Automatic detection of solar flares observed at 45 GHz by the POEMAS telescope

Vanessa Lessa[a], Adriana Valio[a]

[a]*Center for Radio Astronomy and Astrophysics Mackenzie, Mackenzie Presbyterian University, Rua da Consolacao, 930, São Paulo, 01302-907, SP, Brazil*

Abstract

Every 11 years, the Sun goes through periods of activity, with the occurrence of many solar flares and mass ejections, both energetic phenomena of magnetic origin. Due to its effects on Earth, the study of solar activity is of paramount importance. POEMAS (Polarization of Millimeter Emission of Solar Activity) is a system of two telescopes, installed at CASLEO (El Leoncito Astronomical Complex) in Argentina, which monitors the Sun at two millimeter wavelengths (corresponding frequencies of 45 and 90 GHz). The objective of this work is to automatically detect solar flares observed by the polarimeter. First it is necessary to eliminate the background noise, caused mainly by instrumental problems, from the light curves of millimeter solar emission. The methodology used to exclude the noise proposed in this work is to use the tendency of time series. The subtraction of this model from the light curves provides the input to automate the detection of solar flares using artificial intelligence techniques. A Neural Network was trained to recognize patterns and analyze a dataset in order to identify solar flares. Previously, a total of 30 flares had been visually identified and analyzed in the POEMAS database between 2011/11/22 and 2013/12/10. The methodology presented here confirmed 87% of these events, moreover the neural network was able to identify at least 9 new events. As the neural network was trained to detect impulsive events (lasting less than 5 min), long duration bursts were not automatically detected, nor were they detected visually due to the background noise of the telescope. Visual inspection of the POEMAS data, when comparing with microwave data from the RSTN, allowed the identification of an additional 10 long-duration solar flares at 45 GHz. We discuss some problems encountered and possible solutions for future work.

Keywords: Solar Flares, Millimeter Emission, Pattern Recognition, Neural

Contents

1. Introduction

The Sun is an active star with a magnetic cycle of about 11 years (Hathaway, 2015). In periods of maximum activity, an increase in the frequency of solar flares and coronal mass ejections can be observed. Both particles and magnetic fields thrown into interplanetary space by coronal mass from the Sun may impact the Earth. Geomagnetic storms, disruption of telecommunications signals, GPS malfunctions, and blackouts are some of the disruptions affecting Earth.

Over the years, studies on solar activity and the Sun's behavior have been carried out trying to mitigate these effects on Earth (Pulkkinen, 2007). For example, studies involving active regions, magnetic fields, solar flares, coronal mass ejections, and others are all relevant. Since the emission from solar activity is produced at all wavelengths of the eletromagnetic spectrum,

observations at different frequencies is crucial to understand solar phenomena and the mechanisms involved (Dulk, 1985).

In 1859, English astronomers Richard C. Carrington and Richard Hodgson identified the first solar flare (Tsurutani et al., 2003). The explosion was quite intense, and a flash located in a small region was detected in images of the Sun's visible light. Just 17 hours later, a coronal mass ejection hit the Earth, causing one of the largest magnetic storms ever recorded. If a similar storm reaches our planet today, it would cause severe communication and electrical energy problems, among others (Phillips, 2014).

Solar phenomena are usually associated with active regions of the solar atmosphere. When a solar flare occurs, a large amount of energy is released ($10^{28} - 10^{32}$ erg), this energy is used in accelerating particles and heating the plasma, which generate radiation across the entire electromagnetic spectrum (from X rays to radio waves) (Mann et al., 2009).

Observations of solar activity, both from ground and space observatories, have generated a large amount of data. Thus it is necessary to apply artificial intelligence techniques to analyze the data in search of the sudden increases in the emission caused by solar flares. Here we have used a neural network to find patterns and identify solar flares automatically.

This work involves the automatic detection of solar flares in the data of the Polarization Emission of Millimeter Activity at the Sun (POEMAS) telescope (Valio et al., 2013). POEMAS is a polarimeter that observed the Sun daily from December 2011 to December 2013 at the rarely explored frequencies of 45 and 90 GHz.

The paper is organized as follows. In Section 2, we describe the data, and in Section 3, the Neural Network methodology. In Section 4, the results of the Neural Network experiments are detailed. Finally, we conclude, in Section 5, and anticipate future research.

2. POEMAS Telescope

POEMAS is a system of two telescopes, installed at the CASLEO Observatory (El Leoncito Astronomical Complex), in Argentina. The POEMAS telescope provides solar left and right circular polarization measurements at two-millimeter wavelengths (45 and 90 GHz) with a temporal resolution of 10ms. It operated continually for two years (Dec 2011 - Dec 2013) observing the full disk of the Sun, and detected several flares. The data collected from

the Sun every day were written to binary files, converted using the Python programming language, and finally written to FITS files.

2.1. Data Acquisition

The antenna temperature data at both left and right circular polarized emission at 45 and 90 GHz are recorded in the daily binary files (TRK extension). Also the azimuth and elevation angles of the Sun are recorded in the file. For the analyzes, we used the light curve resulting from the sum of the two, right (RCP) and left (LCP), circular polarizations of the antenna temperature at 45 GHz. Using the Python programming language version 3.6, the TRK files were converted to FITS files using the following procedures:

- FITS level 0 - conversion of data from the TRK to FITS file, using the 10ms configuration of the original file.

- FITS level 1 - integration of the temporal resolution of the FITS file level 0 from 1 ms to 1s using the median of the data within the 1s interval.

- FITS level 2 - the merger of all FITS files level 1 of the same day into a single new file

- FITS level 3 - application of the time series (Trend) for all FITS file level 2

The first day of POEMAS observation used in this work was 12/01/2011, while the last day used was 12/10/2013. In this period, we did not have data for 51 days, which resulted in a total of 690 days for analysis.

After converting the POEMAS binary files to Flexible Image Transport System (FITS) files, the data is integrated into the database. Finally, the automatic detection of solar flares is applied using Artificial Intelligence techniques, especially pattern recognition and machine learning (Deep Learning).

