A Statistical Approach to Detecting Polarity Inversion Lines

Hu Sun

Department of Statistics
University of Michigan, Ann Arbor

September 24, 2019

Presentation Outline

- Background and motivation
- Algorithm for finding Polarity Inversion Line(PIL)
- Potential Application
- Limitation of the algorithm

Presentation Outline

- Background and motivation
- Algorithm for finding Polarity Inversion Line(PIL)
- Potential Application
- Limitation of the algorithm

Presentation Outline

- Background and motivation
- Algorithm for finding Polarity Inversion Line(PIL)
- Potential Application
- Limitation of the algorithm

Presentation Outline

- Background and motivation
- Algorithm for finding Polarity Inversion Line(PIL)
- Potential Application
- Limitation of the algorithm

Background: predict solar flares

Data source: Space-Weather HMI-Active Region Patch (SHARP) data files

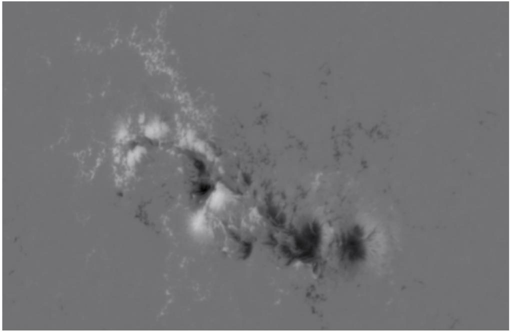


Figure: B-z Component, HARP377, 2011-02-17 14:00:00TAI

Input into LSTM Model

Physical quantities for each HMI-Active Region Patch(HARP):

- Total unsigned flux
- Total unsigned current helicity
- Mean shear angle
- ...

Input Time Series

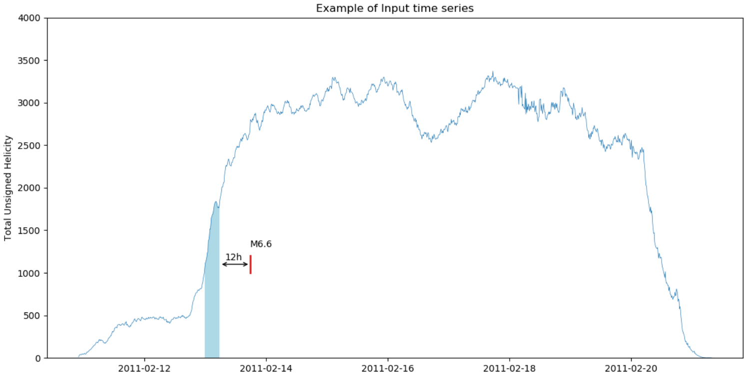


Figure: HARP377 Total Unsigned Helicity Series

Machine Learning Task

Task: With all physical quantities' time series as input, classify whether the flare 12 hours later is an M-X class flare or a B-class flare.

Machine Learning Task

Task: With all physical quantities' time series as input, classify whether the flare 12 hours later is an M-X class flare or a B-class flare.

Class/Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total
X	0	9	7	12	16	2	0	4	0	50
M	0	106	124	97	194	128	15	37	0	701
C	1	1002	1115	1197	1626	1275	294	229	11	6750
B	19	665	475	469	184	446	757	620	207	3842

