IOD and ENSO-Related Time Series Variability and Forecasting of Dengue and Malaria Incidence in Indonesia

Kurnianingsih
Department of Electrical Engineering
Politeknik Negeri Semarang
Semarang, Indonesia
ORCID ID 0000-0001-7339-7449

Anindya Wirasatriya
Department of Oceanography
Diponegoro University
Semarang, Indonesia
ORCID ID 0000-0003-1030-5126

Lutfan Lazuardi
Faculty of Medicine
Universitas Gadjah Mada
Yogyakarta, Indonesia
ORCID ID 0000-0001-5146-8162

Naoyuki Kubota
Graduate School of Systems Design
Tokyo Metropolitan University
Tokyo, Japan
ORCID ID 0000-0001-8829-037X

Nawi Ng
Institution of Medicine
University of Gothernburg
Gothernburg, Sweden
ORCID ID 0000-0003-0556-1483

***Abstract*—Dengue and malaria are mosquito-borne infectious diseases driven by climate change and have an endemic impact in tropical and subtropical regions of the world, particularly in South and South-East Asia. Most studies examined the effect of local climate variability, temperature, and precipitation on dengue and malaria incidence. Nevertheless, the effects of Indian Ocean Dipole (IOD) and El Niño Southern Oscillation (ENSO) on dengue incidence are still rarely discussed. Besides, the study of the influence of IOD and ENSO on malaria incidence remain unexplored. This paper examined the influence of IOD and ENSO on the interannual variability of dengue and malaria incidence in all Indonesia provinces using Pearson correlation. Historical time series data on dengue and malaria incidence in all Indonesia provinces from 2005 to 2018 were investigated for time series prediction using deep Long Short-Term Memory (LSTM).**

Keywords—dengue, malaria, IOD, ENSO, Deep LSTM

I. INTRODUCTION

Changes in human populations, human mobility, landscape, land use, biodiversity, and climate affect the emergence or re-emergence of infectious diseases. The intrinsic vulnerability of mosquitoes to weather and climate raises the likelihood that mosquito transmissions can be the most likely to be affected by climate change in infectious diseases. Mosquito vectors are widely spread all over the world, except in Antarctica. Malaria is a parasite infection spread by anopheline mosquitoes, which causes an estimated 219 million deaths worldwide, resulting in over 400,000 deaths per year [1]. Whilst, dengue is the most prevalent viral infection spread by Aedes mosquitoes and is at risk of contracting dengue with over 3.9 billion people in over 129 countries [1]. Increased climatic variables such as temperatures, wind, humidity, and precipitation will profoundly influence not only population dynamics of mosquito vectors, but also disease transmission dynamics [2]. Mosquito-borne diseases have become a primary global concern for human health, particularly in Indonesia, with the recent rise in the incidence of dengue and malaria.

Among the interannual global climate variabilities, Indian Ocean Dipole (IOD) and El Niño-Southern Oscillation (ENSO) bring the most impact on the Indonesian climate [3]. IOD events are driven by changes in the tropical Indian Ocean which were characterized by the sustained changes in the difference between normal sea surface temperatures in the tropical western and eastern Indian Ocean [4]. ENSO is a natural phenomenon of fluctuating ocean temperatures in the central and eastern equatorial Pacific Ocean, along with atmospheric changes [5]. Both ENSO and IOD events consist of two opposing phases and a neutral phase (N). When El Niño (EN) and IOD positive (IOD+) co-occurred, they cause anomalously hot and dry conditions [6]. Meanwhile, when La Niña (LN) and IOD negative (IOD−) events were concurrently occurred, the temperature decreased, and the precipitation increased significantly [7]. The previous study has been demonstrated the impact of ENSO and IOD on the environment within Indonesia [8].

Dengue and malaria dynamics are driven by complex interactions among hosts, vectors, viruses, and environmental and climatic factors. Several studies examined the influences of local and regional climate variation on dengue incidence [9-11] and malaria incidence [12, 13]. However, the role of Indian Ocean Dipole (IOD) and El Niño Southern Oscillation (ENSO) events on dengue and malaria incidence remain unexplored. This study analyses the effect of IOD and ENSO on dengue and malaria incidence and its forecasted trends in all provinces in Indonesia.

Unlike predictive modelling for regression, time series is a difficult predictive modelling problem. Existing regression methods for predictive modelling are poor in handling time series data. The methods are inadequate in modelling variable length and ignore the long-term dependencies. Deep learning, which has been widely employed in various recognition tasks [14, 15], holds a great potential in constructing end-to-end systems through increasing the depth of neural network. Studies in deep learning have been attempted to address the challenge of long-term dependencies.

Long Short-Term Memory (LSTM), introduced by Hochreiter and Schmidhuber, is a type of recurrent neural network (RNN) used in deep learning [16]. LSTM works well on complex problems unaddressed by previous recurrent network algorithms and can learn long-term dependencies. However, the scarce uses of LSTM in health care research have been limited to predicting life expectancy [17], healthcare usage based on medical records [18], and

978-1-7281-8741-9/20/$31.00 ©2020 IEEE

machine health monitoring [19]. This paper extents the use of LSTM to provide time series forecasting on dengue and malaria incidence in all Indonesian provinces using the deep LSTM architecture. This paper comprises four sections. Section 2 describes related works. Section 3 describes experimental setup. Section 4 presents numerical results and discussion. Finally, conclusion and future works are presented in Section 5.

II. Related Works

A. Influence of IOD and ENSO Indices on Dengue and Malaria Incidences

Climate change has influenced mosquito dynamics and the spread of mosquito-transmitted diseases. Most studies have revealed the influences of local and regional climate on dengue and malaria incidences [9-13]. However, the studies of the El Niño cycle in specific areas associated with changes in the risk of mosquito-transmitted diseases, such as malaria and dengue fever, are still few.

The study of the impact of La Niña and El Niño revealed the increase on dengue incidence and risk level in Bangkok [20]. A strong and important relationship between dengue interannual cycles and ENSO was identified, indicating the El Niño events were partly responsible for the recurrent outbreaks of dengue in northern Venezuela [21]. Dengue occurrence peaks during the warmer and drier years of El Niño had been prevalent more often, which indicated that ENSO was a regional climate driver over such a long-time period, due to local change in temperature and rainfall. Long-lead predictions of El Niño events can be incorporated in a dengue epidemic model to estimate large dengue epidemics in the province of El Oro, Ecuador, where dengue is hyper-endemic [22].