2.2. Data calibration

Unfortunately, there is a misalignment of the telescope support structure due to mechanical problems, causing the signal to abruptly decrease during local noon. This decrease of the antenna temperature is clearly seen in Figure 1, especially between 16 and 17 UT. The red curve on the same plot depicts the expected light curve profile of the observations. Due to variations
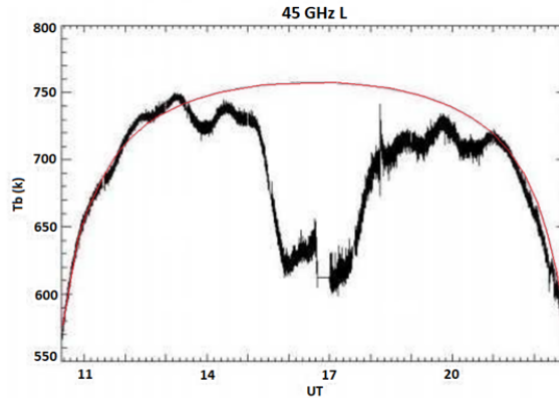
4

Figure 1: Full-day observation of solar energy flux at 45 GHz, left circular polarization (black curve) and attenuated flux (red curve).

in the solar emission caused by this telescope misalignment, it is difficult to identify any increase caused by a solar event, except for the most intense ones, which are usually rare.

Therefore, to minimize the daily variation of the telescope measurements at 45 GHz, a time series subtraction of the signal was performed. This time series is the component trend of the antenna temperature of the same polarization on the day under consideration. In Figure 2, we can see the 3 components of the time series for the observations on 01/27/2012. In the top panel of Figure 2, the light curve observed by POEMAS, that is the result of the sum of the RCP and LCP polarizations, is shown. The second plot from the top is the signal trend for different growth and decrease patterns. In the third plot, the seasonality is presented; in this case, we consider an interval of 50 points to analyze the behavior for every period of 1s. In the last panel, the residuals after subtraction of the effects of seasonality and trend from the data are presented. The residual fluctuations are attributed to random components.

In Figure 3, the original observed data (blue curve), the trend of the time series (red curve), and the result after subtracting the trend from the observed signal (black curve) are shown. A value of 1000 was added to this subtraction residual to place it on the same scale as the original signal. On this day, a solar flare occurred at 18:15 UT, and the impulsive peak can be observed in the original curve (blue). On the same day, between 21:00 UT
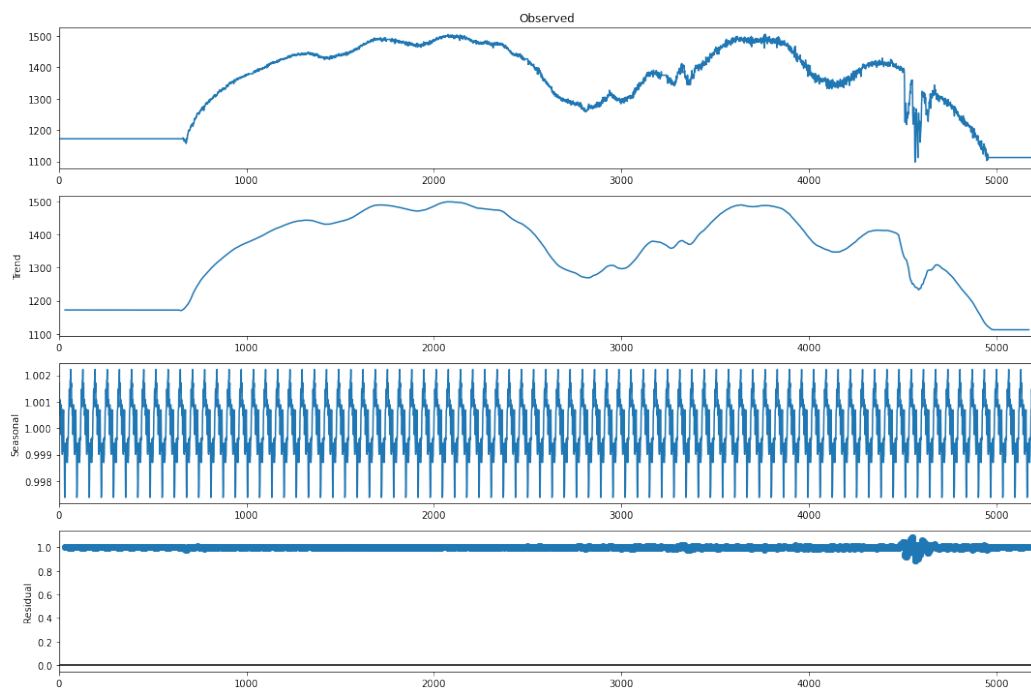
Figure 2: Decomposition of the time series of 2012/01/27

and 22:00 UT, there was a drop in signal due to interference in the Earth's atmosphere, probably due to clouds in front of the Sun. These drops in signal due to clouds are significant noise in the signal, and we have disregarded them from the data analysis.
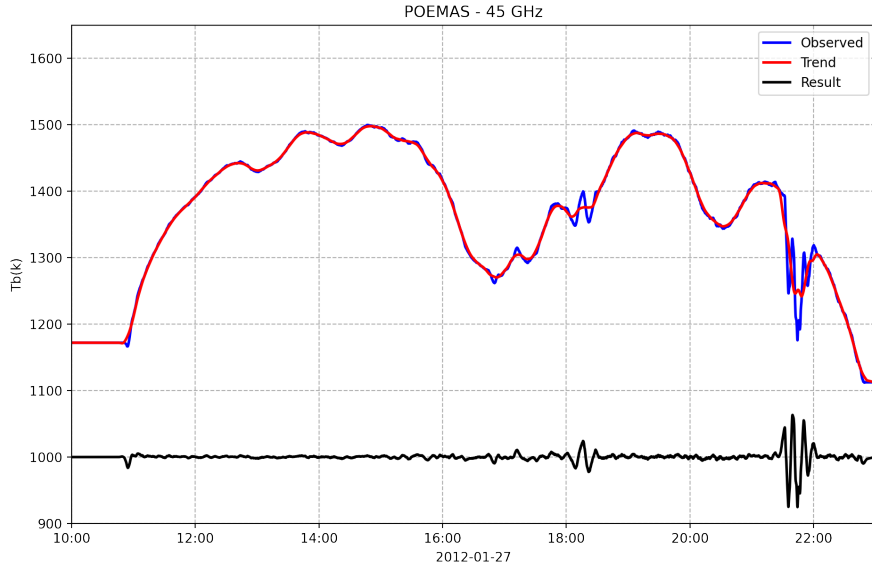

