Modelling of time series of external-driven events with echo state network and its application to substorm activity analysis

S. Nakano (The Institute of Statistical Mathematics)
R. Kataoka (Okinawa Institute of Science and Technology)
M. Nosé (Nagoya City University)

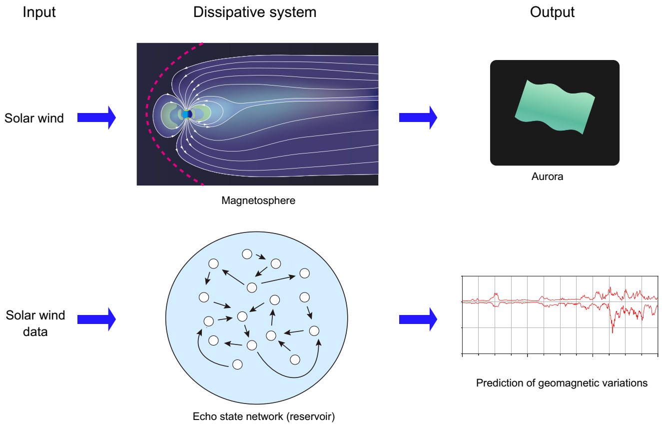
1 May 2025

Predicting auroral activity using a machine-learning model

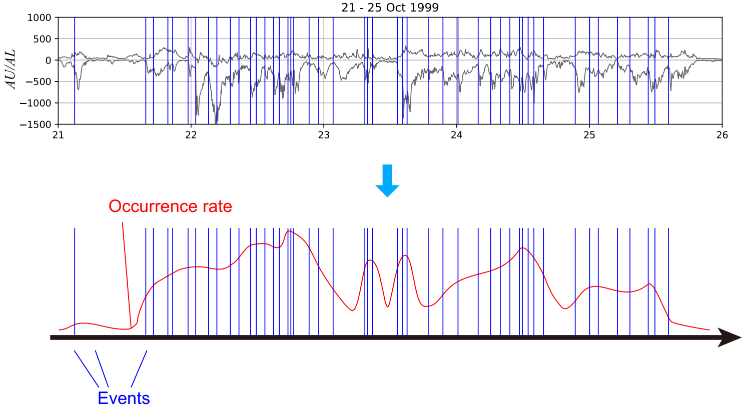
Auroral phenomena result from physical processes in the magnetosphere, which is driven by the solar wind.

We are developing a model to predict the auroral activity under given solar wind conditions.

We employ a machine-learning model called the echo state network (ESN) to approximate the response of the magnetosphere to the solar wind.



Event analysis



- We represent stochastic occurrence of Pi2 events with a non-stationary Poisson process.
- We train an echo state network so that the output well approximates the substorm occurrence rate under given solar wind conditions.

Nonstationary Poisson process

We denote $\nu(t|\boldsymbol{\beta})dt$ as the probability of the occurrence of an event within a small time interval dt, where $\boldsymbol{\beta}$ is a parameter determining the shape of the function ν. Given a sequence of event occurrence times $\tau_{1:N} = \{\tau_1, \tau_2, \ldots, \tau_N\}$, the likelihood of the parameter $\boldsymbol{\beta}$ is written as follows:

$$L(\boldsymbol{\beta}) = p(\tau_{1:N}|\boldsymbol{\beta}) = \prod_{i=1}^{N} \nu(\tau_i|\boldsymbol{\beta}) \exp\left[-\int_{t_0}^{t_K} \nu(t|\boldsymbol{\beta})\, dt\right]. \tag{4}$$

where t denotes time and N is the number of events. The log-likelihood thus becomes

$$\log L(\boldsymbol{\beta}) = \sum_{i=1}^{N} \log \nu(\tau_i|\boldsymbol{\beta}) - \int_{t_0}^{t_K} \nu(t|\boldsymbol{\beta})\, dt. \tag{5}$$

Discretize Eq. (5) in time, we obtain

$$\log L(\boldsymbol{\beta}) = \sum_{i=1}^{N} \log \nu(\tau_i|\boldsymbol{\beta}) - \sum_{k=0}^{K-1} \nu(t_k|\boldsymbol{\beta})\Delta t, \tag{6}$$

where $t_k = t_0 + k\Delta t$.

Use of ESN

We represent the function ν as follows

$$\nu(t_k|\boldsymbol{\beta}) = \exp(\boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{x}_k), \quad (t_k \le t < t_{k+1}), \tag{7}$$

where $\boldsymbol{x}_k$ is the state vector of the ESN.

The value of $\boldsymbol{\beta}$ is obtained with a Bayesian approach. We take the prior distribution of $\boldsymbol{\beta}$ as a Gaussian distribution as

$$p(\boldsymbol{\beta}) = \frac{1}{\sqrt{(2\pi\sigma^2)^m}} \exp\left[-\frac{\boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{\beta}}{2\sigma^2}\right]. \tag{8}$$

where m denotes the dimension of $\boldsymbol{x}_k$ and is 1000.