A strong association between the annual incidence of malaria and the ENSO was measured between 1988 and 1999 by the Southern Oscillation Index (SOI) in five South African countries [23], including Kenya [24]. Below normal malaria incidence associated with a negative SOI (El Niño) and above normal incidence with a positive SOI (La Niña), which contributes to dry and wet weather conditions, respectively. The influences of ENSO events on malaria epidemics also have been mapped and researched in South Asia [25, 26]. The current research revealed that there is a moderate to strong correlation between ENSO and malaria incidences particularly in eastern part of India [26]. The effect of ENSO and malaria epidemics in South America has been studied in 2002 [27].

The Indian Ocean Dipole (IOD) is another global climatic phenomenon that has a major effect on regional precipitation patterns and malaria incidences in the eastern part of the continent. Malaria epidemics in East African highlands are frequently related to the effects of ENSO events. Nonetheless, there are other factors linked to malaria risk and increased concern in the effects of the IOD on East African precipitation [28, 29].

According to the Central of Data and Information, Ministry of Health of the Republic of Indonesia, dengue and malaria outbreaks occur in Indonesia every year [30]. To the best of our knowledge, there is no published research on the influences of IOD and ENSO global climate indices on mosquito-borne diseases in Indonesia at the time of writing this article. Our study will examine the influence of IOD and ENSO indices on the dengue and malaria in all Indonesian provinces from 2005 until 2018.

B. Forecasting on Dengue and Malaria Incidences

Several techniques have been used to forecast the incidences of dengue and malaria. Linear regression prediction on probability of a dengue outbreak within 24 hours in Kaohsiung City, Taiwan was deployed using logistic regression [31]. Similarly, linear regression was used to forecast the age of dengue virus-infected mosquitoes to prevent the spread of dengue disease [32]. Dengue prediction model using k-Nearest Neighbor was developed to indicate the dengue movement patterns in Trinidad [33]. Trend of dengue infection in Malaysia (2000-2010) was developed using regression lines [34]. Ordinary Least Square regression was used to predict dengue cases in Malang, Indonesia [35].

Time series analysis is a common data type occurring in many areas and applications, including health prediction. Time series analysis examines the movement of the chosen data points chosen to find trends or characteristics of interest. Several time series forecasting techniques have been developed for dengue [36-41]. Long Short-Term Memory (LSTM) has shown to achieve state-of-the-art results with time series or sequential data in many applications, including the medical domain [17, 42]. Similarly, an LSTM model was used to predict examination results given previous measurements [43]. Identification of unknown drug-drug interactions (DDIs) is of significant concern for improving drug consumption safety and efficacy. DeepCare is built on Long Short-Term Memory (LSTM), a recurrent neural network equipped with memory cells to read medical records, store previous illness history, infer current illness states and predict future medical outcomes [44].

III. Experimental Setup

A. Data Collection

This study used the national-level data of annual reported dengue and malaria incidence cases for 14 years during 2005-2018 in all provinces in Indonesia. Data were extracted from the annual health profile published by Central of Data and Information, Ministry of Health of the Republic of Indonesia [30]. To investigate the effect of IOD and ENSO on dengue and malaria incidences, we used the Dipole Mode Index (DMI) and the Oceanic Niño Index (ONI), respectively.

The ONI is the sea surface temperature (SST) anomalies in the Niño 3.4 region (5ºN-5ºS, 170ºW-120°W) with the basis period of 1971-2000. We obtained the ONI index from the National Weather Service, Climate Prediction Center [45], which can be downloaded at https://www.cpc.ncep.noaa.gov/data/indices/oni.ascii.txt.

The DMI is determined by the anomaly value of the SST gradient between the western tropical Indian Ocean (10°S-10°N and 50°E-70°E) and the south-eastern tropical Indian Ocean (10°S-0°N and 90°E-110°E) [4]. The anomaly is calculated relative to a monthly climatological seasonal cycle based on the years 1982-2005. The monthly climatology is linearly interpolated to determine weekly anomalies. We obtained the weekly DMI from the Ocean Observations Panel for Climate's webpage at https://stateoftheocean.osmc.noaa.gov/sur/ind/dmi.php [46].

In the present study, we averaged the ONI and DMI indices into yearly data to match with the yearly data of dengue and malaria incidences.

B. Correlation Coeffisien

The Pearson product-moment correlation coefficient is used to measure the degree of linear correlation between two variables, i.e., IOD or ENSO and dengue or malaria. Given a pair of random variables (X, Y), the formula Pearson's correlation which is expressed by ρ is:

$$\rho_{XY} = \frac{E\,[XY] - E[X]E[Y]}{\sqrt{E[X^2] - [E[X]]^2}\sqrt{E[Y^2] - [E[Y]]^2}} \quad (1)$$

Based on the coefficient, the strength of correlation was assessed as very weak (ρ=0 - 0.25), weak (ρ=0.25 - 0.50), strong (ρ=0.50 - 0.75) or very strong (ρ=0.75 – 1).

C. Long Short-Term Memory (LSTM)

The LSTM neural network is a type of a recurrent neural network (RNN). The RNNs use previous time events to inform the later ones. To predict the future incidences of dengue or malaria in each province, the model uses information about previous events. The RNNs work well if the problem requires only recent information to perform the present task. The RNN has limitation to model problem requiring long-term dependencies. The LSTM, on the other hand, was designed to learn long term dependencies. It remembers the information for long periods.

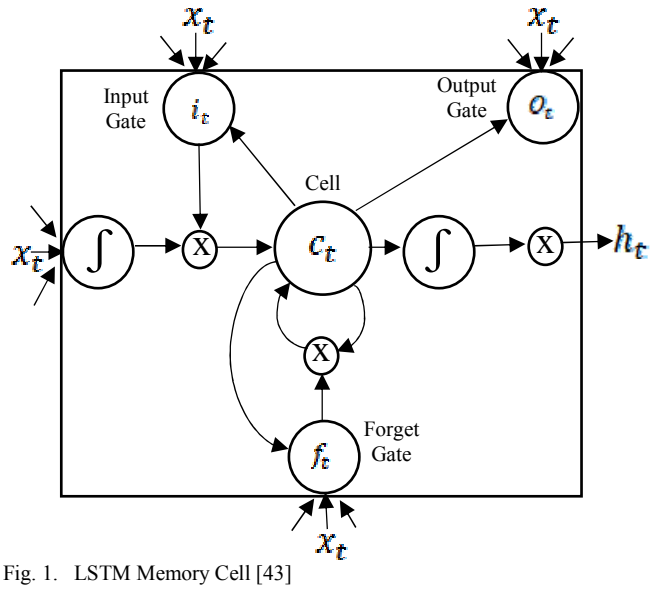

Fig. 1. LSTM Memory Cell [43]

In order to help the LSTM model to converge faster, we normalized our dataset using MinMaxScaler from scikit-learn library, so that all values are within range of 0 and 1. For each value in a feature, MinMaxScaler subtracts the minimum value in the feature and then divides by the range. MinMaxScaler preserves the shape of the original distribution. The scaler is fitted on the training set, and it is used to transform the unseen trade data on validation and test set.

