MACHINE LEARNING MODELS FOR EARTHQUAKE NUMBER FORECASTING IN THE KANTO REGION

H. Huang[(1)], H. Miyake[(2)], H. Tsuruoka[(3)]

[(1)] *Graduate Student, Earthquake Research Institute, University of Tokyo, Tokyo, Japan, huang-hanyuan510@g.ecc.u-tokyo.ac.jp*
[(2)] *Associate Professor, Earthquake Research Institute, University of Tokyo, Tokyo, Japan, hiroe@eri.u-tokyo.ac.jp*
[(3)] *Associate Professor, Earthquake Research Institute, University of Tokyo, Tokyo, Japan, tsuru@eri.u-tokyo.ac.jp*

Abstract

Increases in the size and availability of seismology datasets have resulted in the development of machine learning techniques. These techniques are accelerating the identification of patterns in the spatial and temporal distribution of seismicity. In this study, we implemented machine learning techniques to analyze data from past earthquakes to forecast the number of earthquakes on a given day in the Kanto region. We used the Kanto region earthquake catalog produced by the Japanese Meteorological Agency (JMA) with data from 1 January 1999 to 31 December 2019. We interpreted statistical data from the earthquake catalog, including summation, mean, median, and maximum values of earthquake magnitudes and number in various time windows (1, 3, 5, and 7 days, 1, 2, and 3 months, and 1 year) before the forecast target day. The machine learning models incorporated multiple linear regression (MLR) and a nine-layer recurrent neural network (RNN), containing dense layers, a convolutional layer, and long short-term memory layers. Considering that different patterns may exist at different depths and magnitudes, we built distinct datasets containing earthquakes with JMA magnitudes (M_J) of $\geq$2.0, $\geq$3.0, and $\geq$4.0, and depths in the ranges of 0–30 (0$\leq$depth$\leq$30), 0–100 (0$\leq$depth$\leq$100), and 30–100 (30<depth$\leq$100) km. The ratio of training, validation, and test sets were 6:2:2. Both the MLR and RNN models performed well in the 0–100 km range; the catalog of earthquakes $M_J \geq 2.0$ are almost linearly related to time, except in December 2016 when a shallow M_J 6.3 earthquake occurred. However, while the RNN model remained well fitting in the 0–30 km dataset, the MLR model began to deviate following the M_J 6.3 earthquake. With increases in magnitude, we found it is harder to forecast accurate earthquake numbers. Still, when cumulating the results per day, we found that the long-term forecasting output from the RNN model was more accurate than the epidemic-type aftershock sequence (ETAS) model in all $M_J \geq 3.0$ sets and $M_J \geq 4.0$ sets. Conversely, the MLR model was instable which output low deviation results in some datasets, but fail to output satisfactory results in the other, with the forecasted spike always delayed by 1 day compared to the observed value. To summarize, we applied two machine learning models. We evaluated their results against the established ETAS model to forecast earthquakes' number in the Kanto region on a given day. Although short-term, high magnitude earthquake number forecasting currently appears impossible, our models consistently gave more accurate results for long-term forecasts than the ETAS model.

Keywords: earthquake number forecasting, machine learning, Kanto region, long short-term memory.

1. Introduction

Earthquakes are among the most destructive natural disasters, resulting in concerted efforts to determine their causes and accurately forecast their occurrence. The opinion is divided over whether earthquakes can[1][2] or cannot[3] be accurately predicted. However, researchers have identified occurrence patterns and developed models, such as the Gutenberg-Richter law[4]. The models show the relationship between magnitude and number and the epidemic-type aftershock sequence model (ETAS model)[5], which offers detailed descriptions of aftershock sequences and general seismicity[6]. Japan is an earthquake-prone country that obtains seismic data through national networks, which can then be analyzed using machine learning. Machine learning is a set of related techniques that extracts information directly from data using well-defined optimization rules, which has recently drawn attention due to its wide ranging success in various fields[7]. As solid earth geoscience is both a data-driven and model-driven field by which some researchers have argued that machine learning is likely to play a key role[8]. Considering the success of application of machine learning in seismology, we deem this technique can also be used to analyze recorded earthquake catalog for future earthquake forecasting. Predictions should specify the time window, the spatial window, the magnitude window, the author's level of confidence in the prediction, and the likelihood of the earthquake happening[9]. This study aimed to forecast the number of occurrence of earthquakes with Japanese Meteorological Agency (JMA) magnitudes (M_J) above a defined value (e.g., $M_J \geq 2.0$) on a given day in the Kanto region (longitude 34.5–37.0, latitude 138.5–141.5[10]) through analyzing previous earthquakes using machine learning. Our approach included multiple linear regression (MLR) and a recurrent neural network (RNN). We found that, compared to the ETAS model, our RNN model gave more accurate results in forecasting the number of earthquakes.

2. Dataset

2.1 Earthquake catalog

The Kanto region is a geographical area of Honshu, the largest island of Japan[11], and the main source of data for the Kanto region earthquake catalog is the JMA. In this region, between 1 January 2000 to 31 December 2019, 70,394 earthquakes with $M_J \geq 2.0$ and depths $\leq$100 km were recorded (after TSEIS[12]). Fig. 1**a** shows their spatial and temporal distributions. As shown in Fig. 1**b**, the 2011 Tohoku earthquake (M_J 9.0) increased earthquake number and deviation from the usual linear distribution of further seismicity until 2016. Considering periodic seismic activity[13][14], statistical information was extracted from the raw earthquake catalog, including means, medians, and maximum values for the magnitudes and earthquake number in various time windows, including 1, 3, 5, or 7 days, 1, 2, or 3 months, or 1 year as the features for machine learning's training. "Features" and "labels" are two integral concepts in machine learning[15]. By learning the patterns between features and labels in a dataset, machine learning models can generate functions to describe their relationship. In this study, the features were the statistical data in different time windows before the forecast target day, and the label was the earthquake number of the target day.

