Short term prediction of geomagnetic secular variation with an echo state network

Shin'ya Nakano[1] and Hiroaki Toh[2]

[1]The Institute of Statistical Mathematics
[2]Graduate School of Science, Kyoto University

April 30, 2023

Abstract

A technique for predicting the secular variation (SV) of the geomagnetic field based on the echo state network (ESN) model is proposed. SV is controlled by the geodynamo process in the Earth's outer core. However, it is difficult to model the realistic nonlinear behaviors of the geodynamo due mainly to the very small Ekman number of the actual outer core. This study employs the ESN to represent the temporal evolution of the geomagnetic field on the Earth's surface. The hindcast results of SV demonstrate that the ESN enables us to predict SV for several years with satisfactory accuracy. In particular, the nonlinear behaviors of SV is accurately predicted for the case where accurate geomagnetic data with a 1-year time resolution are available. It is found that an increase in the number of training data does not necessarily improve prediction accuracy. The results suggest that the information on the latest temporal variations is important for the short-term prediciton by the ESN valid for, say 5 years.

Short term prediction of geomagnetic secular variation with an echo state network

S. Nakano[1,2,3]*, **H. Toh**[4]

[1]The Institute of Statistical Mathematics
[2]The Graduate University for Advanced Studies, SOKENDAI
[3]Center for Data Assimilation Research and Applications, Joint Support Center for Data Science
[4]Graduate School of Science, Kyoto University

Key Points:

- A technique for predicting the secular variation of the geomagnetic field based on the echo state network model is proposed.
- The hindcast results show that the secular variation is predicted with satisfactory accuracy.
- It is also sugested that the information on the latest temporal variations is important for predicting the secular variation.

*10–3 Midori-cho, Tachikawa, Tokyo, Japan

Corresponding author: S. Nakano, shiny@ism.ac.jp

Abstract

A technique for predicting the secular variation (SV) of the geomagnetic field based on the echo state network (ESN) model is proposed. SV is controlled by the geodynamo process in the Earth's outer core. However, it is difficult to model the realistic nonlinear behaviors of the geodynamo due mainly to the very small Ekman number of the actual outer core. This study employs the ESN to represent the temporal evolution of the geomagnetic field on the Earth's surface. The hindcast results of SV demonstrate that the ESN enables us to predict SV for several years with satisfactory accuracy. In particular, the nonlinear behaviors of SV is accurately predicted for the case where accurate geomagnetic data with a 1-year time resolution are available. It is found that an increase in the number of training data does not necessarily improve prediction accuracy. The results suggest that the information on the latest temporal variations is important for the short-term prediciton by the ESN valid for, say 5 years.

1 Introduction

The geomagnetic field is gradually and incessantly changing. This change is referred to as secular variation (SV). The magnitude of SV can exceed 10 nT per year, which is comparable to or larger than that of ionospheric and magnetospheric origin. Hence, it is important to predict SV on a time scale of several years. The International Geomagnetic Reference Field (IGRF) model (Alken, Thébault, Beggan, Amit, et al., 2021) includes an SV model for prediction of next 5 years. Since SV sometimes shows nonlinear behaviors such as geomagnetic jerks (e.g., Courtillot & Mouël, 1984; Alexandrescu et al., 1996), its accurate prediction is difficult. Accordingly, various approaches were employed in the 14 SV candidate models which contributed to the latest IGRF model (Alken, Thébault, Beggan, Aubert, et al., 2021 and references therein). Since the geomagnetic main field is thought to be driven by a dynamo process in the Earth's outer core, some candidate models assimilated ground and satellite data into numerical geodynamo models. (e.g, Minami et al., 2020; Fournier et al., 2021). Data assimilation is a straightforward approach to consider the nonlinear dynamics of the outer core. However, a typical geodynamo model represents the state of the geodynamo with millions of variables, whereas the IGRF model represents the geomagnetic main field on the Earth's surface using about 200 parameters. The computational cost of data assimilation using a geodynamo model is thus excessive for predicting the parameters for the geomagnetic field model.

Machine learning approaches for modelling nonlinear systems have recently emerged. The purpose of this study is to explore a machine-learning-based method for predicting SV efficiently. Here, we employ an echo state network (ESN) model (Jaeger & Haas, 2004) for this purpose. The ESN is a kind of reservoir computing framework and it is a recurrent neural network in which the connections and weights between hidden state variables are randomly set and fixed. The ESN is therefore trained by optimizing the weights of only the output layer. Compared to the latest deep neural network models, the degree of freedom of the ESN is small because the weights of only the output layer are made variable. However, for the problem considered here, we have observation data for recent for only the most recent 100 to 1000 years, whereas the convection time scale of the outer core is tens of thousands of years. The available observations are thus insufficient for optimizing the large number of parameters for a deep neural network. Even with its small degree of freedom, the ESN shows satisfactory performance in various geophysical applications (e.g., Kataoka & Nakano, 2021; Nakano & Kataoka, 2022; Walleshauser & Bollt, 2022). Therefore, we apply the ESN for modelling the temporal evolution of the geomagnetic field in the hope of handling the nonlinear behaviors of SV including the geomagnetic jerks.

2 Method

Following many models of the Earth's magnetic field, including the IGRF model, we represent the magnetic field $\boldsymbol{B}$ with a scalar potential V as

$$\boldsymbol{B} = -\nabla V. \tag{1}$$

The potential V is expanded into spherical harmonics:

$$V(r, \theta, \phi, t) = a \sum_{n=1}^{N} \sum_{m=0}^{n} \left(\frac{a}{r}\right)^{n+1} [g_n^m(t) \cos m\phi + h_n^m(t) \sin m\phi]\, P_n^m(\cos\theta) \tag{2}$$

where a denotes the Earth's mean radius. The SV of the geomagnetic field is represented as the first time derivatives of the Gauss coefficients $g_n^m(t)$ and $h_n^m(t)$.

We model their temporal variations by the ESN model. The state of the system at time t_k is represented by state vector $\boldsymbol{x}_k$. The number of state variables M_x is set to 1000 in this study. At time step k, the i-th element of $\boldsymbol{x}_k$, $x_{k,i}$, is updated as follows:

$$x_{k,i} = (1 - \xi)x_{k-1,i} + \xi \tanh\left(\boldsymbol{w}_i^\mathsf{T} \boldsymbol{x}_{k-1} + \boldsymbol{u}_i^\mathsf{T} \boldsymbol{z}_k + \eta_i\right) \tag{3}$$

where $\boldsymbol{z}_k$ denotes the input vector, $\boldsymbol{w}_i$ is a weight vector for connecting among the state variables, $\boldsymbol{u}_i$ is a weight vector for connecting with the input variables, and ξ is the leakage rate (Jaeger et al., 2007; Lukoševičius, 2012). We fixed the value of ξ at 0.5 in this study. The weights $\boldsymbol{w}_i$ and $\boldsymbol{u}_i$ are given in advance and are fixed. We set 90% of the weights $\{\boldsymbol{w}_i\}$ and $\{\boldsymbol{u}_i\}$ (randomly chosen) to zero. The values of the remaining non-zero elements of $\boldsymbol{u}_i$ are drawn randomly from a normal distribution with mean 0 and standard deviation σ_u. The standard deviation σ_u is set to adjust the range of the input variables $\boldsymbol{z}$ as described later. The values of the non-zero elements of $\boldsymbol{w}_i$ are also drawn from a normal distribution. The weights $\{\boldsymbol{w}_i\}$ are then rescaled such that the maximum singular value of the weight matrix, which is defined as

$$W = (\boldsymbol{w}_1\ \boldsymbol{w}_2\ \cdots \boldsymbol{w}_{M_x}), \tag{4}$$

becomes 0.99. This rescaling is applied to satisfy the so-called "echo state property" which guarantees that the state of the ESN is not affected by distant past inputs. The output of the ESN at time t_k, $\boldsymbol{y}_k$, is then obtained from $\boldsymbol{x}_k$ as follows:

$$\boldsymbol{y}_k = \boldsymbol{\Gamma}^\mathsf{T} \boldsymbol{x}_k, \tag{5}$$

where $\boldsymbol{\Gamma}$ denotes the weight matrix. The output $\boldsymbol{y}_k$ corresponds to a prediction of the observation at time t_k.