The key to the LSTM memory cell are the gates, which consists of three gates [43], i.e., input gate (i), output gate (o), and forget gate (f), as shown in Fig. 1, whereas c denotes the cell. The input gate has a role in deciding the input's value to update the memory state, whereas the forget gate is to decide what information to discard from the cell. Both input gate and forget gate are used in updating the internal state. The output gate is a final limiter on what the cell actually outputs.

D. Performance Evaluation

We evaluated the LSTM model's forecasting error using the Mean Absolute Error (MAE), which was calculated as the mean of the forecast error, defined as the difference between the predicted values and the observed values. The Root-Mean-Square-Error (RMSE) was utilized to compare the prediction errors of various models using the same dataset.

$$MAE = \frac{\sum_{i=1}^{n}|A_i - F_i|}{n} \quad (2)$$

$$RMSE = \left(\frac{\sum_{i=1}^{n}(A_i - F_i)^2}{n}\right)^{1/2} \quad (3)$$

Both the MAE and RMSE measure the average prediction error in the same unit of variables, regardless of direction. Their values can range from 0 to ∞. The main difference between MAE and RMSE is that MAE is less sensitive to outliers. Like MAE, RMSE is used as a relative measure to compare different forecasting models developed based on the same data. Its value tends to be larger than the MAE. The RMSE is more appropriate than MAE for the comparison of model performance when the error is normally distributed [47].

IV. Results and Discussion

A. Influence of IOD and ENSO on Dengue and Malaria Incidence

We employed the Geopandas library in Python to generate the polychromatic maps illustrating the strength of correlation between IOD and ENSO with the dengue (Fig. 2. and Fig. 3.) and malaria (Fig. 4. and Fig. 5.) incidence in Indonesia. In general, Fig. 2. shows that the IOD had a negative correlation with the dengue incidence since only five provinces (i.e., Lampung, West Papua, Bangka Belitung Islands, South Sumatera, and West Kalimantan) have a very weak positive correlation. A negative correlation means that positive IOD tends to reduce the dengue incidence in most Indonesian provinces. However, among provinces with a negative correlation, only seven provinces (i.e., South Sulawesi, West Nusa Tenggara, Bali, Southeast Sulawesi, Yogyakarta, East Kalimantan, and Papua) have a strong correlation. On the other hand, ENSO generally has a positive correlation with the dengue incidence, which means that El Nino is associated with higher dengue incidence, as shown in Fig. 3. The strong positive correlations of ENSO and dengue incidence appear in three provinces (i.e., West Sumatera, Papua, and South Kalimantan), while weak to very weak negative correlations appear in eight provinces (i.e., Aceh, West Papua, Jakarta, Central Sulawesi, North Maluku, Lampung, North Sulawesi, and Banten). In the provinces with strong correlation, the dengue incidence shows a strong interannual variability related to IOD and ENSO.

