# Reliability of AI-generated magnetograms from only EUV images

Liu, Jiajia; Wang, Yimin; Huang, Xin; Korsós, Marianna B.; Jiang, Ye; Wang, Yuming; Erdélyi, Robert

*Published in:*
Nature Astronomy

*DOI:*
10.1038/s41550-021-01310-6

*Publication date:*
2021

*Citation for published version (APA):*
Liu, J., Wang, Y., Huang, X., Korsós, M. B., Jiang, Y., Wang, Y., & Erdélyi, R. (2021). Reliability of AI-generated magnetograms from only EUV images. *Nature Astronomy*, *5*(2), 108-110. https://doi.org/10.1038/s41550-021-01310-6

# Reliability of AI-generated magnetograms from only EUV images

Jiajia Liu[1,2,*], Yimin Wang[1,3,*], Xin Huang[4], Marianna Korsos[5], Ye Jiang[6], Yuming Wang[7], Robert Erdélyi[1,8,9,+]

[1] *Solar Physics and Space Plasma Research Centre (SP2RC), School of Mathematics and Statistics, The University of Sheffield, Sheffield, S3 7RH, UK*
[2] *Astrophysics Research Centre, School of Mathematics and Physics, Queen's University, Belfast, BT7 1NN*
[3] *School of Electrical Engineering, University of Jinan, Jinan 250022, China*
[4] *Key Laboratory of Solar Activity, National Astronomical Observatories of Chinese Academy of Sciences, Beijing, China*
[5] *Department of Physics, Aberystwyth University, Penglais, Aberystwyth SY23 3BZ, UK*
[6] *Department of Computer Science, The University of Sheffield, Sheffield S1 4DP, UK*
[7] *CAS Key Laboratory of Geospace Environment, Department of Geophysics and Planetary Sciences, University of Science and Technology of China, Hefei, Anhui 230026, China*
[8] *Department of Astronomy, Eötvös Loránd University, Budapest, Pázmány P. sétány 1/A, H-1117, Hungary*
[9] *Gyula Bay Zoltán Solar Observatory (GSO), Hungarian Solar Physics Foundation (HSPF), Petőfi tér 3., Gyula, H-5700, Hungary*
[*] *These authors contribute equally to this Matters Arising*
[+] *Corresponding Author, email: robertus@sheffield.ac.uk*

Kim et al. 2019 [ref.[1]] (KPL19) proposed an artificial intelligence (AI) model to predict the photospheric magnetograms of the Sun using EUV observations as the only inputs, and concluded that "the model is reliable if the farside active regions conform to Hale's law, as long as the slight overestimation of their total flux and a possible slight difference in their tilt angle are considered". **In this Matters Arising, we present a detailed sensitivity study of the AI algorithm by KPL19.** Despite identifying issues in the inappropriate data preparation process and the possibility of data leakage in KPL19, we also found the physics basis of this idea is problematic. We detail our concerns and analysis below, as well as in the Supplementary Material.

Recently, a number of novel machine learning (ML) and/or AI techniques have been introduced to and used for a variety of purposes in the area of solar physics and space weather forecasting[2]. Using direct images or extracted features from photospheric magnetic field only or combined with solar EUV observations, a number of efforts have been made to predict the occurrence and/or onset time of solar flares employing statistical and/or ML methods[3–9]. In addition, employing algorithms including the support vector machines and convolutional neural networks, the mean absolute error in predicting the arrival time of Corona Mass Ejections (CMEs) has been remarkably reduced to as low as ∼6 hours[10,11], yielding further kudos to using ML/AI in space weather forecasting.

In order to study solar activity and predict space weather, KPL19 employed an AI technique - conditional generative adversal networks (cGANs) - to predict the solar photospheric magnetograms. They fed cGANs with full-disk EUV and photospheric magnetic field observations from the Atmospheric Imaging Assembly[12] (AIA) 304 Å passband and Helioseismic and Magnetic Imager[13] (HMI) onboard the Solar Dynamics Obsservatory (SDO). A model was then built with the SDO/AIA 304 Å images as the input to generate simultaneous SDO/HMI photospheric magnetograms. KPL19 then evaluated the model and found promising correlation coeffecients (CCs) between the total unsigned magnetic flux (TUMF) of the generated and observed magnetograms. KPL19 claimed a conclusion that using their method, the photospheric magnetograms could be well forecasted to greatly improve our current knowledge of the farside active regions. However, there are several vital practical, as well as theoretical, issues in KPL19, detailed as below that to some extent mitigate the success of KPL19.

While pre-processing the SDO/HMI photospheric LOS magnetograms (Supplementary Data and Method) , KPL19 set the upper and lower saturation limits of the magnetic field strength as $\pm$100 G. However, these limits are found to be problematic, especially for active regions, considering one of the main purposes of KPL19 was to predict the farside active regions of the Sun. The average absolute magnetic field (AMF) strength of all 3936 active regions detected from the original observations in the testing set of the data has been found to be 208$\pm$54 G. Only 0.77% of all the active regions reveal an average AMF strength less than 100 G, among which only three active regions has an average AMF strength less than 90 G.

