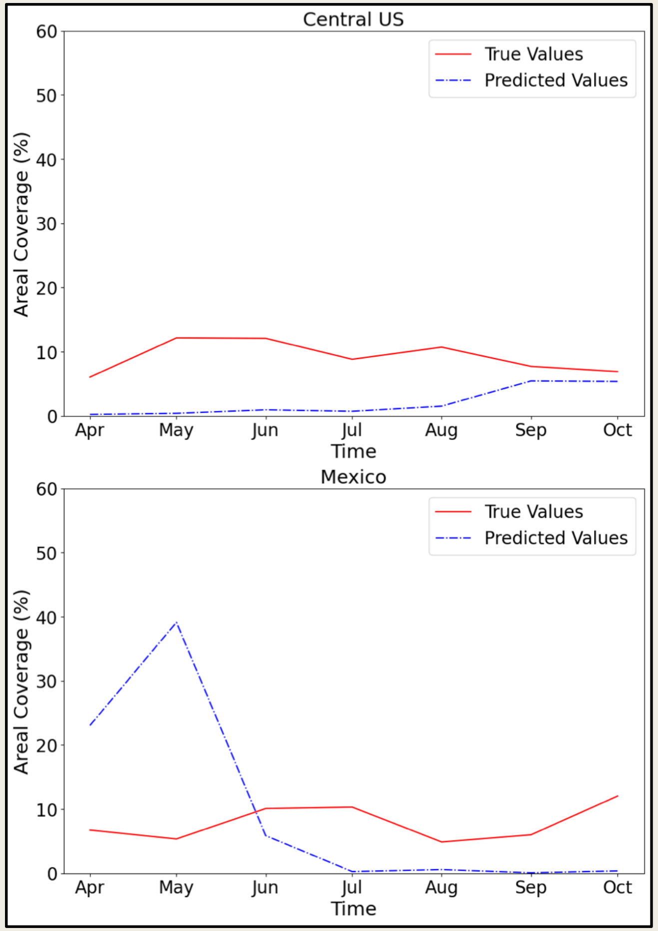


Climatological Predictions

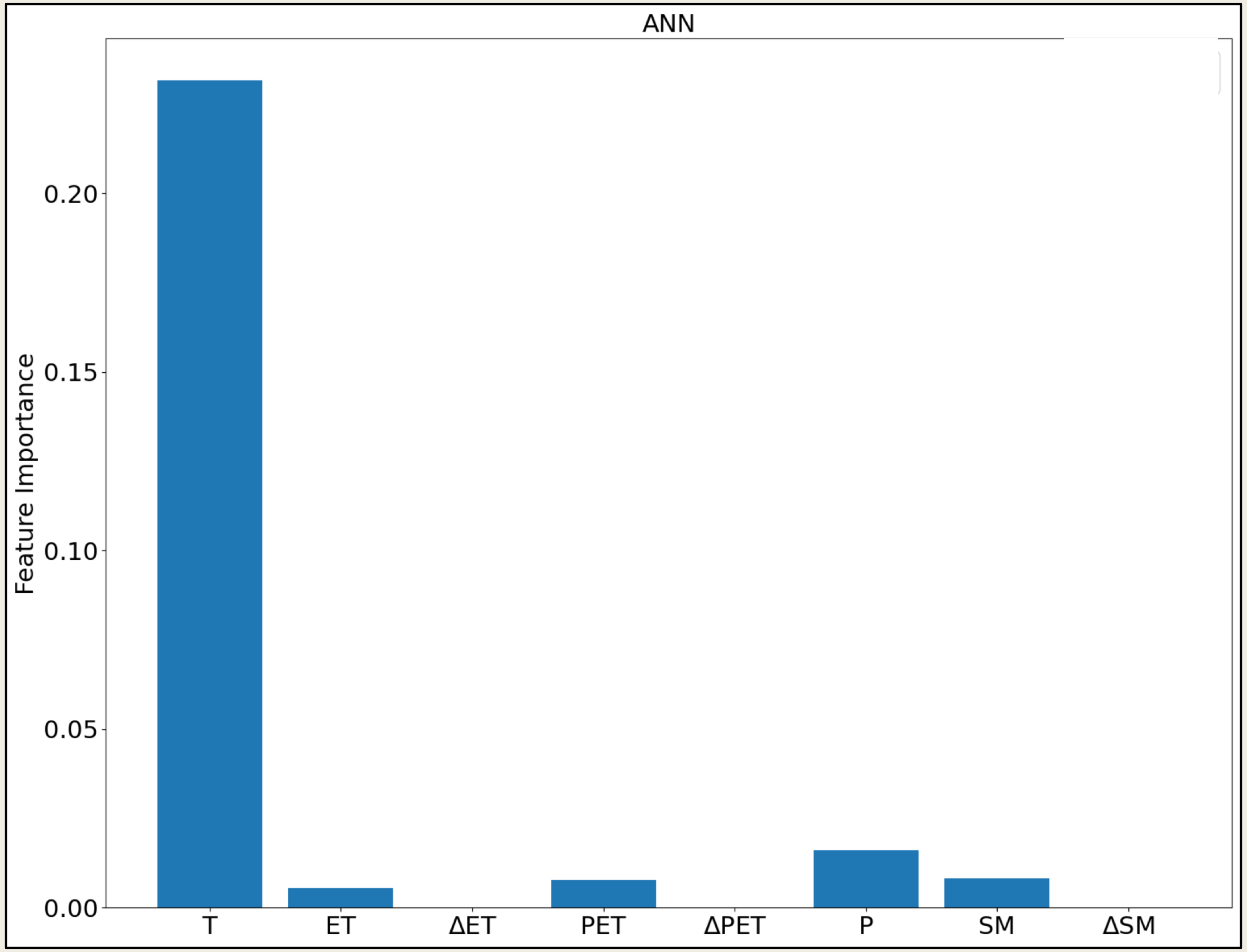


Climatological Predictions

- ANNs struggled with the timing of FD, identifying them either too early or too late
  - *ANN primarily identifies FD primarily at the start or end of the growing season*
  - *Potential reason is the ANN is mainly looking at temperature increases/decreases to capture FD (i.e., it is not learning the surface interactions)*