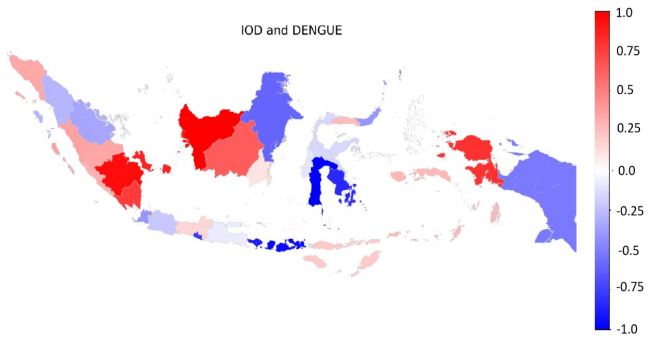

Fig. 2. Correlation of IOD on dengue incidence in all Indonesia provinces

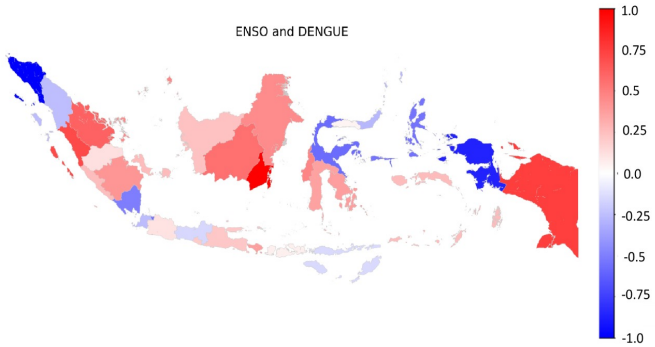

Fig. 3. Correlation of ENSO on dengue incidence in all Indonesia provinces

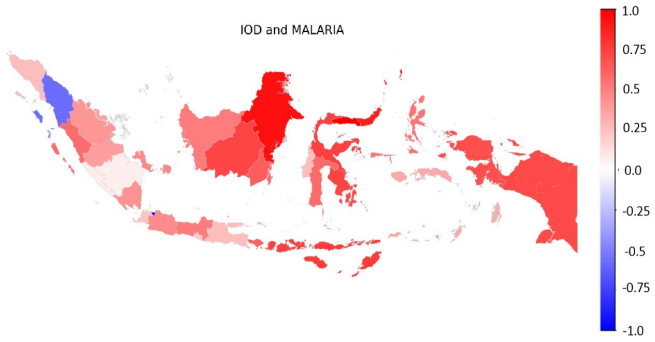

Fig. 4. Correlation of IOD on malaria incidence in all Indonesia provinces

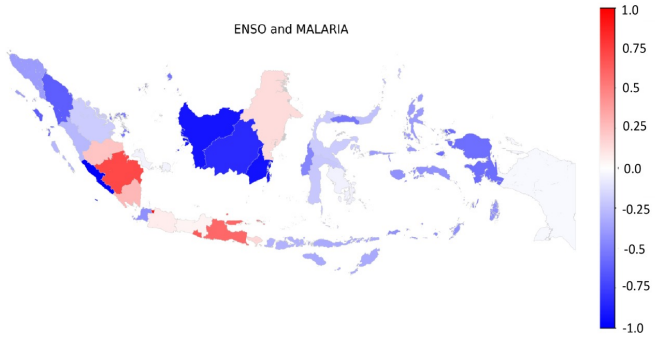

Fig. 5. Correlation of ENSO on malaria incidence in all Indonesia provinces

Fig. 4. and Fig. 5. show the correlation between the incidence of malaria with IOD and ENSO, respectively. In contrast with dengue incidence, the positive (negative) correlation of malaria incidence with IOD (ENSO) appears in most provinces in Indonesia. It shows that positive IOD and La Nina is associated with higher malaria incidence in Indonesia. In contrast, negative IOD and El Nino tend to suppress the malaria incidence in Indonesia. However, it is important to be noted that both correlations (IOD and ENSO) have only a very weak and weak correlation. It indicates that the malaria incidence in Indonesia has less interannual variation related to IOD and ENSO.

B. Time Series Forecasting using Deep LSTM

We developed a stacked LSTM on our small sequence prediction problem for univariate time series forecasting and made a single prediction. The LSTM model consists of a single series of observations and is required to learn from past observations as input to predict the next value in the sequence as an output observation. The stacked LSTM consists of multiple hidden layers. This study defined different models for each province in Indonesia on each dengue and malaria incidence. We used the linear activation function used for neurons in the dense layer, mean squared error (MSE) loss function, and fit the model using Adam optimizer.

The performance results of dengue incidence in all provinces are described in Table I. Table 1 shows the number of the hidden layer (H) of the LSTM model for each province. Each province may have different best LSTM model.

TABLE I. PERFORMANCE RESULTS OF DENGUE INCIDENCE IN ALL INDONESIA PROVINCES

Province	LSTM Model	Train Score		Test Score	
		RMSE	MAE	RMSE	MAE
Aceh	10H	560.8	471.6	801.5	620
North Sumatera	10H	1948.9	1401.2	2162.6	2162.5
West Sumatera	10H	839.5	654.7	911.8	894.5
Riau	10H	1221.3	880	650	530.5
Jambi	10H	610.4	523.5	281.7	270.5
South Sumatera	10H	897	811.6	618	440
Bengkulu	10H	451.7	348	861.8	773
Lampung	10H	2161.3	1673	1380.9	1376
Bangka Belitung Islands	10H	616.3	581.2	312.8	229.5
Riau Islands	10H	432	409.9	204.4	169
Jakarta	10H	11014.6	9492.2	4881.4	4822
West Java	10H	7864	6989.9	3374.2	3194
Central Java	12H	9745.2	8314.5	2376.7	2368.5
Yogyakarta	12H	1517.8	1229.3	612.6	612
East Java	12H	10135	7914.1	552	551
Banten	10H	1475.7	1317.3	557	499.5
Bali	12H	4946.3	3715.9	2018.8	2014.5
West Nusa Tenggara	12H	656.6	528.1	652.7	515
East Nusa Tenggara	10H	441.2	417.6	673.5	538.5
West Kalimantan	12H	2552.7	1889.6	97.7	97
Central Kalimantan	10H	440.8	399.1	1108.9	888.5
South Kalimantan	10H	1212.9	767.6	920	694
East Kalimantan	10H	2302.6	1701.9	436.7	366.5
North Sulawesi	10H	542.3	441.1	704.4	508
Central Sulawesi	12H	548.4	462.8	82.2	66
South Sulawesi	12H	1495.2	1017.1	203.1	156.5
Southeast Sulawesi	10H	851.2	527.3	350	325
Gorontalo	10H	184	130.6	574.2	555.5
West Sulawesi	10H	281.3	220.5	306	228.5
Maluku	10H	110.7	54.3	223.1	193
North Maluku	10H	106.3	88.3	31	31
West Papua	10H	137.7	109	28	23
Papua	10H	323.1	215.7	145.5	137.5

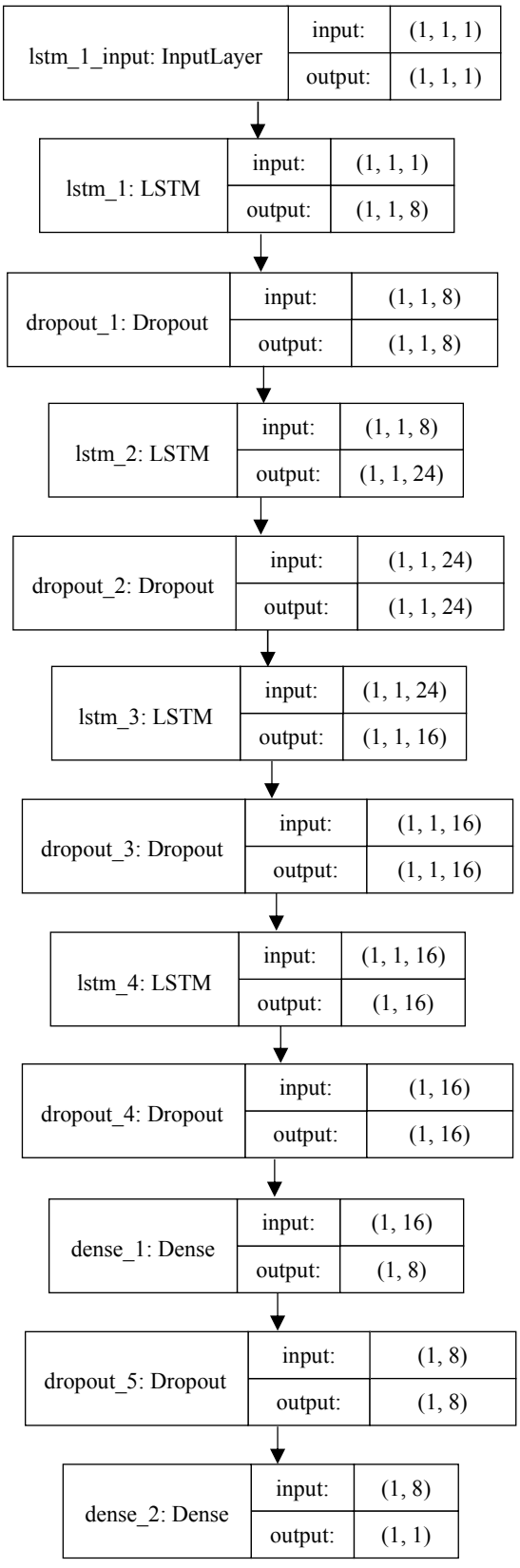

Fig. 6. LSTM model architecture of ten hidden layers

An example of time series forecasting LSTM model on dengue incidence in North Maluku province is shown in Fig. 6., which has ten hidden layers. Table II shows the performance results of training on dengue incidence in North Maluku, Indonesia, from 2006 to 2013 and performance results of validation, from 2014 to 2016. Table III shows the results of testing on dengue incidence for two years (2017-2018). We use the early stopping method to avoid overfitting. Training data on dengue incidence in North Maluku terminates immediately after the fourth epoch with no improvement. Model loss during training is shown in Fig. 7.

