Using Long-Short Term Memory Models to Predict Solar Active Regions Emergence

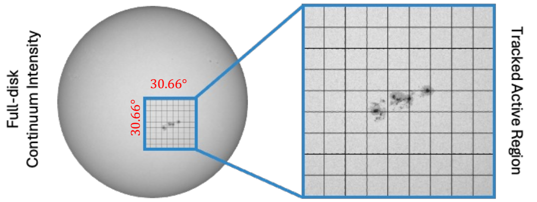
Spiridon Kasapis[1], Irina Kitiashvili[1], Alexander G. Kosovichev[1,2], John T. Stefan[2], Bhairavi Apte[2]

[1]NASA Ames Research Center, [2]New Jersey Institute of Technology

Introduction

We train Long Short-Term Memory (LSTM) models that predict the formation of active regions (ARs), the main source of eruptive solar activity. Using the Doppler shift velocity, the continuum intensity and the magnetic field full-disk maps from SDO/HMI we have created time-series datasets of acoustic power and magnetic flux which are used to train Long Short-Term Memory (LSTM) models on predicting decreases in continuum intensity 12 hours in advance. Testing of the models' performance was done on data from 5 ARs, unseen from the model during training. The model predicted the emergence of AR11726, AR13165 and AR13179, 10, 29 and 5 hours in advance, and variations of this model achieved average RMSE values of 0.11 for both active and quiet parts of the solar disc, showing the ability of the model to capture acoustic power anomalies and predict continuum intensity variations. This work sets the foundations for the very first ML-aided prediction of solar ARs.

Dataset



45 ARs that emerged on the observable solar disc were tracked and analyzed using data from the SDO/HMI. Each selected AR appeared on the solar surface within 30 degrees longitude from the central meridian between March 1st, 2010 and June 1st, 2023, persisted for more than 4 days, and reached a total area of 200 millionths of the solar hemisphere.

For all three HMI data products (the Doppler velocity V_D, the line-of-sight magnetic flux Φ_m and the continuum intensity I_c, we tracked the corresponding 30.66°×30.66° (heliographic coordinates) area around the ARs of the dataset, taking into account the local rotation speed. Although the Φ_m and I_c maps are used-as-is, the Dopplergram (V_D maps) tracked regions are used to generate acoustic power (P_a) maps for four frequency ranges: 2–3, 3–4, 4–5, and 5–6 mHz for the entire 10-day period of the ARs' passage, before and after their emergence.

The resulting P_a maps, along with the corresponding Φ_m and I_c tracked regions are split into a grid of 9 by 9 tiles. By splitting the tracked regions we focus the evolution tracing on a local 30.66°×30.66° area which is then reduced to structured timelines conducive to machine learning analysis by calculating the mean of each tile that belongs to each frame. This ensemble of mean P_a, Φ_m and I_c time series is further processed by removing the solar sphere geometric effect and by normalizing them in order to create ML-ready datasets for each tracked AR.

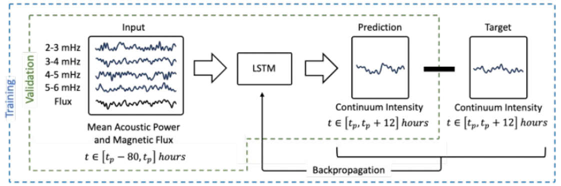
Training

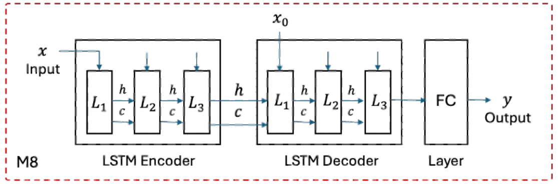
	Mean ϕ	Mean Traveled λ	Mean A_{max}	NOAA Record Span	Number
Training ARs	9.3 ± 16.58	87.6 ± 18.16	318.0 ± 169.91	6.6 ± 1.52	40
Testing ARs	-5.9 ± 17.53	79.8 ± 15.35	424.0 ± 332.08	5.8 ± 1.10	5
Entire Dataset	7.6 ± 17.17	86.7 ± 17.89	329.8 ± 191.70	6.5 ± 1.49	45
Units	degrees	degrees	$\frac{1}{10^6}$ of Hemisphere	days	