2.2 training, validation, and test sets

The training, validation, and test sets were in a ratio of 6:2:2. The training sets contained earthquake information from 1 January 2000 to 31 December 2012; the validation sets from 1 January 2013 to 31 December 2015; and the test sets from 1 January 2016 to 31 December 2019. The most prominent increase in the number of earthquakes happened following the 2011 Tohoku earthquake, which was included in the training sets. Several other moderate increases are observed in the validation sets in early 2012 and 2013, and another relative anomaly was in December 2016, when an M_J 6.3 earthquake occurred in the Kanto region.

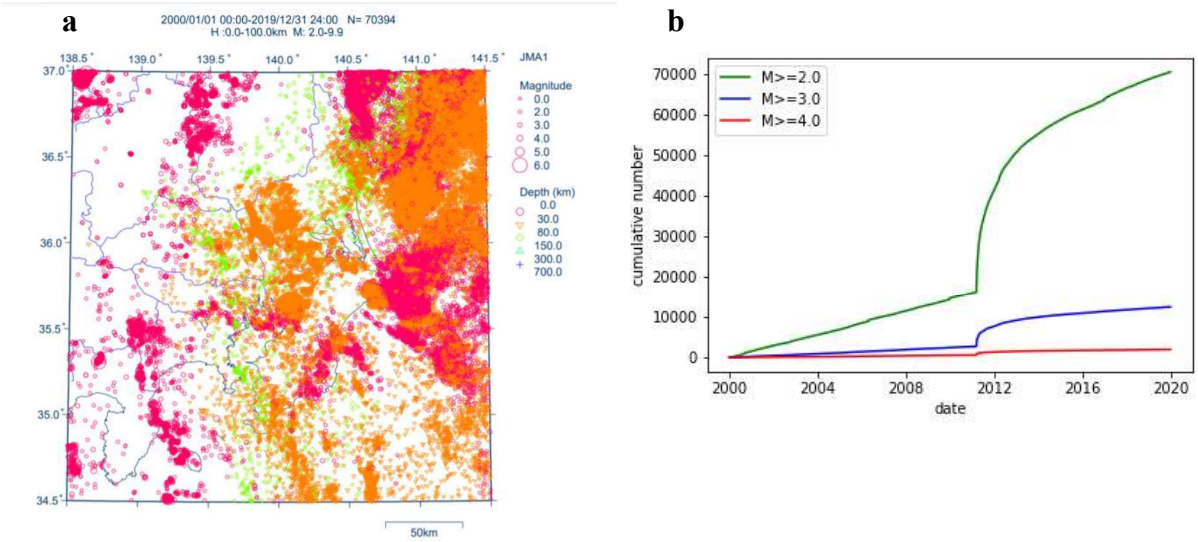

Fig. 1**a** Earthquake distribution with $M_J \geq 2.0$ and depth ≤ 100 km from 1 January 2000 to 31 December 2019 in Kanto region (after TSEIS). **b** Cumulative number of earthquake with JMA magnitude greater or equals to 2.0, 3.0, 4.0 respectively from 1 January 2000 to 31 December 2019 in Kanto region

3. Machine learning techniques for earthquake forecasting

Machine learning is a set of related techniques that extract information directly from data using well-defined optimization rules[7]. In this study, we mainly used two models: MLR and RNN(which is called LSTM model in following parts).

3.1 MLR

MLR is a statistical technique that uses several explanatory variables to predict a response variable's outcome. Its goal is to describe the linear relationship between features and labels through mathematical functions. MLR is different from simple linear regression by having multiple variables in its equation. The whole equation is as follows:

$$y = \beta_1 x_1 + \beta_1 x_1 + ... + \beta_n x_n + C \tag{1}$$

where y represents the output, x_n are the features, β_n are the calculated coefficients, and C is a constant.

3.2 RNN

RNN is a class of artificial neural networks (ANN). Compared with traditional ANN, it allows models to use their internal memory to process variable length sequences of inputs, and outputs result referred to not only the features but also the past outputs. RNN is a type of neural network architecture containing fully recurrent, long short-term memory (LSTM), Independently RNN, etc. In the model of this research, LSTM is being applied.

3.3 LSTM

LSTM is an artificial RNN architecture[16] used in the field of deep learning. LSTM units are composed of a cell, an input gate, an output gate and a forget gate. The cell remembers values at arbitrary time intervals, and the three gates regulate the flow of information into and out of the cell. Fig. 2 illustrates one type of inner structures of a LSTM neuron.

