COMBINATION OF UNIVARIATE LONG-SHORT TERM MEMORY NETWORK AND WAVELET TRANSFORM FOR PREDICTING DENGUE CASE DENSITY IN THE NATIONAL CAPITAL REGION, THE PHILIPPINES

Imee V Necesito[1], John Mark Velasco[2], Jaewon Kwak[3], Jun Hyeong Lee[1], Myung Jin Lee[1], Jong Sung Kim[1] and Hung Soo Kim[1]

[1]Department of Civil Engineering, Inha University, South Korea; [2]Institute of Molecular Biology and Biotechnology, National Institutes of Health and Department of Clinical Epidemiology, College of Medicine, University of the Philippines Manila, the Philippines; [3]Han River Flood Control Office, Ministry of Environment, Seoul, South Korea

Abstract. Use of machine learning algorithms on public health big data has paved the way in helping to understand complex associations between diseases and environment. However, in cases where informative prediction of a pathogen spread or transmission is still equivocal, a rapidly deployable independent mathematical modeling method, which does not require a priori knowledge of the pathogen characteristics or requires input parameters predictive of its transmission is needed. This study shows long-short term memory (LSTM) network with wavelet transform has application as a modeling method for dengue incidence. The information derived from this modeling technique can be used to help guide control and prevention measures during an early stage of a dengue surge. Subsequent application of discrete wavelength transform (DWT) to LSTM output resulted in visualization of dengue surges, the majority of which were observed during the wet season (June-December). Augmenting existing disease surveillance systems with automated mathematical models can help focus intervention programs especially in settings where health care resources and surveillance infrastructure are limited. This modeling method has potential applications to other infectious diseases in a developing country setting.

Keywords: deep learning; dengue; discrete wavelength transform; the Philippines; univariate long-short term memory network

---

Correspondence: Hung Soo Kim, Department of Civil Engineering, Inha University, 100 Inha-ro, Yonghyeon 1(il).4(sa)-dong, Nam-gu, Incheon 22212, South Korea
Tel: +82 328740069 E-mail: sookim@inha.ac.kr

INTRODUCTION

Dengue is caused by a mosquito-transmitted virus of the Flaviviridae family and has four distinct serotypes (DENV-1, -2, -3, -4) (Halstead, 2012). Prior to 1970, less than 10 countries experienced severe dengue epidemics (Gubler, 2002; McLennan-Smith and Mercer, 2014), but currently dengue has spread globally to more than 100 tropical and subtropical countries (Gubler, 2002; WHO, 2012a) with over 100 million recorded cases annually (Bhatt *et al*, 2013). The Philippines experienced its first dengue hemorrhagic fever (DHF) outbreak in the 1950s (Gubler, 1997) with DENV-3 as the predominant serotype (Messer *et al*, 2003). Dengue consistently ranks among the leading causes of morbidity in the Philippines (DOH, 2018).

Previous studies have focused on vector control (Espino *et al*, 2012) and relationship of disease transmission with various climatic factors (Li *et al*, 1985; Promprou *et al*, 2005; Wongkoon *et al*, 2007; Su, 2008; Pham *et al*, 2011; Hii *et al*, 2012; Cheong *et al*, 2013; Hii *et al*, 2016; Lee *et al*, 2017). Studies have shown association of dengue outbreaks (Halide and Ridd, 2008; Sang *et al*, 2014; Ho and Ting, 2015; Li *et al*, 2017) with various meteorological or environmental factors (Focks *et al*, 1995; Barbazan *et al*, 2010; Shang *et al*, 2010; Dela Cruz *et al*, 2012; Buczak *et al*, 2014; Iguchi *et al*, 2018), while other studies used keywords in open source internet data to predict dengue outbreaks (Coberly *et al*, 2013). Focks *et al* (1995), Barbazan *et al* (2010), Shang *et al* (2010), Buczak *et al* (2014), van Panhuis *et al* (2015), and Gbadamosi *et al* (2018) confirmed a correlation of dengue outbreaks with parameters, such as rainfall, temperature (El Nino) and increase in Aedes aegypti vector populations, while Dela Cruz *et al* (2012) and Iguchi *et al* (2018) used seasonal autoregressive moving average (SARIMA) and Poisson regression model respectively, to utilize local meteorological variables for modeling prevalence and predicting dengue outbreaks. These models can be classified into quantitative and qualitative models, with the former for predicting dengue outbreak and the latter determining interaction between human and vector populations (Gbadamosi *et al*, 2018), are established using regression equations (Halide and Ridd, 2008; Potts *et al*, 2010; Anggraeni *et al*, 2017) and time series models (Sang *et al*, 2014; Ho and Ting, 2015; Li *et al*, 2017) respectively.