Figure: Flare Data Summary

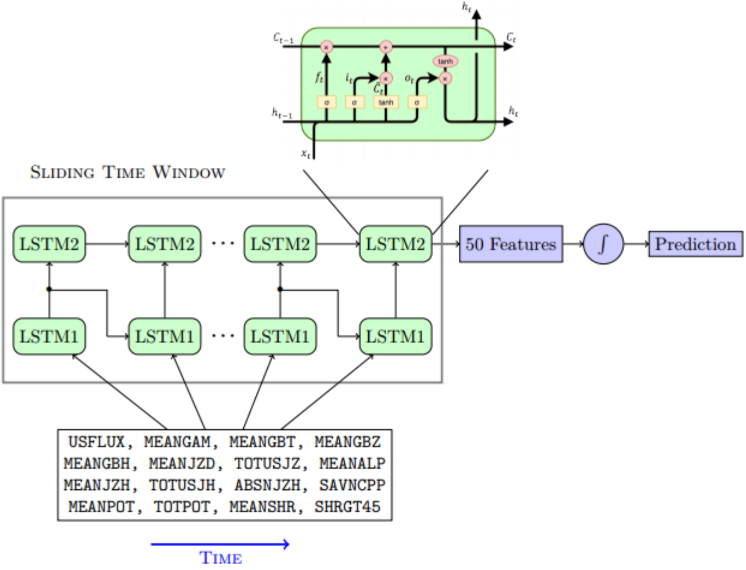
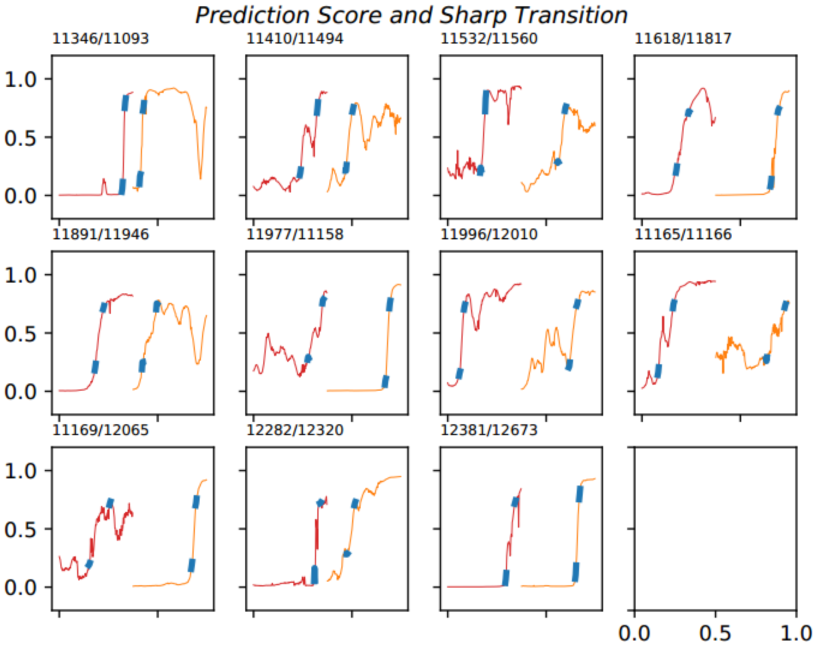


Figure: LSTM Model


Prediction Score and Sharp Transition
11346/11093
11410/11494
11532/11560
11618/11817
11891/11946
11977/11158
11996/12010
11165/11166
11169/12065
12282/12320
12381/12673
1.0
0.5
0.0
0.0
0.5
1.0

LSTM Model Variable Importance

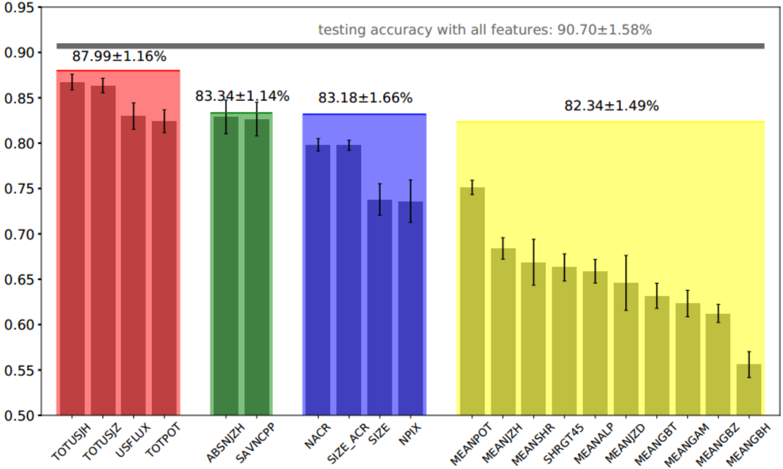


Figure: Feature Importance in trained LSTM Model (Chen et al. 2019)

Motivation

Finding: Many variables are able to single-handedly classfify MX flares against B flares.

