EVALUATION OF FLASH DROUGHT IDENTIFICATION WITH MACHINE LEARNING TECHNIQUES, PART 2: DEEP LEARNING ALGORITHMS

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

[1]School of Meteorology, University of Oklahoma, [2]School of Civil Engineering and Environmental Science, University of Oklahoma, [3] School of Computer Science, University of Oklahoma, [4]Department of Microbiology and Plant Biology, University of Oklahoma

ML IN DROUGHT

- Over the coarse of recent years, ML techniques (most popularly random forest and neural network techniques) have been shown to be able to improve our identification and predictions of drought events

- However, investigation of ML in flash droughts (FD), a subset of drought events, have remained scarce

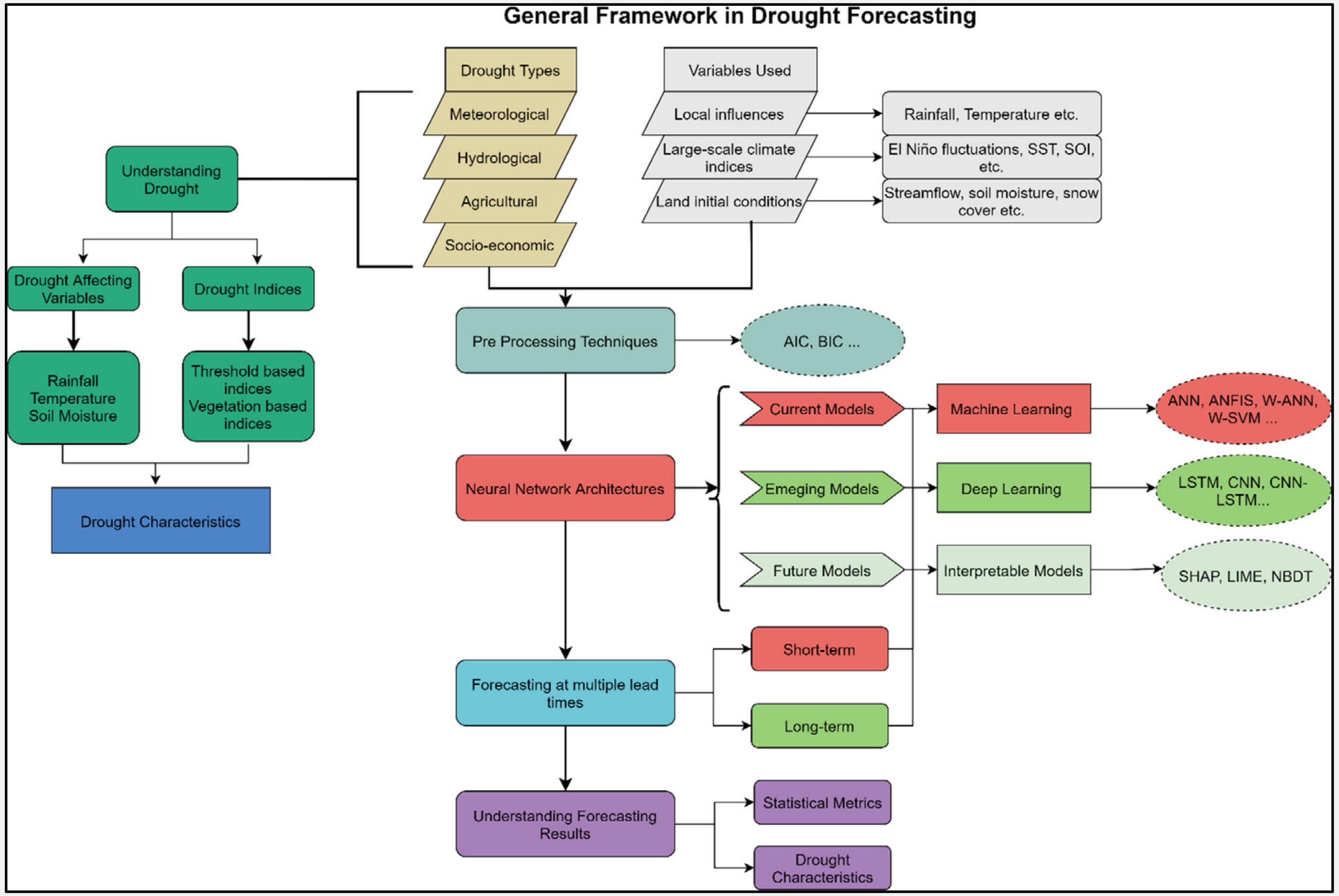


Figure 1 in Dikshit et al. 2022

FLASH DROUGHTS

- Flash droughts (FDs) are droughts that develop over relatively rapid (~1 month) time scales
- These are more complicated than traditional droughts due to their rapid intensification
  - These require variables that change on the seasonal-to-subseasonal timescale

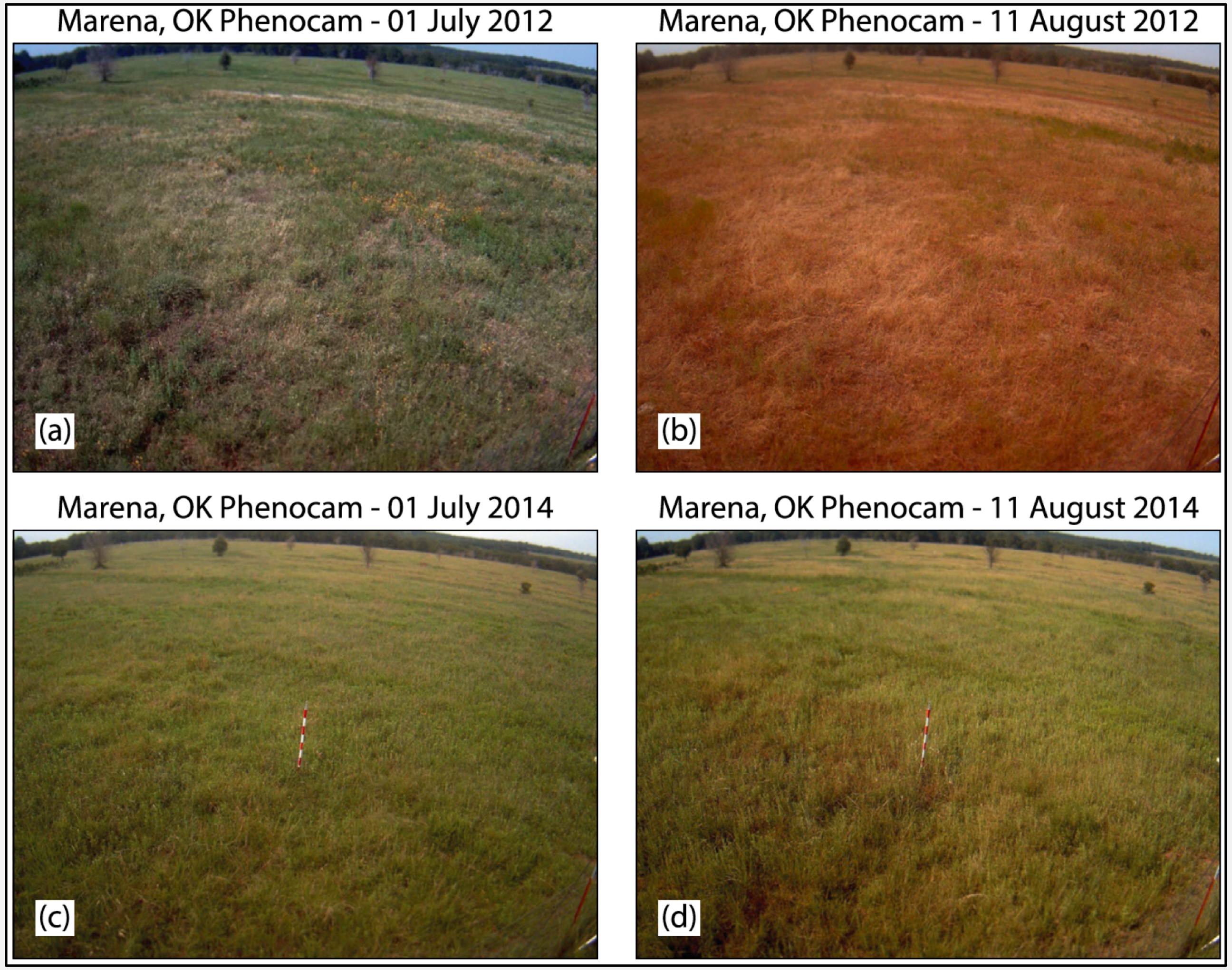


Figure 3 in Otkin et al. 2018

FLASH DROUGHTS

- Flash droughts (FDs) are droughts that develop over relatively rapid (~1 month) time scales

