Poster ID: A21C-2258

Exploring the Climate Variability of Lightning in Alaska with a SOM-RF Model

AGU San Francisco California, 11-15 December 2023

Session A21C: Bridging the Gap from Climate to Extreme Weather

Joshua Hostler[1], Uma S. Bhatt[1], Thomas Ballinger[2], Peter Bieniek[2], Cecilia Borries-Strigle[1], Mitchell Burgard[4], Jonathan Chriest[3], Elizabeth Fischer[2], Rick Lader[2], Zachary Parish[6], Heidi Strader[5], Eric Stevens[5], Richard Thoman[1], and Christine Waigl[2]

[1]Geophysical Institute University Alaska Fairbanks (UAF), [2]International Arctic Research Center UAF, [3]National Weather Service, [4]USDA Forest Service, [5]Alaska Interagency Coordination Center, [6]Unaffiliated

Contact: jphostler@alaska.edu

Photos of poster are welcome.

Main Results:

- Self-Organizing Maps (SOMs) are a useful tool to resolve summertime atmospheric patterns conducive to lightning activity in Alaska.
- Random Forest Model shows skill in classifying lightning-days from SOM patterns.

Motivation

- Log of burned area correlated to number of high tercile lightning days in duff season. Lightning is a key driver of wildland fires in Alaska (Kasischke et al 2010).
- Fire management would benefit from skillful seasonal outlooks on lightning likelihood.
- Lightning has a clear seasonal cycle.

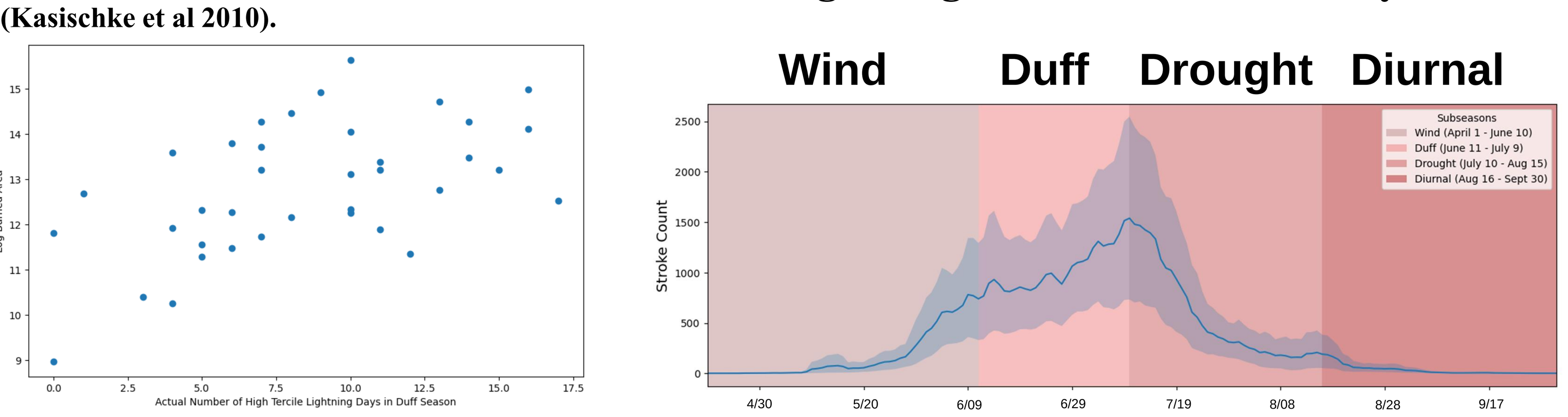

Figure 1. Log of seasonal burned area in vs number of high tercile lightning days in eastern interior Alaska (left) and seasonality of daily mean (1986-2022) lightning stroke count (blue line) with 95 percent confidence interval (shading) (right)

Methods & Data

Data

- ERA5 Reanalysis provided 500hPa geopotential height (500z), 850hPa T CAPE, SLP, 2m temperature, convective precipitation & total column water vapor from 1959-2022
- Historical lightning data from the Alaska Lightning Detection Network (ALDN) for 1986-2022.