But what is the underlying mechanism? Physical quantities in SHARP data are computed using **the whole HARP region magnetic field**...

Motivation

Finding: Many variables are able to single-handedly classfify MX flares against B flares.

But what is the underlying mechanism? Physical quantities in SHARP data are computed using **the whole HARP region magnetic field**...

Find Polarity Inversion Line

Previous methods (Falconer et al. 2002, Bokenkemp 2007) are modeling the spatial distribution of magnetic vector field.

Here we provide an alternative statistical method that does not model the distribution of polars but work on **pixels of the magnetogram** directly.

Find Polarity Inversion Line

Previous methods (Falconer et al. 2002, Bokenkemp 2007) are modeling the spatial distribution of magnetic vector field.

Here we provide an alternative statistical method that does not model the distribution of polars but work on **pixels of the magnetogram** directly.

Step 1: Select PIL candidate pixels

Data Type: Magnetogram with 200,000+ pixels.

Step 1: Select PIL candidate pixels

Data Type: Magnetogram with 200,000+ pixels.

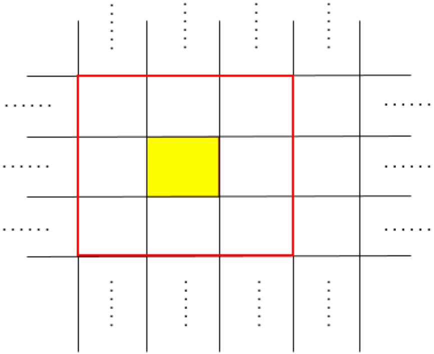


Figure: A pixel's neighborhood, 3×3 window

Step 1: Select PIL candidate pixels

Data Type: Magnetogram with 200,000+ pixels.

In every pixel's neighborhood, we check if there is both a strong positive polar and a strong negative polar. (default threshold ± 100Gauss)

Step 1: Select PIL candidate pixels

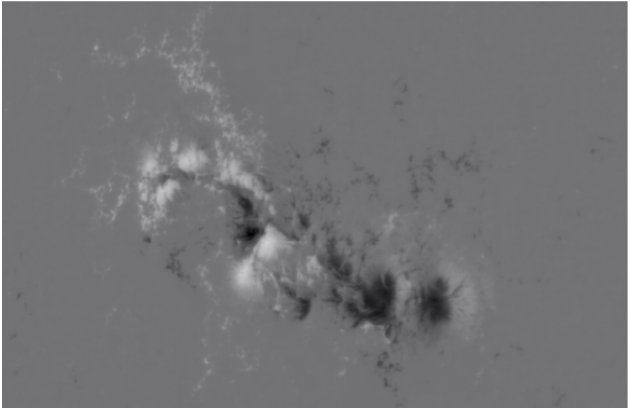


Figure: B-z Component, HARP377, 2011-02-17 14:00:00TAI

Step 1: Select PIL candidate pixels

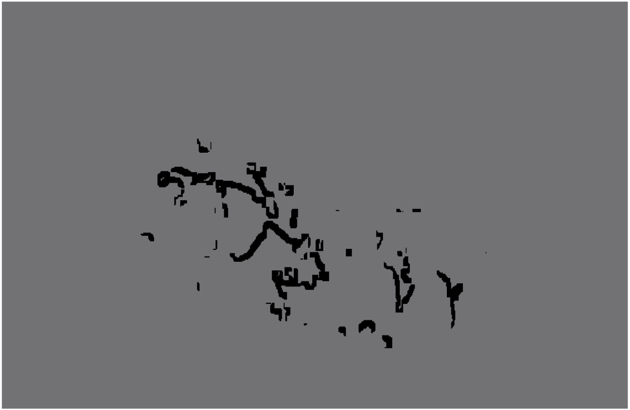


Figure: Pixels remained after neighborhood check (5×5 window)

Step 1: Select PIL candidate pixels

Property of the pixels selected:

- Having shape of multiple PIL segments
- Having miscellaneous clusters of pixels not of our interest
- Pixels on the same PIL are not linked well
- The PILs seem to be too thick.