- These are more complicated than traditional droughts due to their rapid intensification
  - These require variables that change on the seasonal-to-subseasonal timescale

- This is the goal of this paper; to provide a starting point and investigate how well DL methods can represent FD events

- This particular study focuses on investigating deep learning methods developed in Tensorflow; part 1 focuses on some of the methods provided by the scikit-learn package

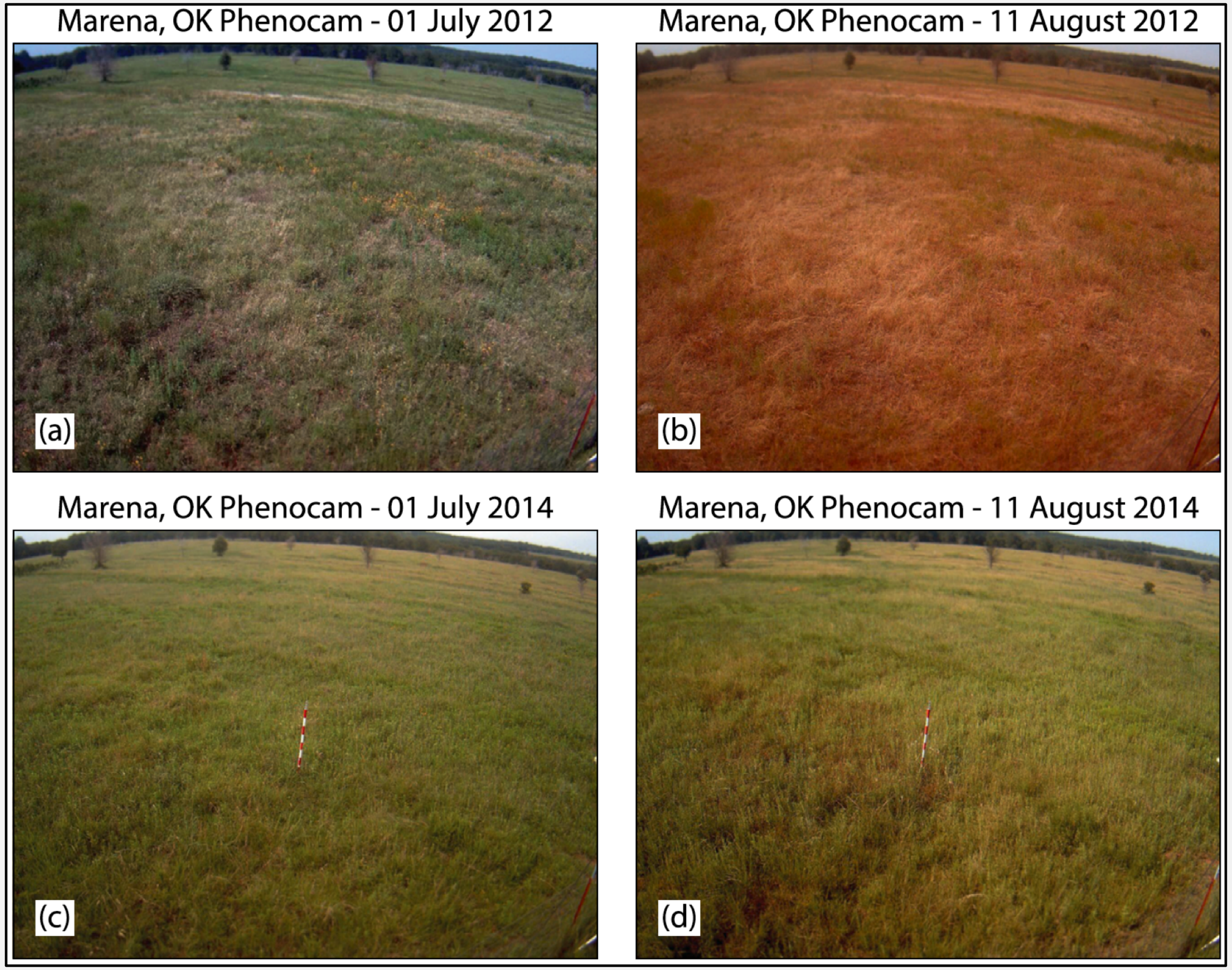
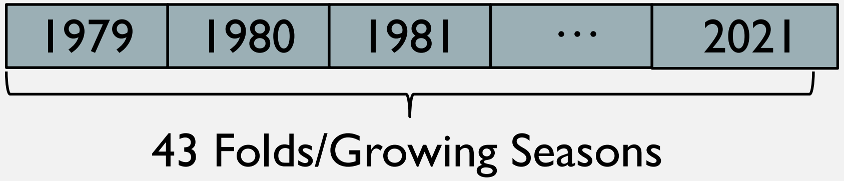


Figure 3 in Otkin et al. 2018

DATA

- Training and validation dataset was collected from the North American Regional Reanalysis
  - Data was focused on the growing seasons of 1979 – 2021
- Input data consists of variables that have been found to drive flash drought events or correlate with them:
  - Temperature (T)
  - Precipitation (P)
  - Evaporation (ET)
  - Average 1 month change in ET (ΔET)
  - Potential evaporation (PET)
  - Average 1 month change in PET (ΔPET)
  - Average 0 – 40 cm soil moisture (SM)
  - Average 1 month change in SM (ΔSM)
- Sea values were ignored (their sample weights were set to 0)