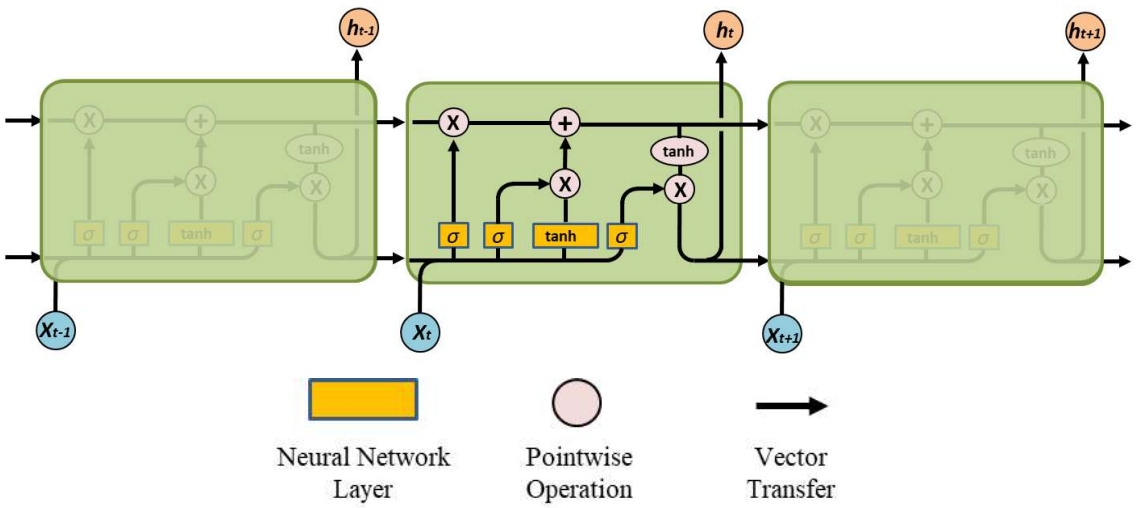

Fig. 2 – Repeating module in LSTM containing four interacting layers

where X_t and h_t represent the temporal input and output, respectively, σ represents a sigmoid layer, *tanh* represents a *tanh* layer, Pointwise Operations include forget and output gates, and the arrow line represents an internal mathematical function. Given that large earthquakes always have a strong influence on future seismicity and their effects decrease as time passes, we chose LSTM to be applied in our model.

3.4 Convolutional Neural Network(CNN)

CNN is a class of deep neural networks usually applied to analyzing images[17] by converting the image into a matrix and analyzing its features. In this study, however, we used it to convolute the one-dimensional (1-d) earthquake seismicity data, which allowed us to accelerate the training process rather than apply a CNN to two-dimensional images.

3.5 The LSTM neural network structure in this research

The neural network, which contained 67,761 parameters that we built in this research, comprises a 1-d convolutional layer, one max-pooling layer, two fully connected layers, and four LSTM layers. The structure is shown in Fig. 3. The CNN layer and max-pooling layer extract the features' significance and accelerate the training process by simplifying the data.

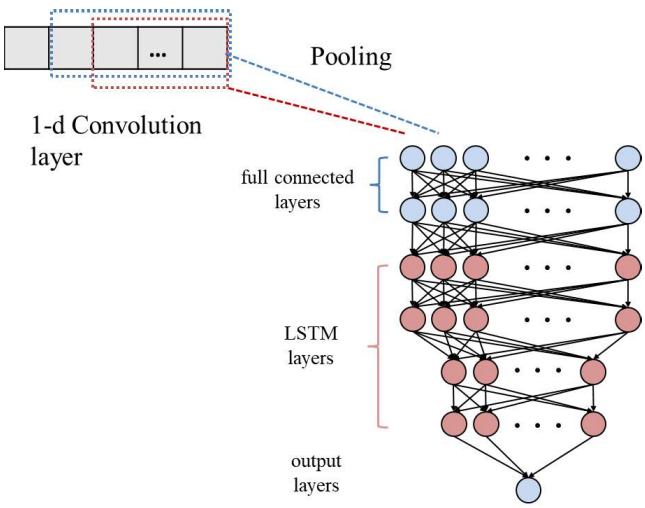

Fig. 3 – Structure of the recurrent neural network applied for number forecasting.

4. ETAS model

The ETAS model, developed by Ogata[5], calculates the short-term probability of an earthquake of magnitude M_Z or larger in a near future period.

$$\lambda_\theta(t \mid H_t) = \mu + \sum_{\{i:0 \le t_i \le t\}} \frac{K_0 e^{\alpha(M_i - M_z)}}{(t - t_i + c)^p} \tag{2}$$

where $t = 0$ is the mainshock time from the aftershock observation, M_z represents the reference magnitude (i.e., the mainshock magnitude) of the earthquake to be treated in the dataset, M_i and t_i indicate the magnitude and the occurrence time of the ith earthquake, respectively, and H_t represents the occurrence series of earthquakes (t_i, M_i) before time t. The parameter set θ consists of five elements (μ, K_0, c, α, p)[6].

5. Results and evaluation

The results are categorized into nine sections, including $M_J \geq 2.0$, ≥ 3.0, and ≥ 4.0, and depths of 0–30, 30–100, and 0–100 km, and the results are shown in ascending order of M_J. Because the ETAS model is not designed to analyze earthquake catalog sizes in excess of 17,777, we decided to focus on comparing the $M_J \geq 3.0$ and ≥ 4.0 sets. In some sets the MLR output was a negative value, and in these cases the value was set as zero.

5.1 $M_J \geq 2.0$ at depths of 0–100, 0–30, and 30–100 km

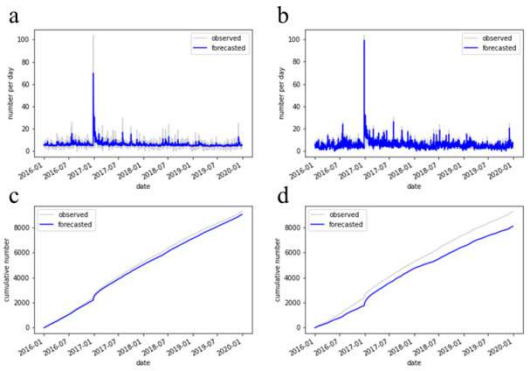

Fig. 4 – Observed and forecasted number of earthquakes in the Kanto region with $M_J \geq 2.0$ at depths between 0–100 km. **a** Daily observed and forecasted numbers of earthquakes from LSTM. **b** Daily observed and forecasted numbers from MLR. **c** Cumulative number from LSTM. **d** Cumulative number from MLR.