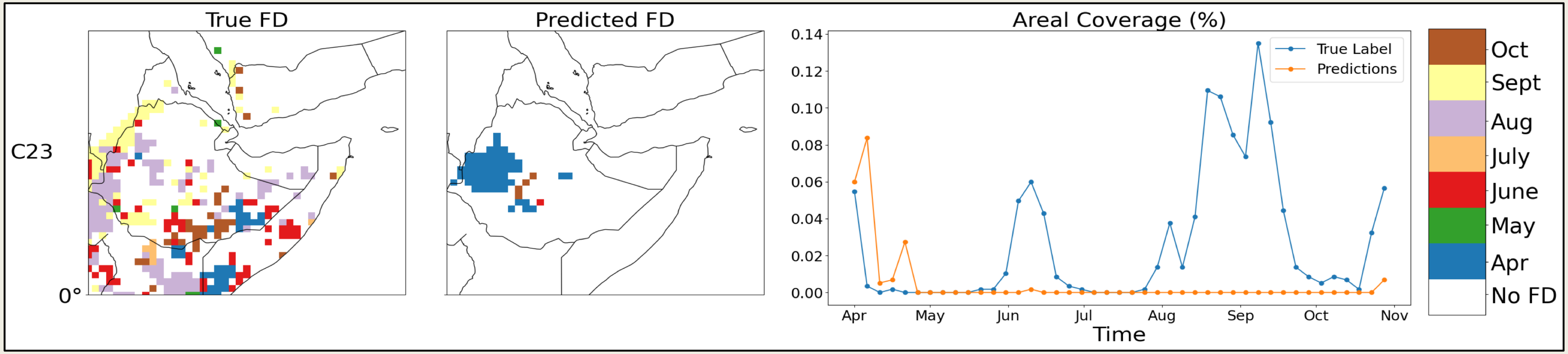


Feature Importance



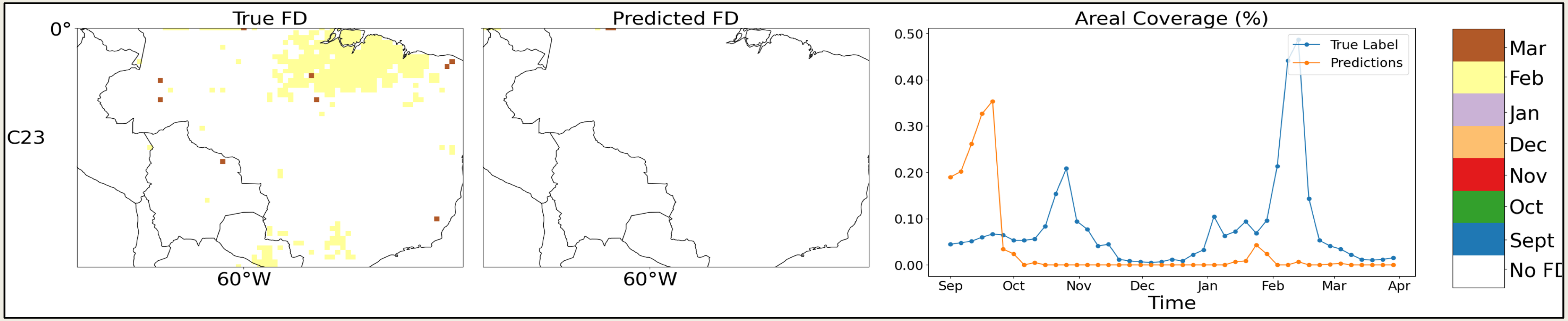
Case Study: Eastern Africa (2016)

- Severe drought that resulted in famine
  - *40% - 80% of the population (depending on country) is employed in and depends on local agriculture*



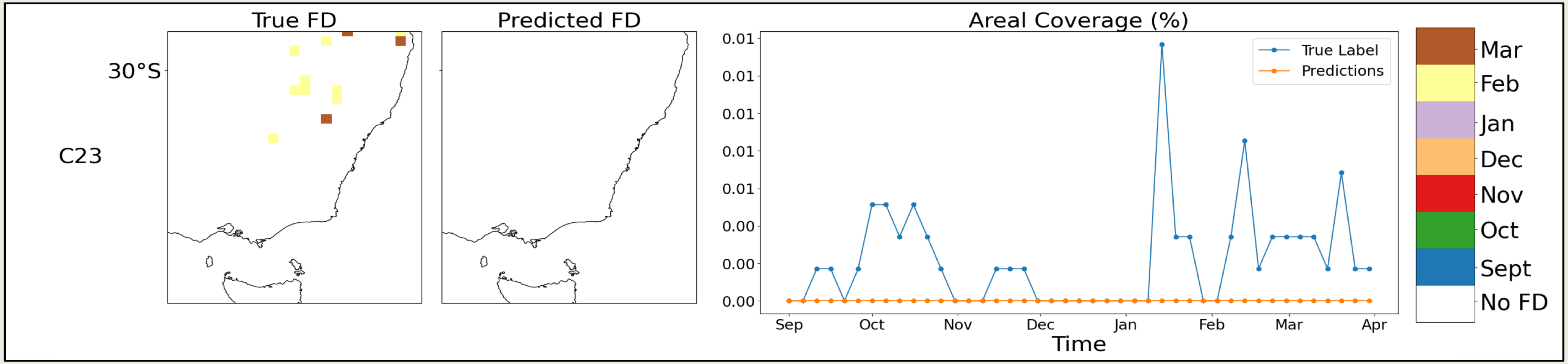
Case Study: Amazon (2015 – 2016)

- Record level drought, primarily impacted vegetations
  - *Estimated CO_2 emissions from killed trees exceed some developed countries*



Case Study: Southern Queensland (2018 – 2019)

