Statistical Properties of Solar Flare Dependency
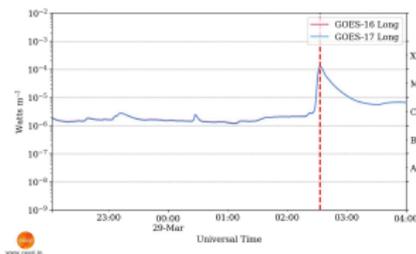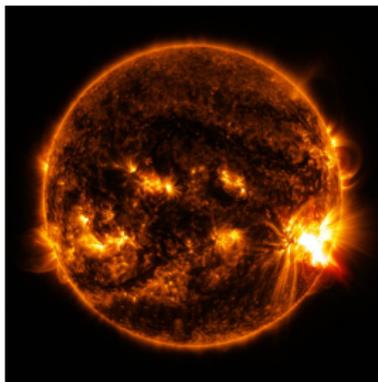
Noah Kochanski   Yang Chen

University of Michigan

June 9, 2023

Contact: noahkoch@umich.edu

What are solar flares?

Solar flares are characterized by a large eruptions of electromagnetic radiation from the Sun. Solar flares can cause:

- Radio blackouts
- Increased radiation exposure to pilots
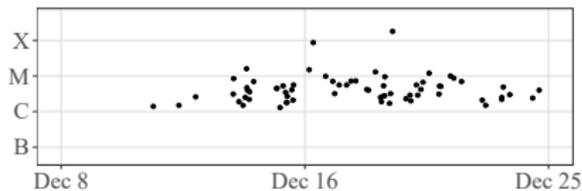- Albeit rare, computer malfunctions

Motivation

Solar flares are difficult phenomena to predict due to their rarity. As a result, machine learning methods are becoming more prevalent within the field of solar flare prediction.
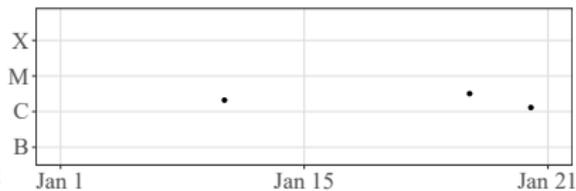
- Regression
- Neural Network
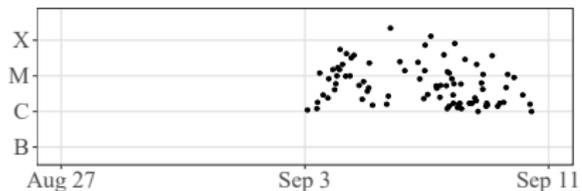- Decision Tree Learning
- Support Vector Machines

Motivation

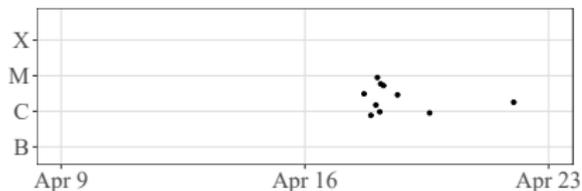- Machine learning models are not oracles.

- Multiple types of models are producing similar accuracy rates.