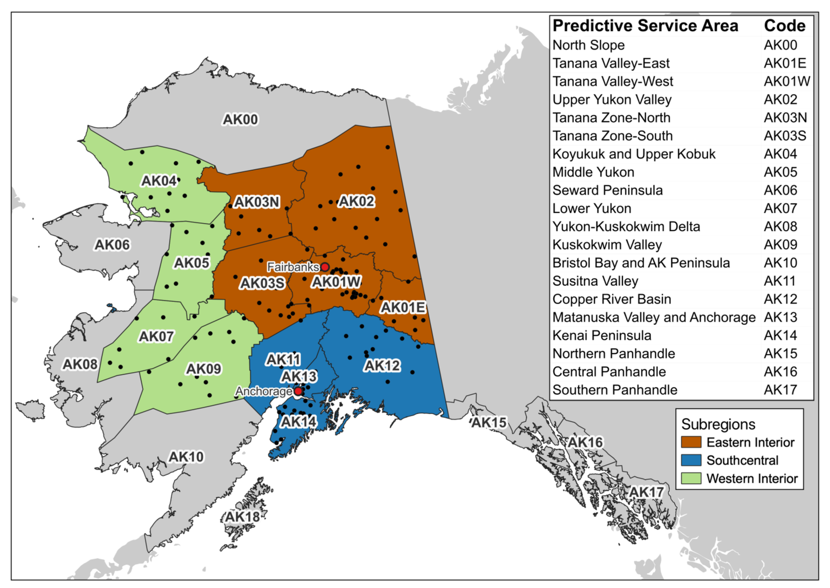

Figure 2. Predictive Service Areas (PSA) for Alaska are based on weather and topography.

Data Preparation

- Compute daily anomalies & reduce dimensions using Self-Organizing Maps (SOMs).
- QA/QC & sum daily ALDN stroke counts over Predictive Service Area (PSA) then label by tercile and lightning day vs non-lightning day.

Methods

- SOMs cluster data and organize clusters by similarity. Daily mean fields are matched to the nearest cluster (Hewitson and Crane, 2002).
- Random forest classifier trained on SOM projections of each anomaly field to predict tercile class of daily lightning counts for Eastern Interior PSA group.
- Model training, performed in scikit-learn (Pedregosa et al, 2011), uses an 80-20 train-test split and 5-fold cross-validation for hyperparameter tuning.
- The model is evaluated with the F_1 score, which measures the model's skill to distinguish classes.