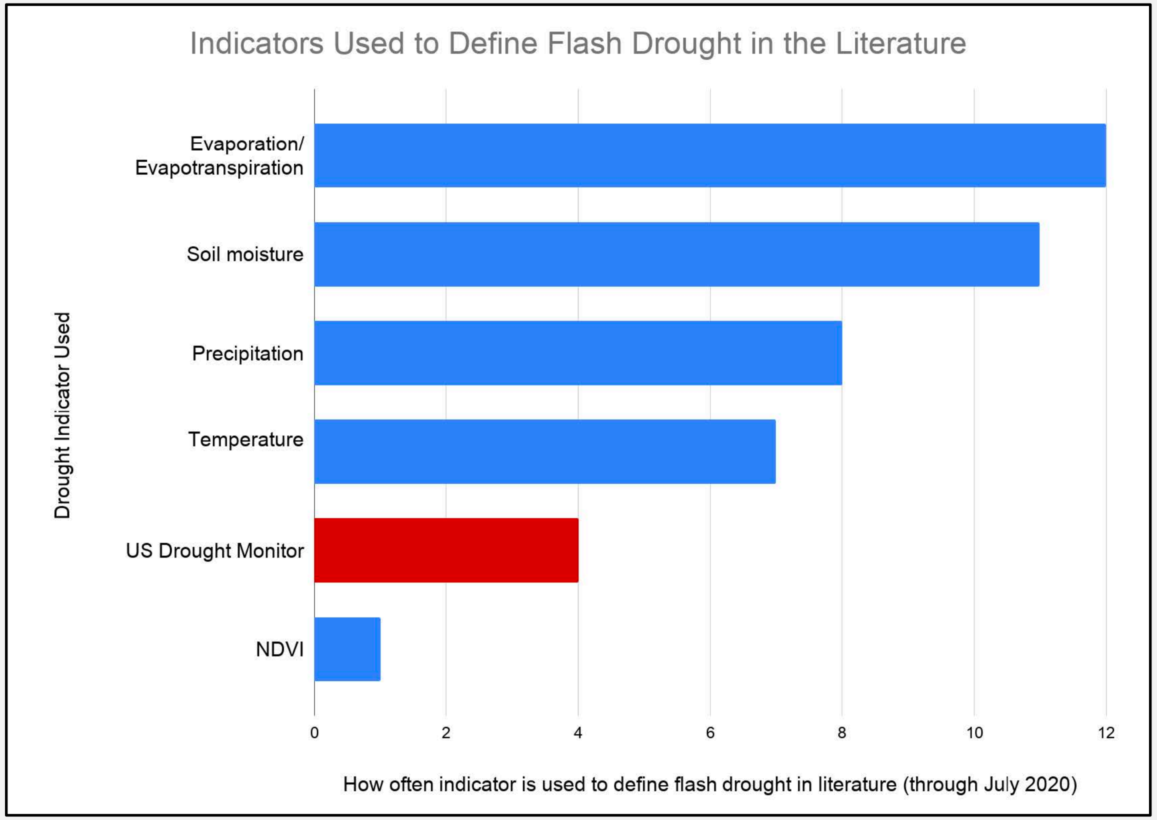


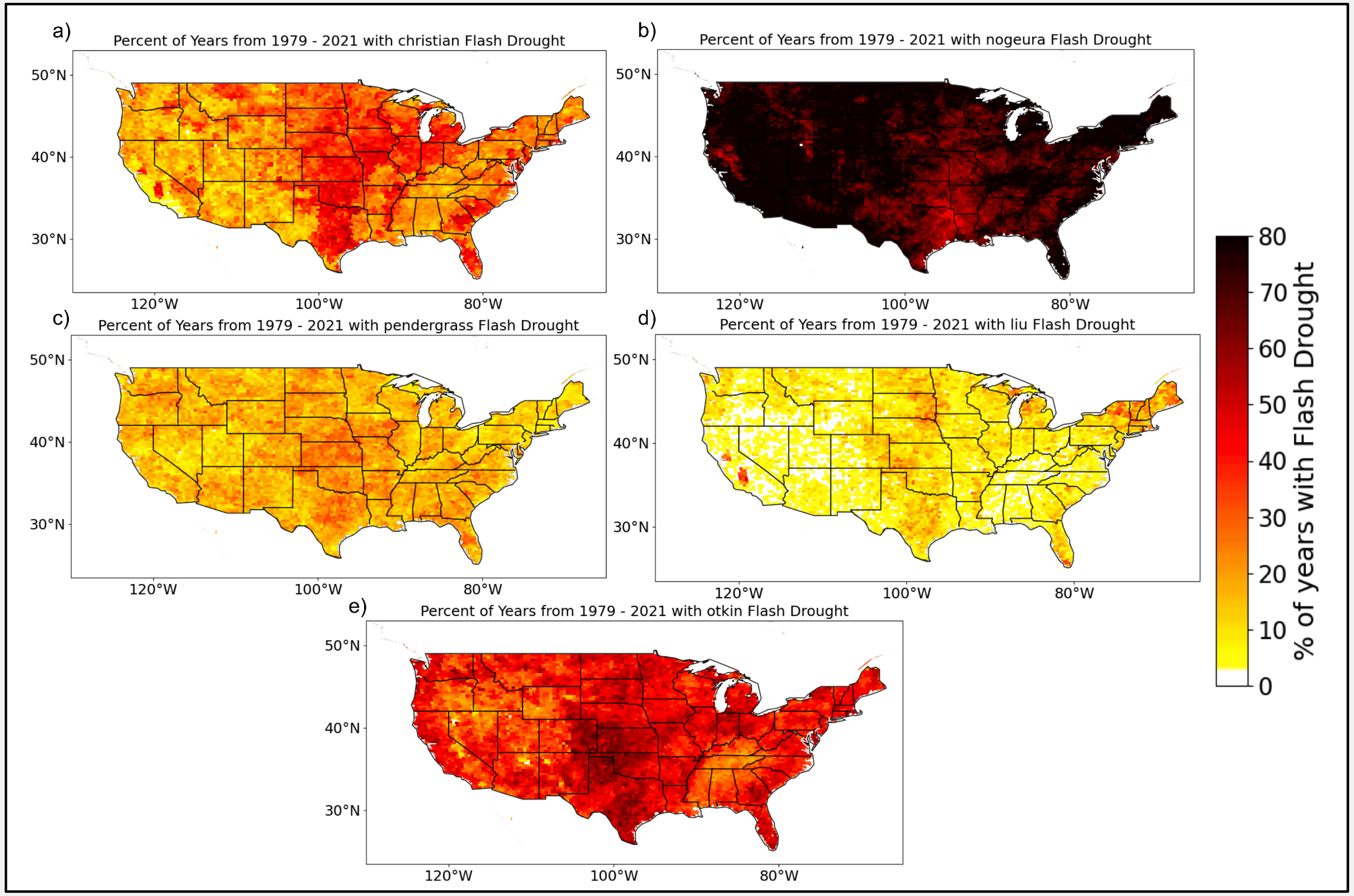
Figure 2 in Lisonbee et al. 2021

DATA

- Similar to drought, FD has no consistent quantitative definition
  - There are multiple FD identification methods focusing on different drivers
  - The DL models were trained to identify FD using 5 different identification methods separately

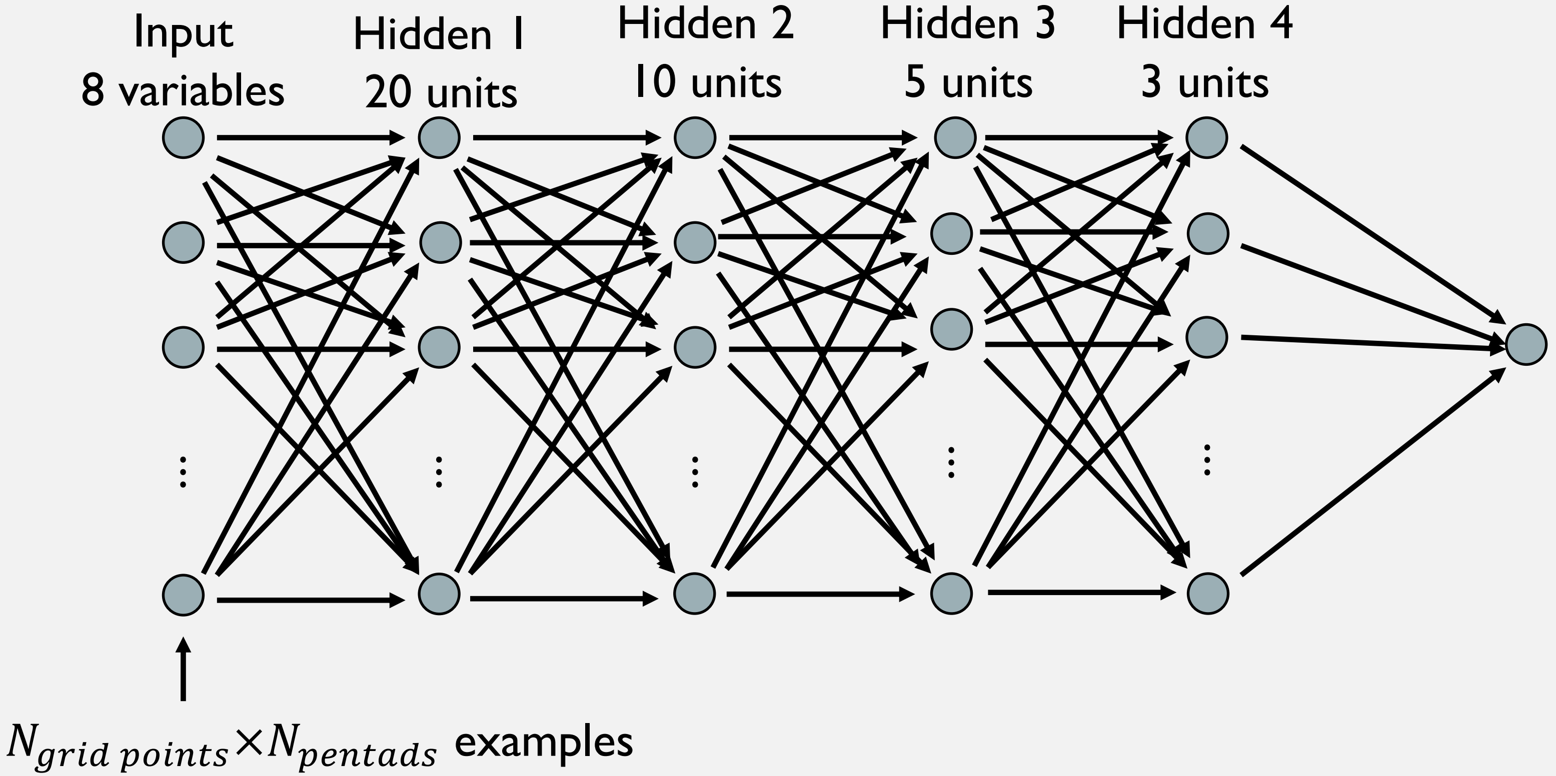
- Vegetative stress:
  - a) SESR (Christian et al. 2019; C19)
- Atmospheric moisture supply/deficit:
  - b) SPEI (Nogeura et al. 2020; N20)
- Increased atmospheric demand:
  - c) EDDI (Pendergrass et al. 2021; P21)
- Increased soil moisture stress:
  - d) SM (Liu et al. 2020; L20)
  - e) FDII (Otkin et al. 2021; O21)