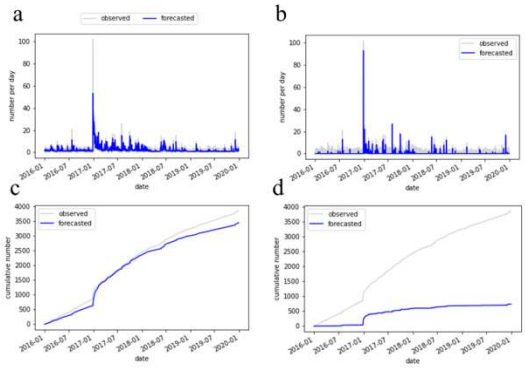

Fig. 5 – Observed and forecasted number of earthquakes in the Kanto region with $M_J \geq 2.0$ at depths between 0–30 km. **a** Daily observed and forecasted numbers of earthquakes from LSTM. **b** Daily observed and forecasted numbers from MLR. **c** Cumulative number from LSTM. **d** Cumulative number from MLR.

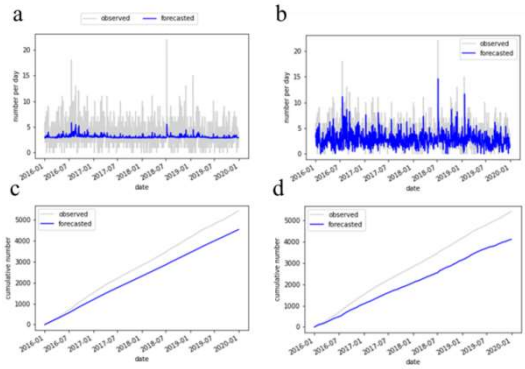

Fig. 6 –Observed and forecasted number of earthquakes in the Kanto region with $M_J \geq 2.0$ at depths between 30–100 km. **a** Daily observed and forecasted numbers of earthquakes from LSTM. **b** Daily observed and forecasted numbers from MLR. **c** Cumulative number from LSTM. **d** Cumulative number from MLR.

5.2 $M_J \geq 3.0$ at depths of 0–100, 0–30, and 30–100 km

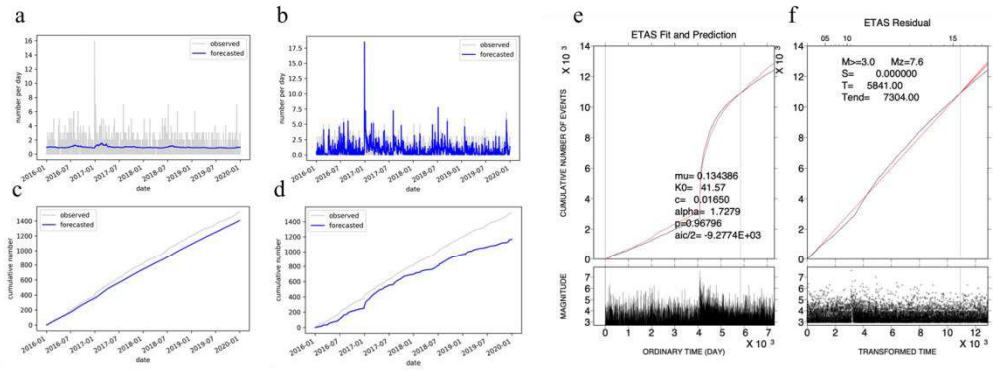

Fig. 7 – Observed and forecasted earthquake number of earthquakes in the Kanto region with $M_J \geq 3.0$ at depths between 0–100 km. **a** Daily observed and forecasted numbers of earthquakes from LSTM. **b** Daily observed and forecasted numbers from MLR. **c** Cumulative number from LSTM. **d** Cumulative number from MLR. **e** ETAS fit and prediction with ordinary time. **f** ETAS residual with transformed time.

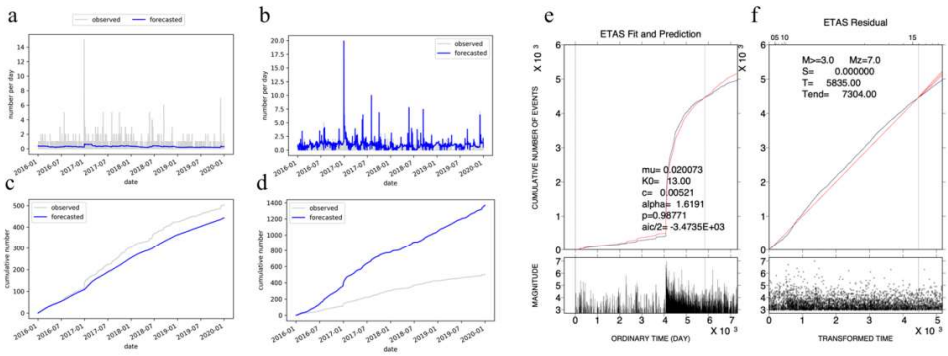

Fig. 8 – Observed and forecasted earthquake number of earthquakes in the Kanto region with $M_J \geq 3.0$ at depths between 0–30 km. **a** Daily observed and forecasted numbers of earthquakes from LSTM. **b** Daily observed and forecasted numbers from MLR. **c** Cumulative number from LSTM. **d** Cumulative number from MLR. **e** ETAS fit and prediction with ordinary time. **f** ETAS residual with transformed time.

