Substorm Onset Prediction using Machine Learning Classified Auroral Images

Pascal Sado[1,1], Lasse Boy Novock Clausen[1,1], Wojciech Jacek Miloch[1,1], and Hannes Nickisch[2,2]

[1]University of Oslo
[2]Philips Research

December 7, 2022

Abstract

We classify all sky images from 4 seasons, transform the classification results into time-series data to include information about the evolution of images and combine these with information on the onset of geomagnetic substorms. We train a lightweight classifier on this dataset to predict the onset of substorms within a 15 minute interval after being shown information of 30 minutes of aurora. The best classifier achieves a balanced accuracy of 59% with a recall rate of 39% and false positive rate of 20%. We show that the classifier is limited by the strong imbalance in the dataset of approximately 50:1 between negative and positive events. All software and results are open source and freely available.

Substorm Onset Prediction using Machine Learning Classified Auroral Images

P. Sado[1], L. B. N. Clausen[1], W. J. Miloch[1], H. Nickisch[2]

[1]Department of Physics, University of Oslo, Oslo, Norway
[2]Philips Research, Hamburg, Germany

Key Points:

- Aurora images are classified and a classifier is trained to predict the onset of substorms within 15 minutes of seeing 30 minutes of images
- A leightweight classifier works reasonably well but is limited in the amount of information that can be processed
- The best classifier recalls 39% of substorms with 59% balanced accuracy and 20% false positive rate

Corresponding author: Pascal Sado, `Pascal.Sado@fys.uio.no`

Abstract

We classify all sky images from 4 seasons, transform the classification results into time-series data to include information about the evolution of images and combine these with information on the onset of geomagnetic substorms. We train a lightweight classifier on this dataset to predict the onset of substorms within a 15 minute interval after being shown information of 30 minutes of aurora. The best classifier achieves a balanced accuracy of 59% with a recall rate of 39% and false positive rate of 20%. We show that the classifier is limited by the strong imbalance in the dataset of approximately 50:1 between negative and positive events. All software and results are open source and freely available.

Plain Language Summary

When charged particle originating from the sun travel into near Earth space, they interact with the Earth's natural magnetic field. These interactions are what leads to the aurora, but can also cause problems with electric installations or satellite communications. Knowing when and where these occur can be used to mitigate negative effects. Such forecasts are also beneficial for research, as rockets could be launched into regions of interest or paths of satellites can be adjusted to arrive at the same time as the occurrence of such events. Our model takes images from ground based cameras to predict the onset of such strong space weather occurrences.

1 Introduction

The solar wind is the driving force of space weather on Earth. Energy can be stored in the Earth's magnetosphere and will subsequently be released. These so called substorms are not only cause for the spectacle we know as the aurora, but have also the potential to cause serious harm to modern technology. Particularly in view of the reliance of today's society on digital communication delivered by satellites has made this a major concern in the last few decades.

Heating and expansion of the atmosphere by the aurora can lead to an increase in drag on satellites, possibly reducing lifespan, warranting course correction or at the very least cause observations of the changed course to avoid collisions (Marcos et al., 2010). Geomagnetically induced currents can affect man-made electrically conducting structures such as the power-grid, under-sea communication cables or pipelines, causing disruption in various services (Pirjola, 2000). GNSS systems can provide exact timing and location services, based on the distance to the satellite calculated from the known position and travel time of the signal to a ground based receiver. However, ionospheric disturbances can change the travel time by several nanoseconds or few microseconds, giving errors in the position by a few meters (Kintner et al., 2007).

Although there is the potential for global events to occur, these are extremely rare and localised events are much more likely. In order to mitigate the risks, it is important to know when and where they will occur.

Originally based on images (Akasofu, 1964 and Akasofu et al., 1965), the study of substorms has moved on to satellite-supported studies (McPherron et al., 1973), giving us the currently used model of substorms. The solar wind has long since been identified as the main driving force behind substorms (Caan et al., 1975). A rapid northward turning of the Interplanetary Magnetic Field (IMF) B_z component was believed to be the main trigger behind substorms, however this has been disproven in recent years (Freeman & Morley, 2009 and P. T. Newell & Liou, 2011 and Johnson & Wing, 2014). During the growth phase of substorms, energy is stored in the Earth's magnetosphere. This energy is released during the expansion phase and the magnetosphere subsequently returns to its steady state in the recovery phase of a substorm.

Different phases during a substorm can trigger different mechanism of energy-release which will in turn have different outcomes on the visible aurora (P. T. Newell et al., 2010 and Akasofu, 2013 and Partamies et al., 2015).

