Machine Learning Approach to Characterize the Postseismic Deformation of the 2011 Tohoku Oki Earthquake Based on Recurrent Neural Network

Yamaga Norifumi, Mitsui Yuta

*Geophysical Research Letters*

**Vol. 46**

**No. 21**

**Pages 11886-11892**

**Year 2019-11-30**

American Geophysical Union

(C)2019. American Geophysical Union. All Rights Reserved.

URL http://hdl.handle.net/10297/00026934
doi: 10.1029/2019GL084578
Geophysical Research Letters

RESEARCH LETTER
10.1029/2019GL084578

Key Points:
- We present a machine learning approach that successfully forecasts the postseismic deformation following a large earthquake.
- We capture the evolution characteristics of postseismic deformation from the differences between the RNN results and the deformation signal.

Supporting Information:
- Supporting Information SI

Correspondence to:
Y. Mitsui, mit@shizuoka.ac.jp

Citation:

Received 16 JUL 2019
Accepted 16 OCT 2019
Accepted article online 9 NOV 2019
Published online 11 NOV 2019

Machine Learning Approach to Characterize the Postseismic Deformation of the 2011 Tohoku-Oki Earthquake Based on Recurrent Neural Network

Norifumi Yamaga1 and Yuta Mitsui1

1Department of Geosciences, Shizuoka University, Shizuoka, Japan

Abstract Postseismic deformation following large earthquakes has generally been analyzed via viscoelastic simulations or regression analyses that employ logarithmic and/or exponential functions. Here we introduce a machine learning approach, the recurrent neural network, to more accurately forecast postseismic deformation and constrain its characteristics. We use Global Navigation Satellite System time-series data (horizontal components) from northeastern Japan since the 2011 Tohoku-oki megathrust earthquake to assess the feasibility of this machine-learning approach. We perform numerical experiment to examine the accuracy of the neural network forecast, compare the results with those from regression analyses, and confirm the improved accuracy of the neural network forecast. The spatiotemporal evolution of the differences between the observation data and forecast results implies alterations in the source of postseismic deformation, which may have occurred in 2013. We can extract detailed information on the spatiotemporal evolution of postseismic signals by implementing this new machine-learning approach.

Plain Language Summary We introduce a state-of-the-art machine-learning algorithm, the recurrent neural network, to characterize the surface displacement after a large earthquake. The accuracy of the forecast results far exceeds those of traditional regression methods. This assessment highlights the feasibility and effectiveness of employing a new machine-learning approach to infer detailed information on the evolution of rock deformation.

1. Introduction

Near-field postseismic slow deformation generally follows large earthquakes, with the rheology of the perturbed rock masses influencing the evolution of the deformation. A prime example of this phenomenon is the observed postseismic surface displacement since the 2011 Tohoku-oki megathrust earthquake (Mw 9.0), which has been captured by the GEONET Global Navigation Satellite System (GNSS) array in northeastern Japan (e.g., Ozawa et al., 2011). Figure 1 illustrates the total postseismic surface displacement field since the 2011 Tohoku-oki earthquake (March 12, 2011, to December 31, 2018).

Viscoelastic simulations (e.g., Agata et al., 2019; Freed et al., 2017; Muto et al., 2016; Sobolev & Muldavesh, 2017; Suito, 2017; Sun et al., 2014; Wang et al., 2012) and regression analyses using logarithmic and/or exponential functions (Nishimura, 2014; Tobita, 2016) have been employed to analyze the postseismic deformation related to this megathrust earthquake, with each possessing key strengths and weaknesses. While viscoelastic simulations are based on continuum mechanics, their strong dependence on assumed constitutive laws and parameter structures yields potential issues. Regression analyses are powerful tools for data fitting, but they are limited by the fitting function, such that a comparison of the estimated results becomes difficult when different fitting functions are required to yield the best fit at each observation station.