Denoting the observation at time t_k as $\boldsymbol{d}_k$, the matrix $\boldsymbol{\Gamma}$ is determined by minimizing the following objective function:

$$J = \sum_{k=1}^{K} \frac{\left\|\boldsymbol{d}_k - \boldsymbol{\Gamma}^\mathsf{T} \boldsymbol{x}_k\right\|_2^2}{\sigma_k^2} + \frac{\|\boldsymbol{\Gamma}\|_F^2}{\lambda^2}, \tag{6}$$

where the second term on the right-hand side of this equation is a regularization term to avoid overfitting and $\|\boldsymbol{\Gamma}\|_F$ denotes the Frobenius norm of the matrix $\boldsymbol{\Gamma}$. The parameters σ_k and λ correspond to the scales of uncertainties for the observations and constraints, respectively. The values of the parameters used in this study are described in the next section. Decomposing $\boldsymbol{d}_k$ and $\boldsymbol{\Gamma}$ as $\boldsymbol{d}_k = \left(d_{k,1}, \ldots, d_{k,M_y}\right)$ and $\boldsymbol{\Gamma} = \left(\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_{M_y}\right)$, respectively, Eq. (6) can be rewritten as:

$$J = \sum_{i=1}^{M_y} \left[\sum_{k=1}^{K} \frac{\left(d_{k,i} - \boldsymbol{\gamma}_i^\mathsf{T} \boldsymbol{x}_k\right)^2}{\sigma_k^2} + \frac{\|\gamma_i\|_2^2}{\lambda^2}\right]. \tag{7}$$

We can thus find the optimal weight matrix $\mathbf{\Gamma}$ by obtaining the optimal value for each $\boldsymbol{\gamma}_k$ that minimizes the following component of J:

$$J_i = \sum_{k=1}^{K} \frac{\left(d_{k,i} - \boldsymbol{\gamma}_i^{\mathsf{T}} \boldsymbol{x}_k\right)^2}{\sigma_k^2} + \frac{\|\boldsymbol{\gamma}_i\|_2^2}{\lambda^2}. \tag{8}$$

For training the ESN, we use the observations as the input. Given a sequence of inputs, the state vector $\boldsymbol{x}_k$ for each step k is deterministically obtained via Eq. (3). The observation $\boldsymbol{d}_k$ is also given. With $\boldsymbol{d}_k$ and $\boldsymbol{x}_k$, the optimal $\boldsymbol{\gamma}_i$ that minimizes J_i is analytically obtained by solving the following equation:

$$\nabla_{\boldsymbol{\gamma}_i} J_i = -\sum_{k=1}^{K} \frac{\boldsymbol{x}_k \left(d_{k,i} - \boldsymbol{x}_k^{\mathsf{T}} \boldsymbol{\gamma}_i\right)}{\sigma_k^2} + \frac{\boldsymbol{\gamma}_i}{\lambda^2} = \mathbf{0}. \tag{9}$$

We obtain the optimal $\boldsymbol{\gamma}_i$ as

$$\hat{\boldsymbol{\gamma}}_i = \left(\frac{\mathbf{I}}{\lambda^2} + \sum_{k=1}^{K} \frac{\boldsymbol{x}_k \boldsymbol{x}_k^{\mathsf{T}}}{\sigma_k^2}\right)^{-1} \sum_{k=1}^{K} \frac{d_{k,i} \boldsymbol{x}_k}{\sigma_k^2}, \tag{10}$$

where $\mathbf{I}$ denotes the identity matrix.

To model the temporal evolution of the geomagnetic field with the ESN, we consider the temporal difference of the Gauss coefficients as follows:

$$\Delta g_n^m(t_k) = g_n^m(t_k) - g_n^m(t_{k-1}), \tag{11}$$

$$\Delta h_n^m(t_k) = h_n^m(t_k) - h_n^m(t_{k-1}). \tag{12}$$

For training the ESN, $\Delta g_n^m(t_{k-1})$ and $\Delta h_n^m(t_{k-1})$ are fed into the ESN as the input $\boldsymbol{z}_k$ in Eq. (3) and $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ are used as the observation $\boldsymbol{d}_k$ in Eq. (6). We derive the time sequence of $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ from the IGRF model and used it for training. As $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ are used as the observation, the trained ESN yields a prediction for $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ as an output $\boldsymbol{y}_k$. When we use the trained ESN for future prediction, the prediction of $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ is fed back into the ESN as the input at the next time step $\boldsymbol{z}_{k+1}$ and we obtain a prediction for $\Delta g_n^m(t_{k+1})$ and $\Delta h_n^m(t_{k+1})$.

3 Hindcast experiments

We conduct hindcast experiments to reproduce the temporal evolution of the geomagnetic main field after training the ESN using the IGRF and Definitive Geomagnetic Reference Field (DGRF) models. The IGRF model as well as the DGRF gives the Gauss coefficients of the scalar potential V for every 5 years. Here, we obtain the Gauss coefficients for every year by interpolating the IGRF and DGRF models with a natural cubic spline. The temporal evolution of $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ for each year is then modelled with the ESN. When $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ are obtained as the temporal difference for a 1-year interval, their typical scale is of the order of 10 nT. To adjust the scale of $\boldsymbol{u}_i^{\mathsf{T}} \boldsymbol{z}_k$ in Eq. (3) to be less than 1, we set the standard deviation of $\boldsymbol{u}_i$, σ_u, to 0.01 when training the ESN. The ESN requires inputs for a sufficient number of time steps before its output can be compared with the observations. Hence, we use the observations of $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ from 1901 to 1920 for spin-up and train the ESN using the observations from 1921 to 2005. We then predict $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ from 2006 to 2015 and obtain the hindcast of $g_n^m(t_k)$ and $h_n^m(t_k)$ accordingly.

To determine $\boldsymbol{\gamma}_i$ using Eq. (10), the parameters σ_k and λ must be given in advance. The parameter σ_k corresponds to the uncertainty of the observation $d_{k,i}$. Until 2000, as the Gauss coefficients of the DGRF may contain errors of ± 0.5 nT, we assume the temporal difference within a 5-year interval, $g_n^m(t_k) - g_n^m(t_{k-5})$, have an uncertainty with the standard deviation of 0.5 nT, which corresponds to an uncertainty with the variance of 0.25. The variance of the uncertainty of $\Delta g_n^m(t_k)$, which is the temporal difference within 1 year, would thus becomes $0.25/5 = 0.05$. We thus estimate that the standard deviation of the uncertainty of $\Delta g_n^m(t_k)$, σ_k, is about $0.22(\approx \sqrt{0.05})$ until 2000. Similarly, after 2000, we assume the temporal difference within a 5-year interval have an uncertainty with the standard deviation of 0.05 nT, and estimate that $\sigma_k = 0.022$. Since the minimization of J_i can be regarded as a Bayesian estimation problem of $d_{k,i}$ with a Gaussian prior for $\boldsymbol{\gamma}_i$, the parameter λ can be determined by the maximization of the marginal likelihood, which is often used in Bayesian estimation (e.g., Morris, 1983; Casella, 1985). We set the value of λ to 0.022 based on the marginal likelihood in this study.

The start time of the hindcast experiments in this section was set to 2005. We prepare input data from the 10th-generation IGRF (IGRF-10) model (Maus et al., 2005), which was released in 2005, in addition to the IGRF and DGRF models from 1900 to 2000. We then obtain the Gauss coefficients for every year since 1900 by interpolating the models. We refer to the product of this interpolation as the interpolated IGRF-10. For reference, we also prepare a model obtained by interpolating the IGRF and DGRF models from 1900 through 2015 plus the 13-th generation IGRF (IGRF-13) (Alken, Thébault, Beggan, Amit, et al., 2021), which we refer to as the interpolated IGRF-13. We train the ESN with the interpolated IGRF-10 and predict the temporal evolution of $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ from 2006. The results of the prediction are then compared with those for the interpolated IGRF-13. Figure 1 shows the results of the hindcast for g_1^0, g_1^1, h_1^1, g_2^0, g_2^1, h_2^1, g_2^2, h_2^2, and g_3^3. In each panel, the blue line indicates results of the hindcast conducted with the ESN, the red line indicates the interpolated IGRF-13, and the green line indicates the prediction of the original IGRF-10. Since the interpolated IGRF-13 is based on the definitive model until 2015, it can be regarded as the actual SV. Since the IGRF-10 was released in 2005, the prediction by the original IGRF-10 indicated by the green line is regarded as a benchmark of the prediction from 2005. Note that the prediction obtained with the ESN shown by the blue line did not use the observations after 2005; it used only the Gauss coefficients obtained by interpolating the DGRF and IGRF-10 models until 2005. Furthermore, the interpolation by the cubic spline treated the epoch 2005 as the end point, which forced the third time derivatives to be nil at 2005. This is the reason why the blue line deviates from the red line even before 2005.

A comparison of the ESN hindcast (blue line) and the IGRF-10 model (green line) indicates that the ESN provides better prediction for g_1^0. For g_1^1, g_2^0, g_2^1, h_2^1, and h_2^2, the performance was comparable between the ESN and the IGRF-10. However, the prediction of the ESN was inferior to that of the IGRF-10 model for h_1^1, g_2^2, and g_3^0. In particular, the prediction obtained by the ESN largely deviates from the actual SV for g_2^2 and g_3^0 which underwent a large change in trend. The IGRF-10 model could not predict these two coefficients, probably because of a problem in the input data from 2000 to 2005. The temporal gradients of g_2^2 and g_3^0 in the interpolated IGRF-13 gradually increased during the period from 2000 to 2005 and the descending trends became less steep in 2005. In contrast, the g_2^2 and g_3^0 from the ESN prediction, which used the interpolated IGRF-10 model as the input, maintained the descending trends in 2005, which made g_2^2 and g_3^0 continue to decrease after 2005.