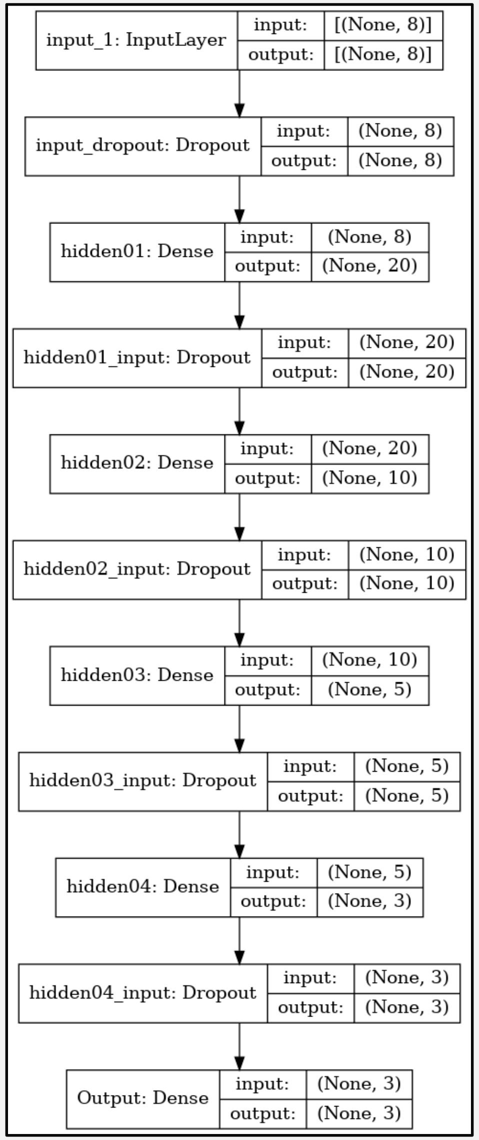
- For comparison, similar ML model parameters were used for all identification method



ANN EXPERIMENT DESIGN

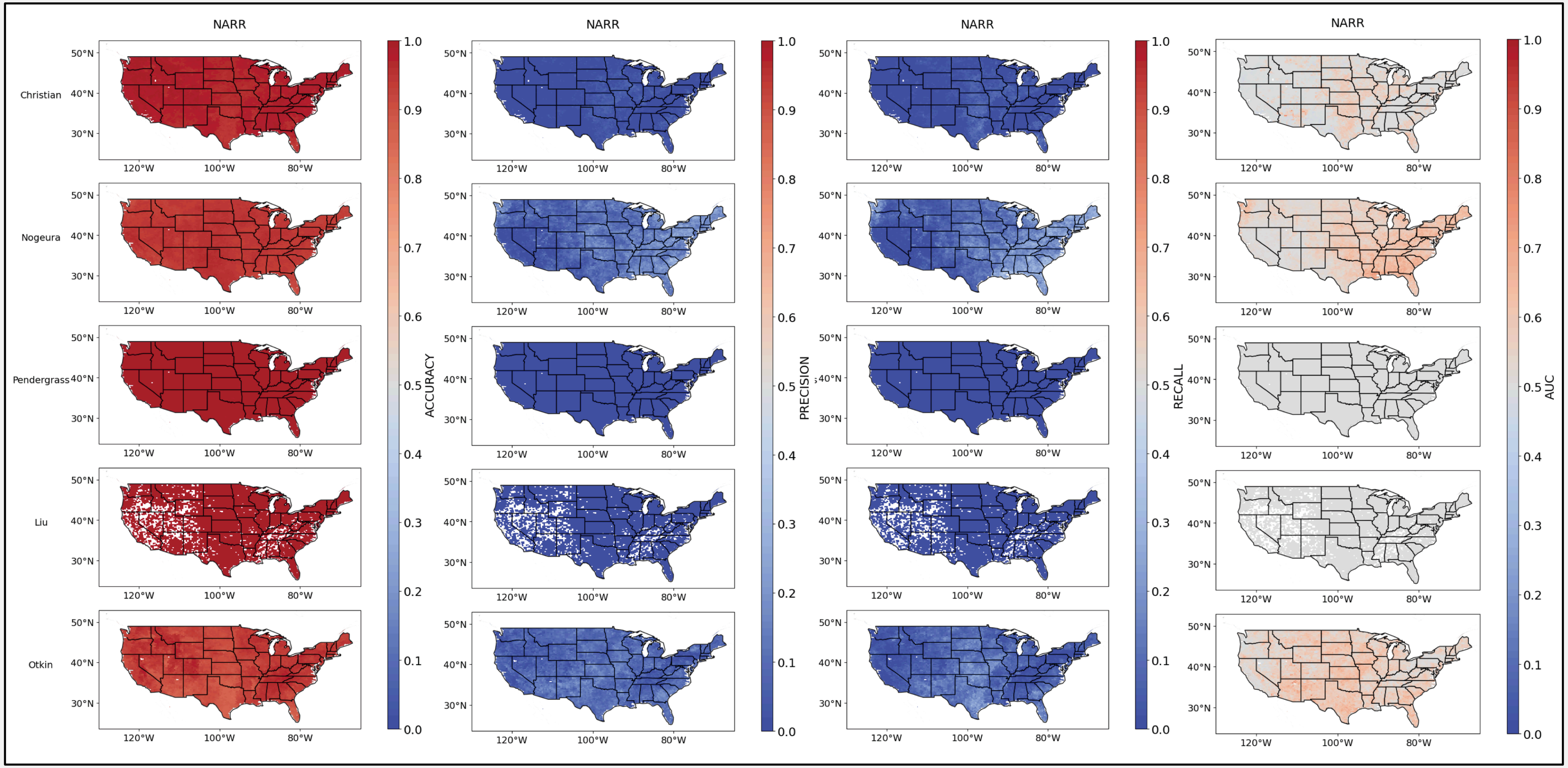
- The fully connected artificial neural networks were set up similar to an scikit-learn experiment
  - Every grid point and pentad was treated as an example
  - Data set up as $(N_{grid\ points} \times N_{pentads}, N_{variables})$