Here we introduce a machine learning approach, the recurrent neural network (RNN), to accurately model the observed postseismic deformation in time-series data. A neural network, which is a collection of connected units between the input and output data that interactively learns to weight the output data according to the input parameters, is employed due to its successful applications in seismology, including P-wave arrival picking (Ross et al., 2018), the discrimination of earthquakes and tremors from spectrograms (Nakano et al., 2019), acceleration of viscoelastic simulations (DeVries et al., 2017), and aftershock location forecasts (DeVries et al., 2018). RNN is a type of neural network, which has internal memories to pass information to a successor. It is useful to process sequential inputs, although it costs more computational time than a
traditional neural network. We construct the RNN and determine the optimal function that fits the sequential time-series data to learn the postseismic deformation characteristics associated with the 2011 Tohoku-oki earthquake, and then assess its feasibility in forecasting the postseismic deformation field for a given time series of previous observations.

2. Methods

2.1. Neural network settings

A neural network is composed of connected layers, each of which includes calculation units that are composed of simple functions. The first layer is the input layer, the last layer is the output layer, and the remaining layers are hidden layers. The RNN learns the characteristics of the sequential time-series data by storing the optimized model weights and biases in the hidden layers, which are then used to process the input sequences. Here we fix the number of hidden layers to one, and use 32 units of long short-term memory (LSTM) (Hochreiter & Schmidhuber, 1997) for the hidden layer. The activation function for the hidden layer is a hyperbolic tangent function, and that for the output layer is a linear function. Parameter optimization is based on a gradient method, Adam (Kingma & Ba, 2015), to determine the mean absolute errors.

We first enter the postseismic deformation time-series data into both the input and output layers, which allows the hidden layer to learn the fitting parameters. We then enter other data that were not used in the learning process into the input layer, and compare the forecasted values from the output layer with the actual data.

Figure 1. Total postseismic deformation since the 2011 Tohoku-oki earthquake, from March 12, 2011, to December 31, 2018, based on the GNSS data provided by the Geospatial Information Authority of Japan. The black arrows indicate the horizontal displacement, and the colormap represents the vertical displacement.
We divide the 191 observation stations into 153 learning points and 38 test points. The test points are selected to avoid bias, with one point selected for each spatial grid cell (0.5° × 0.5°). We only use the first 90% of the time-series data for the learning process. We then normalize the total displacements in each time series to the 0.1–0.9 range (this is just a representative range between 0 and 1). We finally execute the machine-learning runs by iteratively calculating the difference between the forecasted values and actual data at each of the 153 learning points until the differences are minimized. We define the forecasted displacements for a given day as the outputs calculated from the continuous input displacements for the previous 365-day period.

2.3. Forecast experiments

We perform forecast experiments using the above-mentioned RNN model. We forecast the postseismic deformation at the 38 test points from start days when the 365-day periods elapsed from the 2011 Tohoku-oki earthquake (e.g., March 11, 2012) to December 31, 2018, and compare the forecasted values with the actual data.

We remove the offsets due to the coseismic displacements associated with large earthquakes near the study region (Table 1) from the GNSS time-series data. In our main experiment, we perform this correction also for the learning process. Besides, we try additional experiment without this offset correction for the learning process because we expect an improvement in noise immunity via machine learning with noisy data (Isaev & Dolenko, 2018).

We compare the RNN results with those from traditional regression analyses to assess the accuracy of our proposed forecasting method. Referring to Tobita (2016), we perform regression analyses employing logarithmic and/or exponential functions. The following equations “log+exp” (equation (1)), “log+log+exp” (equation (2)), “log+exp+exp” (equation (3)) are used:

\[
Y(t) = a\ln\left(1 + \frac{t}{\tau_a}\right) - b\exp\left(-\frac{t}{\tau_b}\right) + c + v_{at}
\]

(1)

\[
Y(t) = a\ln\left(1 + \frac{t}{\tau_a}\right) + b\ln\left(1 + \frac{t}{\tau_b}\right) - c\exp\left(-\frac{t}{\tau_c}\right) + d + v_{at}
\]

(2)

\[
Y(t) = a\ln\left(1 + \frac{t}{\tau_a}\right) - b\exp\left(-\frac{t}{\tau_b}\right) - c\exp\left(-\frac{t}{\tau_c}\right) + d + v_{at}
\]