TABLE II. TRAINING AND VALIDATION RESULTS OF NORTH MALUKU PROVINCE ON DENGUE INCIDENCE

Year	Expected Incidence (n)	Predicted Incidence (n)	Error	Absolute Error (AE)	Squared Error (SE)
2006	138	276	-138	138	19044
2007	275	273	2	2	4
2008	250	271	-21	21	441
2009	384	268	116	116	13456
2010	347	266	81	81	6561
2011	164	265	-101	101	10201
2012	65	264	-199	199	39601
2013	242	263	-21	21	441
2014	148	262	-114	114	12996
2015	119	261	-142	142	20164
2016	297	261	36	36	1296

TABLE III. TESTING RESULTS OF NORTH MALUKU PROVINCE ON DENGUE INCIDENCE

Year	Expected Incidence (n)	Predicted Incidence (n)	Error	Absolute Error (AE)	Squared Error (SE)
2017	37	66	-29	29	841
2018	110	77	33	33	1089

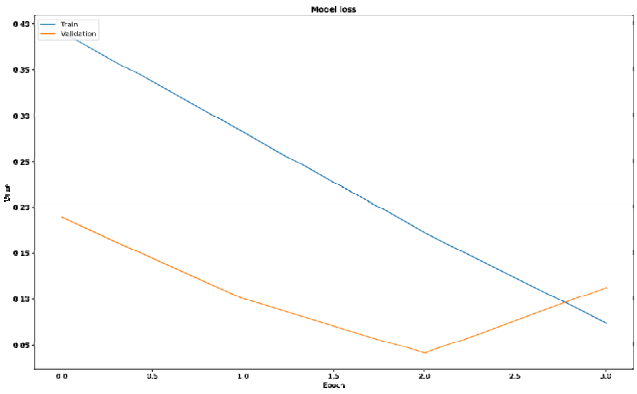

Fig. 7. Model loss of dengue incidence in North Maluku

We analysed the errors between the predicted and the expected dengue incidence based on the performance of training results, as an example of North Maluku province in Table II. We found that the deep LSTM network is good to forecast the dengue incidences for 2006-2011. We found high fluctuation of error started in 2012-2013 in eleven provinces at the national level, i.e., Aceh, Riau, South Sumatera, North Sumatera, Jambi, Bengkulu, West Sumatera, Lampung, Bangka Belitung Islands, Bali, and Central Java. A given example of high fluctuation of error

on dengue incidence in Maluku and West Sulawesi happened in 2012, 2013, and 2016, as shown in Fig. 8. and Fig. 9., respectively. We assume that external factors (such as climate variability) may influence the number of incidences during 2012, 2013, and 2016.

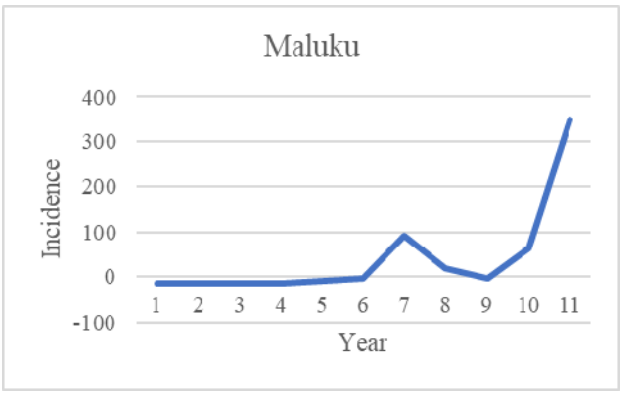

Fig. 8. Error rate on dengue incidence in Maluku

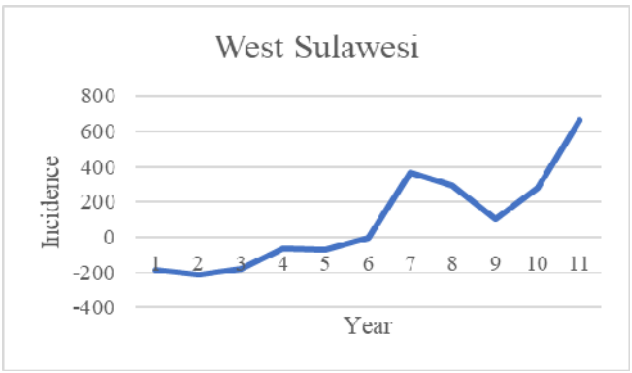

Fig. 9. Error rate on dengue incidence in West Sulawesi

Table IV shows the performance result of malaria incidence estimations in all provinces, with possible different best LSTM models in different provinces. Table V shows the performance results of training on malaria incidence in Banten, Indonesia, from 2006 to 2013 and performance results of validation, from 2014 to 2016. Table VI shows the results of testing on malaria incidence for two years (2017-2018). An example of training data on malaria incidence in Banten terminates immediately after the 299th epoch with no improvement. Model loss during training is shown in Fig. 10.

