Dynamics of Solar and Stellar Convection Zones and Atmospheres
Proceedings IAU Symposium No. 365, 2024
A. V. Getling & L. L. Kitchatinov, eds.
doi:10.1017/S1743921324000127

Detecting coronal mass ejections with machine learning methods

K. Vida[1,2], B. Seli[1,2], T. Szklenár[1,2], L. Kriskovics[1,2], A. Görgei[1,2,3] and Zs. Kővári[1,2]

[1]Konkoly Observatory, HUN-REN Research Centre for Astronomy and Earth Sciences, Budapest, Hungary
email: vidakris@konkoly.hu

[2]CSFK, MTA Centre of Excellence, Budapest, Hungary

[3]Eötvös University, Department of Astronomy, Budapest, Hungary

Abstract. Flares on the Sun are often associated with ejected plasma: these events are known as coronal mass ejections (CMEs). These events, although are studied in detail on the Sun, have only a few dozen known examples on other stars, mainly detected using the Doppler-shifted absorption/emission features in Balmer lines and tedious manual analysis. We present a possibility to find stellar CMEs with the help of high-resolution solar spectra.

Keywords. Sun: coronal mass ejections (CMEs), Stars: activity, methods: data analysis, techniques: spectroscopic, Neural networks

1. Introduction

On our Sun, most of the larger flares are consistently accompanied by a coronal mass ejection (CME). CMEs are important because they release massive amounts of plasma into space, impacting space weather and potentially disrupting technological systems on Earth, such as satellites and power grids. Studying CMEs, understanding their dynamics and – in the future – developing predictive models are crucial to mitigate potential adverse effects on space-based infrastructure and communication systems. Solar CMEs are thoroughly investigated through both observational methods and modeling, as outlined in the reviews by Webb and Howard (2012) and Kilpua et al. (2017), among others. The frequency of solar CMEs ranges from 0.5 to 6 CMEs per day, depending on the phase of the solar activity cycle – these events are observed in detail on the Sun by satellites like the Solar Dynamics Observatory (SDO) or Solar and Heliospheric Observatory (SOHO).

In contrast, the observation of CMEs on other stars poses challenges, and only a handful of such events are known. While flares on these stars can be relatively easily observed through photometry, the observation of CMEs is more challenging. In-situ observations of stellar CMEs will likely remain unfeasible for the foreseeable future.

The primary approach for identifying stellar CMEs involves analyzing their Doppler signature, particularly evident in hydrogen's Balmer lines. The expelled material manifests as a flux increase at the blue wing of the spectral line or, in the case of faster events, may appear as a distinct emission feature, commonly referred to as a "bump" (or absorption if observed against the stellar disk). Odert et al. (2017) extensively discussed the potential observational signatures and constraints.

The majority of reported stellar CME detections in the literature are isolated events discovered serendipitously (e.g., Vida et al. 2016). Until recently, only a limited number

© The Author(s), 2024. Published by Cambridge University Press on behalf of International Astronomical Union.

https://doi.org/10.1017/S1743921324000127 Published online by Cambridge University Press

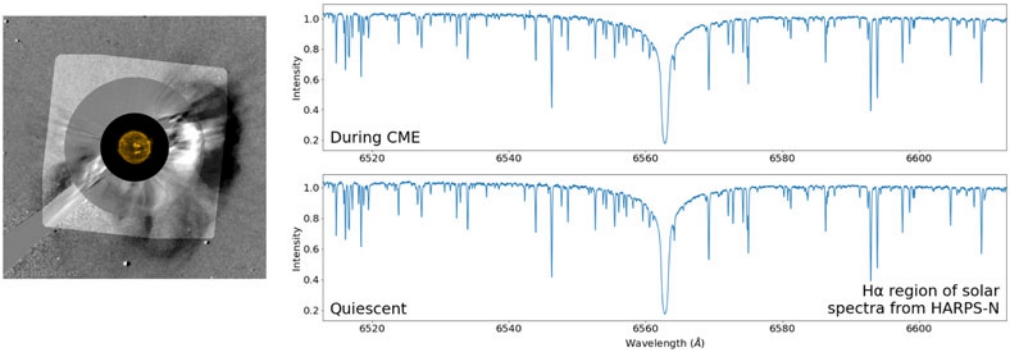


Figure 1. Left: SDO/AIA 171 and SOHO LASCO C2&LASCO C3 observations of the Sun during a massive CME event. Right: the Hα region a solar spectrum during this event compared to the quiescent Sun from HARPS-N observations.

of stellar CME detections existed, and the available data were insufficient for robust statistical analysis. Recent analysis of archive spectral observations has revealed that, although Balmer-line asymmetries are relatively common, the occurrence of successful CME events – those that actually depart from the star – appears to be less frequent than anticipated based on a straightforward scaled-up solar model (Vida et al. 2019; Leitzinger et al. 2020).