- While the physical traits of the flaring process are understood, the distributions governing solar flares are less understood.
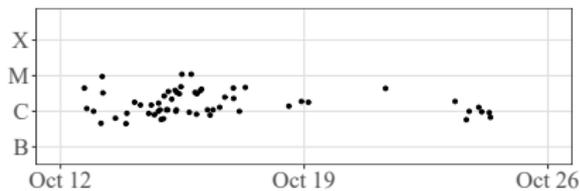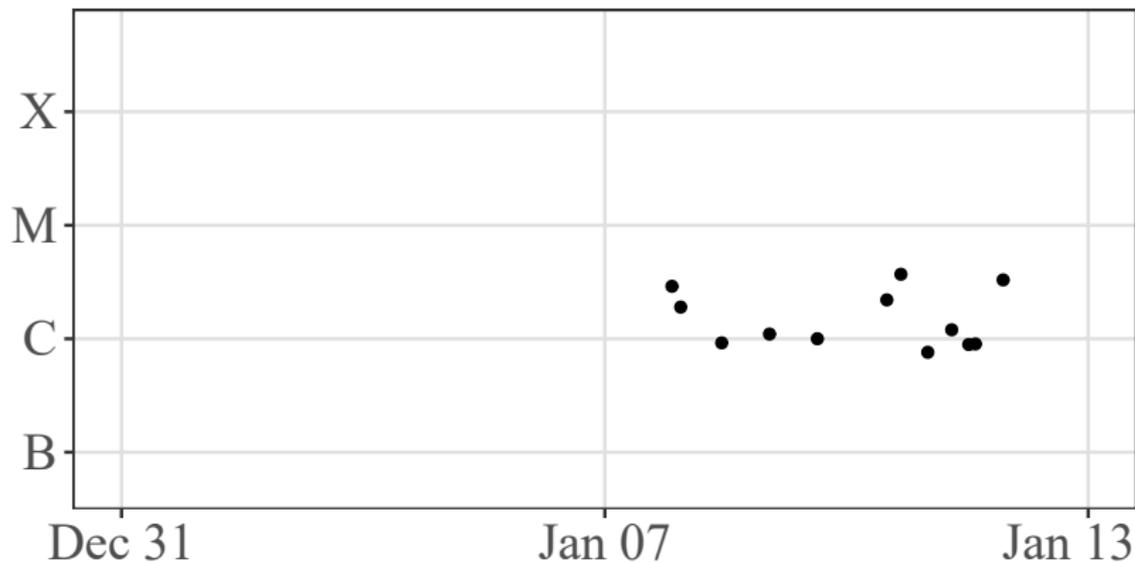

SDO/HMI Tracked AR (HARP)
2013/01/13
00:48

Solar Flares

Main Research Question

What is the dependency structure that solar flares exhibit when coming from a single active region?

AR 1275

H_0 : Flares that occur within 6 hours of each other are **independent**.

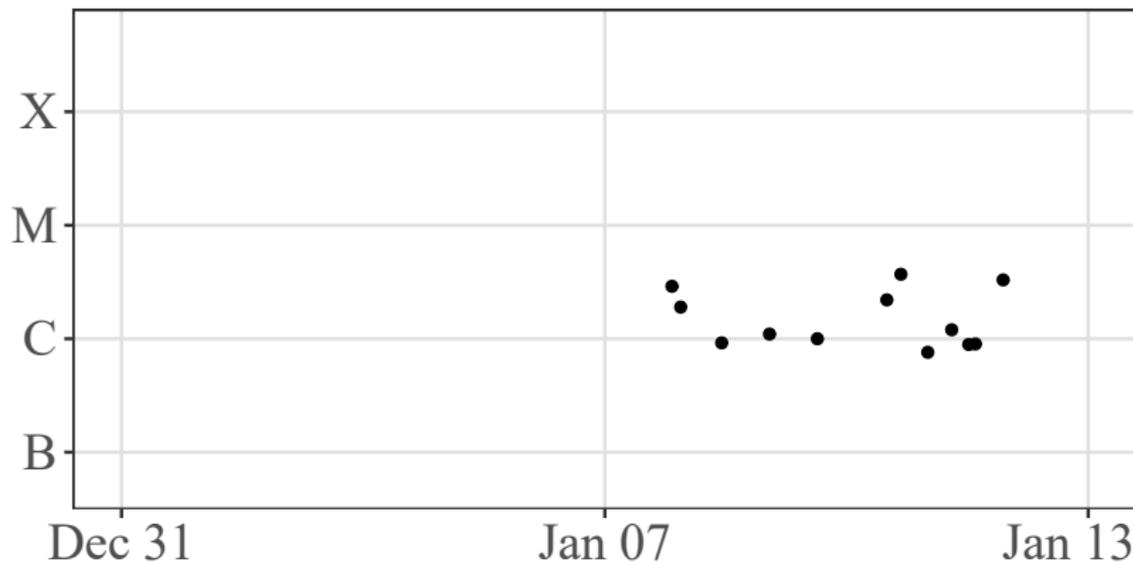H_1 : Flares that occur within 6 hours of each other are **dependent**.