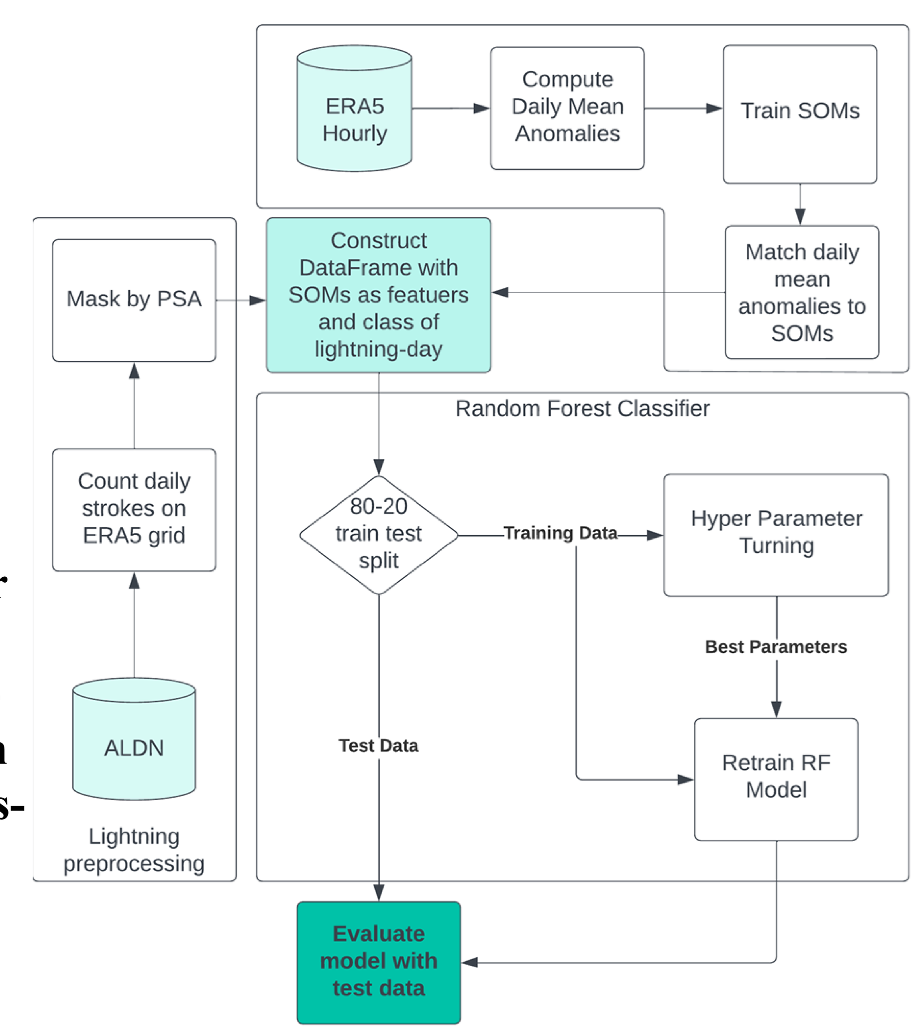

Figure 3. Flow chart for SOM-RF model.