PRELIMINARY RESULTS: ANN

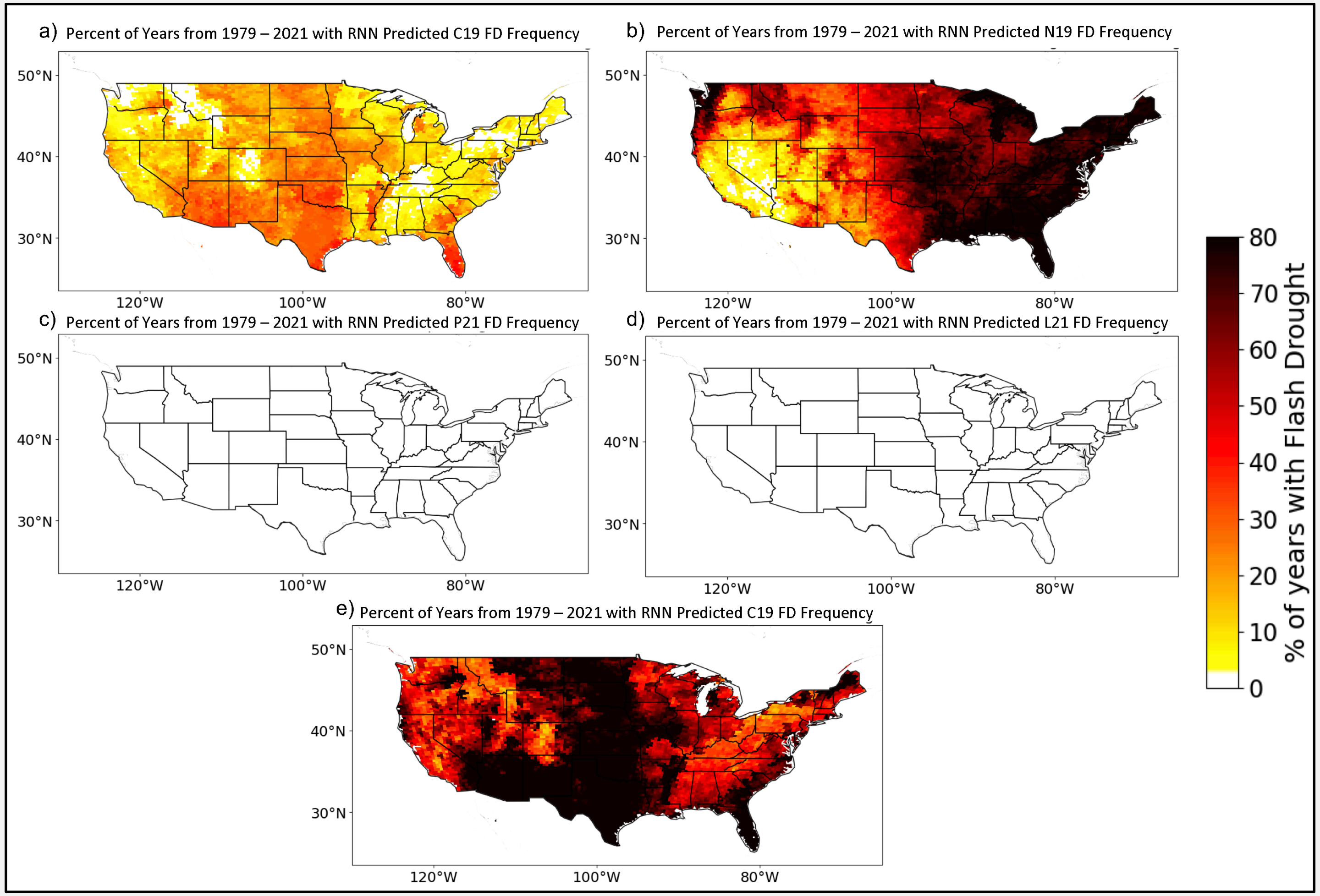
- Results highlight difference of each method; best performance in hotspot regions

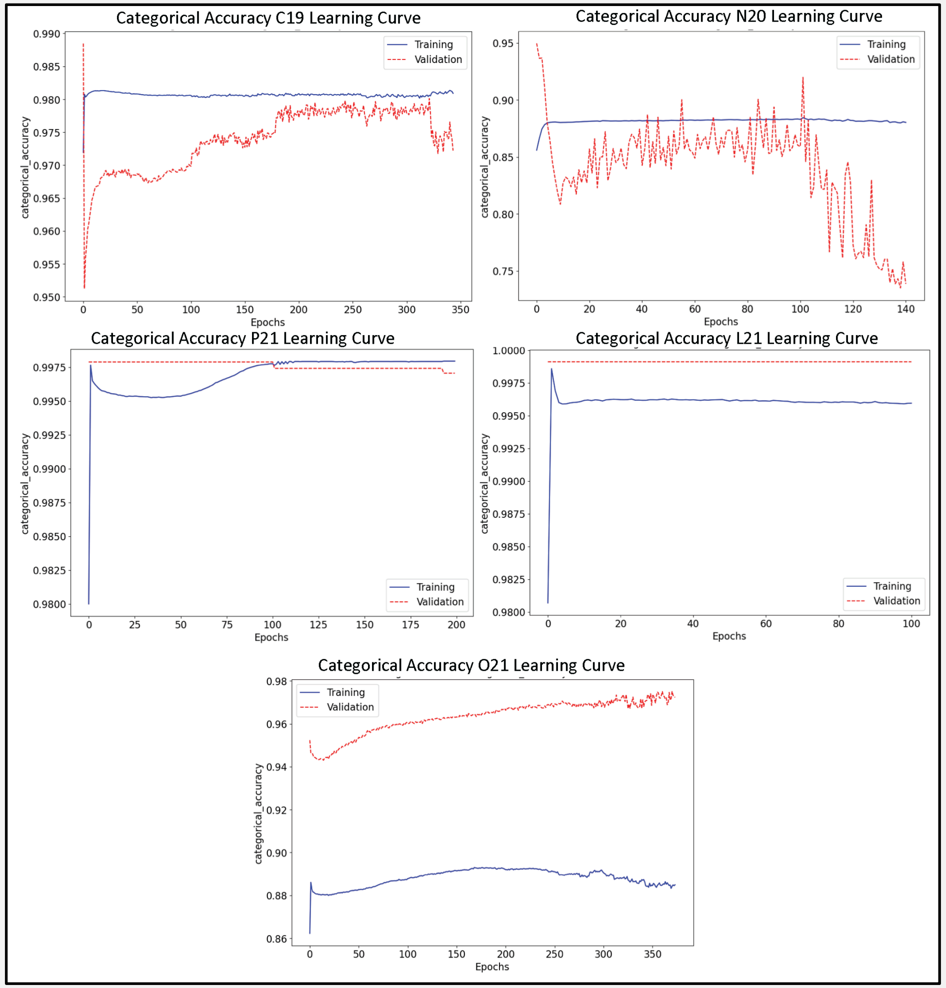


Average ANN test set results

PRELIMINARY RESULTS: ANN

- Model needs to be tuned for each FD identification method; parameters that work for some do not work for others