We estimate $\boldsymbol{\beta}$ such that the following posterior probability density is maximised:

$$p(\boldsymbol{\beta}|\tau_{1:N}) = \frac{p(\tau_{1:N}|\boldsymbol{\beta})\, p(\boldsymbol{\beta})}{p(\tau_{1:N})}. \tag{9}$$

Log posterior density

As $p(\tau_{1:N})$ does not depend on $\boldsymbol{\beta}$, we can obtain the optimal $\boldsymbol{\beta}$ by maximising the following objective function

$$\begin{aligned} J &= \log\left[p(\tau_{1:N}|\boldsymbol{\beta})\,p(\boldsymbol{\beta})\right] \\ &= \log L(\boldsymbol{\beta}) + \log p(\boldsymbol{\beta}) \\ &= \sum_{i=1}^{N} \log \nu(\tau_i|\boldsymbol{\beta}) - \sum_{k=0}^{K-1} \nu(t_k|\boldsymbol{\beta})\Delta t - \frac{\boldsymbol{\beta}^\mathsf{T}\boldsymbol{\beta}}{2\sigma^2} - \frac{m}{2}\log\left(2\pi\sigma^2\right) \\ &= \sum_{i=1}^{N} \boldsymbol{\beta}^\mathsf{T}\boldsymbol{x}_{k_i} - \sum_{k=0}^{K-1} \exp(\boldsymbol{\beta}^\mathsf{T}\boldsymbol{x}_k)\Delta t - \frac{\boldsymbol{\beta}^\mathsf{T}\boldsymbol{\beta}}{2\sigma^2} - \frac{m}{2}\log\left(2\pi\sigma^2\right). \end{aligned} \tag{10}$$

using the Newton–Raphson method.

Analysis of substorm activity

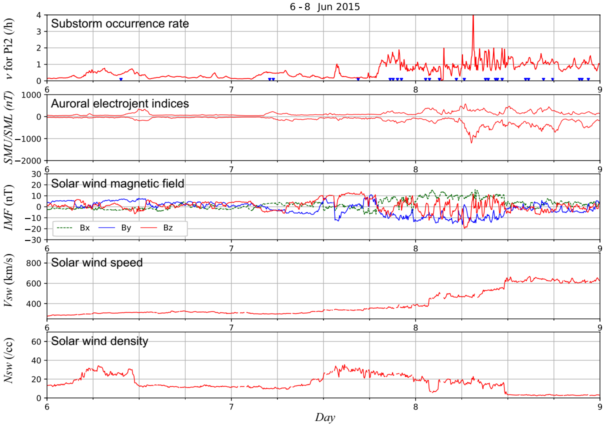
Output:

- Occurrence frequency of substorms defined by Wp index.

Inputs:

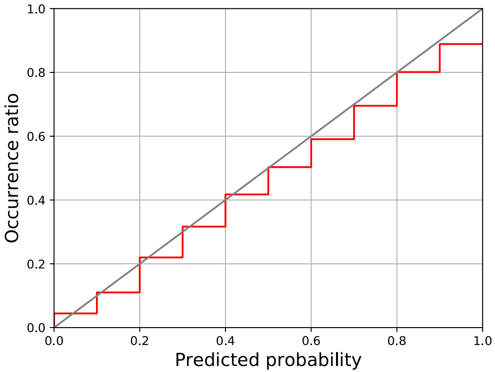
- Solar wind magnetic field ($B_{x,k}$, $B_{y,k}$, $B_{z,k}$), speed V_k, density N_k, and temperature T_k,
- $\cos\left(\frac{H}{12}\right)$ and $\sin\left(\frac{H}{12}\right)$, where H is the universal time (in hour),
- $\cos\left(\frac{D}{365.24}\right)$ and $\sin\left(\frac{D}{365.24}\right)$, where d is the day from the end of 2000,

The data from 2005 to 2014 were used for training the ESN.



(Nakano et al., Ann. Geophys., 2023)

Evaluation



Actual occurrence ratio with respect to the predicted probability for Pi2 substorms from 2015 to 2018 (red). The gray line shows the case where the predicted probability is equal to the actual occurrence ratio.

The ESN is likely to well represent the occurrence frequency.

Summary

- An ESN is a simple recurrent neural network which can be used for representing nonlinear responses to input signals.
- It can be applied for modelling a system controlled by external forcings such as the magnetosphere-ionosphere system.
- It can also be plugged in statistical models for time series data such as a non-stationary Possion model.
- It is therefore useful for analyzing event time series controlled by external forcing.

See our paper for detail (doi: 10.5194/angeo-41-529-2023).