Forecasting SYM-H Index: A Comparison Between Long Short-Term Memory and Convolutional Neural Networks

F. Siciliano[1,3], G. Consolini[2], F. Giannattasio[3]

Presentation for the EGU General Assembly 2022

1: Department of Computer, Control and Management Engineering "Antonio Ruberti", Sapienza University of Rome, Rome, Italy

2: INAF - Istituto di Astrofisica e Planetologia Spaziali, Rome, Italy

3: Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy

DOI 10.1029/2020SW002589
Space Weather - 10/02/2021

Goals

1. Attempt a forecasting of the Earth's magnetosphere response to the solar activity, as monitored by SYM-H index, at least one hour in advance (the time strictly necessary to issue an alert) and with a high temporal resolution (5 minutes).
2. Attempt a modeling/forecasting of the SYM-H index without the aid of solar wind density and speed, which in many situations leads to the non-reliability of prediction models due to missing data.
3. Compare the performance of two conceptually different neural networks: the Long Short-Term Memory (LSTM) and the Convolutional Neural Network (CNN).
4. Understand how far back in time we need to look to get the best performance.

Dataset

42 geomagnetic storms that occurred between 1998 and 2018

- Interplanetary Magnetic Field recorded at L1 from ACE satellite (in GSM coordinates)
- SYM-H Index

➢ 5-min averages

Artificial Neural Networks

Long Short-Term Memory

- A type of Recurrent Neural Network → designed for sequence prediction problems
- Receives data one timestep at a time
- Can distinguish between long-term and short-term memory

Convolutional Neural Network

- Commonly applied to image analysis
- Receives data all at once
- Does not know which inputs are closest in time to the output
- Flexible: conceived to process data stored in the form of multiple arrays

Artificial Neural Networks

- Input (from time t_0 - *lookback** to t_0)
  - Interplanetary Magnetic Field recorded at L1 from ACE satellite (in GSM coordinates)
    - Bz component
    - squared values of IMF magnitude B
    - squared values of IMF By component
  - SYM-H Index (optional)
- Output (at time t_0 + 60 min)
  - SYM-H Index