TABLE IV. PERFORMANCE RESULTS OF MALARIA INCIDENCE IN ALL INDONESIA PROVINCES

Province	LSTM Model	Train Score		Test Score	
		RMSE	MAE	RMSE	MAE
Aceh	12H	14404.7	10913.6	10598.9	10594
North Sumatera	12H	8858.6	6525.6	7149.17	7048.5
West Sumatera	10H	638.7	545.6	125.5	125.5
Riau	10H	5428.6	4917.5	6473.25	6473
Jambi	10H	14737.4	11639.8	16702.5	16702.5
South Sumatera	8H	55074.5	52635	25996.8	25742
Bengkulu	12H	3441.5	2683.7	681	517.5
Lampung	10H	8784.3	7499	9610.7	9522
Bangka Belitung Islands	10H	11945.8	8292.4	16270.7	16267.5
Riau Islands	12H	3876.5	2969.2	4493	4487.5
Jakarta	12H	24.2	11.4	89.2	85.5
West Java	10H	11106	10357.2	7824	7821.5
Central Java	10H	69459.1	59241.8	69136	68942
Yogyakarta	12H	7324	3663.8	19448.9	19204.5
East Java	14H	14569.9	7526.7	46872.2	44966.5
Banten	12H	592.2	356.6	875.8	866.5
Bali	10H	7905	6388.6	6337.9	5197
West Nusa Tenggara	10H	20190.6	15.252.4	36597.8	35382
East Nusa Tenggara	10H	122374	91760.8	144635	144272
West Kalimantan	10H	8472.2	6292.3	15794.3	15025.5
Central Kalimantan	10H	10072	9011.2	13164	11026
South Kalimantan	12H	2548.5	2302.6	8985.1	6859
East Kalimantan	12H	4148.4	2845.8	8166.5	5907
North Sulawesi	10H	8609	6112.1	8870	7684
Central Sulawesi	10H	17951.3	16542.5	14707.4	12343.5
South Sulawesi	10H	1661.3	1264.6	7303.8	6224
Southeast Sulawesi	10H	9710.9	8714.4	5749.1	5058
Gorontalo	10H	2974.7	2389	5305.2	5144
West Sulawesi	10H	3812.7	2792.55	23315.1	17624
Maluku	10H	8257.7	6507	25332	24186
North Maluku	10H	33577.4	23911.3	23303.8	18852.5
West Papua	12H	53372.9	36583.1	64518.4	45805
Papua	10H	108769	84655.6	150718	126240

TABLE V. TRAINING AND VALIDATION RESULTS OF BANTEN ON MALARIA INCIDENCE

Year	Expected Incidence	Predicted Incidence	Error	Absolute Error (AE)	Squared Error (SE)
2006	658	782	-124	124	15376
2007	2692	988	1704	1704	2903616
2008	103	929	-826	826	682276
2009	543	642	-99	99	9801
2010	113	419	-306	306	93636
2011	88	293	-205	205	42025
2012	228	232	-4	4	16
2013	97	210	-113	113	12769
2014	46	208	-162	162	26244
2015	28	213	-185	185	34225
2016	28	222	-194	194	37636

TABLE VI. TESTING RESULTS OF BANTEN ON MALARIA INCIDENCE

Year	Expected Incidence	Predicted Incidence	Error	Absolute Error (AE)	Squared Error (SE)
2017	42	781	-739	739	546121
2018	15	1009	-994	994	988036

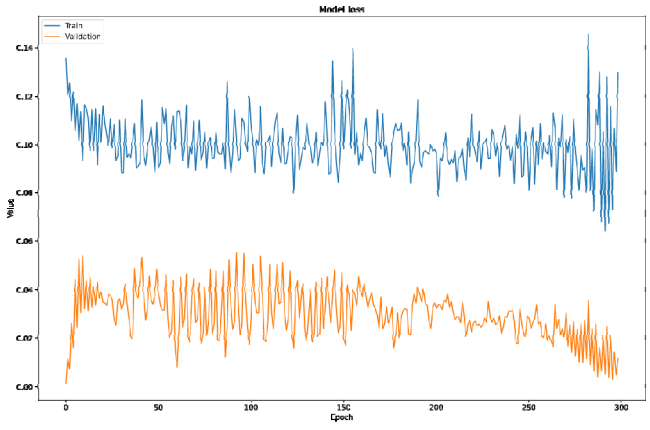

Fig. 10. Model loss of malaria incidence in Banten

We analysed the errors between the predicted and the expected malaria incidence based on the performance of training results. An example of implementation in Banten province, as shown in Table V. We found that the deep LSTM network is useful to forecast the malaria incidence for 2009-2016. However, among thirty-three provinces in Indonesia, we found high fluctuation of error during 2007 in sixteen provinces, i.e., Aceh, North Sumatera, West Sumatera, Riau, Jambi, South Sumatera, Banten, Riau Islands, Bangka Belitung Islands, West Sulawesi, Maluku, North Maluku, Central Kalimantan, West Kalimantan, West Papua, and Papua. An example of a high fluctuation of error on malaria incidence in Banten and West Kalimantan happened in 2007, as shown in Fig. 11, and Fig. 12, respectively. We assume that external factors (such as climate variability) may influence the number of incidence in 2007.

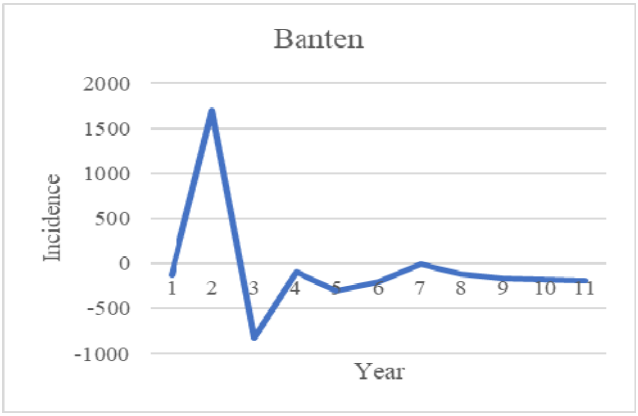

Fig. 11. Error rate on malaria incidence in Banten

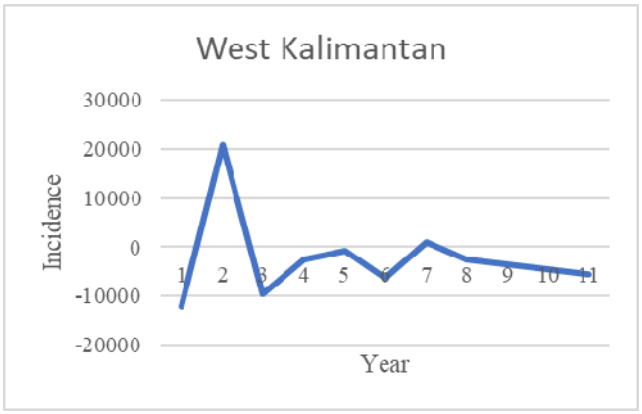

Fig. 12. Error rate on malaria incidence in West Kalimantan

V. Conclusion and Future Work

This study proposes two contributions: (1) we examined the influence of IOD and ENSO on interannual variability on dengue and malaria incidence in all Indonesia provinces, (2) we developed a stacked LSTM model for time series forecasting on dengue and malaria incidence (2005-2018) in all Indonesia provinces.