From the 45 total tracked ARs, only 40 are used for training and 5 are reserved for testing (almost a 90-10 train-test data ratio). The table above outlines the mean values of the ARs latitude, traveled longitude, maximum area and the time of the AR's life on the observable disc for the training and testing datasets but also for the entire dataset. The ARs in our dataset have a mean life on the observable disc of 6.5 days which correspond to 86.7 degrees in heliographic coordinates and a mean maximum area of 329.8 millionths of a hemisphere (moh). The mean latitude of the dataset is 7.6 with a standard deviation of 17.17 degrees, showing that there is a diversity of ARs in terms of latitude but within the range of latitudes that ARs usually emerge.

For the ML application, an LSTM Encoder-Decoder architecture is chosen, which is dependent on two hyperparameters: the number of layers *LN* and the number of units within each layer *U*. More specifically, the models consist of an LSTM Encoder which encodes the input and forwards the cell state *c* and the hidden state *h* to an LSTM Decoder, to decode the information and with the help of a fully connected (FC) layer produce a prediction *y*. For each model, both the encoder and the decoder contain *LN* LSTM layers (nn.LSTM class in Pytorch) and each layer contains *U* LSTM units. In this work, more than 200 models were trained, exploring parameter spaces for different combinations of hyperparameters. Results only for the best performing - based on two testing criteria- model (M8) are presented.

M8 is trained using 110 hours worth of input data and predicts intensity 12 hours ahead, it has three layers with 64 units each and it is trained on 1000 epochs using a decreasing learning rate which starts from 0.01.





Testing

A list of the five ARs included in the testing dataset, their properties and classification is presented below.

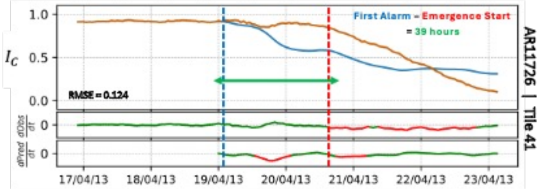
AR#	First Record	Last Record	ϕ	λ_{Carr}	λ_s	λ_e	A_s	A_e	A_{max}	A_{max} Date	McIntosh	Hale
11698	2013.03.15	2013.03.19	-19.5	117.0	29.0	86.0	20	200	200	2013.03.18	Eao	β
11726	2013.04.20	2013.04.27	13.0	327.0	-7.0	93.0	20	600	1000	2013.04.26	Fkc	$\beta\gamma\delta$
13165	2022.12.12	2022.12.18	-20.0	277.5	10.0	88.0	20	150	340	2022.12.16	Ekc	$\beta\delta$
13179	2022.12.30	2023.01.05	13.5	43.0	11.0	92.0	30	80	380	2023.01.03	Dko	β
13183	2023.01.06	2023.01.12	-16.5	309.0	8.0	91.0	30	50	200	2023.01.08	Dso	$\beta\delta$

Two metrics have been used to evaluate the LSTM models' capabilities of predicting the emergence of ARs: the RMSE metric and the emergence criterion. Although the RMSE is a common criterion used in ML learning time-series prediction works, the emergence definition presented in the equation below is a criterion devised specifically for this research.