Although the input for the ESN in Figure 1 was obtained by the interpolation of the models available every 5 years, geomagnetic observations with higher time resolution are actually available. To consider the case where geomagnetic observations can be obtained with high accuracy and high time resolution, we conducted another hindcast with the ESN using the Gauss coefficients of the interpolated IGRF-13 un-

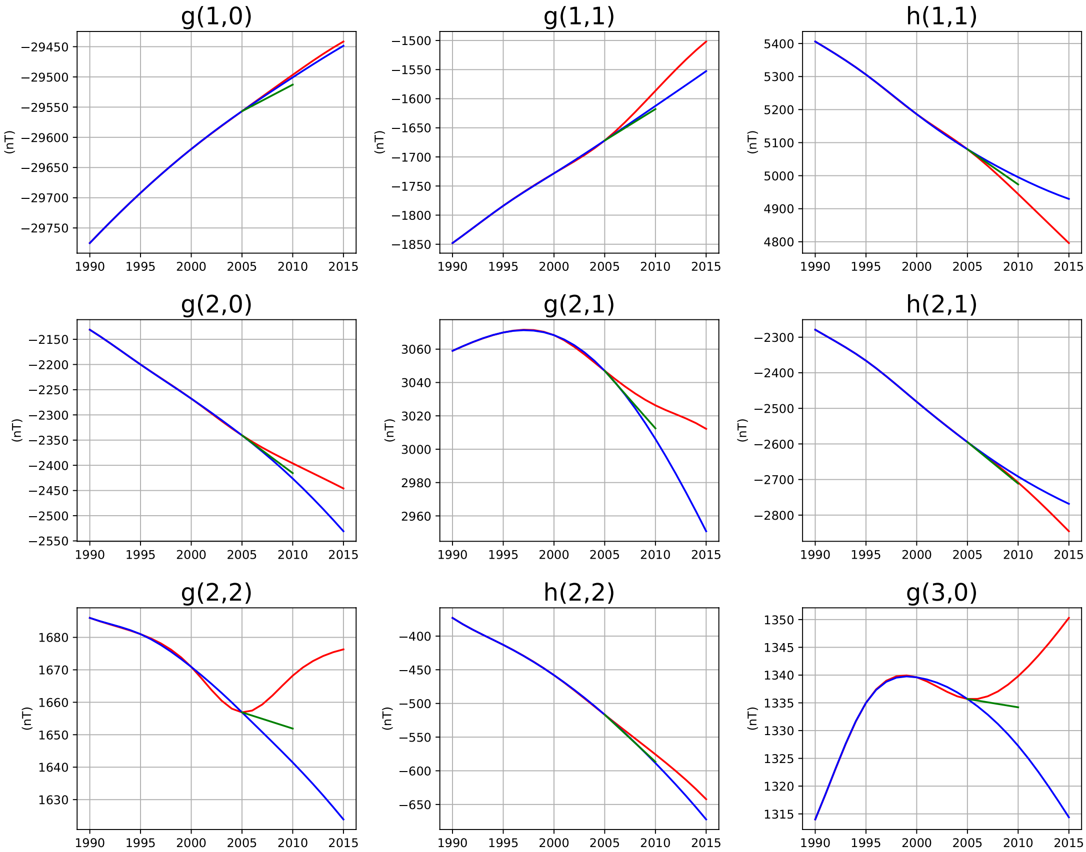

Figure 1. Prediction obtained with ESN (blue), IGRF-13 model (red), and IGRF-10 model (green).

til 2005. In other words, the Gauss coefficients of the IGRF-13 indicated by red lines in Figure 1 were used as the input until 2005 and the temporal evolution after 2005 was predicted. Figure 2 shows the results of the hindcast for the same nine coefficients as those in Figure 1. In each panel, the blue line indicates the prediction obtained with the ESN which used the interpolated IGRF-13 and the red and green lines show the same variations as those in Figure 1. The prediction obtained with the ESN was remarkably improved by using the accurate input with a 1-year time resolution. The change in trend for g_3^0 was successfully reproduced. The ESN also predicted the change in trend for h_1^1, g_2^1, and g_2^2, although the prediction slightly deviated from the actual SV. The performance of the ESN prediction was overall superior to that of the original IGRF-10 indicated by the green line. This result suggests that the ESN has potential for predicting SV with satisfactory accuracy if accurate geomagnetic data with a 1-year time resolution are available.

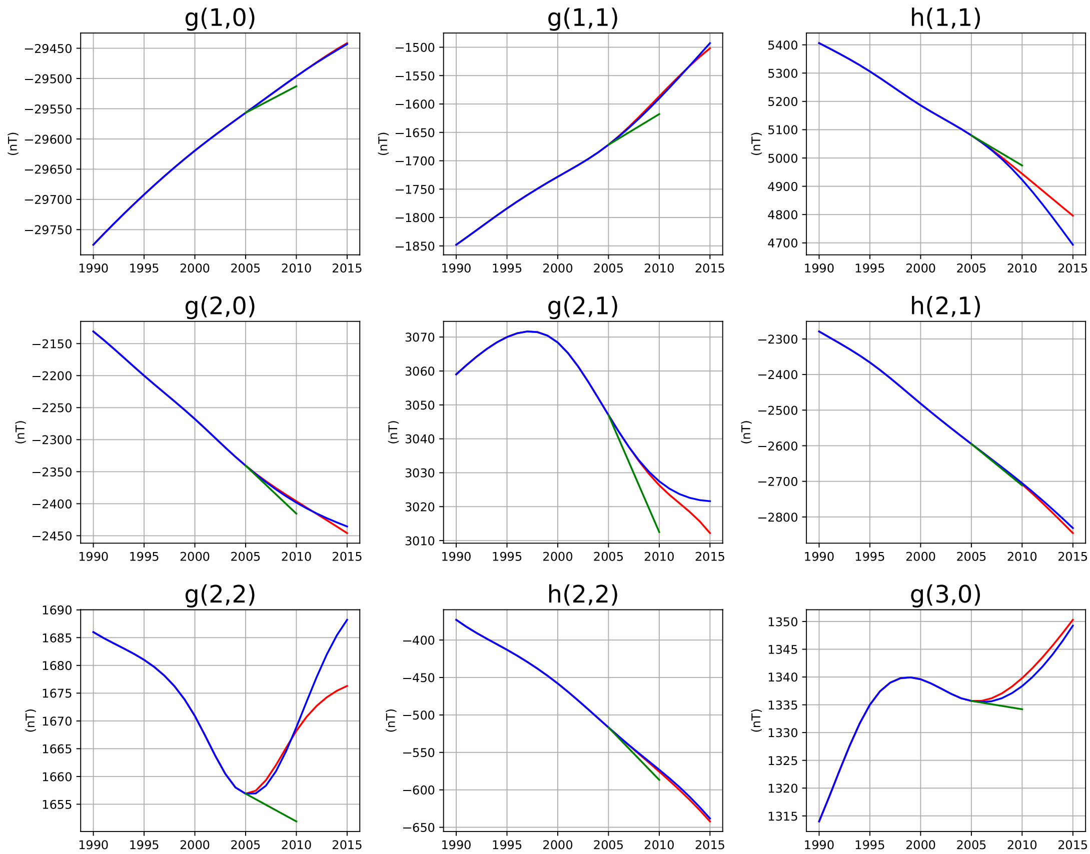

Figure 2. Prediction obtained with ESN using IGRF-13 values until 2005 as input (blue), IGRF-13 (red), and IGRF-10 (green).

The results in Figure 2 were obtained by the ESN trained with the Gauss coefficients for 85 years, from 1921 to 2005. The period of the training data is short compared with the dominant time scales of geodynamo dynamics, which vary on time scales of more than 10,000 years. Although data on the past geomagnetic field are limited, we conducted an experiment using the CALS3k model (Korte & Constable, 2011), which provides the geomagnetic field for about 3000 years from 1000

BCE to 1990 CE. We obtained $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ from the CALS3k model and used them as the observations. We used the observations from 999 BCE to 980 BCE for spin-up and trained the ESN using the observations from 979 BCE to 1990 CE. Although we trained the ESN with the CALS3k model data, the prediction was performed using the interpolated IGRF-13 data until 2005 as the input. Each panel in Figure 3 shows the results of the hindcast conducted with the ESN trained using the CALS3k data with the blue line. While the prediction obtained with the ESN was slightly better than the IGRF-10 (green line) for g_1^0 and g_1^1, the ESN did not predict the change in trend of g_2^2 and g_3^0 even though the interpolated IGRF-13 data were used as the input. A comparison with the ESN trained with the IGRF-13 (Figure 2), indicates that training with CALS3k decreased prediction accuracy.