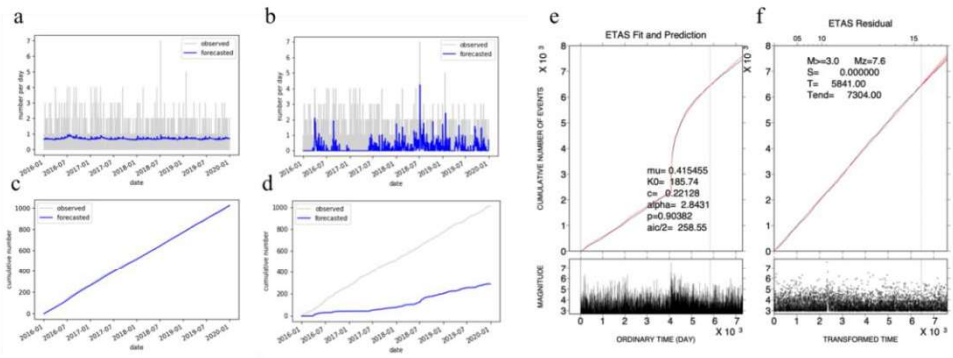

Fig. 9 – Observed and forecasted earthquake number of earthquakes in the Kanto region with $M_J \geq 3.0$ at depths between 30–100 km. **a** Daily observed and forecasted numbers of earthquakes from LSTM. **b** Daily observed and forecasted numbers from MLR. **c** Cumulative number from LSTM. **d** Cumulative number from MLR. **e** ETAS fit and prediction with ordinary time. **f** ETAS residual with transformed time.

5.3 $M_J \geq 4.0$ at depths of 0–100, 0–30, and 30–100 km

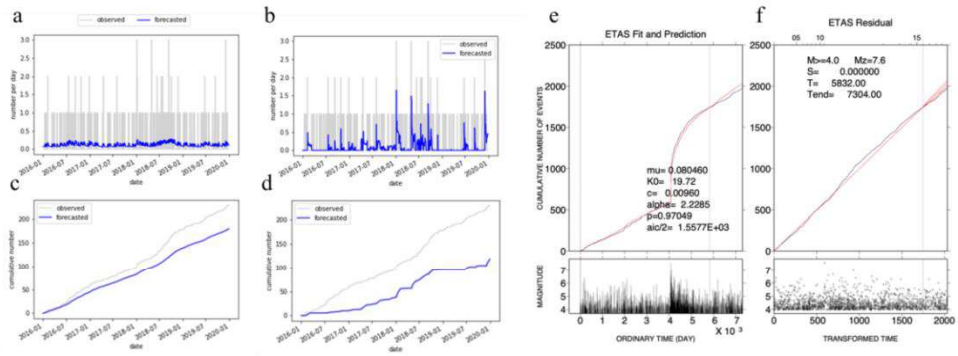

Fig. 10 – Observed and forecasted earthquake number of earthquakes in the Kanto region with $M_J \geq 4.0$ at depths between 0–100 km. **a** Daily observed and forecasted numbers of earthquakes from LSTM. **b** Daily observed and forecasted numbers from MLR. **c** Cumulative number from LSTM. **d** Cumulative number from MLR. **e** ETAS fit and prediction with ordinary time. **f** ETAS residual with transformed time.

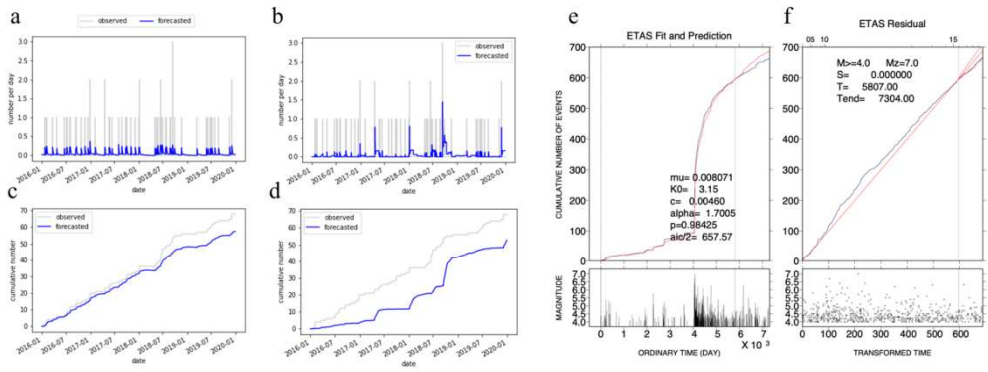

Fig. 11 – Observed and forecasted earthquake number of earthquakes in the Kanto region with $M_J \geq 4.0$ at depths between 0–30 km. **a** Daily observed and forecasted numbers of earthquakes from LSTM. **b** Daily observed and forecasted numbers from MLR. **c** Cumulative number from LSTM. **d** Cumulative number from MLR. **e** ETAS fit and prediction with ordinary time. **f** ETAS residual with transformed time.