$$\frac{dI_c}{dt} < -0.01 \quad \text{for} \quad t_{sus} > 3 \text{ hours}$$

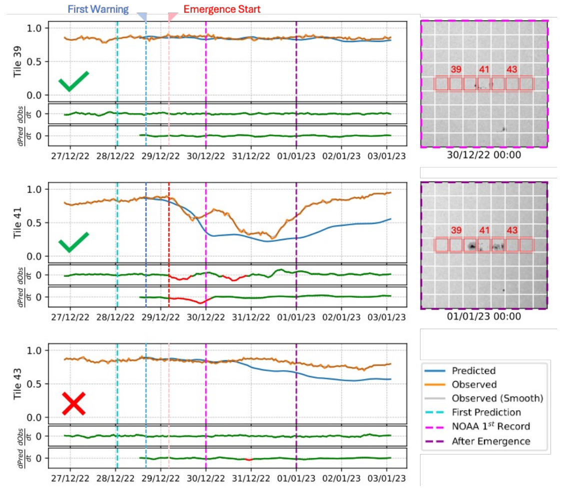
$^{dI_c}/_{dt}$ is the derivative of the I_c (predicted or observed) and t_{sus} is the sustained time, the time for which the derivative is less than -0.01. Therefore in the figures below, marked in red are the parts of the I_c derivative time-series that are less than the threshold of -0.01 for at least 3 hours. This definition of emergence reflects quite accurately what a human observer would define as emergence: a sustained and significant decrease of I_c.

Below are the prediction results of Model 8. The orange line depicts the observed (true) mean normalized I_c of the respective tile and the blue line depicts the predicted values calculated 12 hours ahead, given an input of 96 hours before every prediction. The green and red values on the two smaller subplots for each AR depict the first derivative of the observed (orange) and predicted (blue) I_c values.



Results

The emergence prediction testing includes two different types of tests: one where we consider the emergence of individual tiles (independent tile prediction) and one for which we consider the emergence of the entire AR (Full AR Prediction), where the different tiles are considered as a group rather than individual parts of the solar surface. For the full AR prediction, emergence time is defined as the very first point in any tile of the AR where the Equation 1 criterion is fulfilled.



A comprehensive account of predictions for all five testing ARs and their individual tiles (Independent Tile columns) is presented in the table below. We show results for 4 active (A) and 3 quiet (Q) tiles for each AR, except AR11698 and AR13179, which have only three active tiles. True positives (TP) and true negatives (TN) are marked in green, indicating correct and timely predictions or accurate non-alarms for quiet tiles by Model 8. False positives (FP) and false negatives (FN) are marked in red, indicating missed predictions or false alarms by Model 8.

	Independent Tile Prediction							Full AR Prediction	
Tile Type	Quiet	Quiet	Quiet	Q/A	Active	Active	Active	Experimental	Operational
AR11698	Tile 47 Quiet	Tile 48 Quiet	Tile 50 FA	Tile 53 (Q) Quiet	Tile 49 4h Alarm	Tile 51 12h Alarm	Tile 52 1h Alarm	Tiles 49/49 4h Alarm	 -3h Alarm
AR11726	Tile 38 Quiet	Tile 39 Quiet	Tile 40 Quiet	Tile 41 (A) 39 h Alarm	Tile 42 9h Alarm	Tile 43 39h Alarm	Tile 44 18h Alarm	Tiles 42/41 17h Alarm	 10h Alarm
AR13165	Tile 29 Quiet	Tile 30 Quiet	Tile 35 Quiet	Tile 31 (A) 35h Alarm	Tile 32 36h Alarm	Tile 33 7h Alarm	Tile 34 55h Alarm	Tiles 32/32 36h Alarm	 29h Alarm
AR13179	Tile 38 Quiet	Tile 39 Quiet	Tile 43 FA	Tile 44 (Q) Quiet	Tile 40 20h Alarm	Tile 41 12h Alarm	Tile 42 2h Alarm	Tiles 41/41 12h Alarm	 5h Alarm
AR13183	Tile 38 Quiet	Tile 39 Quiet	Tile 44 Quiet	Tile 40 (A) -2h Alarm	Tile 41 5h Alarm	Tile 42 4h Alarm	Tile 43 11h Alarm	Tiles 41/41 5h Alarm	 -2h Alarm

Spiridon Kasapis | spiridon.kasapis@nasa.gov