In its simplest form during quiet times, aurora are visible in the shape of a single east-west arc, become larger and brighter, expand poleward during a substorm and form rapidly westward travelling folds, before breaking up into smaller structures, becoming more chaotic and returning to their quiet state again towards the end of a substorm (Akasofu, 1964).

Irrespective the origin of substorms, their footprint on Earth stays the same and subsequent identification can be performed either visually through all sky or satellite images of aurora or measurements of the Earth's magnetic field. Visual identification as performed for example by Frey et al. (2004) and Liou (2010) is still based on the definition by Akasofu (1964) consisting of sudden brightening of the aurora followed by poleward motion and increase in intensity of the aurora. Forsyth et al. (2015) and P. T. Newell and Gjerloev (2011) and Ohtani and Gjerloev (2020) use instrument based identification of substorms, where they used the change in Earth's magnetic field.

The lists of substorms originating from this work have found widespread use in the community for prediction of various space weather effects (cf. `https://supermag.jhuapl.edu/publications/`), including the prediction of substorm onsets by Maimaiti et al. (2019) using deep neural networks. With their model, the authors also confirmed the importance of the B_z component of the interplanetary magnetic field (IMF) (P. T. Newell & Liou, 2011) and the solar wind speed (P. Newell et al., 2016) on the occurrence of substorms. Their work shows how well solar wind data can be used to forecast onsets of substorms on a global level. Furthermore, Sado et al. (2022) have shown that all sky images contain sufficient information that can be extracted by a neural network and be used to model the behaviour of the Earth's local magnetic field in vicinity to the imager.

Taking the same approach, in this study we obtained approximately 4 million all sky imager data, classified the images and used a time series of images representing half an hour of data to predict the onset of substorms within the next 15 minutes after the time series.

Our final classifier operates with a recall rate of 39%, a false positive rate of 20% and a balanced accuracy of 59%. We show that the classifier often correctly identifies to occurrence of an event, but fails to pinpoint the exact location in time and therefore either misses or overshoots the target prediction. The classifier itself is as lightweight as possible and makes it therefore necessary to reduce the input information for training to its bare essentials.

In Section 2 we give an overview of which data we use and in Section 3 we detail our preprocessing steps for the images and substorm data. Finally in Section 4 we present our results and give a summary and outlook in Section 5.

2 Description of Data Sources

In this project, we use data from two different sources. Our images are taken from the THEMIS All Sky Imager array's camera in Gillam, Manitoba located at N $56°\,20.24'$, W $94°\,42.36'$. The All Sky Camera takes images every 3 s at a resolution of 256 px by 256 px. The images are taken in the 2009/2010 and 2010/2011 seasons corresponding to conditions of solar minimum and and in the 2014/2015 and 2015/2016 seasons for solar maximum. This gives us a total of approximately 3.7 million images taken over 4 years. The images were taken with a fisheye lense giving a full view of the sky from horizon to horizon. To remove artefacts like trees just above the horizon, a ring 20 px wide was removed. The images were then classified according to the method developed by Sado et al. (2022).

The images are complemented with physical data in the form of substorm occurrences based on the SuperMAG list of substorms. These were created by Forsyth et al. (2015) using the SOPHIE technique, where substorm expansion and growth phases are