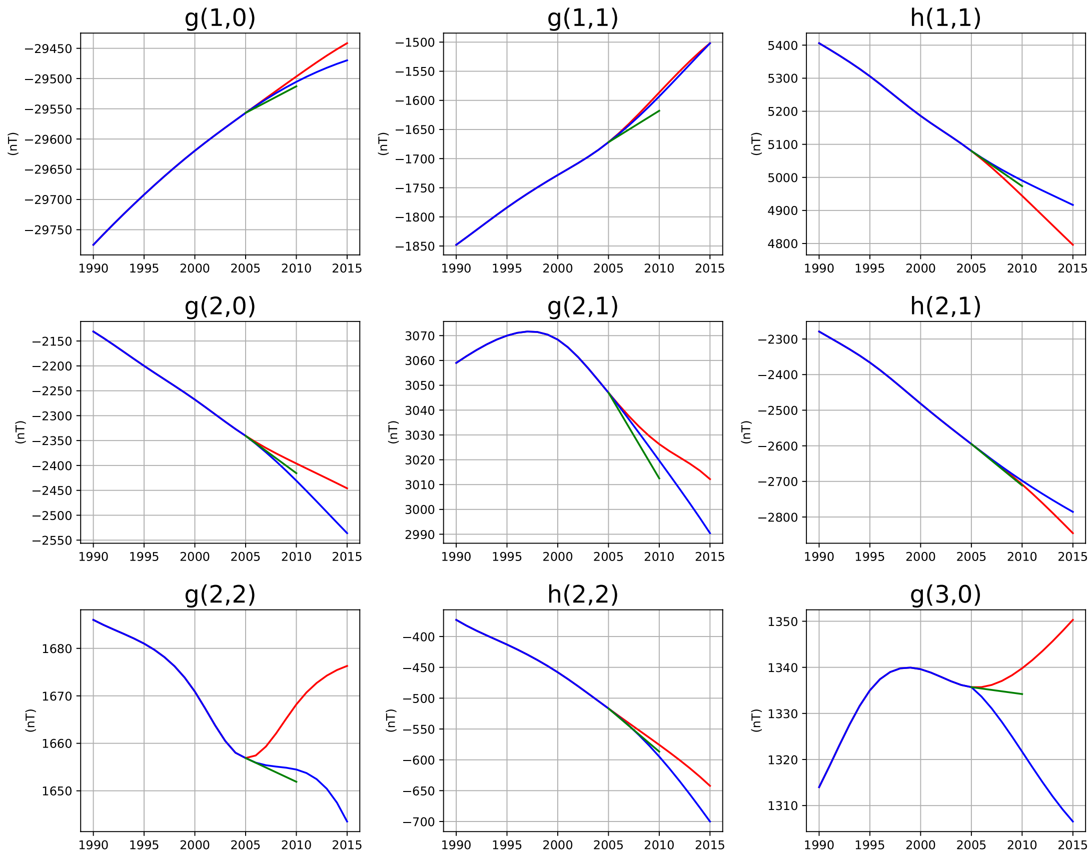

Figure 3. Prediction obtained with ESN trained with CALS3k model using IGRF-13 values until 2005 as input (blue), IGRF-13 (red), and IGRF-10 (green).

4 Discussion

A comparison between Figures 1 and 2 suggests the importance of high-accuracy data with a 1-year or higher time resolution before starting the prediction. The main difference between the interpolated IGRF-10 and the interpolated IGRF-13 is the curvature from 2000 to 2005. The gradients of g_2^2 and g_3^0 in the interpolated IGRF-13 gradually increased from 2000 to 2005, whereas those in the interpolated IGRF-10 remained descending in 2005. It is thus essential to detect such curvatures in the variations for predicting the nonlinear behavior of SV. The experimental results in Figure 3 suggest that prediction accuracy is not necessarily improved when the number of training data is increased. Since the CALS3k model does not include data after 1990, the poorer prediction obtained with the ESN trained with the CALS3k model may indicate the impact of the latest data at the starting point of the prediction.

Data-driven approaches such as machine learning techniques typically require a large number of data to generate a prediction for all possible cases. To predict the evolution of a dynamical system for all possible cases, observation of the global structure of the trajectory in phase space is required. This is not possible for the geodynamo system because the time scale of the observation is much shorter than the convection time scale of the geodynamo. The temporal evolution of the geodynamo is thus difficult to predict. Nevertheless, the results of the hindcasts presented in this article demonstrate that a data-driven approach is applicable for predicting SV for several years even in the occurrence of short-term as well as nonlinear rapid SVs such as the geomagnetic jerks.

As the ESN is likely to learn a local structure of the trajectory in the vicinity of the starting point, simpler methods such as polynomial extrapolation might work for predicting SV for 5 years. However, standard geomagnetic models such as the IGRF model contains more than 100 parameters. It would be difficult to consider a polynomial of 100 variables including cross terms. Hence, the ESN is considered to be a useful tool for the short-term prediction of the geomagnetic field controlled by the geodynamo system.

5 Summary

This study examined the applicability of the ESN, which is a kind of recurrent neural network with fixed connections among hidden state variables, for predicting SV. We trained the ESN using the DGRF model from 1900 to 2000 and IGRF-10 and conducted a hindcast of SV from 2005. The results demonstrate that the ESN can predict SV with satisfactory accuracy. In particular, if accurate geomagnetic data with a 1-year or higher time resolution are available, even the nonlinear behavior of SV such as the geomagnetic jerks is successfully predicted for 5 years. On the other hand, an increase in the number of training data does not necessarily improve prediction accuracy. The availability of a highly accurate temporal evolution of the geomagnetic field, including the curvature in time domain, for the last several years is thus important for predicting SV with the ESN.

Acronyms

SV Secular variation
ESN Echo state network
IGRF International Geomagnetic Reference Field
DGRF Definitive Geomagnetic Reference Field

Open Research Section

The code for the CALS3k model was acquired via EarthRef.org (`https://earthref.org/`).

Acknowledgments

This study was supperted by a research program funded by Joint Support Center for Data Science, Research Organization of Information and Systems (ROIS-DS-JOINT 007RP2022).

References

Alexandrescu, M., Gibert, D., Hulot, G., Le Mouël, J.-L., & Saracco, G. (1996). Worldwide wavelet analysis of geomagnetic jerks. *J. Geophys. Res.*, *101*, 21975–21994. doi: 10.1029/96JB01648

Alken, P., Thébault, E., Beggan, C. D., Amit, H., Aubert, J., Baerenzung, J., ... Zhou, B. (2021). International Geomagnetic Reference Field: the thirteenth generation. *Earth Planets Space*, *73*, 49. doi: 10.1186/s40623-020-01288-x

Alken, P., Thébault, E., Beggan, C. D., Aubert, J., Baerenzung, J., Brown, W. J., ... Wardinski, I. (2021). Evaluation of candidate models for the 13th generation International Geomagnetic Reference Field. *Earth Planets Space*, *73*, 48. doi: 10.1186/s40623-020-01281-4

Casella, G. (1985). An introduction to empirical Bayes data analysis. *Amer. Statistician*, *39*, 83–87.

Courtillot, V., & Mouël, J.-L. L. (1984). Geomagnetic secular variation impulses: A review of observational evidence and geophysical consequences. *Nature*, *311*, 709–716. doi: 10.1186/s40623-020-01313-z

Fournier, A., Aubert, J., Lesur, V., & G., R. (2021). A secular variation candidate model for IGRF-13 based on Swarm data and ensemble inverse geodynamo modelling. *Earth Planets Space*, *73*, 43. doi: 10.1186/s40623-020-01309-9

Jaeger, H., & Haas, H. (2004). Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *Science*, *304*, 78–80. doi: 10.1126/science.1091277

Jaeger, H., Lukoševičius, M., Popovici, D., & Siewert, U. (2007). Optimization and applications of echo state networks with leaky-integrator neurons. *Neural Networks*, *20*, 335–352. doi: 10.1016/j.neunet.2007.04.016

Kataoka, R., & Nakano, S. (2021). Reconstructing solar wind profiles associated with extreme magnetic storms: A machine learning approach. *Geophys. Res. Lett.*, *48*, e2021GL096275. doi: 10.1029/2021GL096275

Korte, M., & Constable, C. (2011). Improving geomagnetic field reconstructions for 0–3 ka. *Phys. Earth Planet. Interiors*, *188*, 247–259. doi: 10.1016/j.pepi.2011.06.017

Lukoševičius, M. (2012). A practical guide to applying echo state networks. In G. Montavon, G. Orr, & K. Müller (Eds.), *Neural networks: Tricks of the trade* (pp. 659–686). Springer.

Maus, S., Macmillan, S., Chernova, T., Choi, S., Dater, D., Golovkov, V., ... Zvereva, T. (2005). The 10th-generation International Geomagnetic Reference Field. *Geophys. J. Int.*, *161*, 561–565. doi: 10.1111/j.1365-246X.2005.02641.x

Minami, T., Nakano, S., Lesur, V., Takahashi, F., Matsushima, M., Shimizu, H., ... Toh, H. (2020). A candidate secular variation model for IGRF-13 based on MHD dynamo simulation and 4DEnVar data assimilation. *Earth Planets Space*, *72*, 136. doi: 10.1186/s40623-020-01253-8

Morris, C. M. (1983). Parametric empirical Bayes inference: theory and applications. *J. Amer. Statist. Assoc.*, *78*, 47–55.