Step 1: Select PIL candidate pixels

Property of the pixels selected:

- Having shape of multiple PIL segments
- Having miscellaneous clusters of pixels not of our interest
- Pixels on the same PIL are not linked well
- The PILs seem to be too thick.

Step 1: Select PIL candidate pixels

Property of the pixels selected:

- Having shape of multiple PIL segments
- Having miscellaneous clusters of pixels not of our interest
- Pixels on the same PIL are not linked well
- The PILs seem to be too thick.

Step 1: Select PIL candidate pixels

Property of the pixels selected:

- Having shape of multiple PIL segments
- Having miscellaneous clusters of pixels not of our interest
- Pixels on the same PIL are not linked well
- The PILs seem to be too thick.

Step 2: Non-maxima Suppression

Use Prewitt filter to calculate local B-z gradient for each of the pixel:

$$\mathbf{G_x} = \begin{bmatrix} +1 & 0 & -1 \\ +1 & 0 & -1 \\ +1 & 0 & -1 \end{bmatrix} * \mathbf{A} \quad \text{and} \quad \mathbf{G_y} = \begin{bmatrix} +1 & +1 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} * \mathbf{A}$$

Step 2: Non-maxima Suppression

With both G_x and G_y, we could calculate the direction of the gradient by $\arctan(\frac{G_y}{G_x})$.

And with the direction of the gradient, we could define the adjacent pixels along the gradient direction of any pixel with non-zero gradient.

Step 2: Non-maxima Suppression

With both G_x and G_y, we could calculate the direction of the gradient by $\arctan(\frac{G_y}{G_x})$.

And with the direction of the gradient, we could define the adjacent pixels along the gradient direction of any pixel with non-zero gradient.

Step 2: Non-maxima Suppression

With both G_x and G_y, we could calculate the direction of the gradient by $\arctan(\frac{G_y}{G_x})$.

And with the direction of the gradient, we could define the adjacent pixels along the gradient direction of any pixel with non-zero gradient.

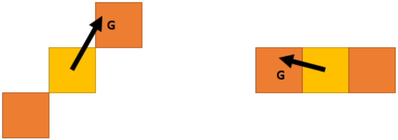


Figure: Adjacent pixels along gradient direction

Step 2: Non-maxima Suppression

With both G_x and G_y, we could calculate the direction of the gradient by $\arctan(\frac{G_y}{G_x})$.

And with the direction of the gradient, we could define the adjacent pixels along the gradient direction of any pixel with non-zero gradient.

Now we only retain the pixels whose gradient norm is local maxima along it gradient direction.

Step 2: Non-maxima Suppression

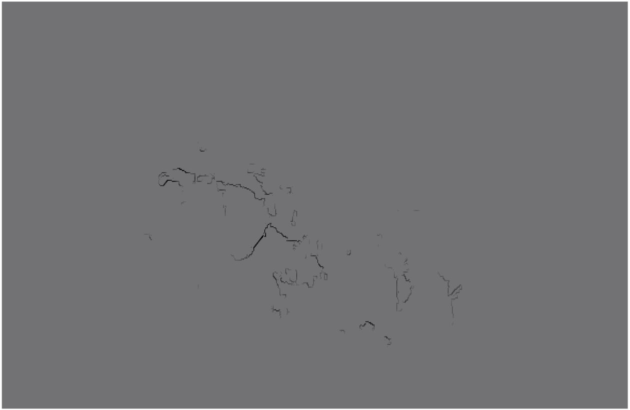


Figure: Thin Edges after Non-maxima Suppression

Step 3: PIL extension

In order to make the PIL drawn to be more connected, we apply a simple continuity extension to pixels on the PIL.