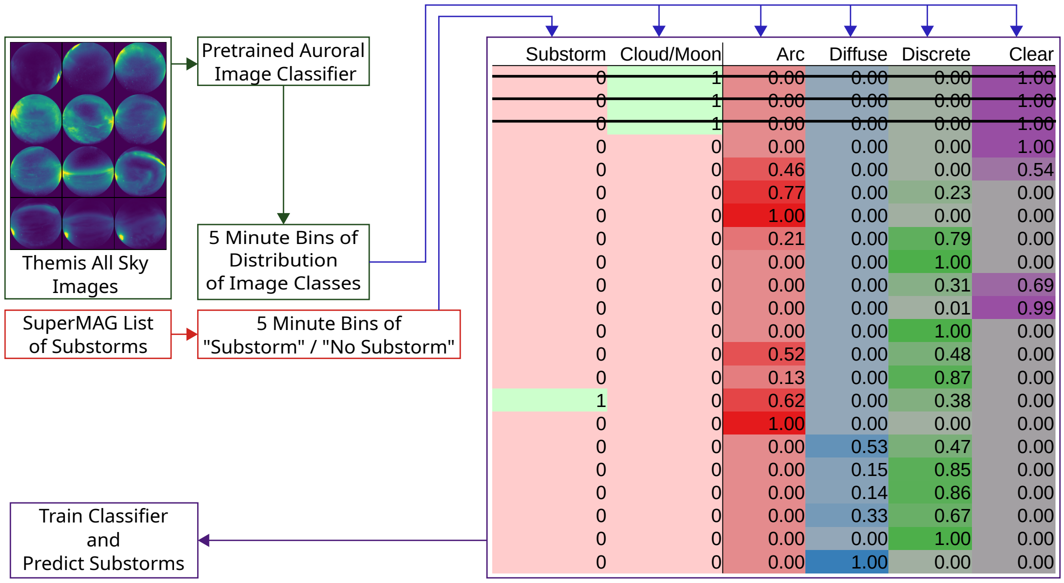

Figure 1: Outline of the workflow. Auroral Images are classified with an established classifier. Based on the classification's result images with clouds or the moon are removed. The predicted images' classes are summarised into 5 minute bins to remove noise and reduce the overall size of the dataset. To these bins, information about whether a substorm has occurred during the interval is added from the SuperMAG list and finally a classifier is trained to predict whether a substorm will occur after a given interval of images.

identified by finding extrema in the derivatives of the SML (Auroral Electrojet Index) and by Ohtani and Gjerloev (2020) who based their identification on the local development of the Earth's magnetic field as it is influenced by a substorm. From these lists of substorms, we use 245 events that occur at a time of image coverage within 10° geographical latitudinal and longitudinal distance to the camera.

3 Methods

3.1 Overview of Dataflow

Figure 1 shows an overview of how the data flows through the system. The all sky images are preprocessed and classified according to the classifier by Sado et al. (2022). This process is detailed in Section 3.2. Those images not showing aurora or a clear night sky are removed. The classified images are condensed into bins where we average over the images' probabilities in regular 5 min intervals. During periods of full camera coverage, a 5-minute-bin will contain 100 images. However, since coverage is not perfect or images have been removed because they were not relevant, bins might contain less images. Each bin is then assigned a binary value based on whether a substorm has occurred during this time or not according to the SuperMAG list.

The processing of substorm data and details on the classifier can be found in sections 3.3 and 3.4 respectively.

3.2 Image Preprocessing

Individual images are classified based on the classifier developed and demonstrated by Sado et al. (2022) and Clausen and Nickisch (2018). In this process, the images are analysed by a pretrained neural network and the image features as defined by this network are extracted. A classifier that has been trained on a labelled set of images that have undergone the same process of feature-extraction is then used to classify the images. This returns a probability for each image to be in either of the following six classes:

- **arc** The image shows mostly a single auroral arc spanning from east to west (left to right in the frame of the image)
- **diffuse** The image shows diffuse aurora without any clear structure
- **discrete** Discrete aurora show structure but not in the form of well-defined arcs. The structures can be of any other shape.
- **cloud** The image shows clouds
- **moon** The image shows the moon
- **clear** The image shows a clear night sky

The probabilities for "cloudy" and "moon" do not contain any physical information and could lead to unforeseen biases with the classifier. Images where the probability to show the moon is above 40% or the probability to show clouds is above 70% are therefore discarded These probabilities are then removed alltogether and we rescale the remaining four classes such that their distribution sums up to 100%.

3.3 Substorms

The list of substorms contains substorms measured and registered all over the world. Because we are only interested in substorms that we will be able to recognise visually based on our images, we remove all substorms outside of a 10° region in geographical coordinates around the location of the camera. We also remove any substorms registered at a time where there is no image data available. Doing this we obtain 245 individual substorms.

3.4 Classification

Our classifier is a simple Linear Ridge Model. As input we use 6 bins of 5 minutes of image data, giving us an input vector containing $6 \times 4 = 24$ cells of input data. As output to be predicted we use a Boolean value whether there will be a substorm within the next 15 minutes after the end of the input interval.

In Figure 2 we demonstrate how the input is prepared for the model. In the upper row the predicted classes for each image up to sixty minutes before and after a substorm has been identified are plotted. In the middle row, the average distribution of classes for each of the 5 minute bins is calculated and shown. Binning the images is an essential part of preprocessing for two reasons. Firstly, there were originally one hundred images taken per interval, the information is therefore reduced by a factor of 100. Secondly, briefly interrupted coverage at for example about 30 min after substorm onset and again about 55 min after onset can be safely ignored. The bottom two panels show a visualisation of the input for the classifier. Each contains a 30-minute-interval of data. The first interval ends more than 15 minutes before the substorm occurs and has therefore been given a negative label. The second interval ends less than 15 minutes before the substorm and has therefore been given a positive label. Of course there are many more times without substorm onset than there are with. In our method of binning the data into 5 minute intervals and looking 15 minutes ahead, 1.80% of our model's input has a positive label. To account for this large imbalance, we adjust hyperparameters for the model's class weight and regularisation strength to avoid overfitting.