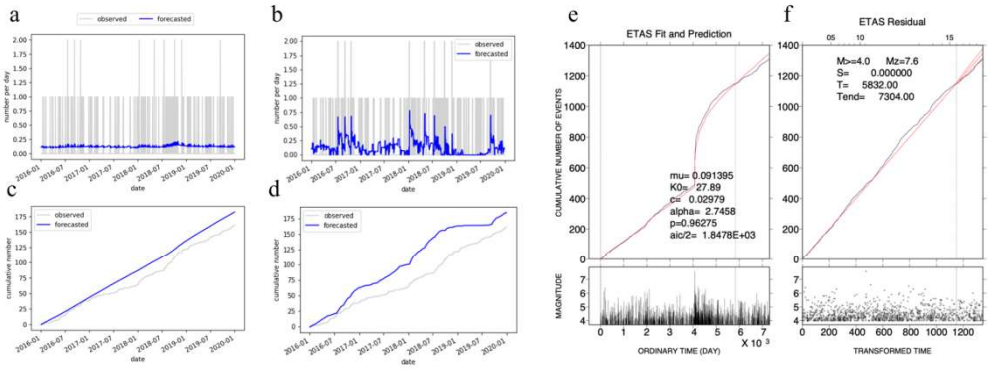

Fig. 12 – Observed and forecasted earthquake number of earthquakes in the Kanto region with $M_J \geq 4.0$ at depths between 30–100 km. **a** Daily observed and forecasted numbers of earthquakes from LSTM. **b** Daily observed and forecasted numbers from MLR. **c** Cumulative number from LSTM. **d** Cumulative number from MLR. **e** ETAS fit and prediction with ordinary time. **f** ETAS residual with transformed time.

5.4 Evaluation

This section shows the performance of the three models in the test sets containing tables and figures where the observed value and the results yielded from different models are compared. The results with the lowest deviation in the tables are in bold. This study mostly focus on the long-term forecast from 1 January 2016 to 31 December 2019 as showed in Fig. 13, but the results and comparison are also illustrated to verify the usability of models and the influence of the 2011 Tohoku earthquake.

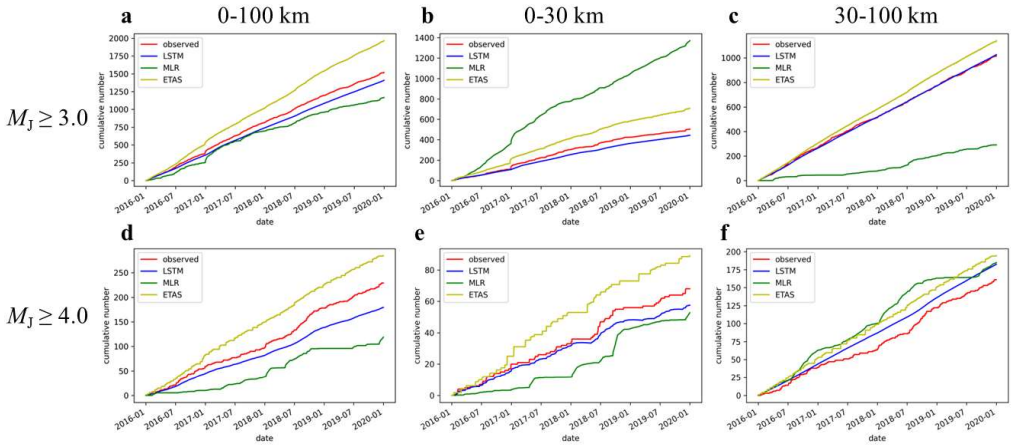

Fig. 13**a–f** Comparison of the cumulative number forecasted in the Kanto region by the three models from 1 January 2016 to 31 December 2019, in various magnitude and depth sets. The red, blue, green, and yellow lines represent the observed number results from LSTM, MLR, and ETAS models, respectively.

Table 1 – Cumulative earthquake number from 1 January 2016 to 31 December 2019

magn-itude	depth (km)	obser-ved	LSTM			MLR			ETAS		
			foreca-sted	deviation		foreca-sted	deviation		foreca-sted	deviation	
				value	%		value	%		value	%
$M_J \geq 2.0$	0–100	9292	9058.8	−233.2	**2.5%**	8109.7	−1182.3	12.7%	-	-	
	0–30	3871	3454.3	−416.7	**10.8%**	739.1	−3131.9	80.9%	-	-	
	30–100	5421	4539.0	−882.0	**16.3%**	4114.5	−1306.5	24.1%	-	-	
$M_J \geq 3.0$	0–100	1519	1407.8	−111.2	**7.3%**	1166.4	−352.6	23.2%	1966.7	447.7	29.5%
	0–30	504	443.6	−60.4	**12.0%**	1370.6	866.6	171.9%	708.0	204.0	40.5%
	30–100	1015	1026.1	11.1	**1.1%**	290.5	−724.5	71.4%	1142.4	127.4	12.6%
$M_J \geq 4.0$	0–100	229	179.5	−49.5	**21.6%**	118.8	−110.2	48.1%	288.6	59.6	26.0%
	0–30	68	57.5	−10.5	**15.4%**	52.8	−15.2	22.4%	92.1	24.1	35.4%
	30–100	161	182.4	21.4	**13.3%**	184.8	23.8	14.8%	197.4	36.4	22.6%

The above models also forecasted earthquake numbers for other time ranges to estimate the 2011 Tohoku earthquake's influence. However, due to space limitations, we are only showing the test sets' comparisons. The earthquake catalog used in the following figures ranges from 1 January 2000 to 10 March 2011, and from 11 March 2011 to 31 December 2019. The ratio of training, validation, and test sets is 6:2:2.