Step 3: PIL extension

In order to make the PIL drawn to be more connected, we apply a simple continuity extension to pixels on the PIL.

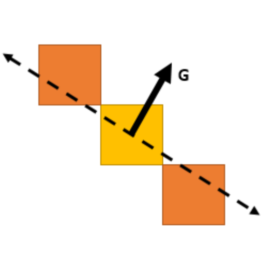


Figure: PIL extension direction

Step 3: PIL extension

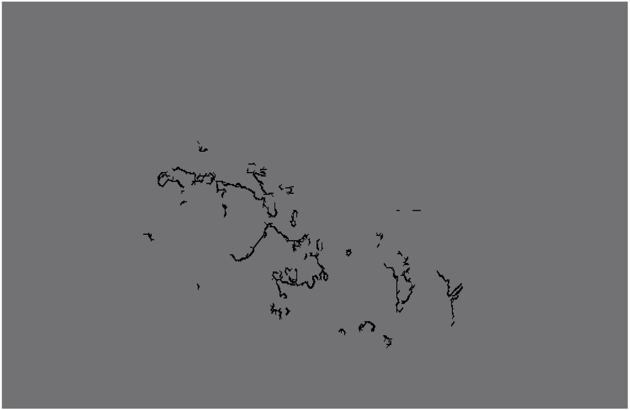


Figure: More connected PILs

Step 3: PIL extension

Such an extension would add more connectivity to each PIL, but will add many pixels with small B-z gradients into the PIL.

Optionally, one could only extend on points with large gradients, and this extension can be conducted **recursively**.

Step 4: Density-Based Clustering

- Generally, the longer the PIL is, the more complex the local magenetic field structure there is.
- We need to retain the longest PILs and erase the miscellaneous PILs.

Step 4: Density-Based Clustering

- Generally, the longer the PIL is, the more complex the local magenetic field structure there is.
- We need to retain the longest PILs and erase the miscellaneous PILs.

Step 4: Density-Based Clustering

How to classify PIL pixels into several clusters?

- Connected-component analysis
- Density-based clustering algorithm

Step 4: Density-Based Clustering

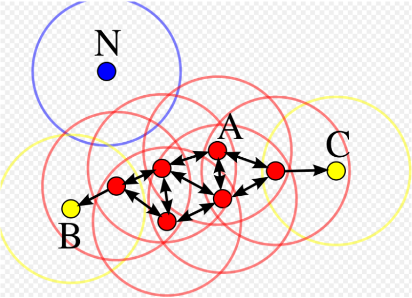
Density-based spatial clustering of applications with noise **(DBSCAN)** :

- Locate some "core" points as the seed for a cluster
- All points within a certain range of core points are in the same cluster as the core points

Step 4: Density-Based Clustering

Density-based spatial clustering of applications with noise **(DBSCAN)** :

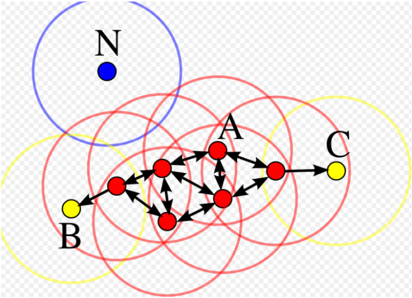
- Locate some "core" points as the seed for a cluster
- All points within a certain range of core points are in the same cluster as the core points



Step 4: Density-Based Clustering

Density-based spatial clustering of applications with noise **(DBSCAN)** :

- Locate some "core" points as the seed for a cluster
- All points within a certain range of core points are in the same cluster as the core points