For evaluation of hyperparameters, we have used 5-fold crossvalidation with an 80:20 split of train to test data. Our final selected model is the one that produces the highest balanced accuracy which is also the model with the highest True Skill Score (True Positive Rate - False Positive Rate).

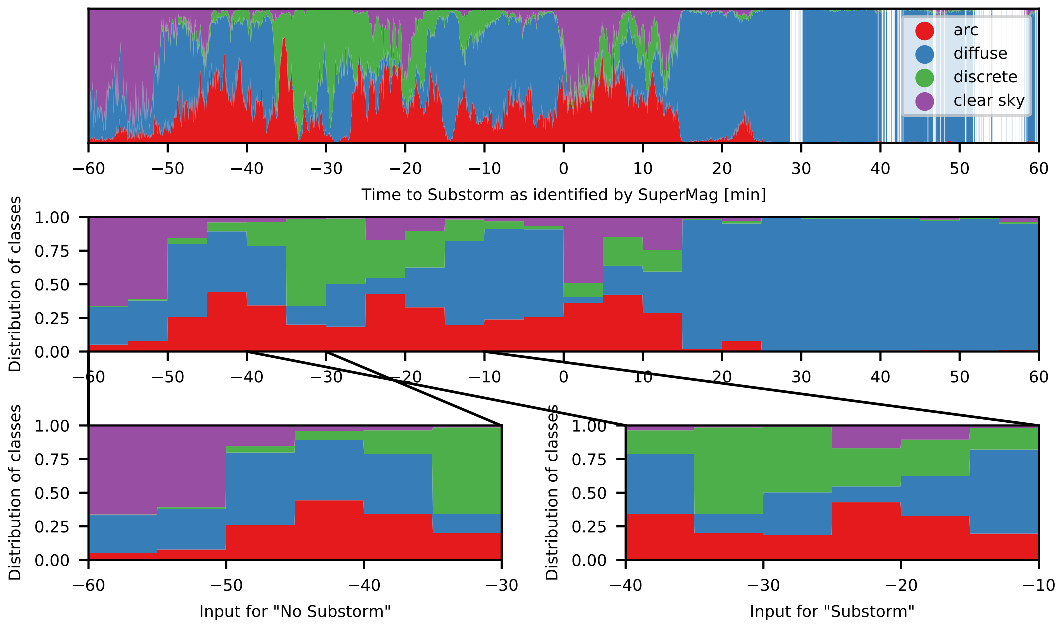

Figure 2: Predicted classes per image (top), binned distribution of classes (middle) and input for "substorm" (bottom left) or "no substorm" (bottom right)

For the final model's training and evaluation we have split train and test data sequentially in such a way that the ratio of positive to negative events in both datasets is as similar as positive.

4 Results

4.1 Distribution of Image Classes around Substorms

In Figure 3, we show the average distribution of predicted image classes up to an hour before and after a substorm has been observed. We can see that before substorm onset, the average probability for "arcs" rises and shortly after onset "discrete" sees a rise. "Diffuse" is a dominant term throughout the whole time series, but strongly rising after substorm onset. This is likely because the classifier tends to default to this value when it is unsure about the classification task. For clouds illuminated from the back, for example when the moon is shining behind cloud cover, or strong aurora that is blanketed by clouds, the classifier will also often classify these cases as diffuse aurora. "Clear sky" is similarly dominant towards the beginning of observation, but decreases over time. There are some substorms that will occur without aurora observation in the field of view of the camera. These will add a baseline value of "clear sky" to the average presented in this Figure.

Overall, we see that the substorms on average follow a pattern that is similar to the observations one would expect when detecting substorms manually on images.

4.2 Prediction

For the prediction task, we prepare the classifier as described in Section 3.4.

Table 1 shows the confusion matrix obtained for this classifier. It illustrates the imbalance in the dataset of approximately 50:1. We manage to correctly identify 41 of the 106 test cases in our dataset, giving us a recall rate of 39%. The imbalance has a large effect on the precision of the prediction, the ratio of true positive predictions to all pos-