Figure 3: Analysis of the emission observed by POEMAS at 45 GHz all day 2012/01/27

To verify the existence of a solar event, we used data from the Radio Solar Telescope Network (RSTN) operated by the Meteorological Agency of the United States Air Force (Guidice, 1979). This network is composed of 4 radio stations located around the globe. Considering the location of POEMAS in Argentina, we will use data from observatories in Palehua, Hawaii (USA), Sagamore Hill in Massachusetts (USA), and San Vito (Italy). The 3 observatories cover the POEMAS observation window depending on the time of year. However, the antenna with the highest time intersection is in Sagamore Hill, Massachusetts (USA). Data from Sagamore Hill and Palehua stations for 01/27/2012 are shown in Figure 4, in the bottom and top panels, respectively.

The solar flare that happened on this day at 18:15 UT is clearly seen in the RSTN data shown in the two panels of Figure 4. The impulsive phase of
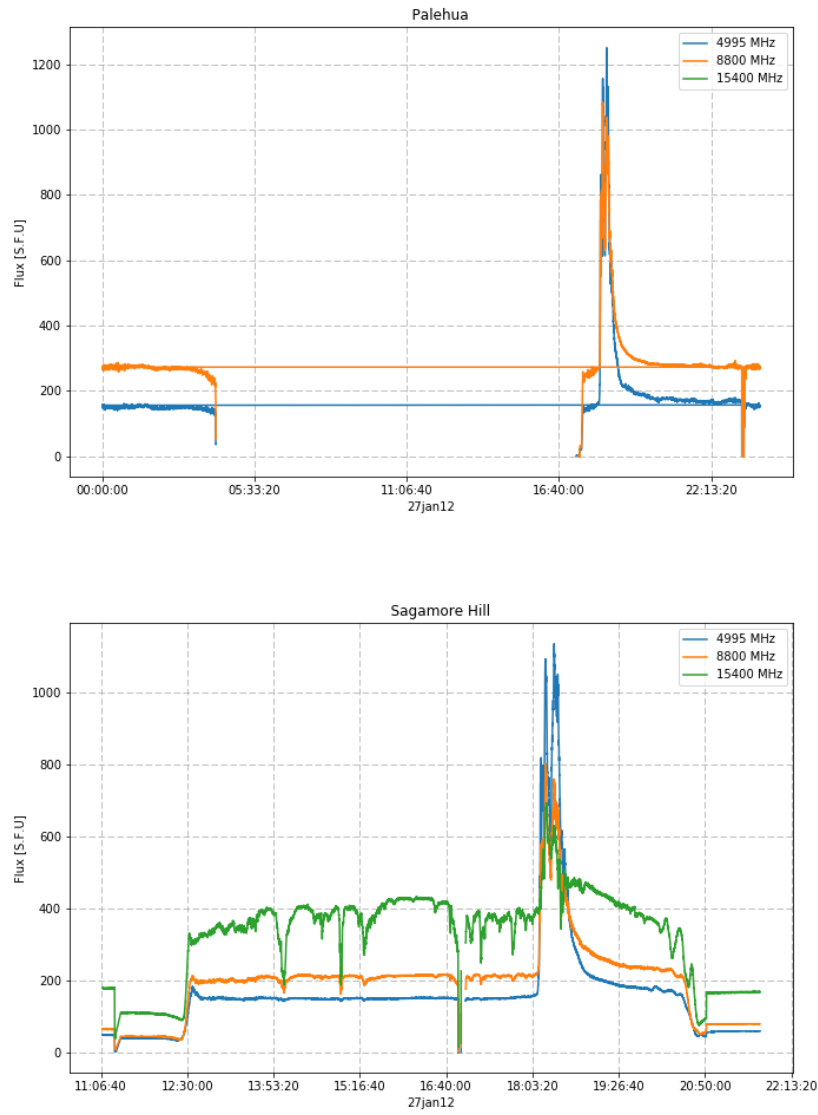
7

Figure 4: RSTN data for 2012/01/27 - Palehua (upper) and Sagamore Hill (bottom)

this event was also detected in the POEMAS signal at 45 GHz (Figure 3). However, to see the gradual phase of the flare, it is necessary to subtract the daily instrumental variation from the signal due to the telescope's misalignment. This can be done by subtracting the signal observed on the day before (or after) the event, since the variation in the signal does not vary significantly in the period of a day or so. However, this is not the procedure performed in this work due to its requirement of human supervision.

3. Neural Network

Based on biological neural networks, that is, on the biological neuron, Artificial Neural Networks (ANNs) are mathematical models that have computational capacity acquired through learning and generalization. This structure attempts to mimic a human brain with connections between neurons (synapses) and input and output signals.

Frank Rosenblatt at the *Cornell Aeronautical Laboratory* developed the first multi-neuron network of the linear discriminator type and named this network the *perceptron*. A *perceptron* is a network with neurons arranged in layers. This proposed model learns concepts and can answer with true (1) or false (0). In the early 1960s, Rosenblatt extended his work by publishing several articles and a book (Rosenblatt, 1962; Tappert, 2019).

The resulting 45 GHz emission signal, after subtraction of the trend of the time series (black curve of Figure 3) was input to a Multilayer Perceptron Neural Network (NN). The temporal resolution of the light curve is 1 second, which would generate a lot of data points for the network. Therefore, the temporal resolution was modified to 10 seconds, using the median in each 10-second interval to reduce the number of data points. Then a 5-minute window was considered to go through the resulting signal extracting chunks every 10 seconds.