Nakano, S., & Kataoka, R. (2022). Echo state network model for analyzing solar-wind effects on the AU and AL indices. *Ann. Geophys.*, *40*, 11–22. doi: 10.5194/angeo-40-11-2022

Walleshauser, B., & Bollt, E. (2022). Predicting sea surface temperatures with coupled reservoir computers. *Nonlin. Processes Geophys.*, *29*, 255–264. doi:

10.5194/npg-29-255-2022

Short term prediction of geomagnetic secular variation with an echo state network

S. Nakano[1,2,3*], **H. Toh**[4]

[1]The Institute of Statistical Mathematics
[2]The Graduate University for Advanced Studies, SOKENDAI
[3]Center for Data Assimilation Research and Applications, Joint Support Center for Data Science
[4]Graduate School of Science, Kyoto University

Key Points:

- A technique for predicting the secular variation of the geomagnetic field based on the echo state network model is proposed.
- The hindcast results show that the secular variation is predicted with satisfactory accuracy.
- It is also sugested that the information on the latest temporal variations is important for predicting the secular variation.

*10–3 Midori-cho, Tachikawa, Tokyo, Japan

Corresponding author: S. Nakano, shiny@ism.ac.jp

Abstract

A technique for predicting the secular variation (SV) of the geomagnetic field based on the echo state network (ESN) model is proposed. SV is controlled by the geodynamo process in the Earth's outer core. However, it is difficult to model the realistic nonlinear behaviors of the geodynamo due mainly to the very small Ekman number of the actual outer core. This study employs the ESN to represent the temporal evolution of the geomagnetic field on the Earth's surface. The hindcast results of SV demonstrate that the ESN enables us to predict SV for several years with satisfactory accuracy. In particular, the nonlinear behaviors of SV is accurately predicted for the case where accurate geomagnetic data with a 1-year time resolution are available. It is found that an increase in the number of training data does not necessarily improve prediction accuracy. The results suggest that the information on the latest temporal variations is important for the short-term prediciton by the ESN valid for, say 5 years.

1 Introduction

The geomagnetic field is gradually and incessantly changing. This change is referred to as secular variation (SV). The magnitude of SV can exceed 10 nT per year, which is comparable to or larger than that of ionospheric and magnetospheric origin. Hence, it is important to predict SV on a time scale of several years. The International Geomagnetic Reference Field (IGRF) model (Alken, Thébault, Beggan, Amit, et al., 2021) includes an SV model for prediction of next 5 years. Since SV sometimes shows nonlinear behaviors such as geomagnetic jerks (e.g., Courtillot & Mouël, 1984; Alexandrescu et al., 1996), its accurate prediction is difficult. Accordingly, various approaches were employed in the 14 SV candidate models which contributed to the latest IGRF model (Alken, Thébault, Beggan, Aubert, et al., 2021 and references therein). Since the geomagnetic main field is thought to be driven by a dynamo process in the Earth's outer core, some candidate models assimilated ground and satellite data into numerical geodynamo models. (e.g, Minami et al., 2020; Fournier et al., 2021). Data assimilation is a straightforward approach to consider the nonlinear dynamics of the outer core. However, a typical geodynamo model represents the state of the geodynamo with millions of variables, whereas the IGRF model represents the geomagnetic main field on the Earth's surface using about 200 parameters. The computational cost of data assimilation using a geodynamo model is thus excessive for predicting the parameters for the geomagnetic field model.

Machine learning approaches for modelling nonlinear systems have recently emerged. The purpose of this study is to explore a machine-learning-based method for predicting SV efficiently. Here, we employ an echo state network (ESN) model (Jaeger & Haas, 2004) for this purpose. The ESN is a kind of reservoir computing framework and it is a recurrent neural network in which the connections and weights between hidden state variables are randomly set and fixed. The ESN is therefore trained by optimizing the weights of only the output layer. Compared to the latest deep neural network models, the degree of freedom of the ESN is small because the weights of only the output layer are made variable. However, for the problem considered here, we have observation data for recent for only the most recent 100 to 1000 years, whereas the convection time scale of the outer core is tens of thousands of years. The available observations are thus insufficient for optimizing the large number of parameters for a deep neural network. Even with its small degree of freedom, the ESN shows satisfactory performance in various geophysical applications (e.g., Kataoka & Nakano, 2021; Nakano & Kataoka, 2022; Walleshauser & Bollt, 2022). Therefore, we apply the ESN for modelling the temporal evolution of the geomagnetic field in the hope of handling the nonlinear behaviors of SV including the geomagnetic jerks.

2 Method

Following many models of the Earth's magnetic field, including the IGRF model, we represent the magnetic field $\boldsymbol{B}$ with a scalar potential V as

$$\boldsymbol{B} = -\nabla V. \tag{1}$$

The potential V is expanded into spherical harmonics:

$$V(r,\theta,\phi,t) = a \sum_{n=1}^{N} \sum_{m=0}^{n} \left(\frac{a}{r}\right)^{n+1} \left[g_n^m(t)\cos m\phi + h_n^m(t)\sin m\phi\right] P_n^m(\cos\theta) \tag{2}$$

where a denotes the Earth's mean radius. The SV of the geomagnetic field is represented as the first time derivatives of the Gauss coefficients $g_n^m(t)$ and $h_n^m(t)$.

We model their temporal variations by the ESN model. The state of the system at time t_k is represented by state vector $\boldsymbol{x}_k$. The number of state variables M_x is set to 1000 in this study. At time step k, the i-th element of $\boldsymbol{x}_k$, $x_{k,i}$, is updated as follows:

$$x_{k,i} = (1-\xi)x_{k-1,i} + \xi \tanh\left(\boldsymbol{w}_i^\mathsf{T}\boldsymbol{x}_{k-1} + \boldsymbol{u}_i^\mathsf{T}\boldsymbol{z}_k + \eta_i\right) \tag{3}$$

where $\boldsymbol{z}_k$ denotes the input vector, $\boldsymbol{w}_i$ is a weight vector for connecting among the state variables, $\boldsymbol{u}_i$ is a weight vector for connecting with the input variables, and ξ is the leakage rate (Jaeger et al., 2007; Lukoševičius, 2012). We fixed the value of ξ at 0.5 in this study. The weights $\boldsymbol{w}_i$ and $\boldsymbol{u}_i$ are given in advance and are fixed. We set 90% of the weights $\{\boldsymbol{w}_i\}$ and $\{\boldsymbol{u}_i\}$ (randomly chosen) to zero. The values of the remaining non-zero elements of $\boldsymbol{u}_i$ are drawn randomly from a normal distribution with mean 0 and standard deviation σ_u. The standard deviation σ_u is set to adjust the range of the input variables $\boldsymbol{z}$ as described later. The values of the non-zero elements of $\boldsymbol{w}_i$ are also drawn from a normal distribution. The weights $\{\boldsymbol{w}_i\}$ are then rescaled such that the maximum singular value of the weight matrix, which is defined as

$$W = \left(\boldsymbol{w}_1 \ \boldsymbol{w}_2 \ \cdots \boldsymbol{w}_{M_x}\right), \tag{4}$$

becomes 0.99. This rescaling is applied to satisfy the so-called "echo state property" which guarantees that the state of the ESN is not affected by distant past inputs. The output of the ESN at time t_k, $\boldsymbol{y}_k$, is then obtained from $\boldsymbol{x}_k$ as follows:

$$\boldsymbol{y}_k = \boldsymbol{\Gamma}^\mathsf{T}\boldsymbol{x}_k, \tag{5}$$

where $\boldsymbol{\Gamma}$ denotes the weight matrix. The output $\boldsymbol{y}_k$ corresponds to a prediction of the observation at time t_k.