The four distinct dengue serotypes, extensive genetic diversity within each serotype and non-linear behavior of dengue outbreaks contribute to its complexity as a disease, resulting in recourse to a variety of models for predicting occurrence of outbreaks. However, in certain situations where factors correlating with disease transmission or spread are unknown, or if data input parameters (*eg* environmental factors) required for the model are not available, applications of these mathematical modelling approaches may be limited.

Univariate long-short term memory (LSTM) network is one of the promising deep learning disease prediction methods that can be used in the absence of a priori knowledge of other factors (Siami-Namini *et al*, 2018). An assumption of dependence on future from past time steps of LSTM can be a representation of the spread of dengue cases. LSTM also has the ability to learn a system's oscillating behavior or seasonality in a time series making it suitable for application to a seasonal disease such as dengue. Commonly employed univariate models such as autoregressive integrated moving average (ARIMA) and SARIMA require calculated parameters [(p, d, q) and (P, D, Q) respectively] and may require lengthy time periods to compute especially if dealing with large datasets. Univariate LSTM, on the other hand, is non-parametric. A susceptible, infected and recovered (SIR) model is also promising (Pandey *et al*, 2013), but is more complex and requires group assignment (susceptible, infectious or recovered) and application of differential equations, which are not necessary with LSTM. In addition, the availability of cloud computing and programming packages, such as Keras, an open-source software library that provides a Python interface for artificial neural networks (https://github.com/keras-team/keras) and MATLAB Deep Learning Toolbox (The Mathworks Inc, Natick, MA) allow conducting of LSTM analysis of large datasets.

Few studies have investigated the use of a univariate dengue prediction model and deep learning combined with discrete wavelet transform (DWT), a method commonly used to denoise signals (Ali *et al*, 2017). Here, we used univariate LSTM combined with DWT to explore and model dengue case incidence in the Philippines from 1994 to 2018. The findings should be of assistance in improving accuracy of dengue outbreak prediction in the National Capital Region (and elsewhere) in the Philippines.

MATERIALS AND METHODS

Data source

We used monthly dengue case density data [case/population of the National Capital Region (NCR), the Philippines] from January 1994 to December 2018 (a total of 300 observations) available from URL: https://www.tycho.pitt.edu/version-2/. NCR is the most densely populated area in the country with 18,651 person/km^2 within an area of 636 km^2 (Naik Singru and Lindfield, 2014).

As publicly available aggregate data with no personal identifiable information were collected, no ethical review was required.

LSTM network protocol

LSTM neural network, first introduced by Hochreiter and Schmidhuber (1997), has three main layers: input, recurrent hidden and output. LSTM does not simply have memory blocks to memorize temporal state but has gating units to adapt and control the flow of information. LSTM is efficient in capturing time series dependencies, particularly in dealing with long time periods (Olah, 2015). LSTM gates can be considered as controls and network as a set or series of LSTM units, each having four main controls or gates: (i) a "forget" control allowing the network to either retain or forget the memory in transit, (ii) a "new memory" control to allow new memories to pass through and later to merge with memories that have passed through the forget control, (iii) a "merge" control to merge a set of new memories being processed by another neural network with existing memories, and (iv) an "output" control to check how much memory should be produced as an output to the next LSTM unit (Fig 1).

The data used to make the forecasts were based on dengue incidence records from 1994-2015, which documented monthly dengue incidence and dengue density data in NCR (with high values observed during months corresponding to the rainy/typhoon season (June-December) in the years of 1996, 1998, 2006, 2010, 2011, and 2015) (Fig 2) and referred as training dataset without application of time lags (0-lag). The first training dataset consisted of monthly dengue incidence from 1994-2015, and the monthly lag (up to 1 month) was employed to consider the transmission period from an infected viremic person up to the combined extrinsic and intrinsic incubation periods that can last up to 21-28 days (CDC, 2018).