The results show that generally, IOD has a negative correlation with the dengue incidence for five provinces in Indonesia and has strong influences for seven provinces. On the other hand, ENSO also has a negative correlation in eight provinces, i.e., Aceh, West Papua, Jakarta, Central Sulawesi, North Maluku, Lampung, North Sulawesi, and Banten. In contrast with dengue incidence, the positive (negative) correlation of malaria incidences with IOD (ENSO) is shown in most provinces in Indonesia. The incidence of dengue shows a strong interannual variability related to IOD and ENSO. On the other hand, the malaria incidence in Indonesia has less interannual variation related to IOD and ENSO.

The deep LSTM network provides a good forecast for the dengue and malaria incidence during the observation period (2006-20011 for dengue and 2009-2016 for malaria). We observed high fluctuation of error on dengue incidence during 2012-2013 in eleven provinces and malaria incidence during 2007 in sixteen provinces. Factors such as climate variability could potentially influence the fluctuation observed in the forecasting models.

The quality of the forecasting results might be compromised by the sparse time-series annual data of dengue and malaria cases. Future studies should attempt to improve the quality of the forecasting by utilizing weekly or monthly data and accounting for climate variabilities in the forecast models.

Acknowledgment

This research is funded by the Fundamental Research Grants 2020 scheme from the Ministry of Research and Technology / National Agency for Research and Innovation of the Republic of Indonesia.

References

[1] World Health Organization, “Vector-borne diseases”, https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases

[2] M. Chan and M. A. Johansson, “The incubation periods of dengue viruses,” PLoS One, vol. 7, pp. 1-7, 2012.

[3] A. A. Rahman, I. W. S Adnyana, M. S. Mahendra, I. W. Arthana, I. N. Merit, I. W. Kasa, N. W. Ekayanti, I. W. Nuarsa, I. N. Sunarta, “Observation of spatial patterns on the rainfall response to ENSO and IOD over Indonesia using TRMM Multisatellite Precipitation Analysis (TMPA),” International Journal of Climatology, vol. 34, no. 15, pp. 3825–3839, 2014.

[4] N. H. Saji, B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, “A dipole mode in the tropical Indian Ocean,” Nature, vol. 401, no. 6751, pp. 360-363, 1999.

[5] K. E. Trenberth, “El Niño Southern Oscillation (ENSO), “ Encyclopedia of Ocean Sciences (Third Edition), vol. 6, pp. 420-432, 2019.

[6] J. R. E. Harger, "ENSO variations and drought occurrence in Indonesia and the Philippines", *Atmospheric Environment*, vol. 29, No. 16, pp. 1943-1955, 1995.

[7] H. H. Hendon, "Indonesian rainfall variability: impacts of ENSO and local airsea interaction," J. Clim., vol.16, pp. 1775-1790, 2003.

[8] R. Y. Setiawan, A. Wirasatriya, U. Hernawan, S. Leung and I. Iskandar, “Spatio-temporal variability of surface chlorophyll-a in the Halmahera Sea and its relation to ENSO and the Indian Ocean Dipole, ” *International Journal of Remote Sensing,* vol. 41, no. 1, pp. 284-299, 2020.

[9] T.W. Chuang, L. F. Chaves, and P. J. Chen, “Effects of local and regional climatic fluctuations on dengue outbreaks in southern Taiwan”, PLoS ONE, vol. 12, no. 6, 2017.

[10] G. Zhu, T. Liu, J. Xiao, B. Zhang, T. Song, Y. Zhang, L. Lin, Z. Peng, A. Deng, W. Ma, and Y. Hao, “Effect of human mobility, temperature and mosquito control on the spatiotemporal transmission of dengue,” *Science of the Total Environment*, vol. 651, pp. 969-978, 2019.

[11] J. Xiang, A. Hansen, Q. Liu, X. Liu, M. X. Tong, Y. Sun, S. Cameron, S. H. Easey, G. S. Han, C. Williams, P. Weinstein, and P.

Bi, “Association between dengue fever incidence and meteorological factors in Guangzhou, China, 2005-2014,” *Environmental Research*, vol. 153, pp. 17-26, 2017.

[12] S. I. Hay, G. D. Shanks, D. I. Stern, R. W. Snow, S. E. Randolph, and D. J. Rogers, “Climate variability and malaria epidemics in the highlands of East Africa,” *TRENDS in Parasitology*, vol. 21, no. 2, 2005.

[13] H. H. Hussien, “Malaria’s association with climatic variables and an epidemic early warning system using historical data from Gezira State, Sudan,” Heliyon, vol. 5, 2019.

[14] Kurnianingsih, H. S. Allehaibi, L. E. Nugroho, Widyawan, L. Lazuardi, A. S. Prabuwono, and T. Mantoro, “Segmentation and Classification of Cervical Cells Using Deep Learning,” *IEEE Access*, vol. 7, pp. 116925-116941, 2019.

[15] S. C. Kalkan and O. K. Sahingoz, “Deep Learning Based Classification of Malaria from Slide Images,” 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science, Istanbul, Turkey, pp. 1-4, 2019.

[16] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.

[17] M. Beeksma, S. Verberne, A. Bosch, E. Das,I. Hendrickx, and S. Groenewoud, “Predicting life expectancy with a long short-term memory recurrent neural network using electronic medical records,” *BMC Medical Informatics and Decision Making*, vol. 19, no. 36, 2019.

[18] T. Pham, T. Tran, D. Phung, and S. Venkatesh, “Predicting healthcare trajectories from medical records: A deep learning approach,” *Journal of Biomedical Informatics*, vol. 69, pp. 218-229, 2017.

[19] R. Zhao, J. Wang, R. Yan, and K. Mao, “Machine Health Monitoring with LSTM Networks,” 10th International Conference on Sensing Technology, 2018

[20] U. Langkulsen, K. P. N. Sakolnakhon, and N. James, "Climate change and dengue risk in central region of Thailand," *International Journal of Environmental Health Research*, 2018.

[21] M. F. Vincenti-Gonzalez, A. Tami, E. F. Lizarazo, and M. E. Grillet, “ENSO-driven climate variability promotes periodic major outbreaks of dengue in Venezuela,” Nature, vol. 8, p. 5727, 2018.