$$\Updownarrow$$

H_0 : $\mathbb{1}_{\text{Flare in next 6 hours}} \perp\!\!\!\perp \mathbb{1}_{\text{Flare in previous 6 hours}}$

H_1 : $\mathbb{1}_{\text{Flare in next 6 hours}} \not\!\perp\!\!\!\perp \mathbb{1}_{\text{Flare in previous 6 hours}}$
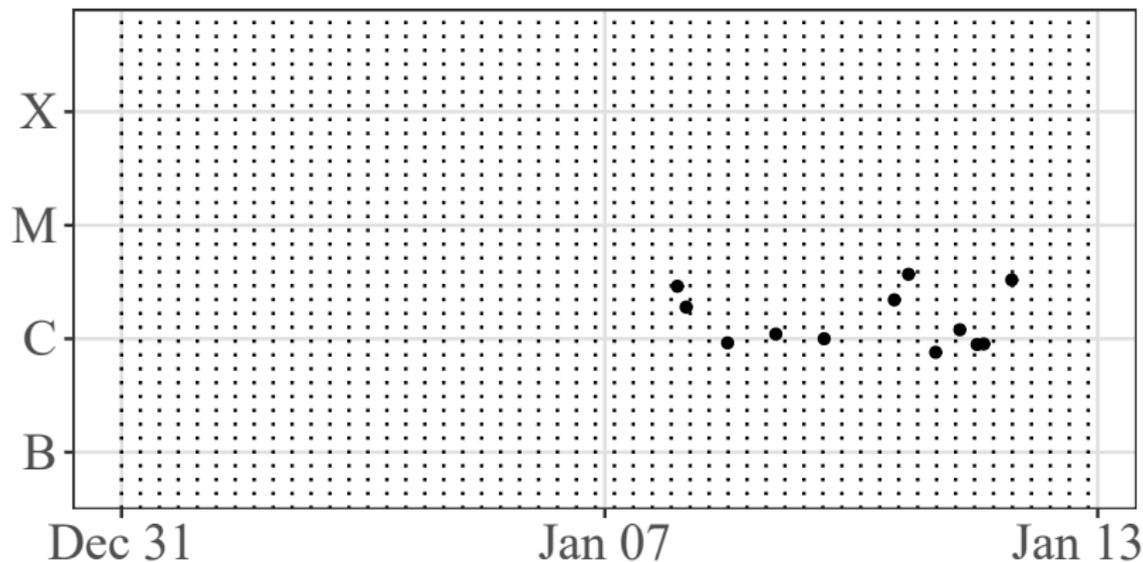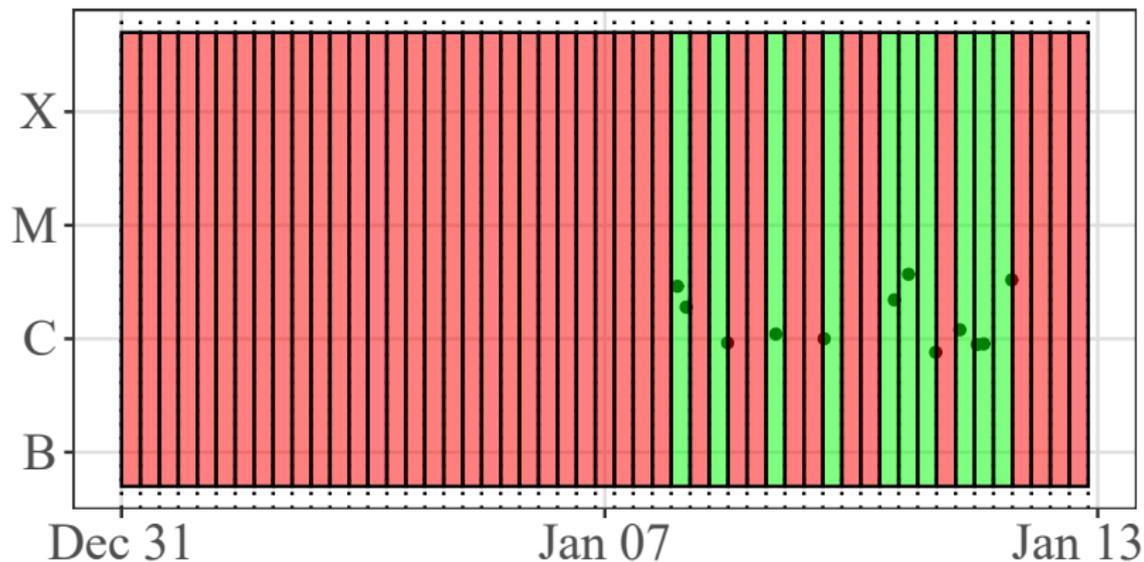
Active Region 1275