To better exemplify the process, in Figure 5 we show the validation of time intervals to be later considered as true signals for input to the neural network. In the example, the supposed event starts at 1:02:00 UT and ends at 1:03:00 UT. The first true interval would be [0:58:00, 1:03:00] UT, the second true interval would be [0:58:10, 1:03:10] UT, and so forth. Therefore, we would have a sliding 5 min window traversing the final signal every 10s. The last true interval of this supposed event would be [1:02:00, 1:07:00] UT. If this was the only event in one day, there would be 25 true intervals.
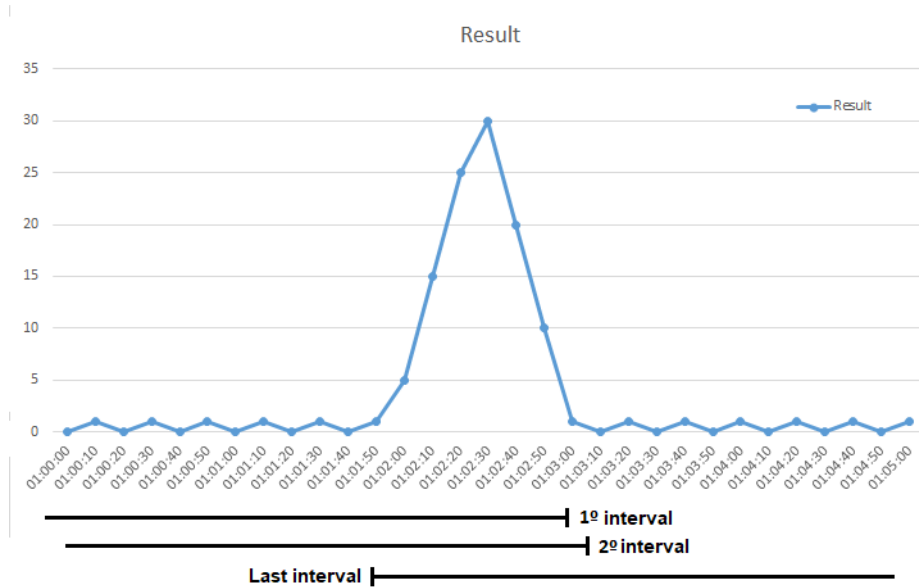
Figure 5: Example of interval extraction

This was the most critical process in evaluating the signal due to the large volume of data. There were a total of 690 days of solar observation by the POEMAS with an average of 10 hours per day. Considering a temporal resolution of 10 s, there are approximately 3600 points in a day; thus resulting in 2.484×10^6 data points.

3.1. Neural Network Training (NNT)

Previously, Hidalgo Ramírez et al. (2019) visually detected and analyzed 30 events observed by POEMAS telescope, which are listed in Table 1. First, we separate these events into two classes: training and classification. The first 14 flares were separated into the classification group. Events 15 through 30 were used for training of the neural network. The model adopted was supervised learning, where we submitted the 228 intervals referring to the 16 events and classified them as positive.

We need a balanced training base, and for that, we use the Nearmiss data balancing algorithm, a subsampling algorithm that randomly reduces most class examples. In the case of negative intervals, however, it selects samples based on distance (Mani and Zhang, 2003).

The next step is to define the NN structure after the training base is

N	Date	Time (UT)	NN
1	2011/12/25	18:14	✓
2	2011/12/25	20:27	✓
3	2011/12/26	20:16	✓
4	2012/01/27	18:15	✓
5	2012/03/05	19:13	✓
5	2012/03/06	21:05	–
7	2012/03/13	17:24	✓
8	2012/05/07	17:23	–
9	2012/05/09	12:30	✓
10	2012/06/03	17:53	✓
11	2012/07/03	17:02	✓
12	2012/07/04	16:36	✓
13	2012/07/05	20:13	✓
14	2012/07/08	16:30	✓
15	2012/07/10	15:30	–
16	2012/11/27	15:56	✓
17	2013/02/17	15:47	✓
18	2013/05/13	16:03	✓
19	2013/10/15	19:00	–
20	2013/10/25	15:00	✓
21	2013/10/26	19:25	✓
22	2013/10/28	15:10	✓
23	2013/10/29	21:48	✓
24	2013/11/05	18:10	✓
25	2013/11/05	22:11	✓
26	2013/11/06	13:43	✓
27	2013/11/07	12:26	✓
28	2013/11/07	14:18	✓
29	2013/11/07	14:23	✓
30	2013/11/07	14:29	✓

Table 1: Events identified in the work of Hidalgo Ramírez et al. (2019). The last column lists the events identified in this work by the Neural Network (NN).

balanced (228 positives and 228 negative intervals). The input layer has 30 nodes that will receive each interval of 5 minutes with a resolution of 10 seconds. The output layer has only 1 node, which reveals if the prediction for the data input is true or false. Several configuration tests were performed for the middle layer, using one and two layers. The best training results were found using only 1 intermediate layer. Thus, we decided to use one layer and varied the number of neurons to compare the results.

Because we are using supervised learning, when the NNT predicts a given input and the result is not expected, the network must adjust the weights to reduce the error and repeat the process until the error rate is 0. The activation function used was ReLu, as it is not linear and has better results compared to Tanh.

ReLU is the most commonly used activation function when designing neural networks today. The function is non-linear, and it does not activate all neurons at the same time, because if input is negative, it will be converted to zero and the neuron will not be activated. The Sigmoid function is a sensitive function and is continuously differentiable. This is not a linear function, and this is an interesting feature because it essentially means that when there are several neurons with sigmoid function as activation function the output too is nonlinear. This function varies between values [0,1]. The Tanh function is very similar to the sigmoid function. In fact, it's just an improved version of the sigmoid function as it varies between values [-1,1] (Burns, 2019).

To assess the quality of training and classification, we consider accuracy as a figure of merit, as it defines the proximity of an experimental result to its actual value. The greater the accuracy, the closer it is to the actual result. For all NNT configurations performed, the training had an accuracy of 100%. That is, they hit 228 positives (VP) and 228 negatives (VN) and had no false positives (FP) and no false negatives (FN).

3.2. Classification

The network efficiency and learning quality depend on its architecture specification, that is, the function of neuronal activation, learning rule, initial values, and training data. We consider a network with 3 layers: one input, one intermediate, and one output. This configuration showed the best performance and results in the training phase. For the primary classification, there were 14 events and 141 actual intervals. The 16 events used in the training were labeled in the classification as negative.