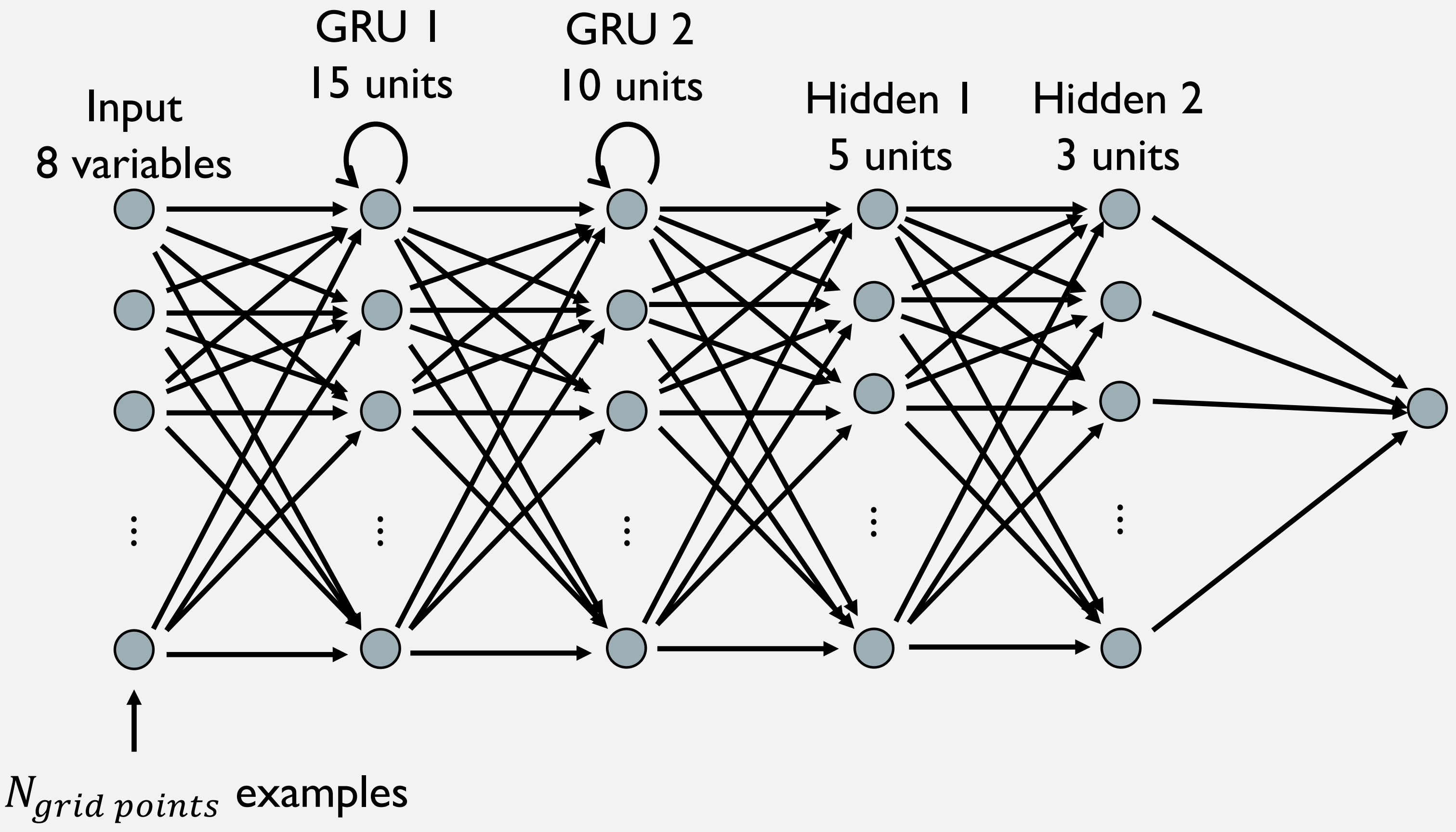


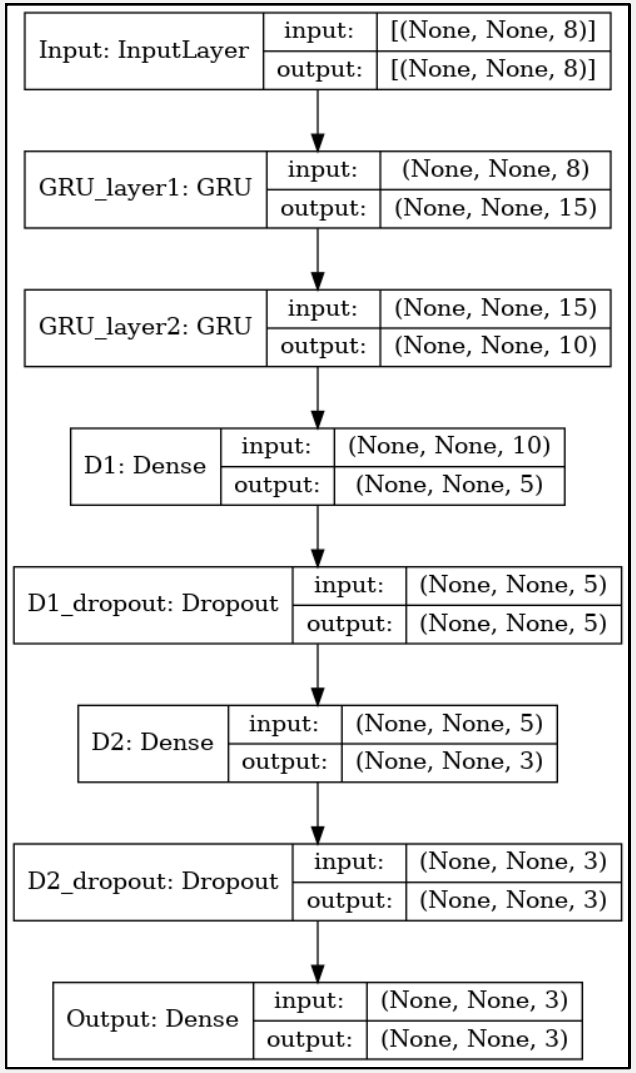


ANN test set predictions

RNN EXPERIMENT DESIGN

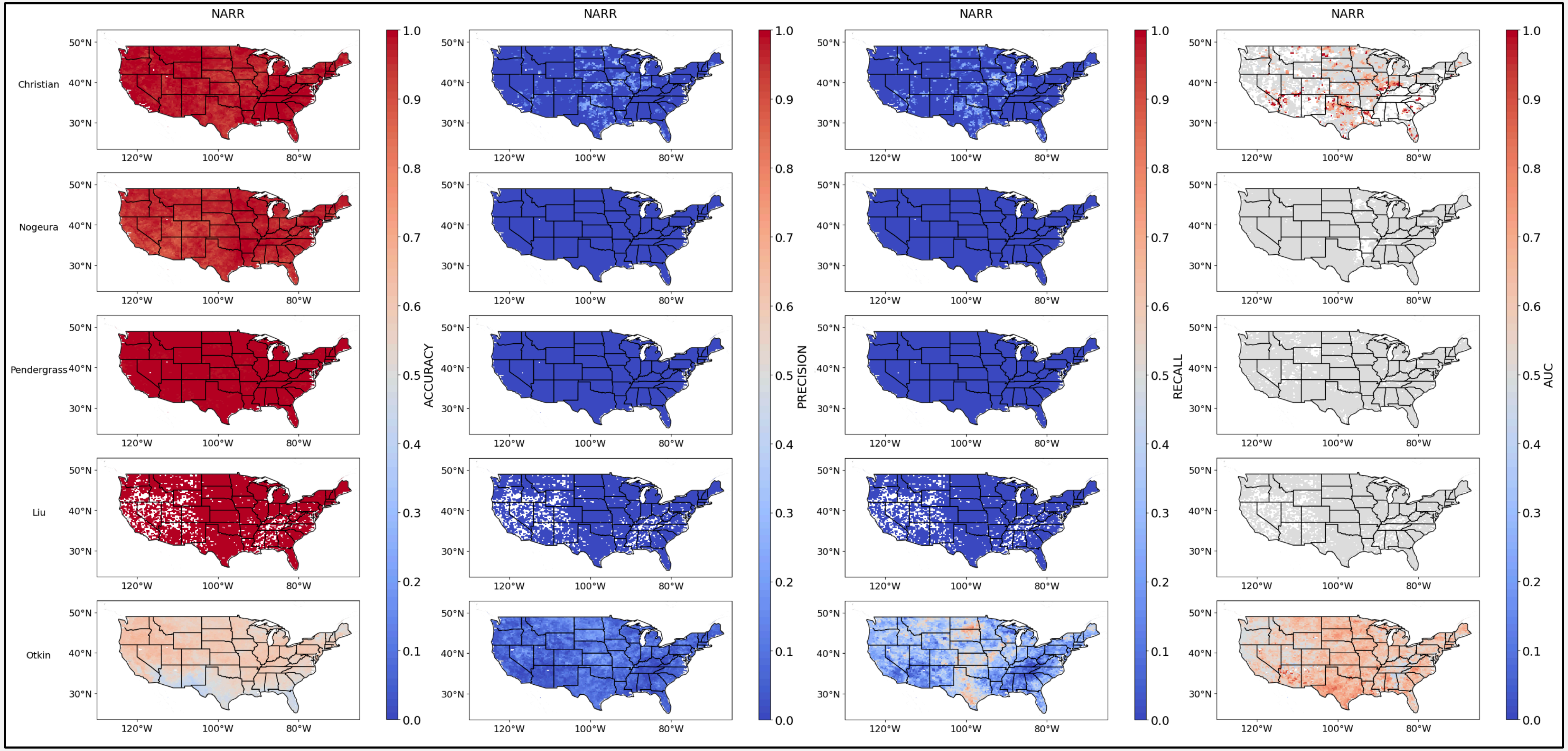
- Recurrent neural networks focused on GRU layers to learn temporal patterns
  - The recurrent layers were evaluated along the temporal axis
  - Each grid point was treated as an example
  - Data set up as $\left(N_{grid\ points}, N_{pentads}, N_{variables}\right)$