AR 1275

Active Region 1275

AR 1275

Active Region 1275

AR 1275

Active Region 1275

Collecting all samples from this active region, we get the following table summarizing the data,

	No flare in next 6 hours	Flare in next 6 hours
No flare in previous 6 hours	36	6
Flare in previous 6 hours	6	4

Using the typical χ^2 Test of Independence, yields a p-value of 0.1592.

U-Statistic Permutation Test

What about other tests? If we un-summarize our data, it would look something like this,

Flare in previous 6 hours	Flare in next 6 hours
0	0
1	1
$\vdots$	$\vdots$
1	1
1	0

Under the null, these two variables are independent.

U-Statistic Permutation Test

- Permute the first column of the data.
- Compute $b_n = \sum_{i,j}(\text{observed}_{i,j} - \text{expected}_{i,j})^2$
- Repeat this process, N times, and compare our observed table to the distributions of $b_1, b_2, \cdots, b_N$

$\Rightarrow$ Doing this for active region 1275 yields a p-value of 0.0128.

Generalizing to δ time interval

For AR 1275, instead of considering a single time interval, we can vary the time considered.

Results from Hypothesis Testing
AR 1275 across different time intervals (δ)
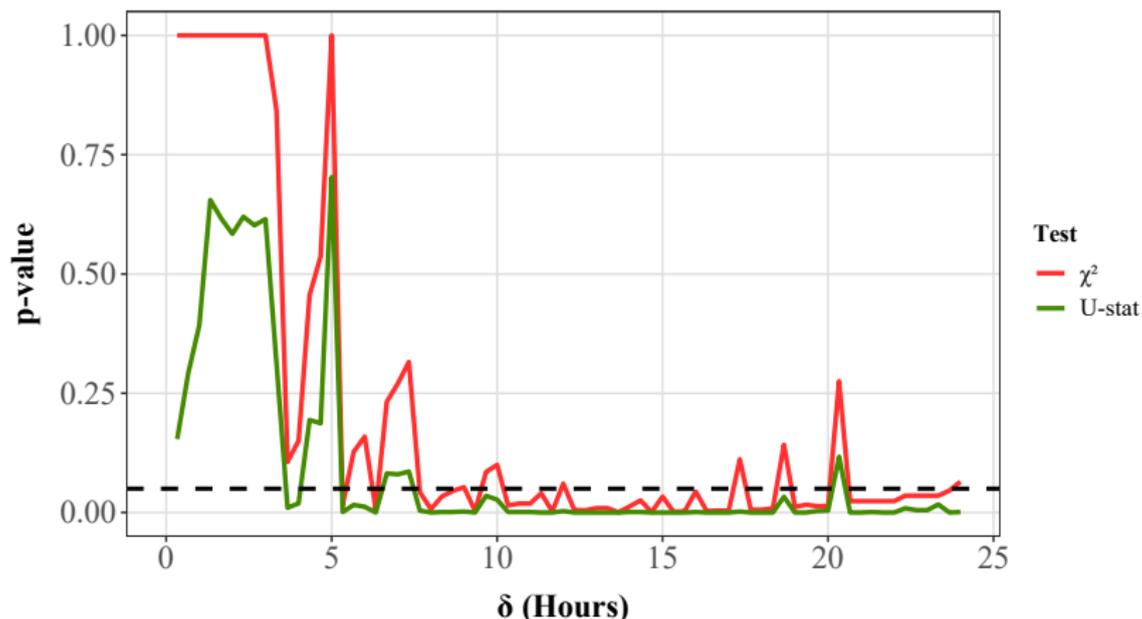
Flare Catalogue