- Large drought in southeastern Australia
  - *Heavily devastated ranching*



Summary and Conclusions

- ANNs showed some limited ability to distinguish spatial patterns of global FD
  - *Heavy overemphasis on tropical regions*
- ANNs struggled to generalize the temporal patterns and seasonality of FD
  - *ANNs oversimplified FDs, using primarily temperature to identify them*
- Some of the case studies provided a look at infrequently examined regions
  - *Different locations, ecosystems, and cultures gave different types of impacts for droughts and drought resilience*
  - *There is a human element in determining how impactful or drought resilient a location is*

Future Work

- Perform ML experiments with other FD identification methods
- Expand to other ML models
  - *RNNs outperformed other ML techniques in parts 1 and 2*
- Recommendations for future work include:
  - *Further investigate drought and FD impacts in different parts of the world to better understand how human response effects that impact*
  - *Improve datasets to have better estimates in all parts of the world and be less biased*
  - *Delve deeper into one or two of these ML techniques (including interpretability) to create a fully applicable, usable, and trustworthy ML model*
    - Goal is create a ML model that can be transferred to S2S ensemble models so that FDs can be identified in real time and predicted

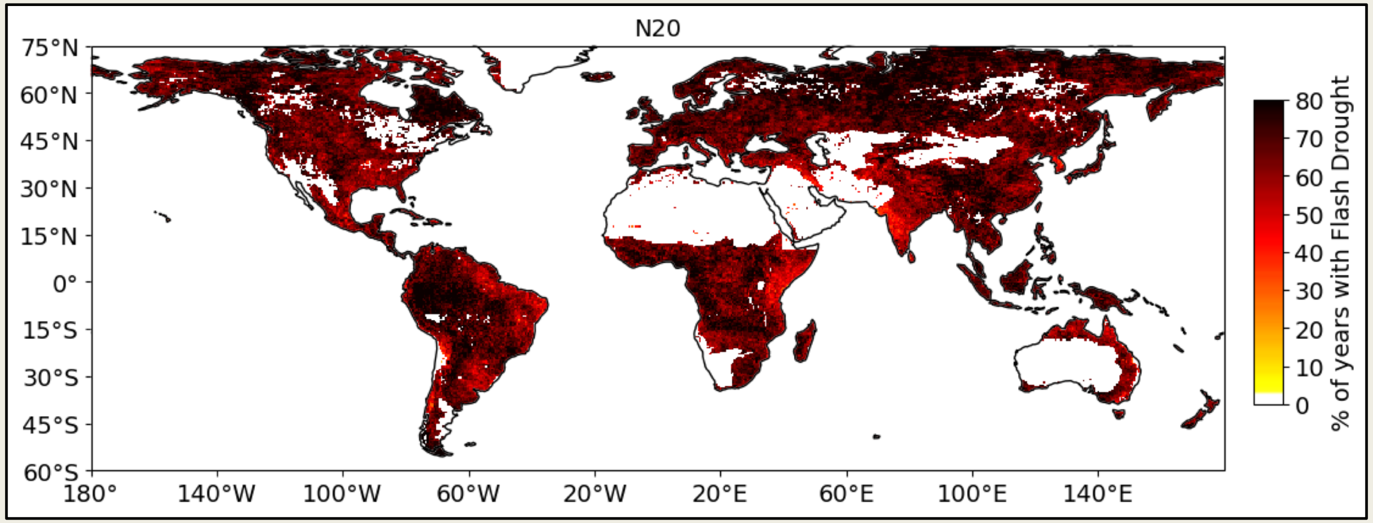
Acknowledgements

- The authors would like to acknowledge and thank the National Science Foundation, US Department of Agriculture Southern Plains Climate Hub, and the Oklahoma EPSCoR for funding this research project.