PRELIMINARY RESULTS: RNN

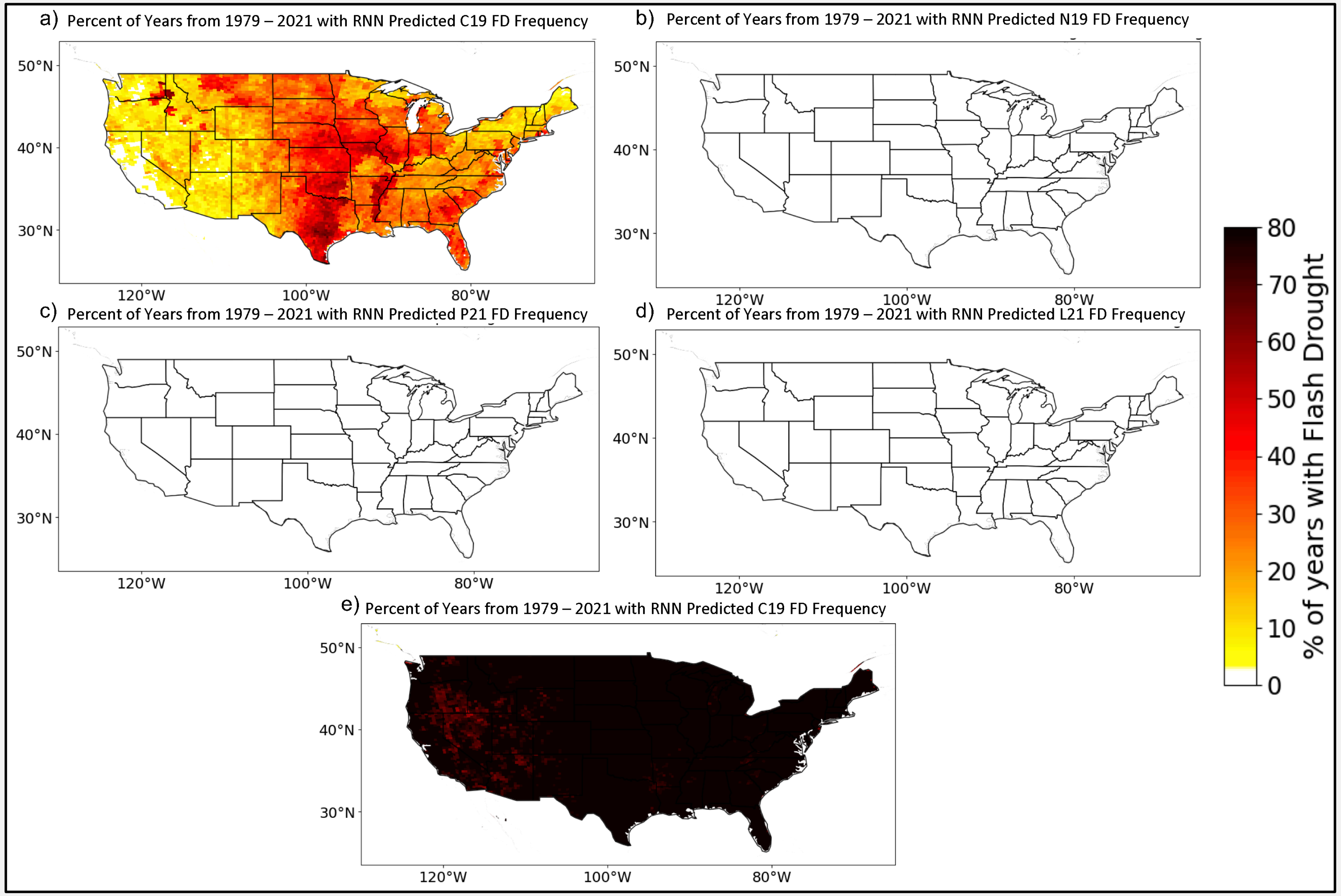
- RNNs are more sensitive to FD identification method



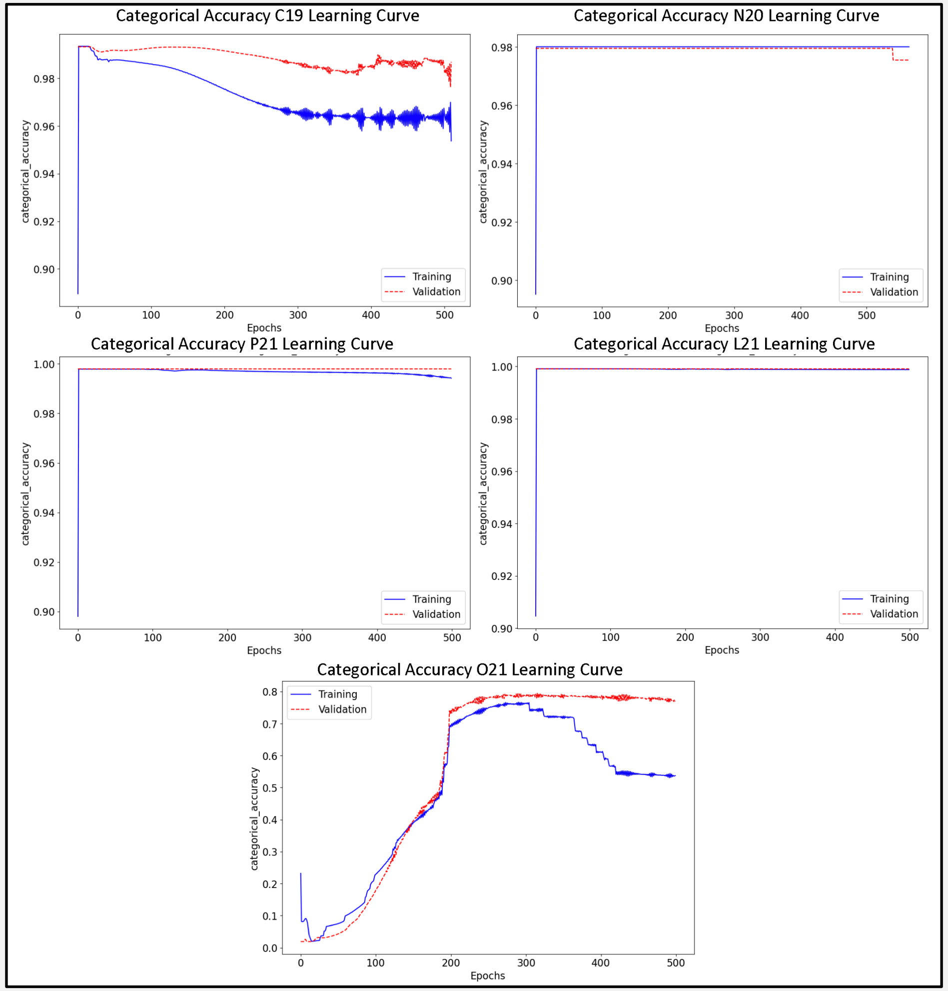
Average RNN test set results

PRELIMINARY RESULTS: RNN

- RNNs are more sensitive to FD identification method;
- RNNs were able to reproduce the climatology of the method it was tuned to well



RNN training set predictions



Background	Data	ANN Experiments	RNN Experiments	Concluding Remarks

CONCLUSIONS AND FUTURE WORK

- Results highlight difference between FD identification methods
  - Each model needs to be individually tuned to the identification method
- Each ML model explored showed potential in being able to identify FD directly
  - They were able to identify climatology hotspots, and some of the overall temporal trends
  - Both models were able to able to outperform random forests for the FD identification method they were tuned for

- Future work:
  - Additional deep learning models (variational autoencoder is in the works and transformer networks are planned)
  - Re-explore results with models tuned to individual FD identification methods
  - Addition of XAI to interpret results and understand predictions
  - Train models in other reanalysis datasets (including global datasets)

REFERENCES

Christian, J. I., J. B. Basara, J. A. Otkin, E. D. Hunt, R. A. Wakefield, P. X. Flanagan, and X. Xiao, 2019: A methodology for flash drought identification: Application of flash drought frequency across the United States. *Journal of Hydrometeorology*, **20** (**5**), 833–846, doi: 10.1175/jhm-d-18-0198.1, URL http://dx.doi.org/10.1175/JHM-D-18-0198.1.

Noguera, I., F. Dom´ınguez-Castro, and S. M. Vicente-Serrano, 2020: Characteristics and trends of flash droughts in Spain, 1961–2018. *Annals of the New York Academy of Sciences*, **1472** (**1**), 155–172, doi:10.1111/nyas.14365, URL http://dx.doi.org/10.1111/nyas.14365.

Pendergrass, A. G., and Coauthors, 2020: Flash droughts present a new challenge for subseasonal-to-seasonal prediction. *Nature Climate Change*, **10** (**3**), 191–199, doi:10.1038/s41558-020-0709-0, URL http://dx.doi.org/10.1038/s41558-020-0709-0.

Liu, Y., Y. Zhu, L. Zhang, L. Ren, F. Yuan, X. Yang, and S. Jiang, 2020: Flash droughts characterization over China: From a perspective of the rapid intensification rate. *Science of The Total Environment*, **704**, 135 373, doi:10.1016/j.scitotenv.2019.135373, URL https://doi.org/10.016%2Fj.scitotenv.2019.135373.