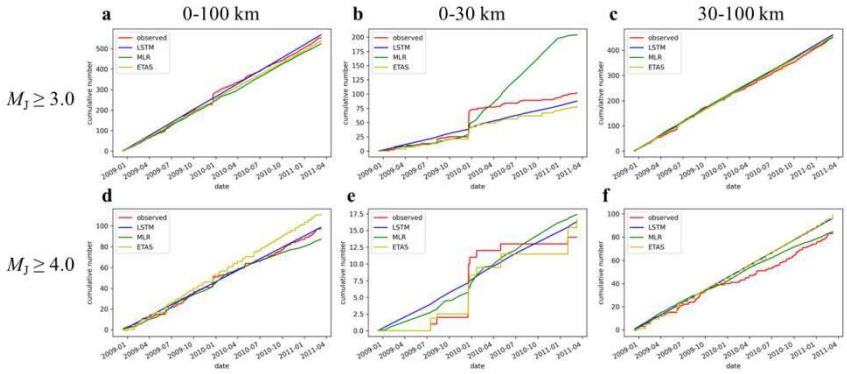

Fig. 14**a–f** Comparison of the cumulative number forecasted in the Kanto region by the three models from 13 December 2008, to 10 March 2011, in various magnitude and depth datasets.

Table 2 – Cumulative earthquake number from 13 December 2008 to 10 March 2011 (before the 2011 Tohoku earthquake)

magn-itude	depth (km)	observed	LSTM			MLR			ETAS		
			forecasted	deviation		forecasted	deviation		forecasted	deviation	
				value	%		value	%		value	%
$M_J \geq 2.0$	0–100	3332	3200.6	−131.4	**3.9%**	2847.8	−479.1	14.4%	-	-	
	0–30	831	648.1	−182.9	22.0%	889.5	58.5	**7.0%**	-	-	
	30–100	2501	2478.7	−22.3	**0.9%**	2463.3	−32.6	1.3%	-	-	
$M_J \geq 3.0$	0–100	553	567.7	14.7	2.7%	523.6	−29.4	5.3%	538.7	−14.2	**2.6%**
	0–30	102	87.5	−14.5	**14.2%**	204.0	102.0	100.0%	78.2	−23.7	23.3%
	30–100	451	461.8	10.8	2.4%	451.6	0.6	**0.1%**	458.1	7.0	1.57%
$M_J \geq 4.0$	0–100	97	98.7	1.7	**1.7%**	87.2	−9.8	10.1%	110.9	13.8	14.3%
	0–30	14	16.3	2.3	**16.6%**	17.4	3.4	24.1%	16.6	2.6	19.1%
	30–100	83	96.3	13.3	**16.0%**	84.6	1.6	1.9%	99.3	16.3	19.6%

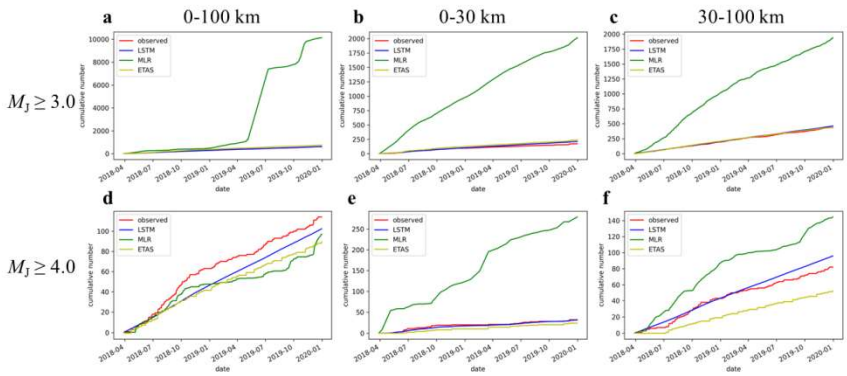

Fig. 15**a–f** Comparison of the cumulative number forecasted in the Kanto region by the three models from 13 December 2008, to 10 March 2011, in various magnitude and depth datasets.

Table 3 – Cumulative earthquake number from 28 March 2018 to 31 December 2019 (after the 2011 Tohoku earthquake)

magn-itude	depth (km)	obse-rved	LSTM			MLR			ETAS		
			foreca-sted	deviation		foreca-sted	deviation		foreca-sted	deviation	
				value	%		value	%		value	%
$M_J \geq 2.0$	0–100	3532	4017.7	485.7	**13.8%**	18320.9	14788.9	418.7%	-	-	
	0–30	1233	1340.9	107.9	**8.8%**	2525.4	1292.4	104.8%	-	-	
	30–100	2299	2114.9	−184.1	**8.0%**	9010.3	6711.3	291.9%	-	-	
$M_J \geq 3.0$	0–100	612	625.8	13.8	**2.3%**	10136.0	9524.0	1556.2%	754.6	142.6	23.3%
	0–30	175	213.6	38.6	**22.0%**	2016.9	1841.9	1052.5%	241.8	66.8	38.2%
	30–100	437	462.0	25.0	5.7%	1939.8	1502.8	343.9%	445.8	8.8	**2.0%**
$M_J \geq 4.0$	0–100	114	102.4	−11.7	**10.2%**	97.2	−16.8	14.8%	89.3	−24.7	21.6%
	0–30	32	31.4	−0.6	**1.8%**	279.0	247.0	772.0%	24.3	−7.7	24.1%
	30–100	82	95.9	13.9	**16.9%**	144.4	62.4	76.1%	52.3	−29.7	36.2%

5. Discussion

This section discusses the forecasted results from our machine learning models and the ETAS model over 4 years in the Kanto region. All three models are based on analyzing the number of occurrence and magnitude of earthquakes, but the machine learning models summarize the catalog by day. According to our results, we deem this distinction to be the cause of the ETAS model producing more accurate forecasts instantly following relatively large earthquakes, as shown in Fig. 13**e**. Conversely, for long-term forecasts, such as our MLR test sets, the LSTM model always outputs the most accurate results. It is worth noting, however, that although the ETAS model was more accurate at times, the deviations in results before and after set 311 from the ETAS and LSTM models with $M_J \geq 3.0$ and depth between 0–100 km were 2.6 and 2.0%, and 2.7 and 5.7%, respectively, which are quite close.