Step 4: Density-Based Clustering

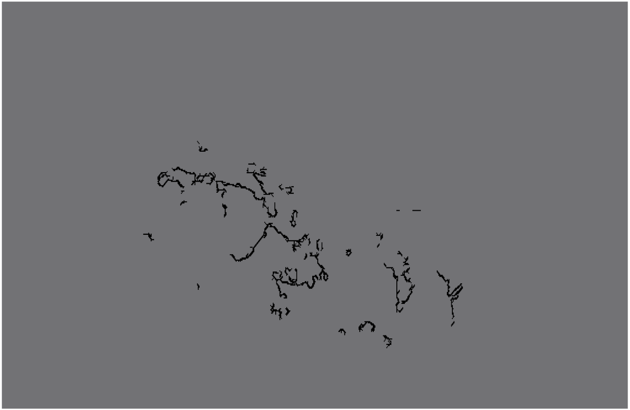


Figure: More connected PILs

Step 4: Density-Based Clustering

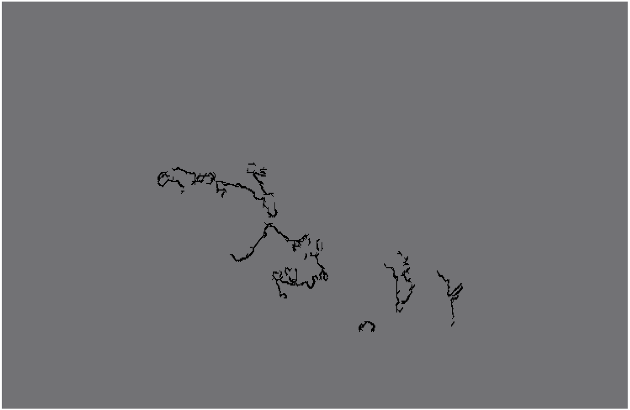


Figure: PILs left after deleting small clusters

Potential Application

Potential Application

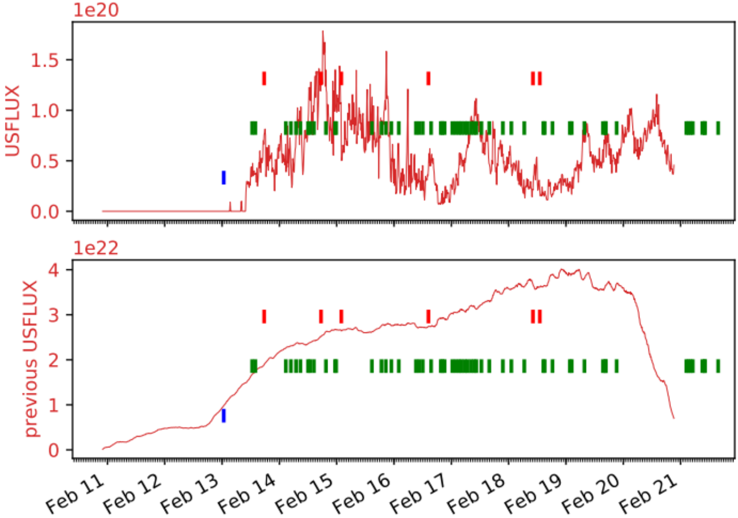


Figure: Recalculate USFLUX for HARP377

Advantage and Limitation

Advantage:

- Computationally Efficient
- Multiple PILs can be compared and tracked
- Accessible for non-physicist

Limitation:

Advantage and Limitation

Advantage:

- Computationally Efficient
- Multiple PILs can be compared and tracked
- Accessible for non-physicist

Limitation:

- All tuning parameters are determined subjectively

Advantage and Limitation

Advantage:

- Computationally Efficient
- Multiple PILs can be compared and tracked
- Accessible for non-physicist

Limitation:

- All tuning parameters are determined subjectively
- Still not very well connected

Advantage and Limitation

Advantage:

- Computationally Efficient
- Multiple PILs can be compared and tracked
- Accessible for non-physicist

Limitation:

- All tuning parameters are determined subjectively
- Still not very well connected

Advantage and Limitation

Advantage:

- Computationally Efficient
- Multiple PILs can be compared and tracked
- Accessible for non-physicist

Limitation:

- All tuning parameters are determined subjectively
- Still not very well connected

Advantage and Limitation

Advantage:

- Computationally Efficient
- Multiple PILs can be compared and tracked
- Accessible for non-physicist

Limitation:

- All tuning parameters are determined subjectively
- Still not very well connected