The actual observed monthly dengue cases of 1994-2015 were used to produce the 0-lagged LSTM model

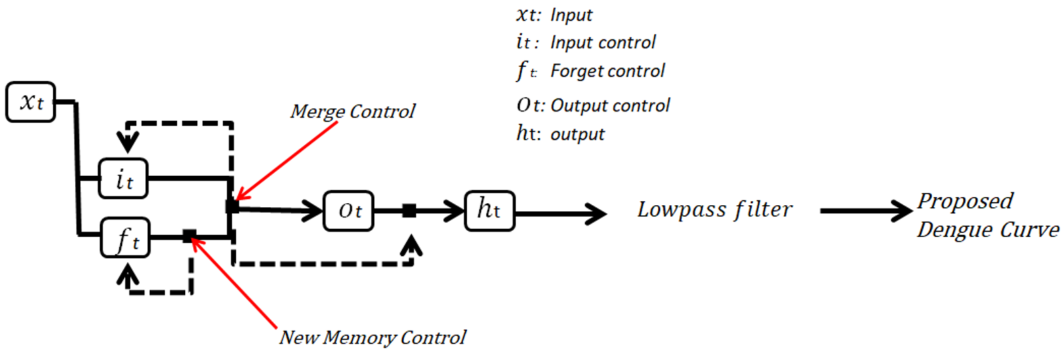

Fig 1 - Schematic diagram of a univariate long-short term memory network to generate a proposed dengue curve

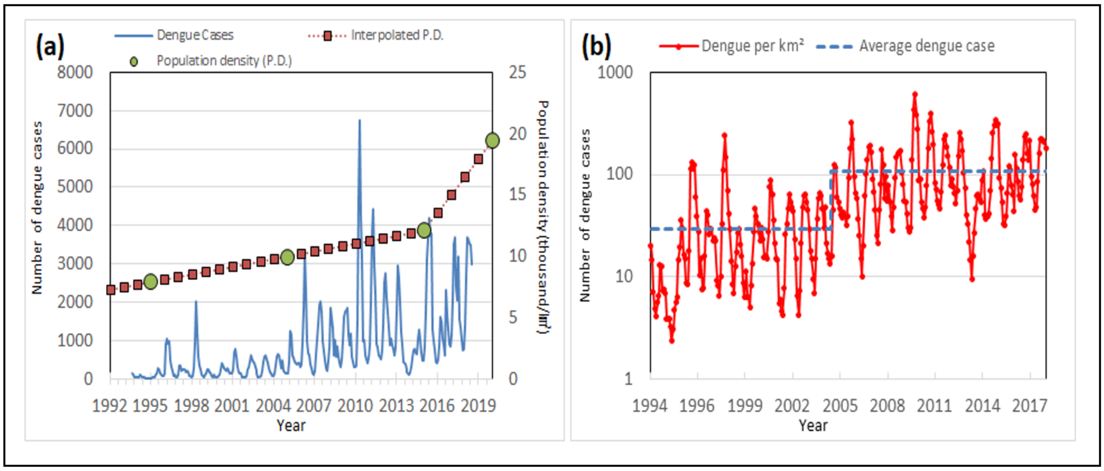

Fig 2 - Population density and number of dengue cases (a) and dengue case density and average dengue case (b) in the National Capital Region, the Philippines (January 1994 - December 2018)

while the 1-month lagged actual observed monthly dengue cases to produce the 1-lagged LSTM model. In order to check LSTM performance, the 1-month lagged model was applied to the actual monthly dengue cases and included both extrinsic and intrinsic incubation periods. As performance indicators favor non-lagged LSTM output, the 0-lagged model was reconstructed and high-frequency noise, which signals dengue surges based on the monthly average dengue cases, was filtered out. Using DWT as a reconstruction technique, it was possible to smooth out the time series and identify probable points of sudden increase in dengue cases. The monthly average of dengue density, calculated from 1994 to 2015, was designated as the threshold value, and if a data point exceeds this value and the preceding data point is below threshold, a surge or noise was considered to have occurred. This surge or noise was assumed to have ended when successive points fell below the threshold value. We employed thresholding modules and NumPy package from Python 3.7 (https://numpy.org/doc/stable/user/whatisnumpy.html).

The study considered one input layer (dengue incidence data) and the hidden units were set at 300 via trial and error. An epoch is an iteration, and for each epoch, each data element will undergo training. If the number of epochs is small, the number of iterations the element will undergo will also be small and potentially can result to underfitting. The maximum number of epochs the model could undergo was set at 400. When validation loss ceased to improve, the system will stop the training even if number of epochs did not reach 400.