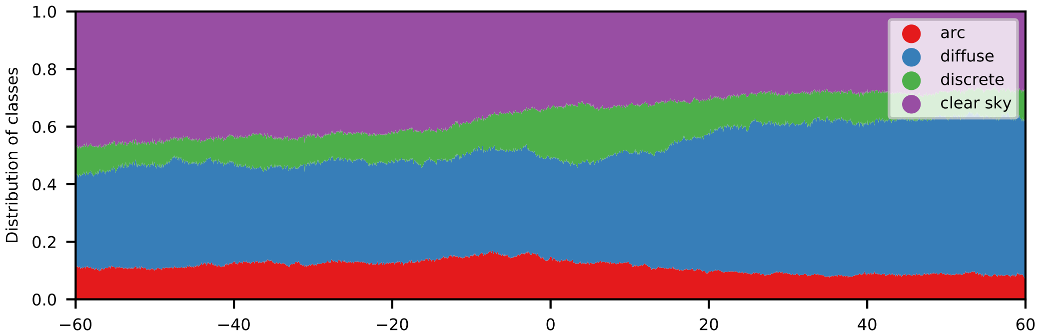

Figure 3: Distribution of predicted Image classes around substorms

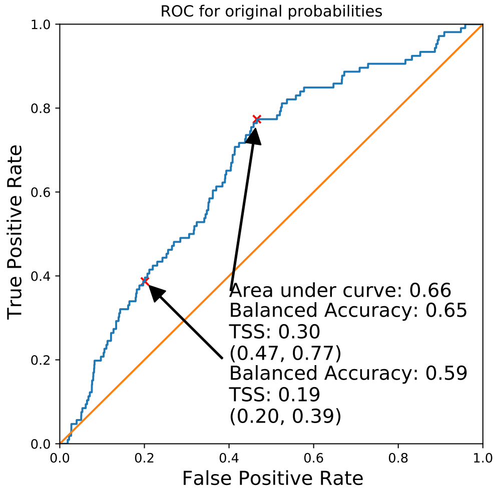

Figure 4: ROC curve for the final classifier.

		Prediction outcome		
		Substorm	**No Substorm**	**total**
actual value	**Substorm**	41	65	106
	No Substorm	1161	4613	5774
	total	1202	4678	

Table 1: Confusion Matrix for the final classifier

itive predictions, which is 3.4% in our case. Accounting for the imbalance, and weighting the accuracy for both cases with their total amount of cases, we achieve a balanced accuracy of 59%.

In Figure 4, we see a ROC-curve for the prediction. A ROC-curve is created by choosing different thresholds for the classifier's output and subsequently plotting the True Positive Rate (TPR) against the False Positive Rate (FPR). It shows how well in a binary classification system positive cases can be separated from negative ones and can be useful to choose a threshold based on the applications. In the two extreme cases, all samples are rejected or all samples are accepted as positive. Between these, the TPR should increase faster than the FPR to make for a good classifier. This threshold is shown as the straight, orange line which would also show the outcome of a classifier than was purely based on chance.

Except for the very extreme cases, our model performs better than guessing and overall, the are under the curve is 0.66. We identify two working regimes that could be useful in real world scenarios. The first at a balanced accuracy of 59% with a True Skill Score (TSS), calculated as the difference between TPR and FPR, of 0.19, a FPR of 20% and a TPR of 39%. The second has a higher balanced accuracy and TSS of 65% and 0.30 respectively, but the higher TPR of 77% comes at the cost of increasing the FPR to 47%. The first case is a more conservative approach and will create less false alarms relative to true positive cases than the second approach. The second approach is more accurate overall but will also give more false alarms.

Figure 5 displays the classification for a specific date. In the top panel we show ground based magnetometer measurements for the evening and the keogram for the timeframe in the panel below. The third panel shows the probabilities of individual images over time on which the substorm prediction has been based. The predicted probability for "substorm" vs "no substorm" is shown in the fourth panel. The horizontal black line denotes our threshold chosen for the final task. It corresponds to the first, more conservative, scenario laid out above. The binary output of this thresholded prediction is shown in panel five and the true result we tested against in the last panel. . The last panel shows the known true test data.

We see that the substorm occurring at 08:56 has been identified correctly, albeit being 5 minutes delayed on the timing. From the keogram we see that at the time there was little to no supporting visual evidence of a substorm occurring in the field of view of the camera. Leading up to the substorm the classifier has increasingly classified images as "arcs" or "discrete", similarly to what we saw in Figure 3. The substorm classifier itself does not have any information about the original images available for its task any more, however it seems that either the context of image class distributions has been enough to identify this substorm where a human would have likely not done so, or - given the low precision from the large imbalance in the dataset and the classifiers tendency to mark too many potential substorms - it has landed a lucky guess. Between 7:10 and 08:05 another event has been identified. As we can see from the magnetometer measurements plotted alongside, another substorm happened earlier with its onset identified at 06:56 by Ohtani and Gjerloev (2020). This substorm is not in our list of true positive data, because it occurred too early after onset of observations. Even if it was, it would not have been identified at the correct time, but the classifier has correctly identified that there was an ongoing event during the time. The substorm was also a longer lasting event, which was picked up by the classifier.