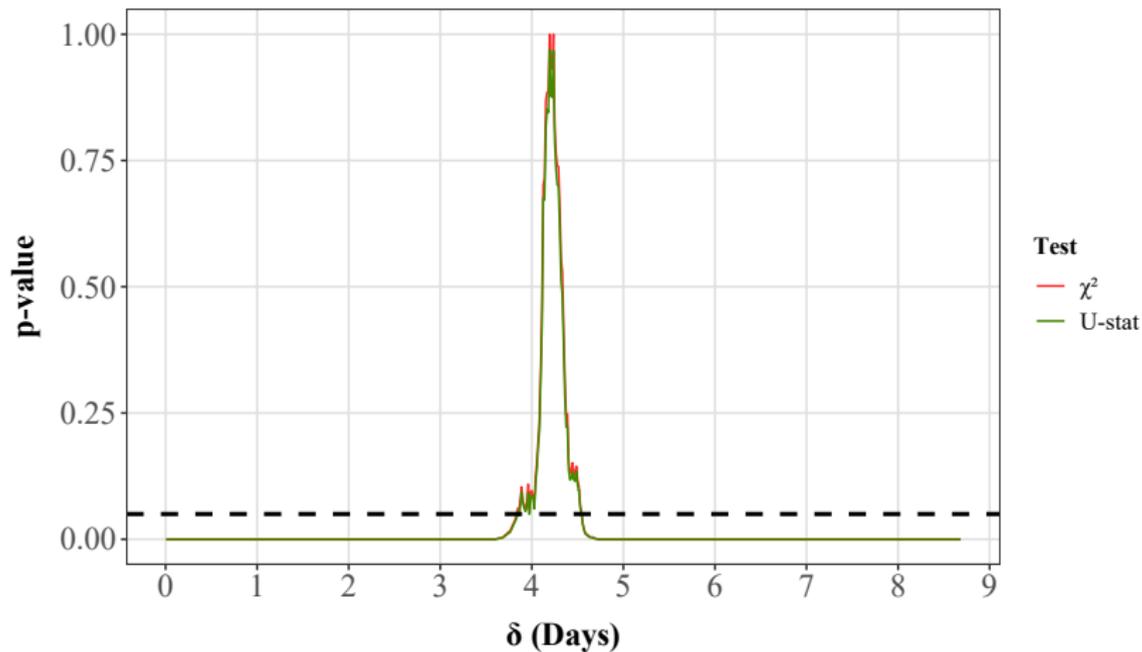The previous process done for a catalogue consisting of:

- Roughly an entire solar cycle
- 12,496 solar flares
- 1,044 active regions

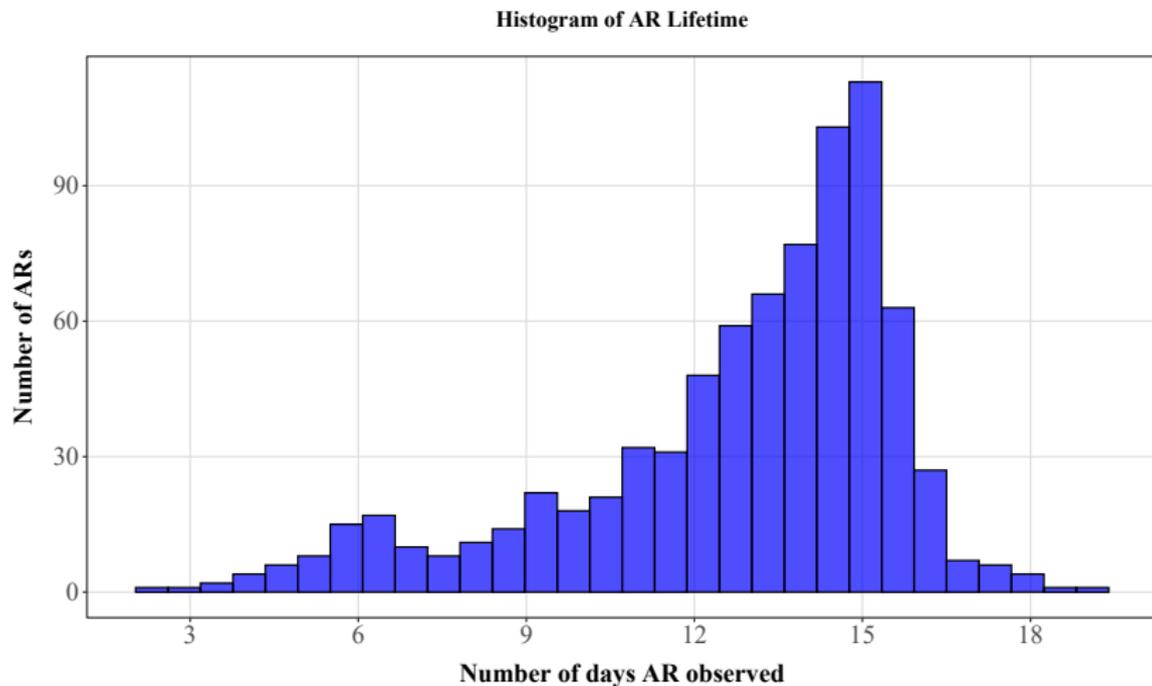Furthermore, we break the testing into homogeneous and heterogeneous results.

Homogeneous Results

Homogeneous Results across different time intervals (δ)

Homogeneous Results

Histogram of AR Lifetime

The previous results treat all solar flares alike, however we can further test the dependency structure that occurs between x-ray classes.

|	Weak Flares (A/B/C)	Strong Flares (M/X)
Weak Flares (A/B/C)	**Hypotheses for B/C flares produce significant results at 0.05 level until ~3 days at which flares act as independent events**	**All hypotheses for M/X vs. C flares produce significant results at 0.05 level**
Strong Flares (M/X)		**All hypotheses for M vs X flares produce significant results at 0.05 level**

Heterogeneous Results

Histogram of Number of Flares within each AR
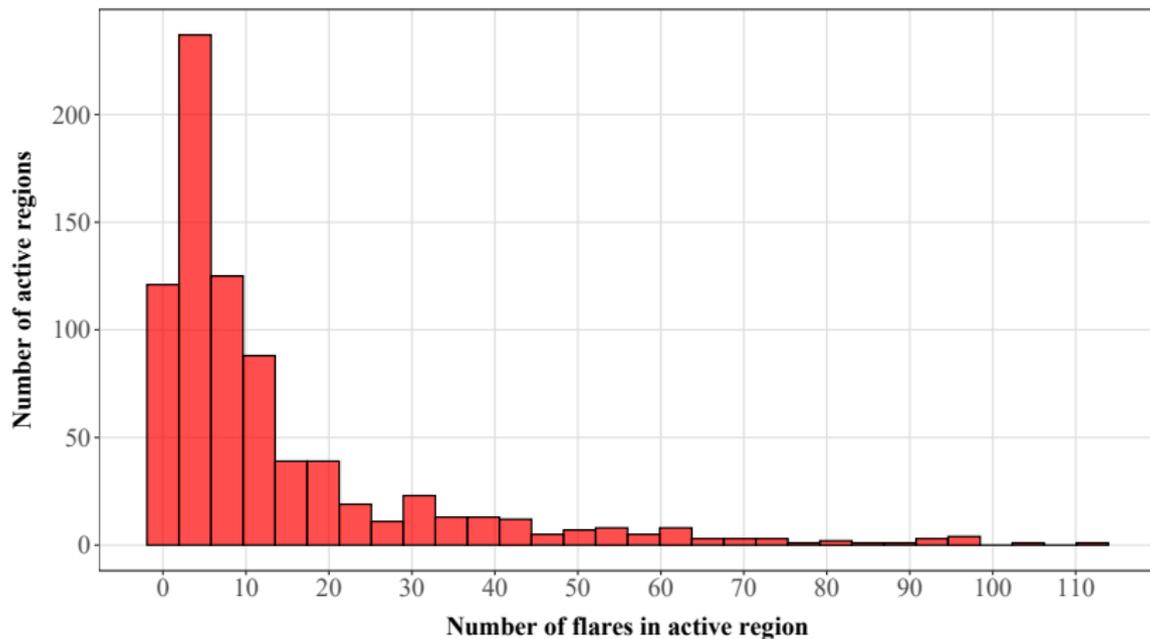
Heterogeneous Results