	# Neurons Occult layer	True Positive	False Negative	False Positive	True Negative	**Acuracy**
Exp. 1	30	126	15	1522161	747732	33%
Exp. 2	60	126	15	1334829	935064	41%
Exp. 3	90	127	14	1338661	931232	42%
Exp. 4	120	127	14	1246748	1023145	45%
Exp. 5	150	126	15	1522170	747732	33%
Exp. 6	160	126	15	1196951	1072942	47%
Exp. 7	170	131	10	1355206	914687	40%
Exp. 8	180	113	28	1764115	505778	22%

Table 2: Results of the Neural Network.

For each experiment, we varied the number of intermediate layer neurons. We started with 30 neurons and added 30 neurons in each experiment afterwards. In Table 2, we present the results of the 8 experiments. In the first column of the Table, the number of the experiment is given, in the second the number of neurons used in the intermediate layer, in the third column the number of positive intervals identified by the network as positive, and in the fourth column the number of positive intervals specified as negative. In the fifth column, the number of negative intervals identified as positive by the NNT are seen. Finally, in the sixth column, the number of negative intervals that are real negative, and in the seventh column, we have the accuracy of each of the experiments.

We started with 33% accuracy in experiment 1 with 30 neurons in the middle layer. In experiments 2, 3, and 4 we had a gradual increase in accuracy, reaching 45%. In experiment 5, with 150 neurons, we had a drop in accuracy to 33%, which is equivalent to the results in experiment 1. As the results were not inferior to any previous investigation, we decided to continue increasing the neurons to check the results. In experiment 6 we achieved an accuracy of 47% with 160 neurons in the hidden layer. In experiment 7, the accuracy decreases to 40%, and experiment 8 reached the lowest accuracy found, 22%. Experiment 7 had the highest number of true positives but a high number of false positives, so we focused on experiment 6 to analyze the results. Experiment 6 with a configuration of 160 neurons had an accuracy of 47%, the best among all the configurations.

To improve the results, we note that each minute has 6 points (10 seconds time resolution) evaluated several times by the neural network. For

experiment 9, if the network assessed less than 3 points within a minute as positive, we would consider that minute as negative. For experiment 10, if less than 4 points are evaluated as positive within one minute, this minute is assumed to be negative. The proposed method improved the accuracy of the neural network to 60%, as seen in Table 3. There are still many false positives and a considerable decrease in true positives, from 126 to 13, when comparing experiments 6 and 10. Also, the false negative increased from 15 to 128 cases.

	# Positive Points	True Positive	False Negative	False Positive	True Negative	**Acuracy**
Exp. 9	< 3	25	116	990033	1279860	56%
Exp. 10	< 4	13	128	906998	1362895	60%

Table 3: Results of the Neural Network for experiments 9 and 10.

Since the results of the neural network presented many false positives, we then compared the neural network results with the RSTN light curves, as well as the POEMAS observations, by visual inspection. First, the daily light curves from December 2011 to December 2013 (years of POEMAS observation) provided by RSTN were checked for events at frequencies $> 4\,$GHz (microwaves). When identifying a microwave event in the RSTN data, we check if the Neural Network identified this event at the same day and time in the 45 GHz data.

4. Results and discussion

In this work, we analyzed the 45 GHz light curves observed by POEMAS telescopes for a total of 690 days, from December 2011 to December 2013. For the analysis, we considered the sum of the signal involving the two circular polarizations, RCP plus LCP. Moreover, visual inspection was performed on the daily microwave light curves observed by the RSTN telescope network during the same period.

The Neural Network (NN) application detected patterns in the 45 GHz light curve of the POEMAs and identified already known events and new ones. We compared the results of the Network with the work of Hidalgo Ramírez et al. (2019), who identified 30 events, as listed in Table 1. In addition, some events identified in the RSTN microwave data were not detected

visually in the POEMAS light curves nor by the NN, given its temporal characteristic of long duration. Below we discuss each of these results in more detail.

4.1. Events identified previously and confirmed by the Neural Network

The Neural Network identified 26 of the 30 events from the work of Hidalgo Ramírez et al. (2019), thus the NN was able to retrieve 87% of the events (last column of Table 1). An example of such an event is shown in Figure 6 that occurred on 2013 May 13th. In the top panel, the POEMAS light curve is shown in blue, whereas the points identified by the NN as positive are depicted in red. In the bottom panel, the microwave light curves of the RSTN data from the Sagamore-Hill telescope (USA) are presented, where the event at 16:03 UT is clearly seen at all frequencies. Thus, the NN correctly identified the event that peaked at approximately 16:03 UT. This was a large event, of GOES X-ray class X2.8, that occurred on the East limb of the Sun.

4.2. Events previously reported and not identified by the Neural Network

In Table 1 there are four events identified visually by Hidalgo Ramírez et al. (2019) but not recognized by the NN, which make up only 13% of the events. The plots in Figure 7 represent the light curves of 2012 May 7th observed at 45 GHz by the POEMAS and in microwaves by RSTN. The event is clearly identified in the RSTN at approximately 17:23 UT. This event is not easily recognized in the POEMAS data and was neither identified by the NN, probably because of a duration longer than 5 min.