Both of these events show the necessity of implementing a loss function that prioritise the correct identification of present events over the precise timings. This could lead to a significant improvement in the model's forecasting abilities. (Guastavino et al., 2022) Both cases lead us to believe that the classifier prefers to identify ongoing substorms instead of the substorm onsets it was trained on. This is most likely due to the fact that the definition of a substorm onset is rather arbitrary with respect to image data and the effect on the images heavily depends on the duration and strength of the substorm.

Nevertheless, the fact that the classifier managed to roughly identify the time both events occurred, is a huge success given the very limited model and training data. It has been trained on data only giving information about the onset of the substorm, resulting in a large imbalance between true and false cases of about 1:50. This means that just by guessing "false" all the time the classifier would achieve an accuracy of about 98%. This would correspond to the top-right corner of the ROC-curve.

Using the correct threshold it is possible to obtain a working regime that is performing better than this trivial case. Given the fact that the original input for half an hour of data has been condensed down from 600 images at 256 px by 256 px giving approximately 40M data points total to just 24 input values, this is a good achievement for a linear classifier.

5 Conclusion & Outlook

We have shown that a simple linear classifier based on the distribution of image classes of auroral images for up to half an hour can predict the onset of a substorm with respectable accuracy. The input data also only contained information directly obtained from images. Replacing the model with a neural network, supplementing the input data with for example solar wind data and implementing a loss function that prioritises the

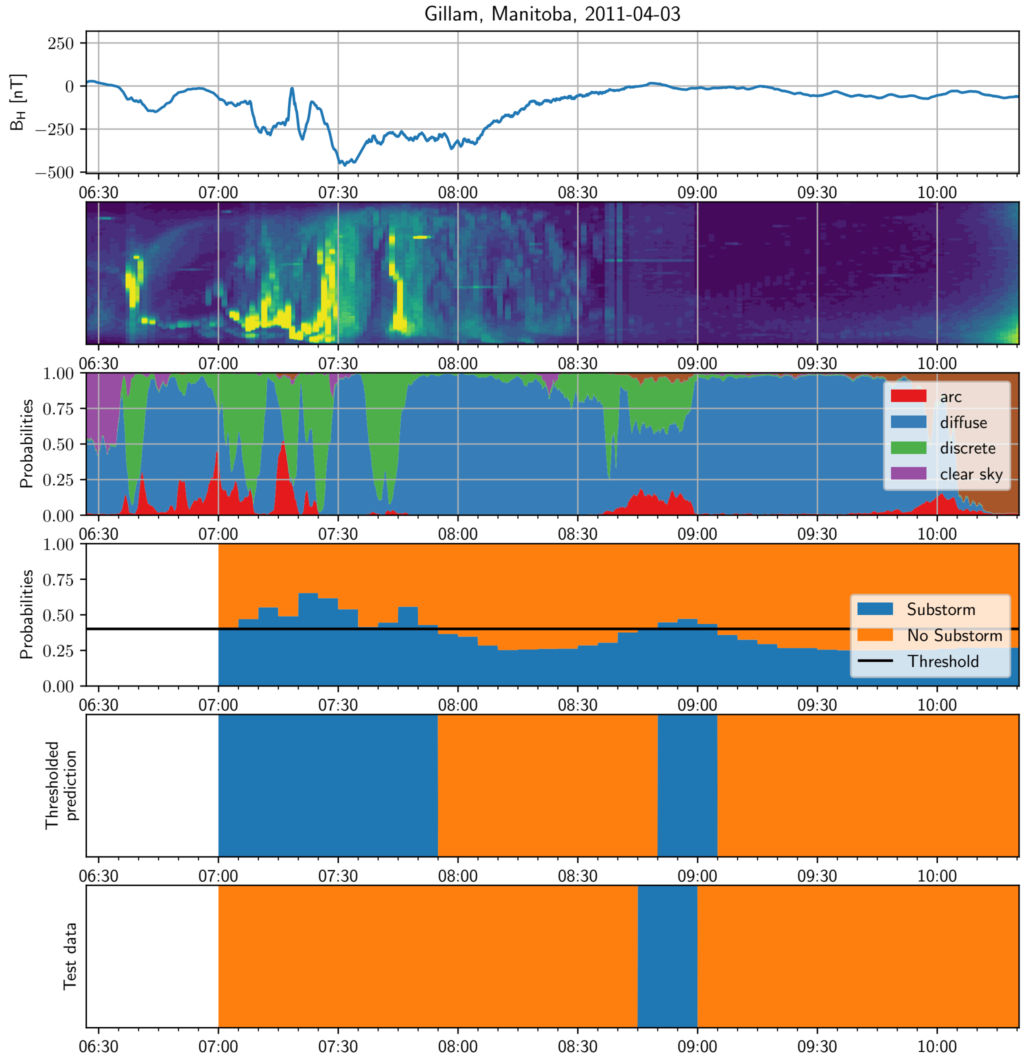

Figure 5: A demonstration of Prediction of a time series. The rows show the following information, from top to bottom: Magnetometer measurement, keogram, per-image classifier output, Substorm prediction probability, thresholded substorm prediction, test data.

forecast result's value over its precision could lead to a more accurate prediction of the local onset and possibly duration of substorms.