Proportion of Significant Active Regions by Number of Flares at 0.05 level
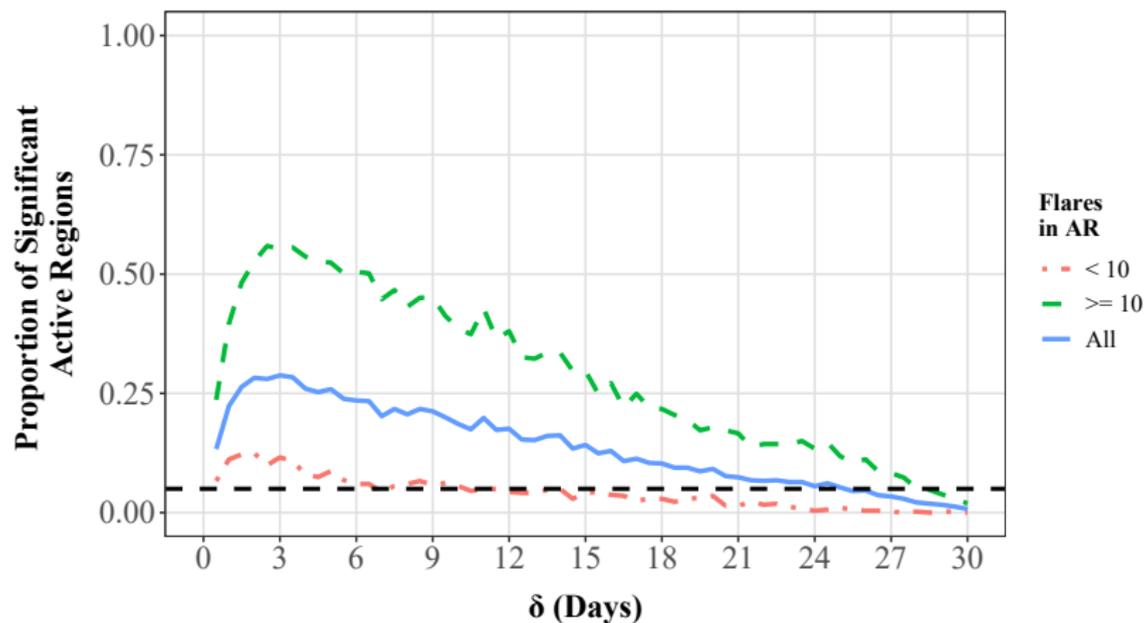
Summary

In summary;

- Solar flares are highly dependent events
- In both 'quiet' and 'active' ARs we see that ~ 3 days is when we see the largest dependency
- Flare dependency is not monotonic in the time interval considered
- For weak flares, after 3 days, is when weak flares appear to act as independent events

Thank you for having me!