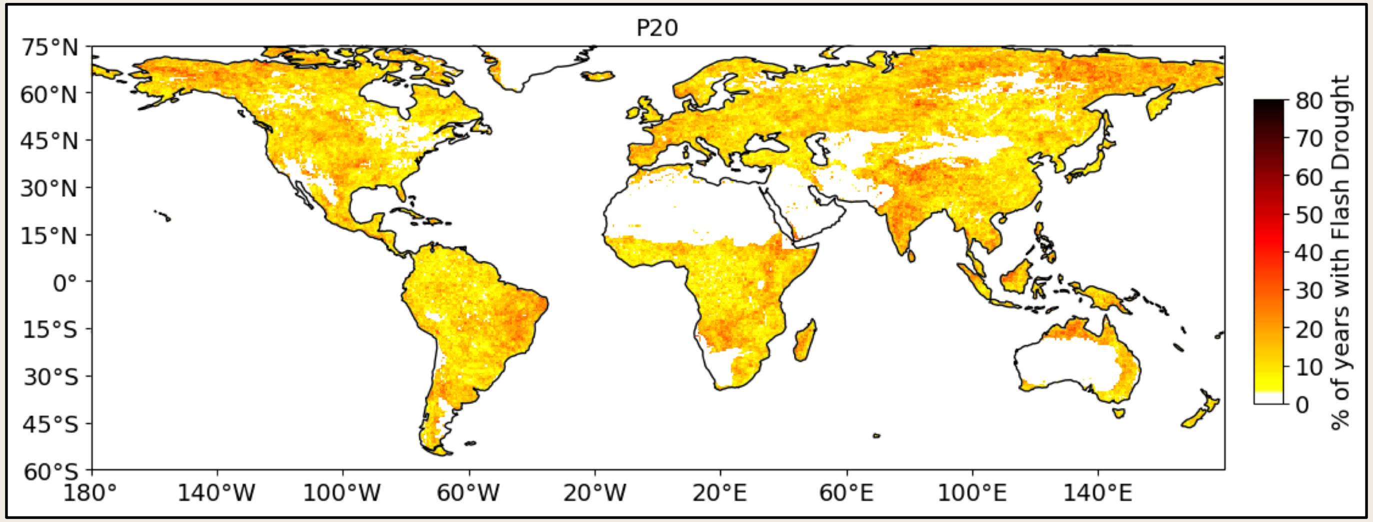
Summary and Conclusions

- ANNs showed some limited ability to distinguish spatial patterns of global FD
  - *Heavy overemphasis on tropical regions*
- ANNs struggled to generalize the temporal patterns and seasonality of FD
  - *ANNs oversimplified FDs, using primarily temperature to identify them*
- Some of the case studies provided a look at infrequently examined regions
  - *Different locations, ecosystems, and cultures gave different types of impacts for droughts and drought resilience*
  - *There is a human element in determining how impactful or drought resilient a location is*