SOM Arrangement of 500hPa Geopotential

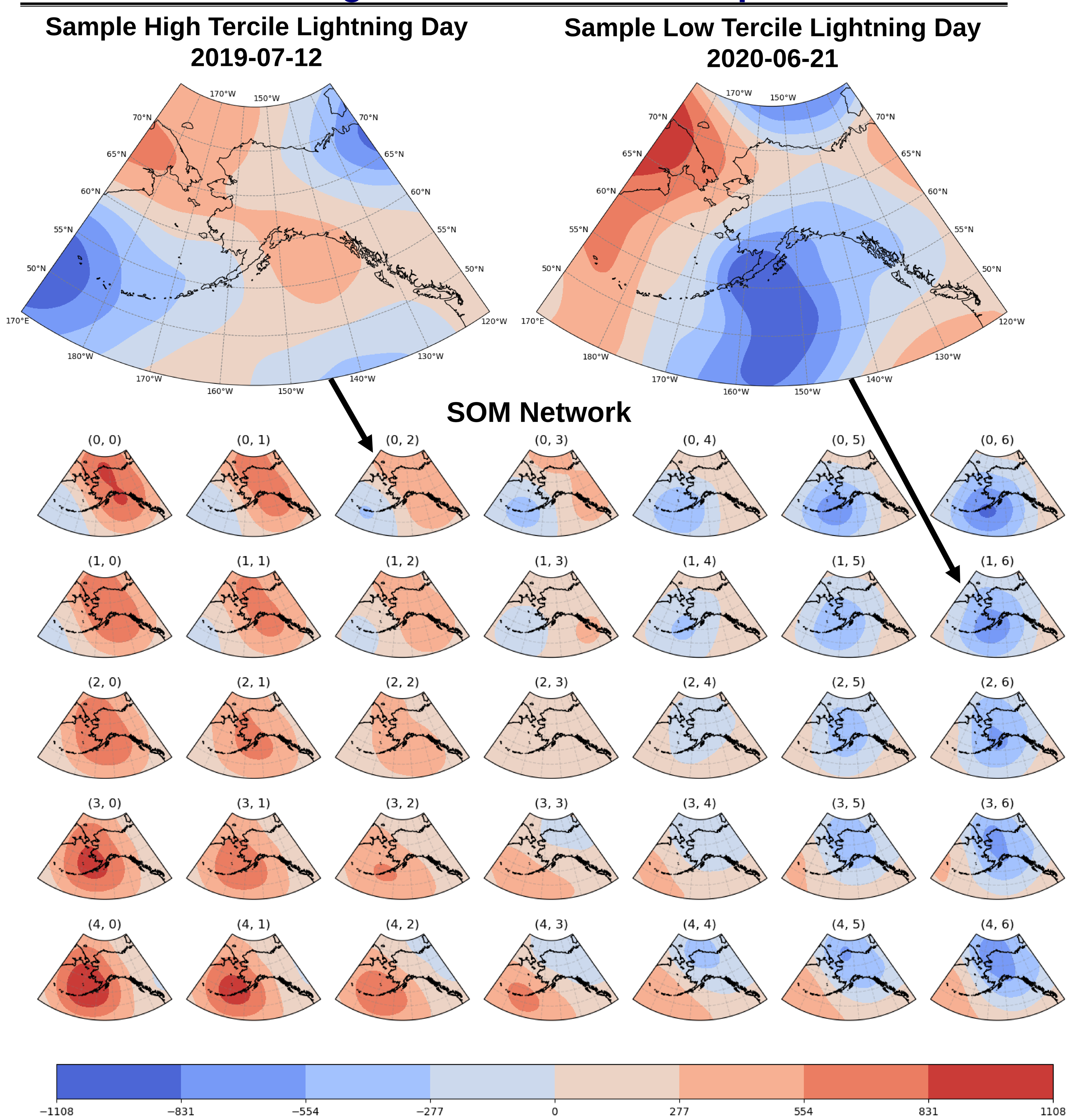

Figure 4. Examples of daily mean 500hPa geopotential anomalies for high and low tercile lightning days with arrows representing the projection into the space defined by the SOM network

Considerations for Extreme Valued Distributions

- A p-value of 0.748 under the modified null Kolmogorov-Smirnov distribution does not reject the hypothesis that the observations arise from an inverse Weibull distribution. However, a power analysis is needed to rule out other EV distributions (future work).

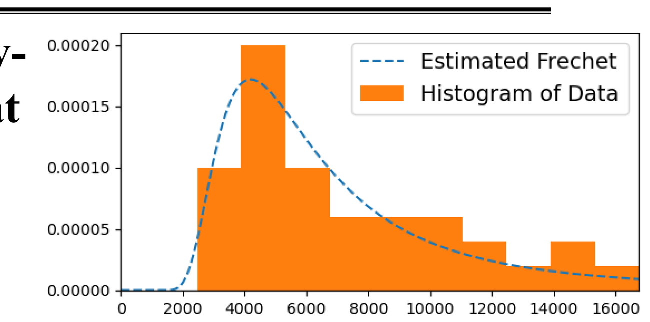

Figure 5. Histogram and inverse Weibull distribution fit of annual daily maximum stroke counts.

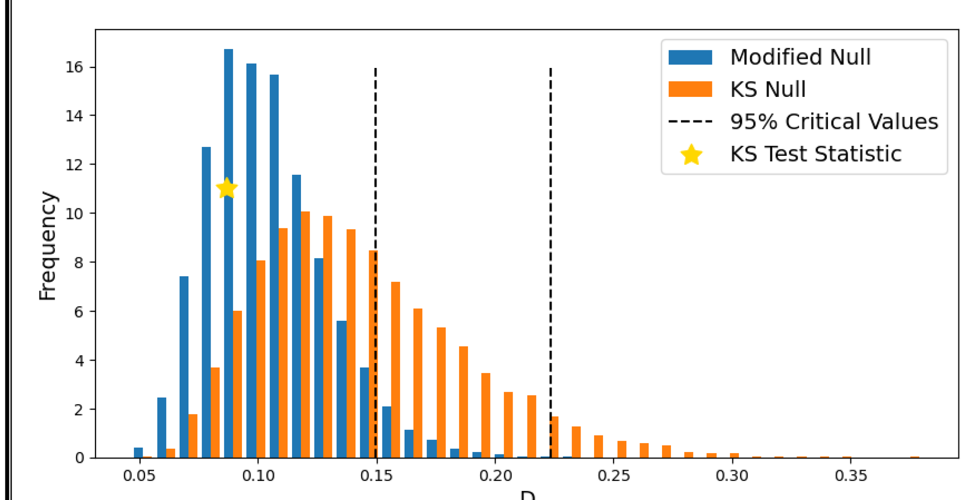

Figure 6. Modified null distribution for the Kolmogorov-Smirnov goodness of fit test – computed via Monte Carlo simulation.

- Quantifying the role of large scale dynamics, which arise from internal modes of variability in the climate system, may inform the selection of an appropriate EV distribution for lightning in Alaska.

SOM-RF model shows skill in predicting daily lightning magnitudes from large scale patterns/predictors

- Two RF models are trained for each region of Alaska – the first classifies non-lightning and lightning days and the second predicts the intensity of lightning day.
- F_1-scores range from 0 to 1 where scores above 0.5 (0.33) indicate outperforming uniform sampling of class frequencies for binary (ternary) classification. AUROC ranges from 0.5 to 1 for binary and multiclass problems.
- Tercile classification loses some skill over the binary classifier, indicating that the large scale predictors have limits resolving magnitude of events.
- Reconstruction of duff season high tercile lightning events demonstrates SOM-RF captures interseasonal variance of lightning.

