SYM-H Prediction Notes

Daniel Iong

At time t, let

- $y(t)$ denote SYM-H
- $X(t)$ denote the solar wind parameters that are available to us (i.e. what is in the dataset)
- $\tilde{X}(t)$ denote the actual solar wind parameters.

1 Propagation time/Time shift

- Let $\alpha(t) = D/V_x(t)$, where $D \approx 1.5 \times 10^6$ km.
- Our initial goal is to fit a model with the following form:

$$y(t + \alpha(t)) = f\Big(y(t), \ldots, y(t - L\Delta t), \tilde{X}(t + \alpha(t)), \ldots, \tilde{X}(t + \alpha(t) - p\Delta t)\Big)$$

2 Data processing Steps

1. Time resolution = 5 minutes (Δt)
2. Solar wind parameters used: B_y, B_z, V_x, Density
3. Deleted storms with too much missing (15, 69, 124)
4. Split data into training and testing.
    - For now, I chose storm 27 as test storm. When I do tune the hyper-parameters, I will take out a few storms as testing and tune with the rest.

Let $I(t)$, $O(t)$ denote the inputs and output at time t.

2.1 Features processing

1. Transformed each feature to be in the same range (For now, I'm following Cai et. al and using (-0.8,0.8)).
2. $I(t) = \{y(t), \ldots, y(t - L\Delta t), \tilde{X}(t + D/V_x(t)), \ldots, \tilde{X}(t + D/V_x(t) - p\Delta t)\}$

2.2 Target processing

1. $O(t) = y\big(t + D/V_x(t)\big)$