(3)

where \(Y\) is the displacement (m); \(t\) is the time (day); \(\tau_a, \tau_b,\) and \(\tau_c\) are characteristic relaxation times (day); \(a, b, c,\) and \(d\) are constants. \(v_{at}\) represents linear trend terms. We estimate \(v_{at}\) following Tobita (2016) on the basis of
The regression analyses are divided into two processes for the comparison with the RNN results. First, we estimate the relaxation times ($\tau_a$, $\tau_b$, and $\tau_c$) from the first 90% of the time-series data by a non-linear least-square method, which corresponds to the learning process of the RNN. Second, using the above relaxation times, we estimate the constants ($a$, $b$, $c$, and $d$) from the initial 365-days data immediately after the 2011 Tohoku-oki earthquake, with the extension of the regression line treated as the forecast for the 23 comparison points. In addition, because Tobita (2016) proposed the reference values of the relaxation times ($\tau_{a0}$, $\tau_{b0}$, and $\tau_{c0}$) for the Tohoku region, we also use those values to estimate the constants ($a$, $b$, $c$, and $d$) for the above second process as trials (Supporting Information).

### 3. Results

The RNN forecasts the postseismic deformation for the 38 test points since the 365-day periods had elapsed from the 2011 Tohoku-oki earthquake. The regression analyses fit the time-series data using the initial 365 days immediately after the 2011 Tohoku-oki earthquake, with the extension of the regression line treated as the forecast for the 23 comparison points. Examples of the forecast results from these methods are compared with actual observations from a comparison point (950233) in Figure 2. The RNN outputs the best result for the forecasted E–W displacement, whereas the “log+exp” regression analysis outputs the best result for the forecasted N–S displacement. The results of several other stations (950170, 950178, 950182) are shown in Supporting Information (Figures S1–S3). The RNN result generally wins. Occasionally one of the regression result wins, although which type of the regression wins varies with the data.

We use the mean absolute errors between the forecasted values and actual data as an indicator of the forecasting accuracy. The average values of the mean absolute errors for each forecasting method are listed in Table 2, with the accuracy of the RNN results far exceeding those of regression analyses. The poor accuracies of the relatively complex (“log+log+exp” and “log+exp+exp”) regression results are due to overfitting to the initial 365 days. Furthermore, the method that yields the most accurate forecasting results at each comparison point is shown in Figures 3(a) and (b). The RNN results for the E–W displacements exceed those of the regression analyses at 17 comparison points (73.9% winning rate). The RNN results for the N–S displacements exceed those of the regression analyses at 18 comparison points (78.3% winning rate). Supporting Information (Figure S4 and Table S1) further presents the results with the regression equations using the reference relaxation times provided by Tobita (2016). The RNN results for the E–W displacements exceed those of the regression analyses at 22 comparison points (95.7% winning rate). The RNN results for the N–S displacements exceed those of the regression analyses at 18 comparison points (78.3% winning rate). The overfitting of the “log+log+exp” and “log+exp+exp” regression models are restrained in comparison with the above results from our original relaxation times, probably because the reference relaxation times provided by Tobita (2016) are universal values throughout eastern Japan.

In additional experiment without the offset correction for the learning process, we did not find evidence of the noise immunity improvement. On the contrary, the average values of the mean absolute errors got worse than the RNN results of the main experiment (Table 2): 0.017 m and 0.006 m for the E–W and N–S displacements, respectively. Therefore, we will discuss the results of the main experiment.

Besides, we also examined the vertical displacements of the GNSS data but the RNN forecasts did not work well. It was likely due to low signal-to-noise (S/N) ratio of the vertical displacements. In order to confirm this
point, we approximately estimate the S/N ratio for the E–W, N–S, and vertical displacements. The mean values of the postseismic deformation signals for the 191 stations are 0.93 m, 0.40 m, and 0.11 m for the E–W, N–S, and vertical displacements, respectively. The mean values of noise for the 191 stations, where we regard a standard deviation of the last 30 time-series data as noise amount, are 2.5 mm, 1.8 mm, and 8.8 mm respectively for the E–W, N–S, vertical displacements. Thus, the S/N ratio for the vertical displacements, about 13, is one order of magnitude lower than that for the E–W displacements, about 370, and that for the N–S displacements, about 220. It implies that S/N ratio should be larger than $10^2$ for the RNN forecasts.