Because this method and underlying source code is made freely available, it can be used to forecast substorms live. While we have not undertaken such steps, the time-limiting factor in a project like this would be the image preprocessing. Since our methods operate much faster on commercial hardware than the limit of one image every three seconds, an optimised implementation should be possible.

Acknowledgements

This work is funded in part by the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (ERC CoG grant agreement No 866357). The All Sky Image Classifier was made available by Sado et al. (2022) on `http://tid.uio.no/TAME`. We provide the data and code for this project openly and freely on `https://doi.org/10.11582/2022.00070` and `http://tid.uio.no/SOP` respectively. We acknowledge NASA contract NAS5-02099 and V. Angelopoulos for use of data from the THEMIS Mission. Specifically: S. Mende and E. Donovan for use of the ASI data, the CSA for logistical support in fielding and data retrieval from the GBO stations, and NSF for support of GIMNAST through grant AGS-1004736. We acknowledge the substorm timing list identified by the SOPHIE technique (Forsyth et al., 2015), the SMU and SML indices (P. T. Newell & Gjerloev, 2011), the Ohtani and Gjerloev technique (Ohtani & Gjerloev, 2020), the SMU and SML indices (P. T. Newell & Gjerloev, 2011); and the SuperMAG collaboration (Gjerloev, 2012).

References

Akasofu, S.-I. (1964, April). The development of the auroral substorm. *Planetary and Space Science*, *12*(4), 273–282. Retrieved 2022-07-28, from `https://linkinghub.elsevier.com/retrieve/pii/0032063364901515` doi: 10.1016/0032-0633(64)90151-5

Akasofu, S.-I. (2013, March). Auroral Morphology: A Historical Account and Major Auroral Features During Auroral Substorms. In A. Keiling, E. Donovan, F. Bagenal, & T. Karlsson (Eds.), *Geophysical Monograph Series* (pp. 29–38). Washington, D. C.: American Geophysical Union. Retrieved 2022-08-22, from `http://doi.wiley.com/10.1029/2011GM001156` doi: 10.1029/2011GM001156

Akasofu, S.-I., Chapman, S., & Meng, C.-I. (1965, November). The polar electrojet. *Journal of Atmospheric and Terrestrial Physics*, *27*(11-12), 1275–1305. Retrieved 2022-07-28, from `https://linkinghub.elsevier.com/retrieve/pii/0021916965900875` doi: 10.1016/0021-9169(65)90087-5

Caan, M. N., McPherron, R. L., & Russell, C. T. (1975). Substorm and interplanetary magnetic field effects on the geomagnetic tail lobes. *Journal of Geophysical Research*, *80*(1), 191–194.

Clausen, L. B. N., & Nickisch, H. (2018). Automatic classification of auroral images from the oslo auroral themis (oath) data set using machine learning. *Journal of Geophysical Research: Space Physics*, *123*(7), 5640-5647. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JA025274` doi: https://doi.org/10.1029/2018JA025274

Forsyth, C., Rae, I. J., Coxon, J. C., Freeman, M. P., Jackman, C. M., Gjerloev, J., & Fazakerley, A. N. (2015). A new technique for determining substorm onsets and phases from indices of the electrojet (sophie). *Journal of Geophysical Research: Space Physics*, *120*(12), 10,592-10,606. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JA021343` doi: https://doi.org/10.1002/2015JA021343

Freeman, M. P., & Morley, S. K. (2009). No evidence for externally triggered substorms based on superposed epoch analysis of imf bz. *Geophysical Research Letters*, *36*(21). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009GL040621 doi: https://doi.org/10.1029/2009GL040621

Frey, H., Mende, S., Angelopoulos, V., & Donovan, E. (2004). Substorm onset observations by image-fuv. *Journal of Geophysical Research: Space Physics*, *109*(A10).