4.3. New events identified by the Neural Network

A total of 9 events were identified by the Neural Network, but went unnoticed in the visual inspection of Hidalgo Ramírez et al. (2019). These flares are listed in Table 4. An example of such is the burst that occurred on 2012 July 28th, identified by the NN at approximately 21:00 UT. The time profile of the flare is shown in the top right panel of Figure 8, where the light curve of the previous day was subtracted to eliminate the variation due to the misalignment of the telescope, and better show the event.
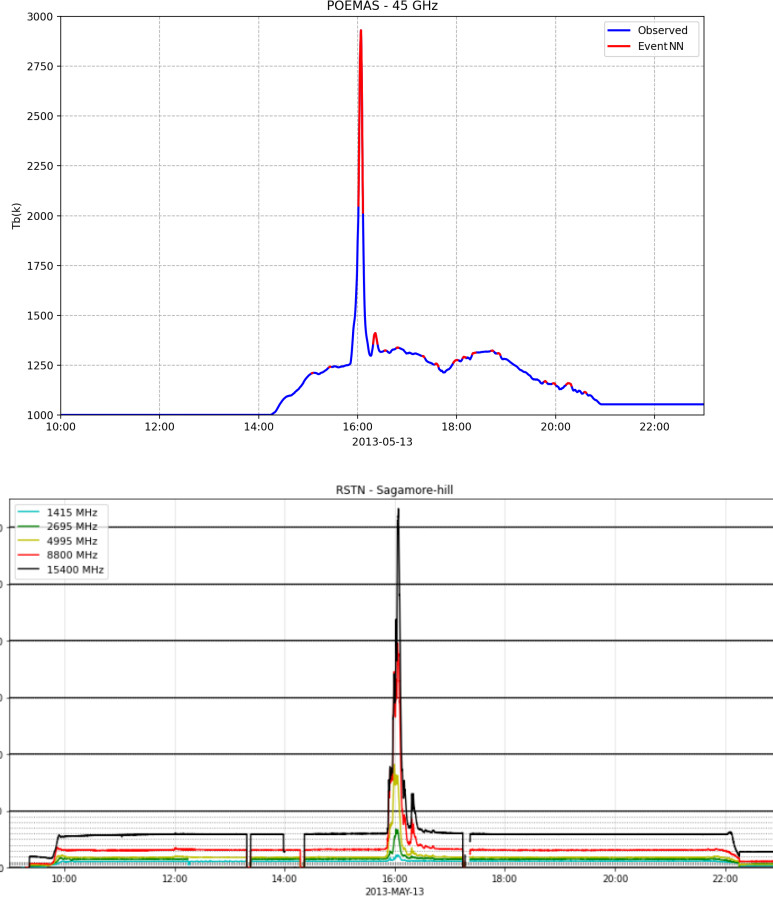
Figure 6: Light curves from the event on 2013/05/13 at 16:03 UT. **Upper:** 45 GHz data from POEMAS (blue) and Neural Network result (red) and **Lower:** data from RSTN.

4.4. New long duration events visually identified

From the visual inspection of the RSTN microwave light curves, we identified 11 events not identified by Hidalgo Ramírez et al. (2019) nor by the Neural Network, which are listed in Table 5. An example of such event is shown in Figure 9 for the 2012 July 12th flare, a long-duration event that lasted for more than 2 hours. The top panel shows the POEMAS 45 GHz light curve in blue, withe the light curve of the day before depicted in green. The subtraction of the emission from the previous day is shown in black on the top panel and highlighted in the middle panel of Figure 9. The flare that
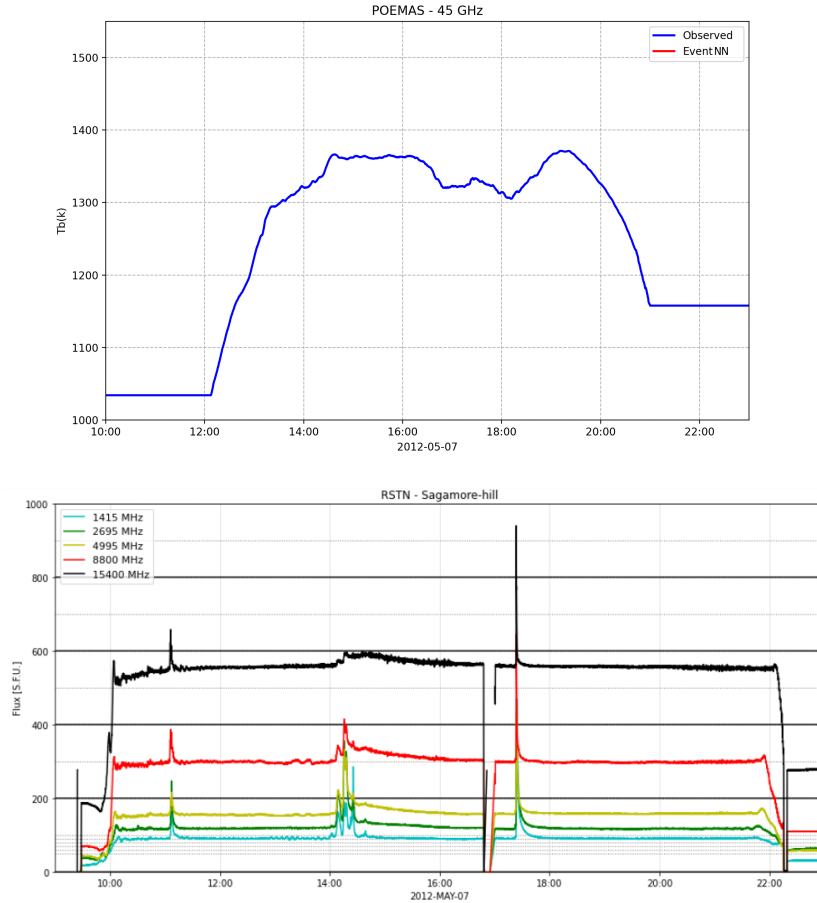
16

Figure 7: Light curves from the event on 2012/05/07 at 17:23 UT. **Upper:** 45 GHz data from POEMAS (blue) and Neural Network result (red) and **Lower:** data from RSTN.