Otkin, J. A., and Coauthors, 2021: Development of a flash drought intensity index. *Atmosphere*, **12** (**6**), 741, doi:10.3390/atmos12060741, URL http://dx.doi.org/10.3390/atmos12060741.

Otkin, J. A., M. Svoboda, E. D. Hunt, T. W. Ford, M. C. Anderson, C. Hain, and J. B. Basara, 2018: Flash droughts: A review and assessment of the challenges imposed by rapid-onset droughts in the United States. *Bulletin of the American Meteorological Society*, **99** (**5**), 911–919, doi:10.1175/bams-d-17-0149.1, URL http://dx.doi.org/10.1175/BAMS-D-17-0149.1.

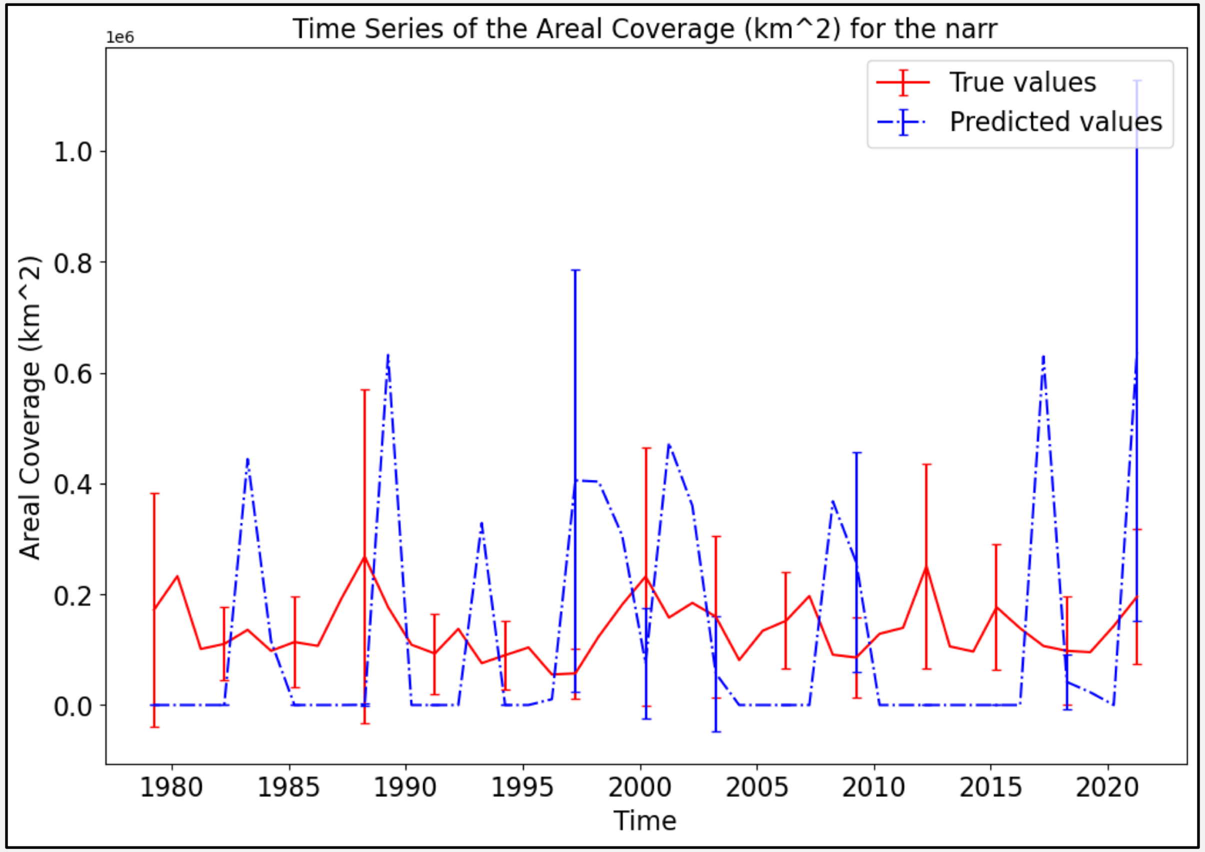
Lisonbee, J., M. Woloszyn, and M. Skumanich, 2021: Making sense of flash drought: definitions, indicators, and where we go from here. *Journal of Applied and Service Climatology*, **2021** (**1**), 1–19, doi:10.46275/joasc.2021.02.001, URL http://dx.doi.org/10.46275/JOASC.2021.02.001.

Dikshit, A., B. Pradhan, and M. Santosh, 2022: Artificial neural networks in drought prediction in the 21st century–a scientometric analysis. *Applied Soft Computing*, **114**, 108 080, doi:10.1016/j.asoc.2021.108080, URL https://doi.org/10.1016%2Fj.asoc.2021.108080.

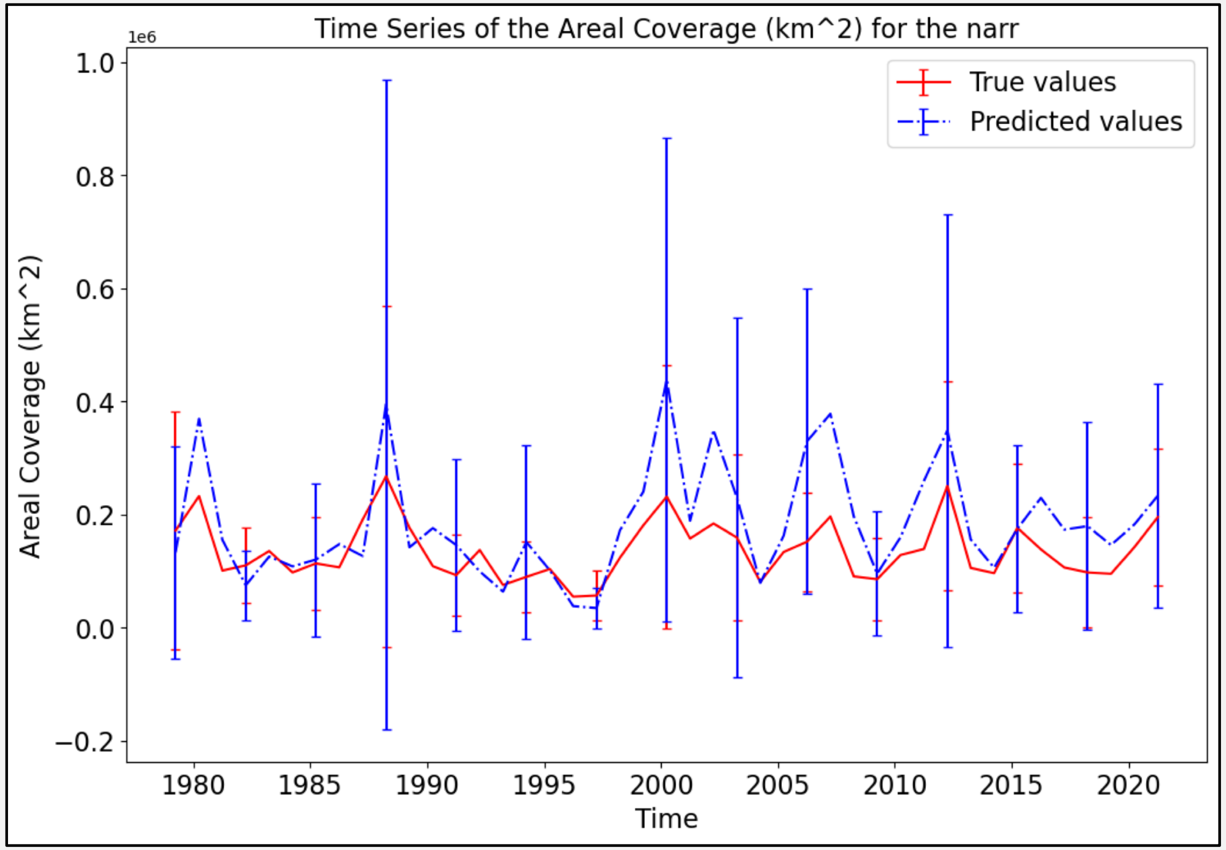
EXTRA FIGURES

- Time Series Figures

ANN Areal Coverage Time Series prediction for the C19 FD



RNN Areal Coverage Time Series prediction for the C19 FD



EXTRA FIGURES

- Case Study Figures