A large number of observed earthquakes in the catalog at low magnitude thresholds ($M_J \geq 2.0$) leads that the MLR and LSTM models perform relatively well in cumulative number forecasting. However, the MLR's performance is quite unstable, with notable deviation from the observed values in the $M_J \geq 2.0$, 30–100 km set, and the $M_J \geq 3.0$, 0–30 km, and 30–100 km sets, leading to consistently delayed spike forecasts by 1 day. These results suggest that the MLR model is unsuitable for ubiquitous seismicity forecasting. The LSTM model, however, is more robust and yields more stable forecasting values. The ETAS model can recognize the mainshock and influential aftershocks to produce forecasts for aftershock numbers by analyzing the background earthquake catalog, which is also why the ETAS model produces better short-term forecasting results in the shallow $M_J \geq 3.0$ set after the December 2016 M_J 6.3 earthquake. Table 1 shows that the LSTM can produce a more satisfactory daily earthquake number forecasting value. The main disadvantages of the LSTM model are: first, both the MLR and ETAS models can be described specifically by several mathematical functions, yet with 67,761 parameters, the LSTM cannot, and second, the time required for LSTM analysis is much longer than other options and requires attention to avoid overfitting. LSTM's deep neural network structure is also a reason why the LSTM model's forecasting is not as sensitive as the others following a large earthquake but still leads to an improved performance for long-term forecasts. Regarding the training sets, datasets from 1 January 2000 to 31 December 2012, were used to train the machine learning models, and the vacant part before 1 January 2016, was applied to validate in the fear of overfitting while the ETAS model analysis of the earthquake catalog from 1 January 2000 to 31 December 2015. Given that the LSTM model can forecast cumulative number more accurately using older earthquake catalog data, we conclude that periodicity exists in the temporal distribution of seismicity. It can be better understood using machine learning models.

6. Conclusions

Two machine learning models were applied to daily earthquake number forecasting in the Kanto region. With increased magnitude thresholds, the daily forecasting was inaccurate, but concerning long-term forecasts, our LSTM model always performed more accurately than the ETAS model in $M_J \geq 3.0$ and ≥ 4.0 datasets. This study showed that different characteristics of seismicity exist across various depths so that mathematical models have different offsets between forecasted and observed numbers at various range of depth, and, compared with traditional aftershock number forecasting methods, our machine learning models may provide more accurate results than established techniques in long-term earthquake analysis. These results may also support that periodicity exists in the temporal distribution of earthquakes, which machine learning models could better describe.

7. References

[1] Asim, K. M., Martínez-Álvarez, F., Basit, A., & Iqbal, T. (2017). Earthquake magnitude prediction in Hindukush region using machine learning techniques. *Natural Hazards*, **85**(1), 471-486.

[2] Brehm, D. J., & Braile, L. W. (1998). Intermediate-term earthquake prediction using precursory events in the New Madrid seismic zone. *Bulletin of the Seismological Society of America*, 88(2), 564-580.

[3] Geller, R. J., Jackson, D. D., Kagan, Y. Y., & Mulargia, F. (1997). Earthquakes cannot be predicted. *Science*, **275**(5306), 1616-1616.

[4] Gutenberg, B., & Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, **34**(4), 185-188.

[5] Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical association*, **83**(401), 9-27.

[6] Ogata, Y., & Tsuruoka, H. (2016). Statistical monitoring of aftershock sequences: a case study of the 2015 Mw7. 8 Gorkha, Nepal, earthquake. *Earth, Planets and Space*, **68**(1), 44.

[7] Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., & Gerstoft, P. (2019). Machine learning in seismology: Turning data into insights. *Seismological Research Letters*, **90**(1), 3-14.

[8] Bergen, K. J., Johnson, P. A., Maarten, V., & Beroza, G. C. (2019). Machine learning for data-driven discovery in solid Earth geoscience. *Science*, **363**(6433).

[9] Geller, R. J. (1997). Earthquake prediction: a critical review. *Geophysical Journal International*, **131**(3), 425-450.

[10] Ogata, Y., Imoto, M., & Katsura, K. (1991). 3-D spatial variation of b-values of magnitude-frequency distribution beneath the Kanto District, Japan. *Geophysical Journal International*, 104(1), 135-146.

[11] Nussbaum, L. F., & Roth, K. (2005). Japan encyclopedia. *Cambridge: Harvard University*.

[12] Tsuruoka, H. (1998). Development of seismicity analysis software on workstation. *Tech. Res. Rep. ERI, Univ. Tokyo*, *2*, 34-42.

[13] Utsu, T. (2001). *Seismology*, Kyoritsu Shuppan, ISBN978-4-320-04637-5 (in Japanese)

[14] Morgan, W. J., Stoner, J. O., & Dicke, R. H. (1961). Periodicity of earthquakes and the invariance of the gravitational constant. *Journal of Geophysical Research*, 66(11), 3831-3843.

[15] Blum, A. L., & Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial intelligence*, 97(1-2), 245-271.

[16] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, **9**(8), 1735-1780.

[17] Wu, J. (2017). Introduction to convolutional neural networks. *National Key Lab for Novel Software Technology. Nanjing University. China*, 5, 23.