The slopes of black dots in Supplementary Figures 1(a) and (b) suggest that the rescaled magnetograms with saturation limits of $\pm$100 G give on average 0.45 and 0.67 of the original TUMF and net magnetic flux (NMF). In addition, the degree of the scattering of the dots yields the R2 scores (Supplementary Eq. 1) of -0.07 and 0.77, respectively. The percentage of instances where the rescaled magnetograms yield the opposite signs of the NMF to the original observations is about 19.4%. The above evaluation suggests, again, that the generated magnetograms could still be significantly different from the original observations even if the model were perfect, if saturation limits of $\pm$100 G are used when preparing the dataset. For a comparison, blue dots and lines in Supplementary Figures 1(a) and (b) show the corresponding results for saturation limits of $\pm$625 G. We note that setting inappropriate large saturation limits might also be problematic as that could introduce too much noise that might then severely impact the ability of the generative models to capture the prior distribution. Thus we encourage researchers to evaluate carefully before choosing the saturation limits for normalization purposes. In addition, KPL19 have used observations in September and October in each year as the testing set and the rest as the training set, risking a possibility of a data leakage considering that the Sun rotates at a period of $\sim$27.3 days (see Supplementary Potential Data Leakage).

Despite the above practical issues, we foresee the model to be not successful based on the theoretical fact that EUV observations of the chromosphere and corona do not provide any information about the photospheric magnetic field polarities. We trained the neural network to build an optimistic AI model using the code provided by KPL19, fed with the same dataset preprocessed with the same parameter settings (Supplementary Data and Method). Figures 1 shows a comparison of the observations (panel b) and the generated magnetograms by the model we built (panel c, first run), which could be directly compared with the one generated from the model KPL19 built (Figure 1 therein). Overall, the AI models could successfully identify the active regions presented in the original observation. However, the shape of the active regions and the distribution of the positive and negative polarities are poorly reconstructed (see the two active regions enclosed in the green rectangle and blue square boxes in panels b to c). Further, we ran the same procedure twice and built two new models (best models built at 128 and 212 epochs for the second and third runs respectively). The generated magnetogram from one of these two models (the second run) is shown in Figure 1 (d). Significant differences can be seen between the generated active regions in panels (c) and (d), which should not happen to a robust and reliable model. **We shall note that though we have used the same architecture, hyper-parameters and training data in all our three models, they are different models and are exactly the same as that in KPL19, because they all have different trained parameters due to factors including different weight initialisation, the stochastic character of the optimizer and specificities of the loss hypersurface, etc. All the further analysis shown below and in the Supplementary Material is based on the first model we built.** Detailed evaluations of the correlation between the generated (from the first run) and rescaled (with saturation limits of $\pm 100$ G) magnetograms (see Supplementary Full-disk Parameters) yield that the proposed AI model is only successful in reproducing the TUMF of the global magnetic field, but fails in reconstructing the relative relations between the positive and negative polarities, suggested by the low pixel-to-pixel cross correlation and the low correlation between the NMFs. Integrating the TUMF information and area of the farside active regions into dedicated models together with the frontside magnetograms could in some cases improve the performance in predicting the in-situ solar wind speed[14].

We employ an automated detection system[15,16] to automatically extract active regions and their parameters from the rescaled and the generated full-disk magnetograms. Supplementary Figure 2 depicts a direct comparison between the active regions detected from the rescaled and generated magnetograms at 00 UT on 28 September 2011 as an example. One can clearly observes significant differences between the sizes, shapes and polarity inversion lines (PILs) of the active regions in the northern hemisphere, especially in the two big active regions with one close to the disk center and the other on the right. In addition, there are missed active regions in the northern hemisphere and one extra active region in the southern hemisphere of the generated magnetogram compared to the rescaled one.

Statistical analysis (Supplementary Active Region Parameters and Supplementary Figure 3) reveals that, on average, the model only reproduces less than half of the active regions in each of the observations. The centres of the detected active regions in the AI-generated magnetograms are on average $\sim$1.3° away in heliographic coordinates from the real ones. Detailed evaluations on a number of key parameters of the detected active regions (Supplementary Active Region Parameters and Supplementary Figure 4) suggest that the AI model performs fairly well in predicting the areas of the active regions, but poorly in reproducing the NMF of the active regions, and the total number, the length and the average magnetic gradient across PILs. **To conclude, our sensitivity study suggests that the AI model proposed by KPL19 shall be rather far from providing scientifically reliable magnetograms.**

Figure 1: **Example of the observations and the AI-generated magnetograms.** The four panels are the SDO/AIA 304 Å observation (a), the SDO/HMI photospheric magnetogram (b), **the AI-generated magnetogram of the Sun by two of our independent verification processes with (c) for the first run and (d) for the second run, respectively, at 12 UT on September 5 2017.**

## Data Availability

SDO/AIA and SDO/HMI data are publicly available from NASA's SDO website (`https://sdo.gsfc.nasa.gov/data/`). Information of the dataset we have used is available at `https://github.com/yiminking/pix2pix_EUV2HMI_datasets`.