Extra

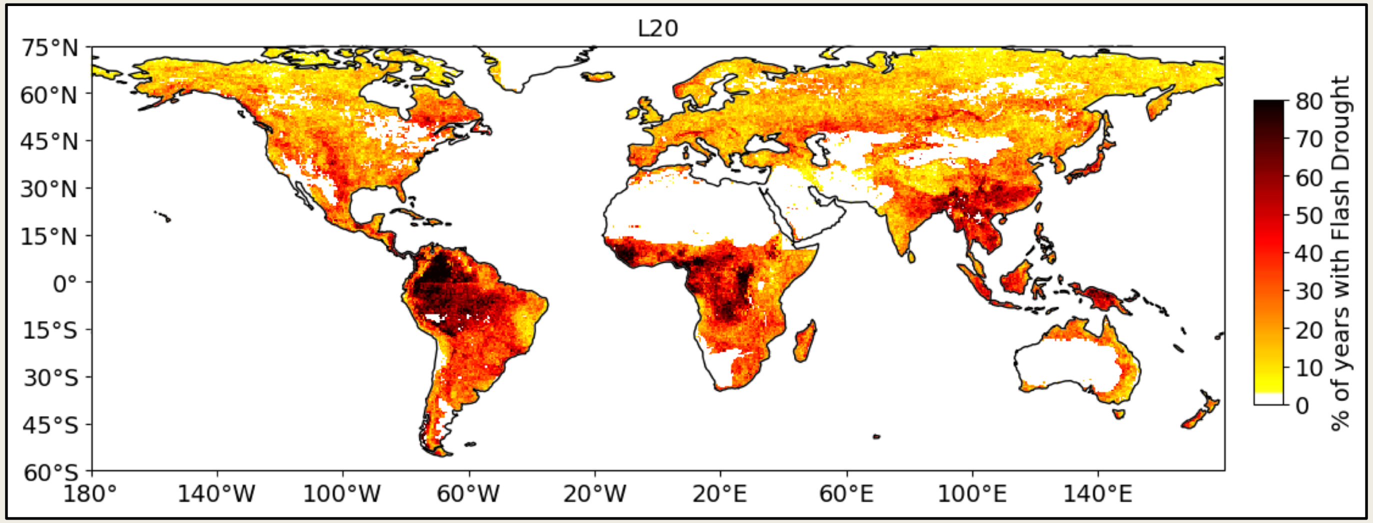


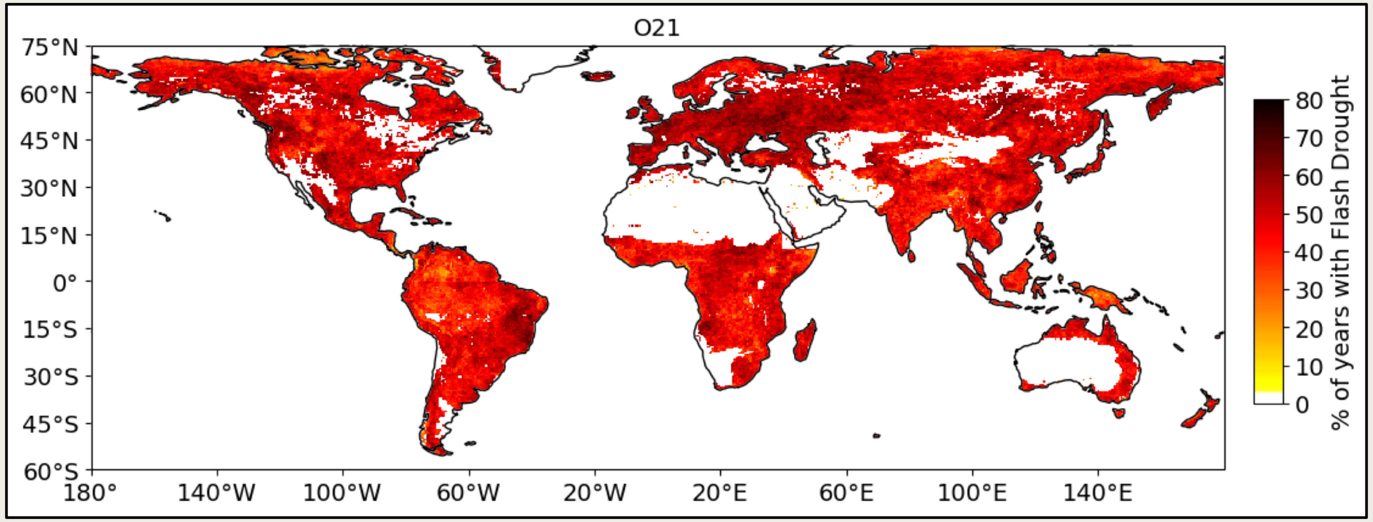
N20
75°N
60°N
45°N
30°N
15°N
0°
15°S
30°S
45°S
60°S
180°
140°W
100°W
60°W
20°W
20°E
60°E
100°E
140°E
80
70
60
50
40
30
20
10
0
% of years with Flash Drought



P20
75°N
60°N
45°N
30°N
15°N
0°
15°S
30°S
45°S
60°S
180°
140°W
100°W
60°W
20°W
20°E
60°E
100°E
140°E
80
70
60
50
40
30
20
10
0
% of years with Flash Drought

Extra


L20
75°N
60°N
45°N
30°N
15°N
0°
15°S
30°S
45°S
60°S
180°
140°W
100°W
60°W
20°W
20°E
60°E
100°E
140°E
80
70
60
50
40
30
20
10
0
% of years with Flash Drought


O21
75°N
60°N
45°N
30°N
15°N
0°
15°S
30°S
45°S
60°S
180°
140°W
100°W
60°W
20°W
20°E
60°E
100°E
140°E
80
70
60
50
40
30
20
10
0
% of years with Flash Drought