Denoting the observation at time t_k as $\boldsymbol{d}_k$, the matrix $\boldsymbol{\Gamma}$ is determined by minimizing the following objective function:

$$J = \sum_{k=1}^{K} \frac{\left\|\boldsymbol{d}_k - \boldsymbol{\Gamma}^\mathsf{T}\boldsymbol{x}_k\right\|_2^2}{\sigma_k^2} + \frac{\|\boldsymbol{\Gamma}\|_F^2}{\lambda^2}, \tag{6}$$

where the second term on the right-hand side of this equation is a regularization term to avoid overfitting and $\|\boldsymbol{\Gamma}\|_F$ denotes the Frobenius norm of the matrix $\boldsymbol{\Gamma}$. The parameters σ_k and λ correspond to the scales of uncertainties for the observations and constraints, respectively. The values of the parameters used in this study are described in the next section. Decomposing $\boldsymbol{d}_k$ and $\boldsymbol{\Gamma}$ as $\boldsymbol{d}_k = \left(d_{k,1},\ldots,d_{k,M_y}\right)$ and $\boldsymbol{\Gamma} = \left(\boldsymbol{\gamma}_1,\ldots,\boldsymbol{\gamma}_{M_y}\right)$, respectively, Eq. (6) can be rewritten as:

$$J = \sum_{i=1}^{M_y}\left[\sum_{k=1}^{K} \frac{\left(d_{k,i} - \boldsymbol{\gamma}_i^\mathsf{T}\boldsymbol{x}_k\right)^2}{\sigma_k^2} + \frac{\|\gamma_i\|_2^2}{\lambda^2}\right]. \tag{7}$$

We can thus find the optimal weight matrix $\mathbf{\Gamma}$ by obtaining the optimal value for each $\boldsymbol{\gamma}_k$ that minimizes the following component of J:

$$J_i = \sum_{k=1}^{K} \frac{\left(d_{k,i} - \boldsymbol{\gamma}_i^\mathsf{T} \boldsymbol{x}_k\right)^2}{\sigma_k^2} + \frac{\|\boldsymbol{\gamma}_i\|_2^2}{\lambda^2}. \tag{8}$$

For training the ESN, we use the observations as the input. Given a sequence of inputs, the state vector $\boldsymbol{x}_k$ for each step k is deterministically obtained via Eq. (3). The observation $\boldsymbol{d}_k$ is also given. With $\boldsymbol{d}_k$ and $\boldsymbol{x}_k$, the optimal $\boldsymbol{\gamma}_i$ that minimizes J_i is analytically obtained by solving the following equation:

$$\nabla_{\boldsymbol{\gamma}_i} J_i = -\sum_{k=1}^{K} \frac{\boldsymbol{x}_k \left(d_{k,i} - \boldsymbol{x}_k^\mathsf{T} \boldsymbol{\gamma}_i\right)}{\sigma_k^2} + \frac{\boldsymbol{\gamma}_i}{\lambda^2} = \mathbf{0}. \tag{9}$$

We obtain the optimal $\boldsymbol{\gamma}_i$ as

$$\hat{\gamma}_i = \left(\frac{\mathbf{I}}{\lambda^2} + \sum_{k=1}^{K} \frac{\boldsymbol{x}_k \boldsymbol{x}_k^\mathsf{T}}{\sigma_k^2}\right)^{-1} \sum_{k=1}^{K} \frac{d_{k,i} \boldsymbol{x}_k}{\sigma_k^2}, \tag{10}$$

where $\mathbf{I}$ denotes the identity matrix.

To model the temporal evolution of the geomagnetic field with the ESN, we consider the temporal difference of the Gauss coefficients as follows:

$$\Delta g_n^m(t_k) = g_n^m(t_k) - g_n^m(t_{k-1}), \tag{11}$$

$$\Delta h_n^m(t_k) = h_n^m(t_k) - h_n^m(t_{k-1}). \tag{12}$$

For training the ESN, $\Delta g_n^m(t_{k-1})$ and $\Delta h_n^m(t_{k-1})$ are fed into the ESN as the input $\boldsymbol{z}_k$ in Eq. (3) and $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ are used as the observation $\boldsymbol{d}_k$ in Eq. (6). We derive the time sequence of $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ from the IGRF model and used it for training. As $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ are used as the observation, the trained ESN yields a prediction for $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ as an output $\boldsymbol{y}_k$. When we use the trained ESN for future prediction, the prediction of $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ is fed back into the ESN as the input at the next time step $\boldsymbol{z}_{k+1}$ and we obtain a prediction for $\Delta g_n^m(t_{k+1})$ and $\Delta h_n^m(t_{k+1})$.

3 Hindcast experiments

We conduct hindcast experiments to reproduce the temporal evolution of the geomagnetic main field after training the ESN using the IGRF and Definitive Geomagnetic Reference Field (DGRF) models. The IGRF model as well as the DGRF gives the Gauss coefficients of the scalar potential V for every 5 years. Here, we obtain the Gauss coefficients for every year by interpolating the IGRF and DGRF models with a natural cubic spline. The temporal evolution of $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ for each year is then modelled with the ESN. When $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ are obtained as the temporal difference for a 1-year interval, their typical scale is of the order of 10 nT. To adjust the scale of $\boldsymbol{u}_i^\mathsf{T} \boldsymbol{z}_k$ in Eq. (3) to be less than 1, we set the standard deviation of $\boldsymbol{u}_i$, σ_u, to 0.01 when training the ESN. The ESN requires inputs for a sufficient number of time steps before its output can be compared with the observations. Hence, we use the observations of $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ from 1901 to 1920 for spin-up and train the ESN using the observations from 1921 to 2005. We then predict $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ from 2006 to 2015 and obtain the hindcast of $g_n^m(t_k)$ and $h_n^m(t_k)$ accordingly.

To determine $\boldsymbol{\gamma}_i$ using Eq. (10), the parameters σ_k and λ must be given in advance. The parameter σ_k corresponds to the uncertainty of the observation $d_{k,i}$. Until 2000, as the Gauss coefficients of the DGRF may contain errors of ± 0.5 nT, we assume the temporal difference within a 5-year interval, $g_n^m(t_k) - g_n^m(t_{k-5})$, have an uncertainty with the standard deviation of 0.5 nT, which corresponds to an uncertainty with the variance of 0.25. The variance of the uncertainty of $\Delta g_n^m(t_k)$, which is the temporal difference within 1 year, would thus becomes $0.25/5 = 0.05$. We thus estimate that the standard deviation of the uncertainty of $\Delta g_n^m(t_k)$, σ_k, is about $0.22(\approx \sqrt{0.05})$ until 2000. Similarly, after 2000, we assume the temporal difference within a 5-year interval have an uncertainty with the standard deviation of 0.05 nT, and estimate that $\sigma_k = 0.022$. Since the minimization of J_i can be regarded as a Bayesian estimation problem of $d_{k,i}$ with a Gaussian prior for γ_i, the parameter λ can be determined by the maximization of the marginal likelihood, which is often used in Bayesian estimation (e.g., Morris, 1983; Casella, 1985). We set the value of λ to 0.022 based on the marginal likelihood in this study.

The start time of the hindcast experiments in this section was set to 2005. We prepare input data from the 10th-generation IGRF (IGRF-10) model (Maus et al., 2005), which was released in 2005, in addition to the IGRF and DGRF models from 1900 to 2000. We then obtain the Gauss coefficients for every year since 1900 by interpolating the models. We refer to the product of this interpolation as the interpolated IGRF-10. For reference, we also prepare a model obtained by interpolating the IGRF and DGRF models from 1900 through 2015 plus the 13-th generation IGRF (IGRF-13) (Alken, Thébault, Beggan, Amit, et al., 2021), which we refer to as the interpolated IGRF-13. We train the ESN with the interpolated IGRF-10 and predict the temporal evolution of $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ from 2006. The results of the prediction are then compared with those for the interpolated IGRF-13. Figure 1 shows the results of the hindcast for g_1^0, g_1^1, h_1^1, g_2^0, g_2^1, h_2^1, g_2^2, h_2^2, and g_3^3. In each panel, the blue line indicates results of the hindcast conducted with the ESN, the red line indicates the interpolated IGRF-13, and the green line indicates the prediction of the original IGRF-10. Since the interpolated IGRF-13 is based on the definitive model until 2015, it can be regarded as the actual SV. Since the IGRF-10 was released in 2005, the prediction by the original IGRF-10 indicated by the green line is regarded as a benchmark of the prediction from 2005. Note that the prediction obtained with the ESN shown by the blue line did not use the observations after 2005; it used only the Gauss coefficients obtained by interpolating the DGRF and IGRF-10 models until 2005. Furthermore, the interpolation by the cubic spline treated the epoch 2005 as the end point, which forced the third time derivatives to be nil at 2005. This is the reason why the blue line deviates from the red line even before 2005.

A comparison of the ESN hindcast (blue line) and the IGRF-10 model (green line) indicates that the ESN provides better prediction for g_1^0. For g_1^1, g_2^0, g_2^1, h_2^1, and h_2^2, the performance was comparable between the ESN and the IGRF-10. However, the prediction of the ESN was inferior to that of the IGRF-10 model for h_1^1, g_2^2, and g_3^0. In particular, the prediction obtained by the ESN largely deviates from the actual SV for g_2^2 and g_3^0 which underwent a large change in trend. The IGRF-10 model could not predict these two coefficients, probably because of a problem in the input data from 2000 to 2005. The temporal gradients of g_2^2 and g_3^0 in the interpolated IGRF-13 gradually increased during the period from 2000 to 2005 and the descending trends became less steep in 2005. In contrast, the g_2^2 and g_3^0 from the ESN prediction, which used the interpolated IGRF-10 model as the input, maintained the descending trends in 2005, which made g_2^2 and g_3^0 continue to decrease after 2005.