➢ **Lookback* = number of time instants in the past

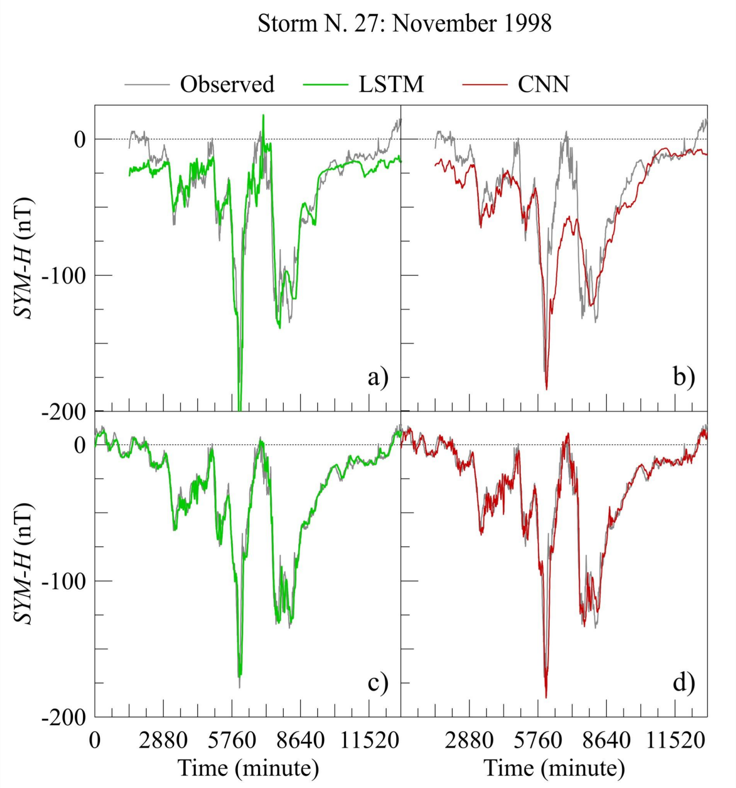
Results

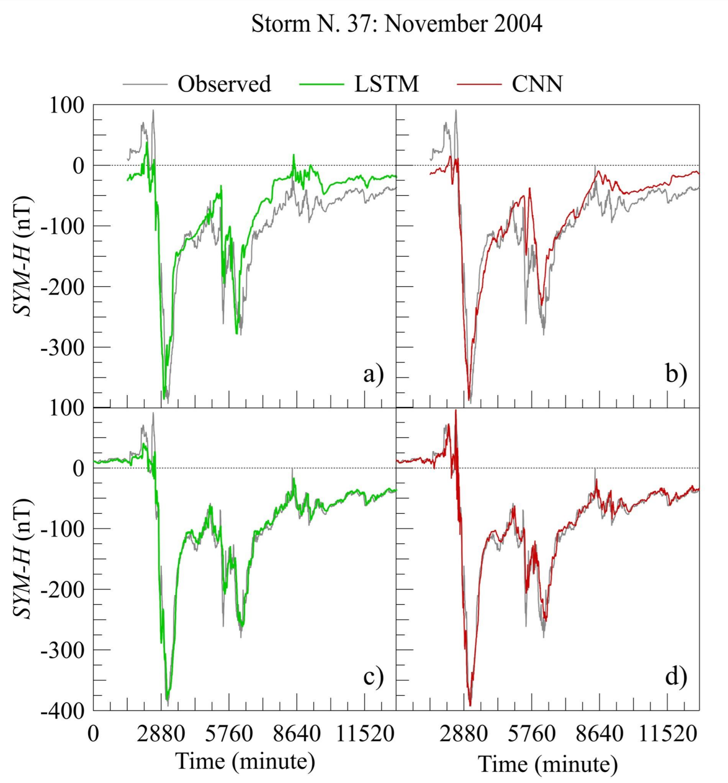
First row

Without SYM-H in input

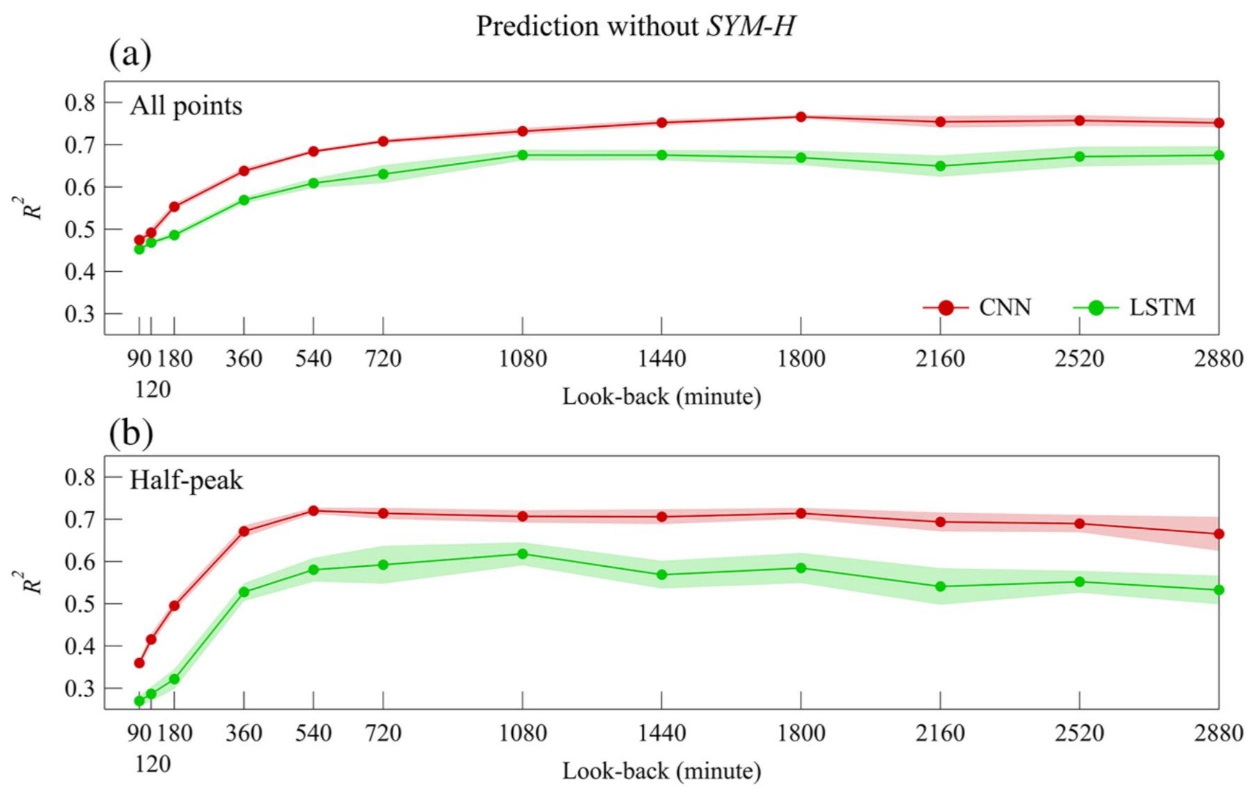
Second row

With SYM-H in input



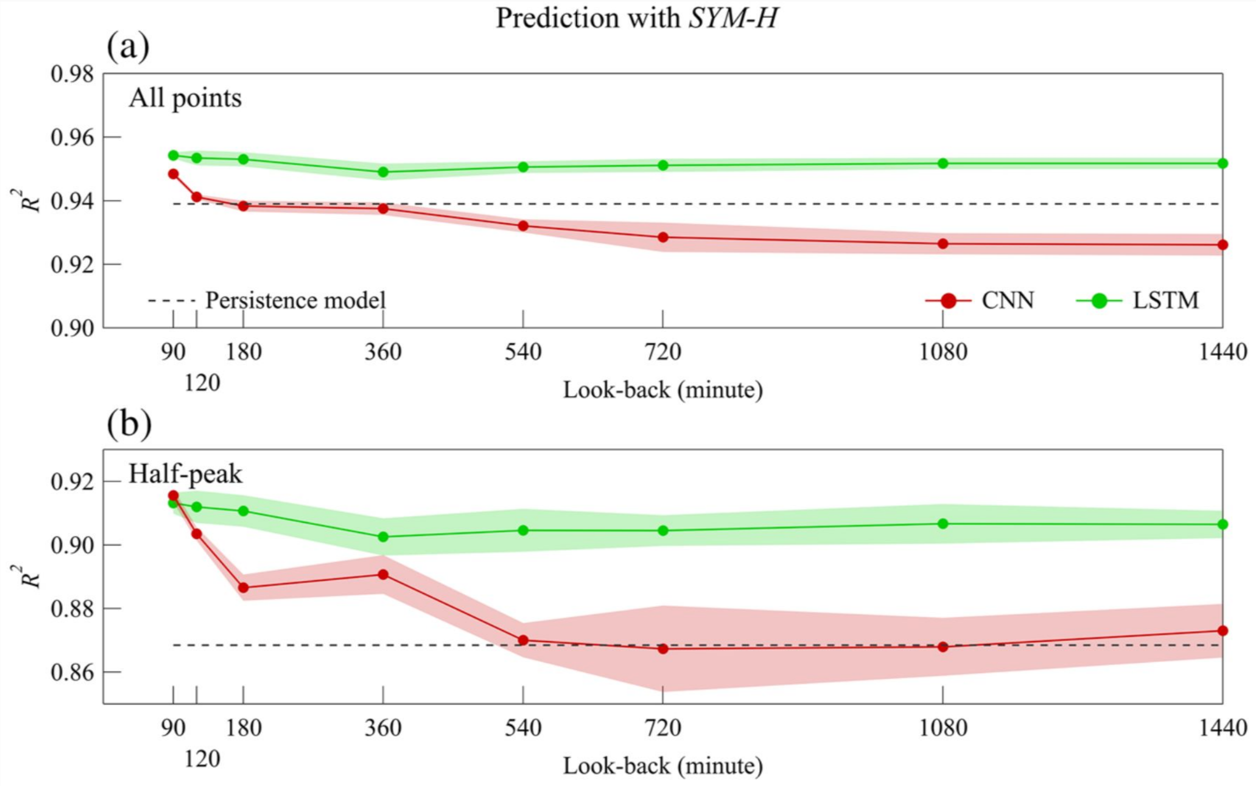


Lookback



- When not using SYM-H as an input, many hours back are needed to infer the state of the magnetosphere
- CNN clearly outperforms LSTM
- To predict all phases well, at least 1440 minutes (24h) are needed
- To predict the storm phase well, 540 minutes (9h) seems sufficient
  - ➢ Half-peak: metrics calculated around the peak of the storm (in particular, when SYMH takes a value equal to half the value of the peak)

Lookback



- When using SYM-H as input, performance deteriorates as lookback increases. This means that the value of SYM-H already contains enough information about the current situation of the magnetosphere.
- LSTM slightly outperforms CNN

➢ *Half-peak: metrics calculated around the peak of the storm (in particular, when SYM-H takes a value equal to half the value of the peak)

Goals vs Results

1. Attempt a forecasting of the Earth's magnetosphere response to the solar activity, as monitored by SYM-H index, at least one hour in advance (the time strictly necessary to issue an alert) and with a high temporal resolution (5 minutes).
2. Attempt a modeling/forecasting of the SYM-H index without the aid of solar wind density and speed, which in many situations leads to the non-reliability of prediction models due to missing data.

1. Our results suggest that both LSTM and CNNs are able to well forecast a high-resolution index like SYM-H one hour in advance
2. Since magnetospheric dynamics is characterized by both directly driven processes and externally triggered internal ones, our idea is that IMF data do not contain the necessary information to predict short timescale fluctuations of geomagnetic indices, and additional data, such as solar wind data, are needed.

Goals vs Results

3. Compare the performance of two conceptually different neural networks: the Long Short-Term Memory (LSTM) and the Convolutional Neural Network (CNN).
4. Understand how far back in time we need to look to get the best performance.

3. LSTM performs better than CNN when using SYM-H as input, probably by exploiting the temporal dependency. CNN on the other hand performs better than LSTM when SYM-H is not used as input, probably by being able to capture long-term relationships better. CNN therefore seems worthy of further investigation
4. When previous SYM-H values are not available, 540 minutes before the storm appears to be the time strictly necessary to obtain a good prediction of the storm phase, whereas in the quiet phase correlations can also be found with data from 24 hours before.

Thank for your attention

For further questions contact me at:

siciliano@diag.uniroma1.it