Region	Precision	Recall	F1-Score	AUROC
East Interior	0.768	0. 901	0. 829	0.822
West Interior	0.710	0.693	0.701	0.696
South Central	0.728	0.744	0.736	0.798
Combined	0.819	0.940	0.876	0.830

Region	Precision	Recall	F1-Score	AUROC
East Interior	0.609	0.610	0.608	0.786
West Interior	0.538	0.535	0.527	0.706
South Central	0.440	0.430	0.436	0.627
Combined	0.627	0.627	0.616	0.790

Figure 7. Precision, recall, F1-score, and AUROC for binary (Top), and tercile (Bottom) models.

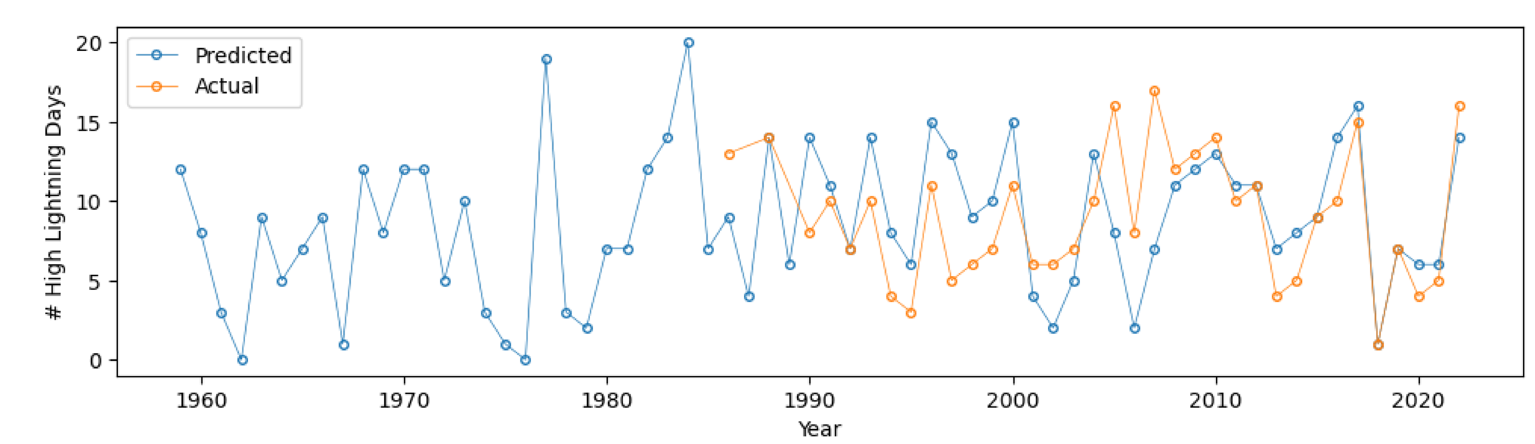

Figure 8. Count of observed (blue) and constructed (yellow) high tercile lightning days in duff subseason by year

Summary

- We are working with fire managers on predictions of daily lightning activity at the subseasonal to seasonal scale.
- The SOM bridges the gap from climate variability to daily lightning activity.
- The RF shows skill in predicting lightning activity, in terms of occurrence and intensity, given the state of the atmosphere represented as a position in the space of modes of variability (SOM space).
- The SOM-RF has limited resolution in space and quantitative predictions due to the nature of the predictors, and the structure of the model.
- Non-local features result in low skill at small spatial scales as local variances dominate.
- Future work includes:
  - evaluation of the SOM-RF over dynamical forecasts,
  - an analysis of related teleconnection patterns,
  - and a power analysis for the modified KS test for EV distributions.

References

- Hewitson, B. C., & Crane, R. G. (2002). Self-organizing maps: applications to synoptic climatology. *Climate Research*, *22*(1), 13–26. https://doi.org/10.3354/CR022013
- Kasischke, E. S., et al., 2010: Alaska's changing fire regime—Implications for the vulnerability of its boreal forests. Can. J. For. Res., 40, 1313–1324, https://doi.org/10.1139/X10-098.
- Pedregosa et al., 2011: Scikit-learn: Machine Learning in Python, *JMLR* 12, 2825-2830.

Acknowledgements

This work was made possible through financial support from the State of Alaska and NSF EPSCoR Grant OIA-1757348.