Gjerloev, J. W. (2012). The supermag data processing technique. *Journal of Geophysical Research: Space Physics*, *117*(A9). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012JA017683 doi: https://doi.org/10.1029/2012JA017683

Guastavino, S., Piana, M., & Benvenuto, F. (2022). Bad and good errors: Value-weighted skill scores in deep ensemble learning. *IEEE Transactions on Neural Networks and Learning Systems*, 1-10. doi: 10.1109/TNNLS.2022.3186068

Johnson, J. R., & Wing, S. (2014). External versus internal triggering of substorms: An information-theoretical approach. *Geophysical Research Letters*, *41*(16), 5748-5754. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014GL060928 doi: https://doi.org/10.1002/2014GL060928

Kintner, P. M., Ledvina, B. M., & de Paula, E. R. (2007, September). GPS and ionospheric scintillations: GPS AND IONOSPHERIC SCINTILLATIONS. *Space Weather*, *5*(9), n/a–n/a. Retrieved 2022-08-22, from http://doi.wiley.com/10.1029/2006SW000260 doi: 10.1029/2006SW000260

Liou, K. (2010). Polar ultraviolet imager observation of auroral breakup. *Journal of Geophysical Research: Space Physics*, *115*(A12).

Maimaiti, M., Kunduri, B., Ruohoniemi, J., Baker, J., & House, L. L. (2019). A deep learning-based approach to forecast the onset of magnetic substorms. *Space Weather*, *17*(11), 1534–1552.

Marcos, F., Lai, S., Huang, C., Lin, C., Retterer, J., Delay, S., & Sutton, E. (2010, August). Towards Next Level Satellite Drag Modeling. In *AIAA Atmospheric and Space Environments Conference.* Toronto, Ontario, Canada: American Institute of Aeronautics and Astronautics. Retrieved 2022-08-22, from https://arc.aiaa.org/doi/10.2514/6.2010-7840 doi: 10.2514/6.2010-7840

McPherron, R. L., Russell, C. T., & Aubry, M. P. (1973, June). Satellite studies of magnetospheric substorms on August 15, 1968: 9. Phenomenological model for substorms. *Journal of Geophysical Research*, *78*(16), 3131–3149. Retrieved 2022-07-28, from http://doi.wiley.com/10.1029/JA078i016p03131 doi: 10.1029/JA078i016p03131

Newell, P., Liou, K., Gjerloev, J., Sotirelis, T., Wing, S., & Mitchell, E. (2016). Substorm probabilities are best predicted from solar wind speed. *Journal of Atmospheric and Solar-Terrestrial Physics*, *146*, 28–37.

Newell, P. T., & Gjerloev, J. W. (2011). Evaluation of supermag auroral electrojet indices as indicators of substorms and auroral power. *Journal of Geophysical Research: Space Physics*, *116*(A12). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JA016779 doi: https://doi.org/10.1029/2011JA016779

Newell, P. T., Lee, A. R., Liou, K., Ohtani, S.-I., Sotirelis, T., & Wing, S. (2010, September). Substorm cycle dependence of various types of aurora: SUBSTORM DEPENDENCE OF AURORAL TYPES. *Journal of Geophysical Research: Space Physics*, *115*(A9), n/a–n/a. Retrieved 2022-08-22, from http://doi.wiley.com/10.1029/2010JA015331 doi: 10.1029/2010JA015331

Newell, P. T., & Liou, K. (2011). Solar wind driving and substorm triggering. *Journal of Geophysical Research: Space Physics*, *116*(A3). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JA016139 doi: https://doi.org/10.1029/2010JA016139

Ohtani, S., & Gjerloev, J. W. (2020). Is the substorm current wedge an ensemble of wedgelets?: Revisit to midlatitude positive bays. *Journal of Geophysical Research: Space Physics*, *125*(9), e2020JA027902. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JA027902 (e2020JA027902 2020JA027902) doi: https://doi.org/10.1029/2020JA027902

Partamies, N., Juusola, L., Whiter, D., & Kauristie, K. (2015, July). Substorm evolution of auroral structures. *Journal of Geophysical Research: Space Physics*, *120*(7), 5958–5972. Retrieved 2022-08-03, from https://onlinelibrary.wiley.com/doi/10.1002/2015JA021217 doi: 10.1002/2015JA021217

Pirjola, R. (2000, December). Geomagnetically induced currents during magnetic storms. *IEEE Transactions on Plasma Science*, *28*(6), 1867–1873. Retrieved 2022-08-22, from http://ieeexplore.ieee.org/document/902215/ doi: 10.1109/27.902215

Sado, P., Clausen, L. B. N., Miloch, W. J., & Nickisch, H. (2022). Transfer learning aurora image classification and magnetic disturbance evaluation. *Journal of Geophysical Research: Space Physics*, *127*(1), e2021JA029683. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021JA029683 (e2021JA029683 2021JA029683) doi: https://doi.org/10.1029/2021JA029683