### 4. Discussion

We focus on the spatiotemporal characteristics of the RNN forecast. Systematic deviations in the actual GNSS observations from the RNN forecasted values suggest a change in the physical signals that shaped the observations. The spatiotemporal evolution of the differences between the GNSS data and RNN forecast is shown in Figure 4. A westward deviation in the GNSS observations from the RNN forecasted values is observed along the eastern coast of northeastern Japan over the length of the time series, especially in Figure 4(e)–(g). This deviation may represent spatiotemporal changes in both the postseismic deformation and state of the plate boundary fault.

A physical interpretation is as follows. Postseismic deformation is generally controlled by two primary mechanisms: transient afterslip of the mainshock fault and viscoelastic relaxation of the asthenosphere (e.g., Wang et al., 2012). We find that the zone with the westward deviation in Figure 4(e)–(g) well corresponds to north–south extension of the near-coast afterslip area (almost 37.5°–38°N and 38.5°–40°N) from April 23, 2011, to December 10, 2011, estimated by Iinuma et al. (2016). It may suggest that the westward deviation represented the decay of the afterslip and dominance of the viscoelastic relaxation.

**Table 2**

<table>
<thead>
<tr>
<th>Direction</th>
<th>RNN (m)</th>
<th>“log+exp” (m)</th>
<th>“log+log+exp” (m)</th>
<th>“log+exp+exp” (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E–W</td>
<td>0.012</td>
<td>0.038</td>
<td>0.611</td>
<td>0.160</td>
</tr>
<tr>
<td>N–S</td>
<td>0.005</td>
<td>0.026</td>
<td>0.431</td>
<td>0.086</td>
</tr>
</tbody>
</table>

**Figure 3.** Comparison of the RNN and regression analysis forecast results at the observation stations. The red squares denote points where the RNN yielded the best accuracies, and the triangles denote where one of the regression analyses (the green color represents the “log+exp” model; the violet color represents the “log+exp+exp” model) yielded the best accuracies. The six digits indicate the locations of several stations in this study (Figure 2, Figure 4, and Supporting Information). (a) E–W displacement. (b) N–S displacement.
Figure 4. Spatiotemporal evolution of the differences between the GNSS data and RNN forecast (data minus forecast) for the E–W displacements. (a) Forecast start day to December 31, 2012. (b) January 1 to December 31, 2013. (c) January 1 to December 31, 2014. (d) January 1 to December 31, 2015. (e) January 1 to December 31, 2016. (f) January 1 to December 31, 2017. (g) January 1 to December 31, 2018. (h) Temporal differences (data minus forecast) at an observation station (950170; 141.798°E, 39.253°N). The blue dotted line indicates the zero value.
The temporal evolution of the deviation at a test point (950170; 141.798°E, 39.253°N) indicates that this westward deviation began in mid-2013, as shown in Figure 4(h). This point is consistent with a recent study which proposed stress field changed about two years after the 2011 Tohoku-oki earthquake on the basis of focal mechanism transition of low-frequency earthquakes (Oikawa et al., 2019). Our RNN forecasting approach therefore allows us to detect the spatiotemporal characteristics associated with the evolution of postseismic signals.

5. Conclusions

We adopt a machine learning algorithm, the recurrent neural network, to assess the feasibility of forecasting the postseismic deformation following the 2011 Tohoku-oki earthquake. We obtain high-accuracy results compared with traditional regression analyses. We can extract detailed information on the evolution of the postseismic signals from the spatiotemporal distribution of the differences between the time-series data and the forecasted values.

References


Acknowledgments

We thank Masaru Nakano and Takahiko Uchide for insightful discussions. Comments by Masayuki Kano and Jeff Freymueller were fruitful. The figures were made using Generic Mapping Tools (Wessel et al., 2013). We used the daily displacement solutions from the GEONET array (available at http://datahouse1.gsi.go.jp/terras/terras_english.html), provided by the Geospatial Information Authority of Japan (GSJ). We acknowledge financial support from the Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant Numbers JP16H06477 and JP19K04036).