started at approximately 16:00 UT, peaked just before 17 UT and ended after 18:30 UT, was not readily visually identified in the POEMAS light curve. In the bottom panel of the figure, the microwave light curves observed by RSTN clearly show the event at all frequencies, where the temporal profile of the 15 GHz emission closely resembles that of the 45 GHz from POEMAS. The non-identification of this and the other 9 events probably occurred due to their gradual temporal profile, lasting from 30 minutes to more than an hour. We point out that the data input for NNT consisted of 5 min intervals.
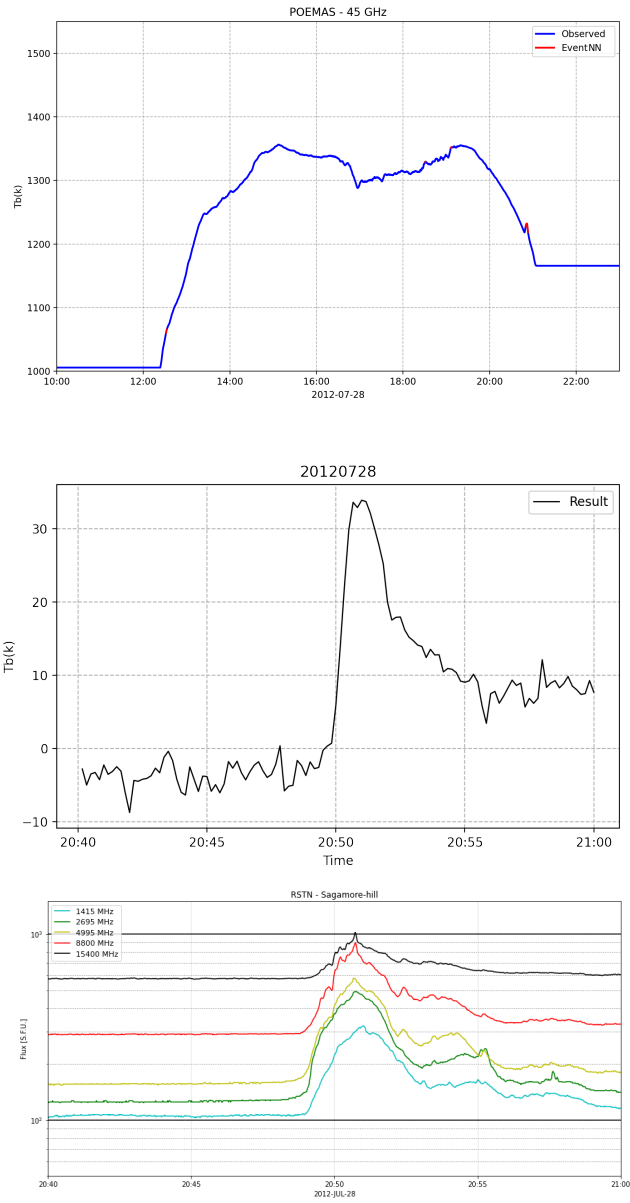
Figure 8: **Top:** Poemas light curve on 2012 July 28th, with the event at 20:50 UT identified by the NN depicted in red. **Middle:** Highlight of the 45 GHz event after subtraction of the emission from the previous day. **Bottom:** Microwave light curve of the same day observed by RSTN.
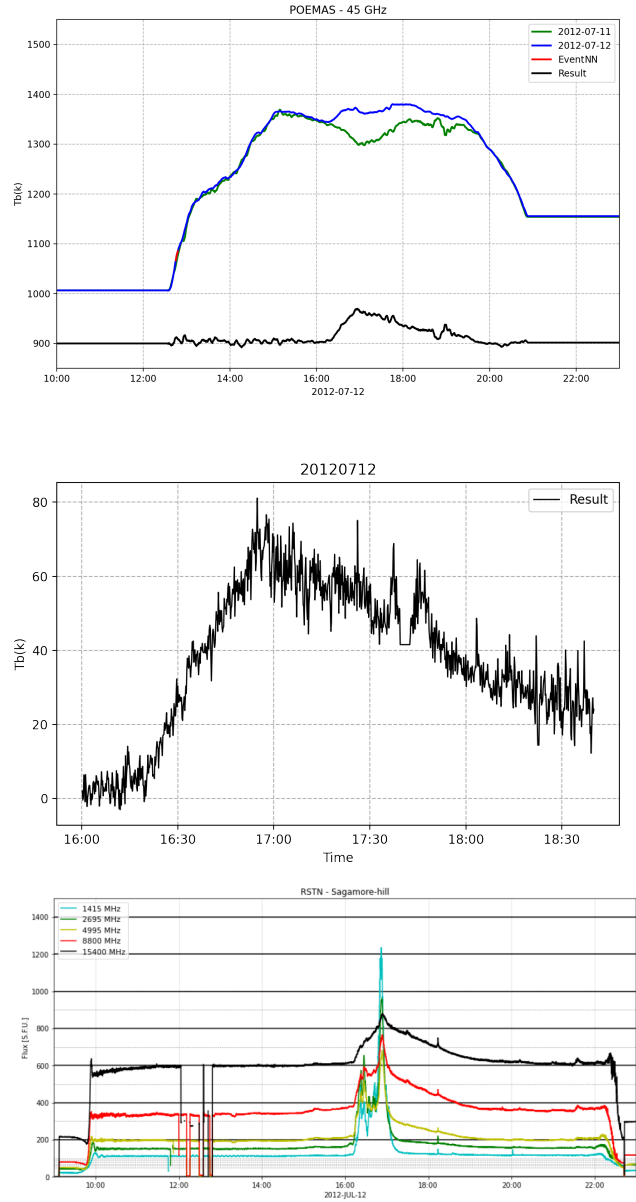
Figure 9: Solar flare of 2012/07/12. **Top:** Light curve mission at 45 GHz from POEMAS for the day of the event, 2012/07/12 (blue curve) and the previous day, 2012/07/11 (green curve). The result of the subtraction of the emission on the 12th by the 11th of July 2012 is shown by the black curve (shifted by 900 to fit the scale of the figure). **Middle:** Blow up of the subtracted light curve to better identify the solar flare detected at 45 GHz. **Bottom:** Microwave light curves observed by RSTN $(1-15 \text{ GHz})$ for the whole day.

19

N	Date	Time (UT)
1	2012/01/28	11:50
2	2012/05/08	13:00
3	2012/07/28	21:00
4	2012/09/02	18:10
5	2012/10/20	18:15
6	2012/10/21	20:00
7	2013/05/03	16:30
8	2013/05/03	17:30
9	2013/07/02	17:50

Table 4: Events identified only by the Neural Network

N	Date	Time (UT)
1	2012/03/02	17:40
2	2012/03/03	18:00
3	2012/03/04	11:00
4	2012/03/17	20:50
5	2012/06/06	20:00
6	2012/06/14	13:30-15:00
7	2012/07/12	16-17:00
8	2012/07/27	17:15
9	2013/07/03	20:00
10	2013/08/17	18:20-19:30

Table 5: Long-term events not identified in the work of Hidalgo Ramírez et al. (2019) nor by the Neural Network

5. Summary and conclusions

In this work, we analyzed two years of light curves from POEMAS telescopes at 45 GHz, from December 2011 through December 2013. The main objective was to automatically detect solar events in these light curves. The detection of solar flares in the light curves of POEMAS was hindered due to problems with the telescope's pointing, causing daily variations in the signal. Therefore, it was necessary to apply initial computational techniques to calibrate and reduce the POEMAS data. We created and used a Neural Network (NN) to identify solar flares in the data automatically. The application of this Neural Network was later compared with the microwave emission ($1-15$

Event	Total	Old*	Neural Network	Table
Old*	30	30 (100%)	26 (87%)	1
New	9	0 (0%)	9 (100%)	4
Long duration	10	0 (0%)	0 (0%)	5
Total	49	30 (61%)	35 (71%)	

Table 6: Summary of the events detected, or not, by the Neural Network.
* refers to the events identified visually by Hidalgo Ramírez et al. (2019).