2. Using the solar paradigm to understand stellar CMEs

While a recent investigation of solar-like stars (Leitzinger et al. 2020) revealed an absence of CME events on these targets, CMEs on solar-like stars do exist (Namekata et al. 2023). Their rarity could be attributed to either the lower activity levels of these stars, or potential observational biases.

Explanations for phenomena related to stellar magnetism often draw from the solar paradigm, i.e., assuming that activities on stars are analogous to those observed on the Sun. Unlike on stars, on our Sun, we have a detailed knowledge when CMEs occur from direct satellite observations. Could high-resolution, high-cadence Sun-as-a-star observations help us to move forward? Currently, our Sun is entering a period of increased activity in Cycle 25, anticipated to reach its peak around 2025. During this phase, we expect heightened numbers of flares and increased CME activity, providing a unique opportunity to scrutinize the Sun as a star and refine our understanding of stellar CMEs.

The HARPS-N Solar Telescope is a low-cost solar telescope connected to the HARPS-N spectrograph, one of the most precise spectroscopic instruments. The instrument started observing the Sun every clear day with a 5-minute cadence from the Telecopio Nationale Galileo (TNG) facility in La Palma, Spain. The high-resolution, high-cadence Sun-as-a-star data of the first years of observations are publicly available (see Fig. 1). Using these data with the observations from the STEREO/SOHO solar satellites featuring in-situ CME measurements providing the precise timing and physical properties of CMEs we started developing a neural network to better understand these phenomena. If successful, we can use the network to analyze archival spectra possibly revealing events that would not be necessarily found by humans, on other stars.

3. The first experiments

First, we checked the optical spectra at the times of the largest solar CME events. Unfortunately, the spectra show no obvious signs of any ejected material as we observe

https://doi.org/10.1017/S1743921324000127 Published online by Cambridge University Press

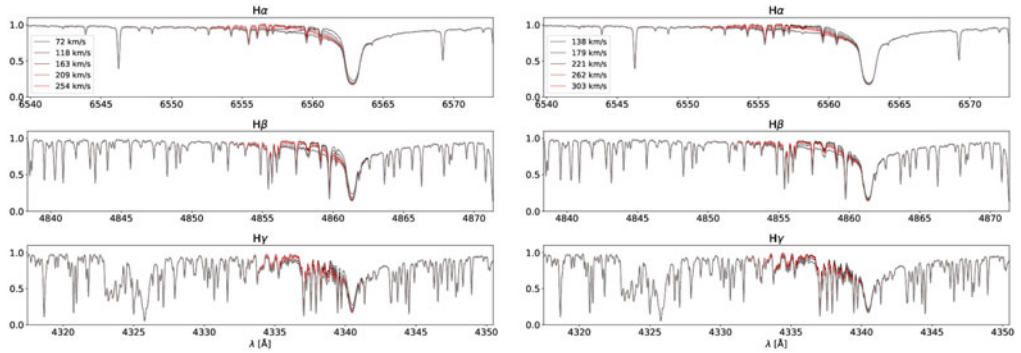


Figure 2. Two examples of injected artificial CME events for training.

these on other stars (Fig. 1) – the most significant changes in the spectra happen in the telluric lines due to airmass and humidity variations.

In our first experiments, we tried various network architectures using Keras (Chollet et al. 2015) using both the full spectral range (≈3900–6900Å) and a smaller region around the Hα line as training data, but these returned random predictions and failed to converge.

To find out what kind of neural network architecture could be optimal for this problem, we created a training set with artificial coronal mass ejections to have obvious signals. These have a typical mass of 5×10^{19}g, 5000 times larger than realistic values. The velocities of the artificial CMEs are realistic, with a maximum velocity of $\approx$400 km s^{-1} (see Fig. 2). This training set contained 50–50% spectra with/without CME signals. After testing different neural network architectures, the ones based on 1D Convolution performed the best. These types of neural networks "scan" through the data with very high accuracy and great performance. After a set of trials, one of the best-performing architecture – to our surprise – was one of the most basic runs, consisting only two `Conv1D` layers with 16 and 32 filters, respectively, followed by two `Dense` layers with 32 neurons each (see Fig. 3). During the preparation of the training, validation, and test datasets, we made sure that there was no overlap between these sets. Our test showed that the accuracy of the deeper 1D Convolution networks was similar but at the expense of speed and performance. Training our neural network was very quick, one epoch took about 2 seconds. For spectral windows of 10 000 data points (a wavelength range of ≈180Å) the whole training finished in less than 2 minutes.

While the best runs achieved over 99% accuracy in both the train and validation set, this happened only on artificial data with events 5000 times larger than realistic values where 50% of the spectra contained events and using only 50 000 data points of the spectrum (the full dataset consists of more than 200 000 points). However, with a network that works consistently on the artificial training set, we can use transfer learning in the future to use that weight file on the original scientific observations. If our experiment is successful, and we can detect coronal mass ejections on the solar spectra, we can possibly find new CME indicators, and determine what magnitude of CMEs are observable on other stars. The new algorithm can be also used to detect CMEs in archive stellar spectroscopic data.