Although the input for the ESN in Figure 1 was obtained by the interpolation of the models available every 5 years, geomagnetic observations with higher time resolution are actually available. To consider the case where geomagnetic observations can be obtained with high accuracy and high time resolution, we conducted another hindcast with the ESN using the Gauss coefficients of the interpolated IGRF-13 un-

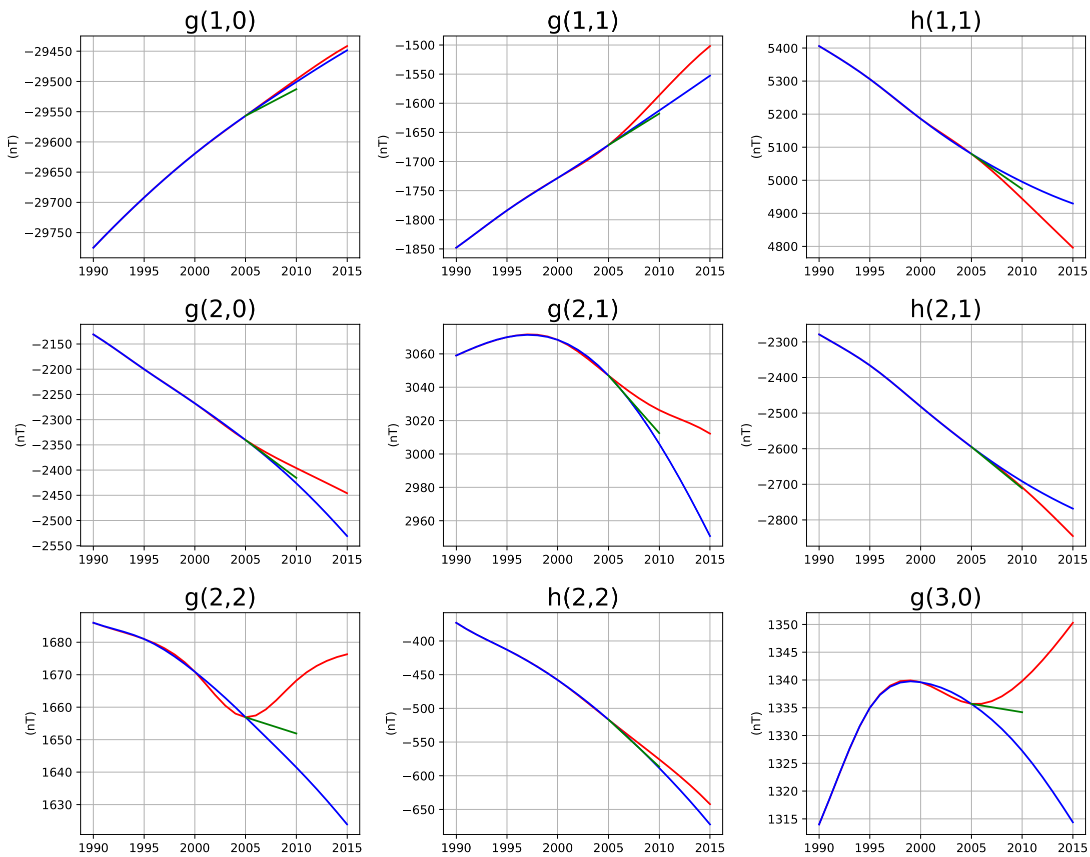

Figure 1. Prediction obtained with ESN (blue), IGRF-13 model (red), and IGRF-10 model (green).

til 2005. In other words, the Gauss coefficients of the IGRF-13 indicated by red lines in Figure 1 were used as the input until 2005 and the temporal evolution after 2005 was predicted. Figure 2 shows the results of the hindcast for the same nine coefficients as those in Figure 1. In each panel, the blue line indicates the prediction obtained with the ESN which used the interpolated IGRF-13 and the red and green lines show the same variations as those in Figure 1. The prediction obtained with the ESN was remarkably improved by using the accurate input with a 1-year time resolution. The change in trend for g_3^0 was successfully reproduced. The ESN also predicted the change in trend for h_1^1, g_2^1, and g_2^2, although the prediction slightly deviated from the actual SV. The performance of the ESN prediction was overall superior to that of the original IGRF-10 indicated by the green line. This result suggests that the ESN has potential for predicting SV with satisfactory accuracy if accurate geomagnetic data with a 1-year time resolution are available.

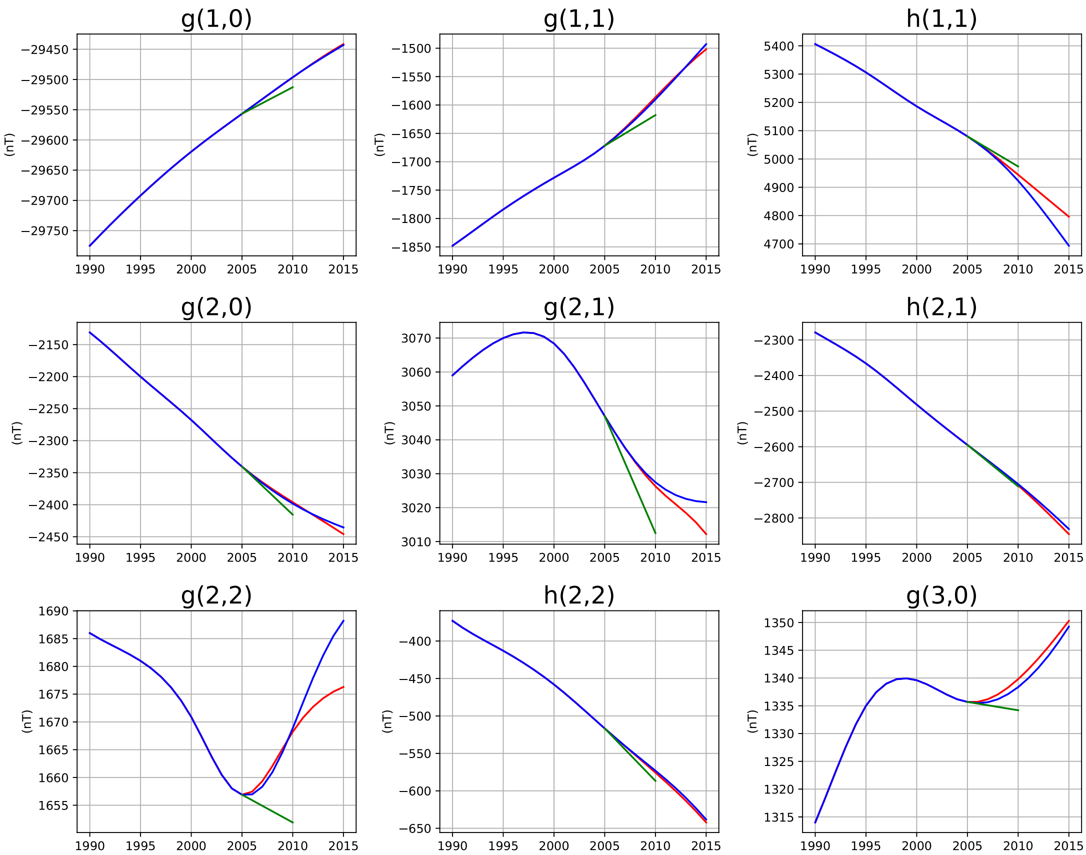

Figure 2. Prediction obtained with ESN using IGRF-13 values until 2005 as input (blue), IGRF-13 (red), and IGRF-10 (green).

The results in Figure 2 were obtained by the ESN trained with the Gauss coefficients for 85 years, from 1921 to 2005. The period of the training data is short compared with the dominant time scales of geodynamo dynamics, which vary on time scales of more than 10,000 years. Although data on the past geomagnetic field are limited, we conducted an experiment using the CALS3k model (Korte & Constable, 2011), which provides the geomagnetic field for about 3000 years from 1000

BCE to 1990 CE. We obtained $\Delta g_n^m(t_k)$ and $\Delta h_n^m(t_k)$ from the CALS3k model and used them as the observations. We used the observations from 999 BCE to 980 BCE for spin-up and trained the ESN using the observations from 979 BCE to 1990 CE. Although we trained the ESN with the CALS3k model data, the prediction was performed using the interpolated IGRF-13 data until 2005 as the input. Each panel in Figure 3 shows the results of the hindcast conducted with the ESN trained using the CALS3k data with the blue line. While the prediction obtained with the ESN was slightly better than the IGRF-10 (green line) for g_1^0 and g_1^1, the ESN did not predict the change in trend of g_2^2 and g_3^0 even though the interpolated IGRF-13 data were used as the input. A comparison with the ESN trained with the IGRF-13 (Figure 2), indicates that training with CALS3k decreased prediction accuracy.