GHz) detected by the RSTN radio-telescope network.

The first challenge of this work was the data transformation, due to noise and augmented by the misalignment of the telescope. Moreover, clouds obstructed the observation of the Sun and interference in the Earth's atmosphere even in the absence of clouds, such as increased water vapor or ice crystals, also precluded the detection of the flare signal by creating several spurious peaks in the light curves.

Using a Neural Network with supervised learning, we reached an accuracy of 47%. This value is low, however it is due to the few samples used for training and the intrinsic noise of POEMAS' light curves. Later the accuracy was improved to 60% by using a constraint on the False Positives. Nevertheless, despite the problems in the data mentioned above, we confirmed 26 previously known events and could identify 9 new events detected in the light curves of POEMAS, thanks to the NN.

Comparing the RSTN data for the two years of 2012 and 2013, we visually identified 10 long-term events not previously identified in the POEMAS light curve by visual inspection nor by the NN. As the Neural Network was not supplied with any long-term events for training, it is not capable of detecting this type of event. Thus the Neural Network constructed here can detect only short-term impulsive events, with duration less than 5 minutes.

In summary, with the aid of artificial intelligence, in this work we have identified 35 solar events in the 45 GHz emission from 2012 to 2013 from a total of 49 bursts, or 71%. If we consider, that the NN was not trained to detect events with duration longer than 5 min, then the accuracy of the NN increases to 90%. From the total of 49 events, 19 solar flares are unprecedented at 45 GHz, not having been identified in previous works that analyzed these data (Hidalgo Ramírez et al., 2019). The statistics of the NN is summarized in Table 6.

The use of artificial intelligence is innovating Astronomy. Especially in

Solar Physics, we can mention the works by Hou et al. (2020); Ishikawa et al. (2021); Neira et al. (2020), who obtained an accuracy of approximately 90%. To leverage the study of solar flares, such techniques must be explored.

The search for solar flares using Neural Network was a first step to automating the process. Several challenges were encountered, such as clouds, periodic signal variations, and non-detection of long-term events, and their solution will be proposed in future work.

Acknowledgements

References

Burns, S., 2019. Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn and Tensorflow. Independently Published.

Dulk, G.A., 1985. Radio emission from the sun and stars. Annual Rev. Astronomy and Astrophysics 23, 169–224.

Guidice, D.A., 1979. Sagamore Hill Radio Observatory, Air Force Geophysics Laboratory, Hanscom Air Force Base, Massachusetts 01731. Report., in: Bulletin of the American Astronomical Society, pp. 311–312.

Hathaway, D., 2015. The solar cycle. Living Reviews in Solar Physics doi:10.12942/lrsp-2010-1.

Hidalgo Ramírez, R.F., Morosi, A., Silva, D., Simões, P.J.A., Valio, A., 2019. Center-to-Limb Variation of Solar Bursts Polarization at Millimeter Wavelengths. Solar Physics 294, 108.

Hou, Y.C., Zhang, Q.M., Feng, S.W., Du, Q.F., Gao, C.L., Zhao, Y.L., Miao, Q., 2020. Identification and Extraction of Solar Radio Spikes Based on Deep Learning. Solar Physics 295, 146.

Ishikawa, S.n., Matsumura, H., Uchiyama, Y., Glesener, L., 2021. Automatic Detection of Occulted Hard X-Ray Flares Using Deep-Learning Methods. Solar Physics 296, 39.

Mani, I., Zhang, I., 2003. knn approach to unbalanced data distributions: a case study involving information extraction, in: WORKSHOP ON LEARNING FROM IMBALANCED DATASETS.

Mann, G., Warmuth, A., Aurass, H., 2009. Generation of highly energetic electrons at reconnection outflow shocks during solar flares. Astronomy & Astrophysics 494, 669–675. doi:10.1051/0004-6361:200810099.

Neira, M., Gómez, C., Suárez-Pérez, J.F., Gómez, D.A., Reyes, J.P., Hoyos, M.H., Arbeláez, P., Forero-Romero, J.E., 2020. MANTRA: A Machine-learning Reference Light-curve Data Set for Astronomical Transient Event Recognition. The Astrophysical Journal Supplements 250, 11.

Phillips, T., 2014. Near miss: The solar superstorm of july 2012. URL: https://science.nasa.gov/science-news/science-at-nasa/2014/23jul_superstorm/.

Pulkkinen, T., 2007. Space weather: Terrestrial perspective. Living Reviews in Solar Physics doi:10.12942/lrsp-2007-1.

Rosenblatt, R., 1962. Principles of Neurodynamics. Spartan.

Tappert, C.C., 2019. Who is the father of deep learning?, in: 2019 International Conference on Computational Science and Computational Intelligence (CSCI), pp. 343–348. doi:10.1109/CSCI49370.2019.00067.

Tsurutani, B.T., Gonzalez, W.D., Lakhina, G.S., Alex, S., 2003. The extreme magnetic storm of 1–2 september 1859. Journal of Geophysical Research: Space Physics 108. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JA009504, doi:https://doi.org/10.1029/2002JA009504, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2002JA009504.

Valio, A., Kaufmann, P., Giménez de Castro, C.G., Raulin, J.P., Fernandes, L.O.T., Marun, A., 2013. POlarization Emission of Millimeter Activity at the Sun (POEMAS): New Circular Polarization Solar Telescopes at Two Millimeter Wavelength Ranges. Solar Physics 283, 651–665. doi:10.1007/s11207-013-0237-4.