Our results based on the artificial data are promising, but of course, there are many other tests ahead of us. Not just the size of the artificial events need to be reduced significantly, but we must utilize the performance of the 1D Convolution to be able to recognize real solar events as well.

https://doi.org/10.1017/S1743921324000127 Published online by Cambridge University Press

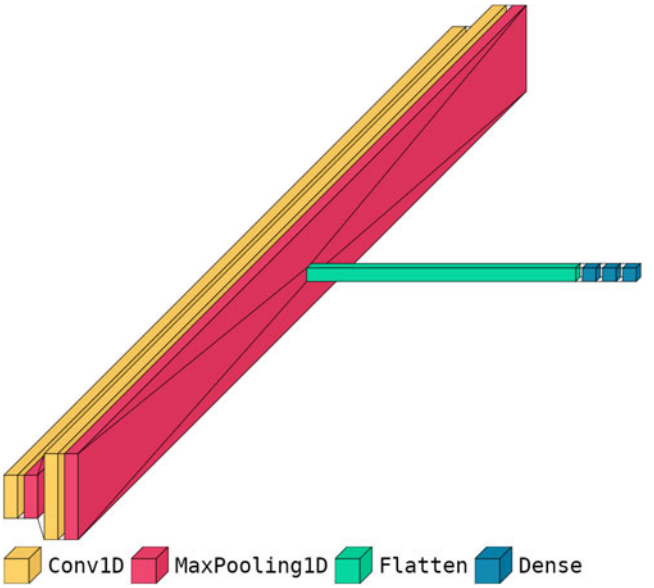


Figure 3. One of the best-performing network architectures for the artificial CME events.

Acknowledgements Authors from Konkoly Observatory acknowledge the Hungarian National Research, Development and Innovation Office (NKFIH) grants OTKA K-131508 and KKP-143986. LK acknowledges the NKFIH grants OTKA PD-134784. LK and KV are Bolyai János research fellows. KV was supported by the Bolyai+ grant ÚNKP-22-5-ELTE-1093, BS was supported by the ÚNKP-22-3 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund.

References

Chollet, F. *et al.* 2015,. Keras. https://keras.io.

Kilpua, E., Koskinen, H. E. J., & Pulkkinen, T. I. 2017, Coronal mass ejections and their sheath regions in interplanetary space. *Living Reviews in Solar Physics*, 14(1), 5.

Leitzinger, M., Odert, P., Greimel, R., Vida, K., Kriskovics, L., Guenther, E. W., Korhonen, H., Koller, F., Hanslmeier, A., Kővári, Z., & Lammer, H. 2020, A census of coronal mass ejections on solar-like stars. *Monthly Notices of the Royal Astronomical Society*, 493(3), 4570–4589.

Namekata, K., Airapetian, V. S., Petit, P., Maehara, H., Ikuta, K., Inoue, S., Notsu, Y., Paudel, R. R., Arzoumanian, Z., Avramova–Boncheva, A. A., Gendreau, K., Jeffers, S. V., Marsden, S., Morin, J., Neiner, C., Vidotto, A. A., & Shibata, K. 2023, Multiwavelength Campaign Observations of a Young Solar-type Star, EK Draconis. I. Discovery of Prominence Eruptions Associated with Superflares. *arXiv e-prints*,, arXiv:2311.07380.

Odert, P., Leitzinger, M., Hanslmeier, A., & Lammer, H. 2017, Stellar coronal mass ejections - I. Estimating occurrence frequencies and mass-loss rates. *Monthly Notices of the Royal Astronomical Society*, 472(1), 876–890.

Vida, K., Kriskovics, L., Oláh, K., Leitzinger, M., Odert, P., Kővári, Z., Korhonen, H., Greimel, R., Robb, R., Csák, B., & Kovács, J. 2016, Investigating magnetic activity in very stable stellar magnetic fields. Long-term photometric and spectroscopic study of the fully convective M4 dwarf V374 Pegasi. *Astronomy & Astrophysics*, 590, A11.

https://doi.org/10.1017/S1743921324000127 Published online by Cambridge University Press

Vida, K., Leitzinger, M., Kriskovics, L., Seli, B., Odert, P., Kovács, O. E., Korhonen, H., & van Driel–Gesztelyi, L. 2019, The quest for stellar coronal mass ejections in late-type stars. I. Investigating Balmer-line asymmetries of single stars in Virtual Observatory data. *Astronomy & Astrophysics*, 623, A49.

Webb, D. F. & Howard, T. A. 2012, Coronal Mass Ejections: Observations. *Living Reviews in Solar Physics*, 9(1), 3.

https://doi.org/10.1017/S1743921324000127 Published online by Cambridge University Press