[22] D. Petrova, R. Lowe, A. Stewart-Ibarra, J. Ballester, S. J. Koopman, and X. Rodó, “Sensitivity of large dengue epidemics in Ecuador to long-lead predictions of El Niño,” *Climate Services*, vol. 15, no. 100096, 2019.

[23] M. L. H. Mabaso, I. Kleinschmidt, B. Sharp, and T. Smith, “El Niño Southern Oscillation (ENSO) and annual malaria incidence in Southern Africa,” *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 101, no. 4, pp. 326-330, 2007.

[24] V. Brown, M. A. Issak, M. Rossi, P. Barboza, and A. Paugam, “Epidemic of malaria in north-eastern Kenya,” *Lancet*, vol. 352, pp. 1356–7, 1998.

[25] M. J. Bouma, “Epidemiology and control of malaria in northern Pakistan,” Dordrecht: ICG Printing; 2005.

[26] R. C. Dhiman and S. Sarkar, “El Niño Southern Oscillation as an early warning tool for malaria outbreaks in India,” *Malar J*, vol. 16, no. 122, 2017.

[27] A. S. Gagnon, K. E. Smoyer-Tomic, and A. B. G. Bush, “The El Niño Southern Oscillation and malaria epidemics in South America,” *Int J Biometeorol*, vol. 46, pp. 81–9, 2002.

[28] M. Hashizume, T. Terao, and N. Minakawa, “The Indian Ocean Dipole and malaria risk in the highlands of western Kenya,” Proc. Natl Acad Sci USA, vol. 106, no. 6, pp. 1857-1862, 2009.

[29] L. F. Chaves, A. Satake, M. Hashizume, and N. Minakawa, “Indian Ocean Dipole and Rainfall Drive a Moran Effect in East Africa Malaria Transmission,” *The Journal of Infectious Diseases*, vol. 205, no. 12, pp. 1885-1891.

[30] Ministry of Health of Republic of Indonesia, “Indonesian Health Profile”, Central of Data and Information, available online: https://pusdatin.kemkes.go.id/folder/view/01/structure-publikasi-pusdatin-profil-kesehatan.html

[31] T.-C. Chan, T.-H. Hu, and J.-S. Hwang, ‘‘Daily forecast of dengue fever incidents for urban villages in a city,’’ *Int. J. Health Geograph*., vol. 14, no. 1, p. 9, 2015.

[32] L. E. Hugo et al., ‘‘Adult survivorship of the dengue mosquito Aedes aegypti varies seasonally in central Vietnam,’’ *PLoS Neglected Tropical Diseases*, vol. 8, no. 2, p. e2669, 2014.

[33] K. D. Sharma, R. S. Mahabir, K. M. Curtin, J. M. Sutherland, J. B. Agard, and D. D. Chadee, ‘‘Exploratory space-time analysis of dengue incidence in Trinidad: A retrospective study using travel hubs as dispersal points, 1998–2004,’’ *Parasite Vectors*, vol. 7, no. 1, p. 341, 2014.

[34] M. S. Mia, R. A. Begum, A. C. Er, R. D. Z. R. Z. Abidin, and J. J. Pereira, “Trends of dengue infections in Malaysia, 2000-2010,” *Asian Pacific Journal of Tropical Medicine*, pp. 462-466, 2013.

[35] W. Anggraeni, R. Nurmasari, E. Riksakomara, F. Samopa, R. P. Wibowo, L. Condro, Pujiadi, “Modified Regression Approach for Predicting Number of Dengue Fever Incidents in Malang Indonesia,” Proc. of ISICO, pp. 142-150, 2017.

[36] Y. L. Hii, H. Zhu, N. Ng, L. C. Ng, and J. Rocklöv, ‘‘Forecast of dengue incidence using temperature and rainfall,’’ *PLoS Neglected Tropical Dis*., vol. 6, no. 11, p. e1908, 2012.

[37] M. A. Johansson, N. G. Reich, A. Hota, J. S. Brownstein, and M. Santillana, ‘‘Evaluating the performance of infectious disease forecasts: A comparison of climate-driven and seasonal dengue forecasts for Mexico,’’ *Sci. Rep*., vol. 6, Sep. 2016, Art. no. 33707.

[38] M. Gharbi et al., ‘‘Time series analysis of dengue incidence in Guadeloupe, French West Indies: Forecasting models using climate variables as predictors,’’ *BMC Infectious Diseases*, vol. 11, no. 1, p. 166, 2011.

[39] S. Bhatnagar, V. Lal, S. D. Gupta, and O. P. Gupta, ‘‘Forecasting incidence of dengue in Rajasthan, using time series analyses,’’ Indian J. Public Health, vol. 56, no. 4, pp. 281–285, 2012.

[40] C. C. Ho and C.-Y. Ting, ‘‘Time series analysis and forecasting of dengue using open data,’’ in Advances in Visual Informatics. Cham, Switzerland: Springer, 2015, pp. 51–63.

[41] A. Lal, T. Ikeda, N. French, M. G. Baker, and S. Hales, ‘‘Climate variability, weather and enteric disease incidence in New Zealand: Time series analysis,’’ PLoS ONE, vol. 8, no. 12, p. e83484, 2013.

[42] J. Xia, S. Pan, M. Zhu, G. Cai, M. Yan, Q. Su, J. Yan, and G. Ning, “A Long Short-Term Memory Ensemble Approach for Improving the Outcome Prediction in Intensive Care Unit,” vol. 2019, 2019.

[43] A. Graves, “Generating Sequences With Recurrent Neural Networks,” arXiv:1308.0850, 2014, online: https://arxiv.org/abs/1308.0850

[44] T. Pham, T. Tran, D. Phung, and S. Venkatesh, “DeepCare: A Deep Dynamic Memory Model for Predictive Medicine,” arXiv:1602.00357, 2017, online: https://arxiv.org/abs/1602.00357

[45] National Weather Service, Monthly Atmospheric & SST Indices,” Climate Prediction Center, available [online] https://www.cpc.ncep.noaa.gov/data/indices/oni.ascii.txt

[46] National Oceanic and Atmospheric Administration, “The state of the ocean climate,” Ocean Observations Panel for Climate, available [online] https://stateoftheocean.osmc.noaa.gov/sur/ind/dmi.php

[47] T. Chai and R. R. Draxler, ‘‘Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature,’’ *Geosci. Model Develop*., vol. 7, no. 3, pp. 1247–1250, 2014.