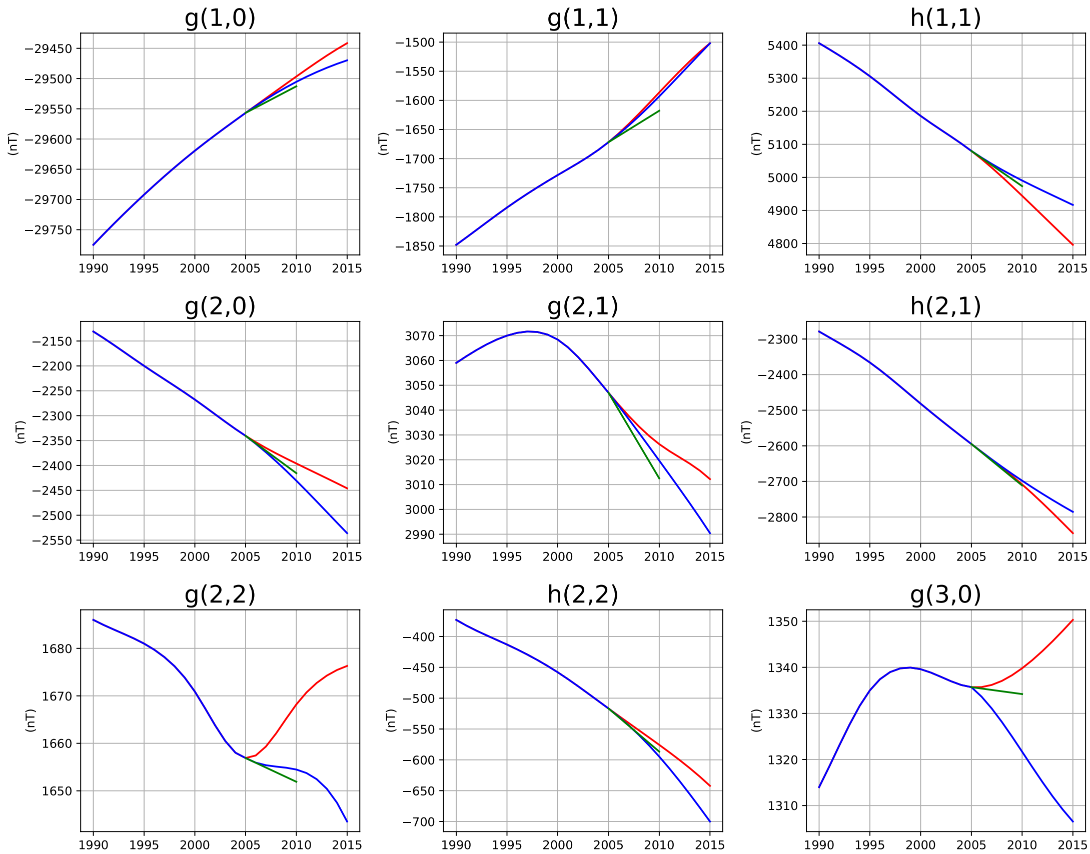

Figure 3. Prediction obtained with ESN trained with CALS3k model using IGRF-13 values until 2005 as input (blue), IGRF-13 (red), and IGRF-10 (green).

4 Discussion

A comparison between Figures 1 and 2 suggests the importance of high-accuracy data with a 1-year or higher time resolution before starting the prediction. The main difference between the interpolated IGRF-10 and the interpolated IGRF-13 is the curvature from 2000 to 2005. The gradients of g_2^2 and g_3^0 in the interpolated IGRF-13 gradually increased from 2000 to 2005, whereas those in the interpolated IGRF-10 remained descending in 2005. It is thus essential to detect such curvatures in the variations for predicting the nonlinear behavior of SV. The experimental results in Figure 3 suggest that prediction accuracy is not necessarily improved when the number of training data is increased. Since the CALS3k model does not include data after 1990, the poorer prediction obtained with the ESN trained with the CALS3k model may indicate the impact of the latest data at the starting point of the prediction.

Data-driven approaches such as machine learning techniques typically require a large number of data to generate a prediction for all possible cases. To predict the evolution of a dynamical system for all possible cases, observation of the global structure of the trajectory in phase space is required. This is not possible for the geodynamo system because the time scale of the observation is much shorter than the convection time scale of the geodynamo. The temporal evolution of the geodynamo is thus difficult to predict. Nevertheless, the results of the hindcasts presented in this article demonstrate that a data-driven approach is applicable for predicting SV for several years even in the occurrence of short-term as well as nonlinear rapid SVs such as the geomagnetic jerks.

As the ESN is likely to learn a local structure of the trajectory in the vicinity of the starting point, simpler methods such as polynomial extrapolation might work for predicting SV for 5 years. However, standard geomagnetic models such as the IGRF model contains more than 100 parameters. It would be difficult to consider a polynomial of 100 variables including cross terms. Hence, the ESN is considered to be a useful tool for the short-term prediction of the geomagnetic field controlled by the geodynamo system.

5 Summary

This study examined the applicability of the ESN, which is a kind of recurrent neural network with fixed connections among hidden state variables, for predicting SV. We trained the ESN using the DGRF model from 1900 to 2000 and IGRF-10 and conducted a hindcast of SV from 2005. The results demonstrate that the ESN can predict SV with satisfactory accuracy. In particular, if accurate geomagnetic data with a 1-year or higher time resolution are available, even the nonlinear behavior of SV such as the geomagnetic jerks is successfully predicted for 5 years. On the other hand, an increase in the number of training data does not necessarily improve prediction accuracy. The availability of a highly accurate temporal evolution of the geomagnetic field, including the curvature in time domain, for the last several years is thus important for predicting SV with the ESN.

Acronyms

SV Secular variation
ESN Echo state network
IGRF International Geomagnetic Reference Field
DGRF Definitive Geomagnetic Reference Field

Open Research Section

The code for the CALS3k model was acquired via EarthRef.org (`https://earthref.org/`).

Acknowledgments

This study was supperted by a research program funded by Joint Support Center for Data Science, Research Organization of Information and Systems (ROIS-DS-JOINT 007RP2022).

References

Alexandrescu, M., Gibert, D., Hulot, G., Le Mouël, J.-L., & Saracco, G. (1996). Worldwide wavelet analysis of geomagnetic jerks. *J. Geophys. Res.*, *101*, 21975–21994. doi: 10.1029/96JB01648

Alken, P., Thébault, E., Beggan, C. D., Amit, H., Aubert, J., Baerenzung, J., ... Zhou, B. (2021). International Geomagnetic Reference Field: the thirteenth generation. *Earth Planets Space*, *73*, 49. doi: 10.1186/s40623-020-01288-x

Alken, P., Thébault, E., Beggan, C. D., Aubert, J., Baerenzung, J., Brown, W. J., ... Wardinski, I. (2021). Evaluation of candidate models for the 13th generation International Geomagnetic Reference Field. *Earth Planets Space*, *73*, 48. doi: 10.1186/s40623-020-01281-4

Casella, G. (1985). An introduction to empirical Bayes data analysis. *Amer. Statistician*, *39*, 83–87.

Courtillot, V., & Mouël, J.-L. L. (1984). Geomagnetic secular variation impulses: A review of observational evidence and geophysical consequences. *Nature*, *311*, 709–716. doi: 10.1186/s40623-020-01313-z

Fournier, A., Aubert, J., Lesur, V., & G., R. (2021). A secular variation candidate model for IGRF-13 based on Swarm data and ensemble inverse geodynamo modelling. *Earth Planets Space*, *73*, 43. doi: 10.1186/s40623-020-01309-9

Jaeger, H., & Haas, H. (2004). Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *Science*, *304*, 78–80. doi: 10.1126/science.1091277

Jaeger, H., Lukoševičius, M., Popovici, D., & Siewert, U. (2007). Optimization and applications of echo state networks with leaky-integrator neurons. *Neural Networks*, *20*, 335–352. doi: 10.1016/j.neunet.2007.04.016

Kataoka, R., & Nakano, S. (2021). Reconstructing solar wind profiles associated with extreme magnetic storms: A machine learning approach. *Geophys. Res. Lett.*, *48*, e2021GL096275. doi: 10.1029/2021GL096275

Korte, M., & Constable, C. (2011). Improving geomagnetic field reconstructions for 0–3 ka. *Phys. Earth Planet. Interiors*, *188*, 247–259. doi: 10.1016/j.pepi.2011.06.017

Lukoševičius, M. (2012). A practical guide to applying echo state networks. In G. Montavon, G. Orr, & K. Müller (Eds.), *Neural networks: Tricks of the trade* (pp. 659–686). Springer.

Maus, S., Macmillan, S., Chernova, T., Choi, S., Dater, D., Golovkov, V., ... Zvereva, T. (2005). The 10th-generation International Geomagnetic Reference Field. *Geophys. J. Int.*, *161*, 561–565. doi: 10.1111/j.1365-246X.2005.02641.x

Minami, T., Nakano, S., Lesur, V., Takahashi, F., Matsushima, M., Shimizu, H., ... Toh, H. (2020). A candidate secular variation model for IGRF-13 based on MHD dynamo simulation and 4DEnVar data assimilation. *Earth Planets Space*, *72*, 136. doi: 10.1186/s40623-020-01253-8

Morris, C. M. (1983). Parametric empirical Bayes inference: theory and applications. *J. Amer. Statist. Assoc.*, *78*, 47–55.

Nakano, S., & Kataoka, R. (2022). Echo state network model for analyzing solar-wind effects on the AU and AL indices. *Ann. Geophys.*, *40*, 11–22. doi: 10.5194/angeo-40-11-2022

Walleshauser, B., & Bollt, E. (2022). Predicting sea surface temperatures with coupled reservoir computers. *Nonlin. Processes Geophys.*, *29*, 255–264. doi:

10.5194/npg-29-255-2022