## Code Availability

Codes for the AI models built in this paper are accessible from Kim et al. 2019 (`https://github.com/tykimos/SolarMagGAN`). Codes used for the active region detection are available upon requests.

## References


1. Kim, T. *et al.* Solar farside magnetograms from deep learning analysis of STEREO/EUVI data. *Nat. Astron.* **3**, 397–400 (2019).

2. Camporeale, E. The Challenge of Machine Learning in Space Weather: Nowcasting and Forecasting. *Space Weather* **17**, 1166–1207 (2019).

3. Wang, H. N., Cui, Y. M., Li, R., Zhang, L. Y. & Han, H. Solar flare forecasting model supported with artificial neural network techniques. *Adv. in Space Res.* **42**, 1464–1468 (2008).

4. Bobra, M. G. & Couvidat, S. Solar Flare Prediction Using Sdo /Hmi Vector Magnetic Field Data With a Machine-Learning Algorithm. *Astrophys. J.* **798**, 135–145 (2015).

5. Korsós, M. B., Ludmány, A., Erdélyi, R. & Baranyi, T. On Flare Predictability Based on Sunspot Group Evolution. *Astrophys. J. Lett.* **802**, L21–L26 (2015).

6. Nishizuka, N. *et al.* Solar Flare Prediction Model with Three Machine-learning Algorithms using Ultraviolet Brightening and Vector Magnetograms. *Astrophys. J.* **835**, 156–165 (2017).

7. Liu, C., Deng, N., Wang, J. T. L. & Wang, H. Predicting Solar Flares Using SDO/HMI Vector Magnetic Data Products and the Random Forest Algorithm. *Astrophys. J.* **843**, 104–117 (2017).

8. Florios, K. *et al.* Forecasting solar flares using magnetogram-based predictors and machine learning. *Sol. Phys.* **293**, 28–69 (2018).

9. Huang, X. *et al.* Deep Learning Based Solar Flare Forecasting Model . I . Results for Line-of-sight Magnetograms. *Astrophys. J.* **856**, 7–17 (2018).

10. Liu, J., Ye, Y., Shen, C., Wang, Y. & Erdélyi, R. A New Tool for CME Arrival Time Prediction using Machine Learning Algorithms: CAT-PUMA. *Astrophys. J.* **855**, 109–118 (2018).

11. Wang, Y., Liu, J., Jiang, Y. & Erdélyi, R. CME Arrival Time Prediction Using Convolutional Neural Network. *Astrophys. J.* **881**, 15–24 (2019).

12. Lemen, J. R. *et al.* The Atmospheric Imaging Assembly (AIA) on the Solar Dynamics Observatory (SDO). *Sol. Phys.* **275**, 17–40 (2012).

13. Scherrer, P. *et al.* The Helioseismic and Magnetic Imager (HMI) Investigation for the Solar Dynamics Observatory (SDO). *Sol. Phys.* **275**, 207–227 (2012).

14. Arge, C. N. *et al.* Modeling the corona and solar wind using ADAPT maps that include far-side observations. In Zank, G. P. *et al.* (eds.) *Solar Wind 13*, vol. 1539 of *American Inst. of Phys. Conf. Ser.*, 11–14 (2013).

15. Wang, Y. & Zhang, J. A Statistical Study of Solar Active Regions That Produce Extremely Fast Coronal Mass Ejections. *Astrphys. J.* **680**, 1516–1522 (2008).

16. Zhang, J., Wang, Y. & Liu, Y. Statistical Properties of Solar Active Regions Obtained from an Automatic Detection System and the Computational Biases. *Astrophys. J. Lett.* **723**, 1006–1018 (2010).


## Acknowledgements


We acknowledge the use of the data from the Solar Dynamics Observatory (SDO). SDO is the first mission for the NASA's Living With a Star (LWS) program. JL and RE are grateful to STFC (UK, grant number ST/M000826/1) and EU H2020 (SOLARNET grant nr 158538). JL also acknowledges the support from STFC under grant No. ST/P000304/1 and from the Leverhulme Trust via grant RPG-2019-371. RE also acknowledges the support from the Chinese Academy of Sciences President's International Fellowship Initiative (PIFI, grant number 2019VMA0052) and The Royal

Society (grant nr IE161153). Yimin Wang thanks for the warm hospitality and support received as an MSRC Visiting Research Fellow while carrying out this research at the Solar Physics and Space Plasma Research Centre (SP2RC), School of Mathematics and Statistics (SoMaS), The University of Sheffield. MBK is grateful to the STFC under grant No ST/S000518/1.

## Author Contributions

JL conducted the data preparation, result analysis and drafted the manuscript. Yimin Wang performed the machine learning approach with the help of YJ. RE, XH and JL recognised the core problems. RE led the overall research. All authors joined the discussion and participated in the interpretation of the results. All authors reviewed the manuscript.

## Competing Interests

The authors declare no competing interests.

## Additional Information

**Supplementary information** is linked to the online version of this Matters Arising at https://www.nature.com/natastron/.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Correspondence and requests for materials** should be addressed to R.E.