The learning rate of a deep learning model is inversely proportional to the model reliability (*ie* the lower the learning rate, the more reliable the training process). Typically, the learning rate ranges 0-1. Learning rate is an important hyperparameter in neural networks as it determines the iteration's step size before reaching the minimum loss function. A long training process is attributed to a small learning rate but creates a more stable training process. In the study, initial learning rate was established at 0.005 and was finalized at 0.00004. To control overfitting, a dropout factor was established, which is a form of regularization that minimizes interdependent learning during the training process by literally dropping out some units in the network. In the study, we used a 20% drop rate, the usual minimum dropout for deep learning models. A MATLAB Deep Learning Toolbox (The Mathworks Inc, Natick, MA) was applied to model disease incidence and to calibrate and validate the model based on monthly dengue incidence from 1994-2015 and 2016-2018 respectively.

DWT protocol

DWT (Percival and Walden, 2000) is often applied in signal processing to transform signals into a frequency (Tsui, 1996). DWT was used in surveillance or tracking of moving objects (Hsia *et al*, 2011) and in disease classification (Patil *et al*, 2009). Shmueli (2013) and Alimohamadi *et al* (2020) investigated the use of DWT to detect outbreaks using biosurveillance systems and the latter on pertussis aberrations but, to date, there is no report of its application to dengue. The study applied DWT to generate a dengue curve from 1994-2018 using LSTM from both training and validation datasets as inputs in the DWT process. Only the low pass filter (Fig 1) was utilized in the time series to recognize surge boundary as the high frequency pass filter is considered as a high-frequency noise of dengue incidence.

Performance indicators employed

Root mean square error (RMSE) was utilized as a metric for evaluation. RMSE is the standard deviation of residual, the difference between simulated and observed values, and is opted as an indicator of error between LSTM simulated results compared to actual number of dengue case density. In addition to comparing the highest and lowest values of simulated and observed values, Nash-Sutcliffe efficiency (NSE), correlation coefficient (CC), Kling-Gupta efficiency (KGE) and index of agreement (IA) methods were used to assess LSTM model performance.

Model accuracy determination

Analysis of dengue data from the Philippines Department of Health (DOH) used "thresholds": an "epidemic threshold" where an outbreak is declared and an "alert threshold" where public health officials are in state of readiness for the upcoming potential outbreak. NCR dengue alert and epidemic threshold were calculated and compared to NCR generated dengue density curve. Alert and epidemic thresholds are calculated using the following equations:

$$ETi = Ave_{\text{past 5 yrs}} \times 2SD_{\text{past 5 yrs}} \quad (1)$$

$$ATi = Ave_{\text{past 5 yrs}} \times 2SD_{\text{past 5 yrs}} \quad (2)$$

where ETi = epidemic threshold for the year concerned
ATi = alert threshold for the year concerned
$Ave_{\text{past 5 yrs}}$ = average value of the 5 years prior to the year concerned
$SD_{\text{past 5 yrs}}$ = standard deviation of the 5 years prior to the year concerned

RESULTS

Applying the LSTM model to forecast dengue case density in NCR, the Philippines from January 1994 to December 2018 (a total of 300 observations), there was a good correlation between observed and forecast values (Fig 3). A variety of statistical parameters (CC, IA, KGE, NSE, and RMSE) was determined to evaluate the model performance, with the 0-month lagged model resulting in value of 0.84, 0.91, 0.84, 0.68, and 48.04, respectively while in the 1-lagged model, the statistical value was 0.55, 0.73, 0.55, 0.11, and 79.84, respectively. Overall, the 0-lagged model outperformed the 1-lagged model and was chosen as the input signals to be reconstructed by DWT. The calibrated LSTM dengue case incidence showed persistently good agreement with the reconstructed LSTM output (Fig 4, red curve), which was also the dengue incidence baseline. In our study, LSTM dengue incidence time series was smoothed out to eliminate surges or noises as the average monthly values were used as thresholds.

The threshold values of epidemic and alert levels obtained from the Philippines DOH for 2016, 2017 and 2018 in the NCR were compared with forecast and actual dengue case density, showing low RMSE value (indicating low error) for each year evaluated (Fig 5). CC, KGE and IA also showed satisfactory results while NSE of the actual dengue curve vs the generated dengue curves for 2017 and 2018 showed satisfactory performance.

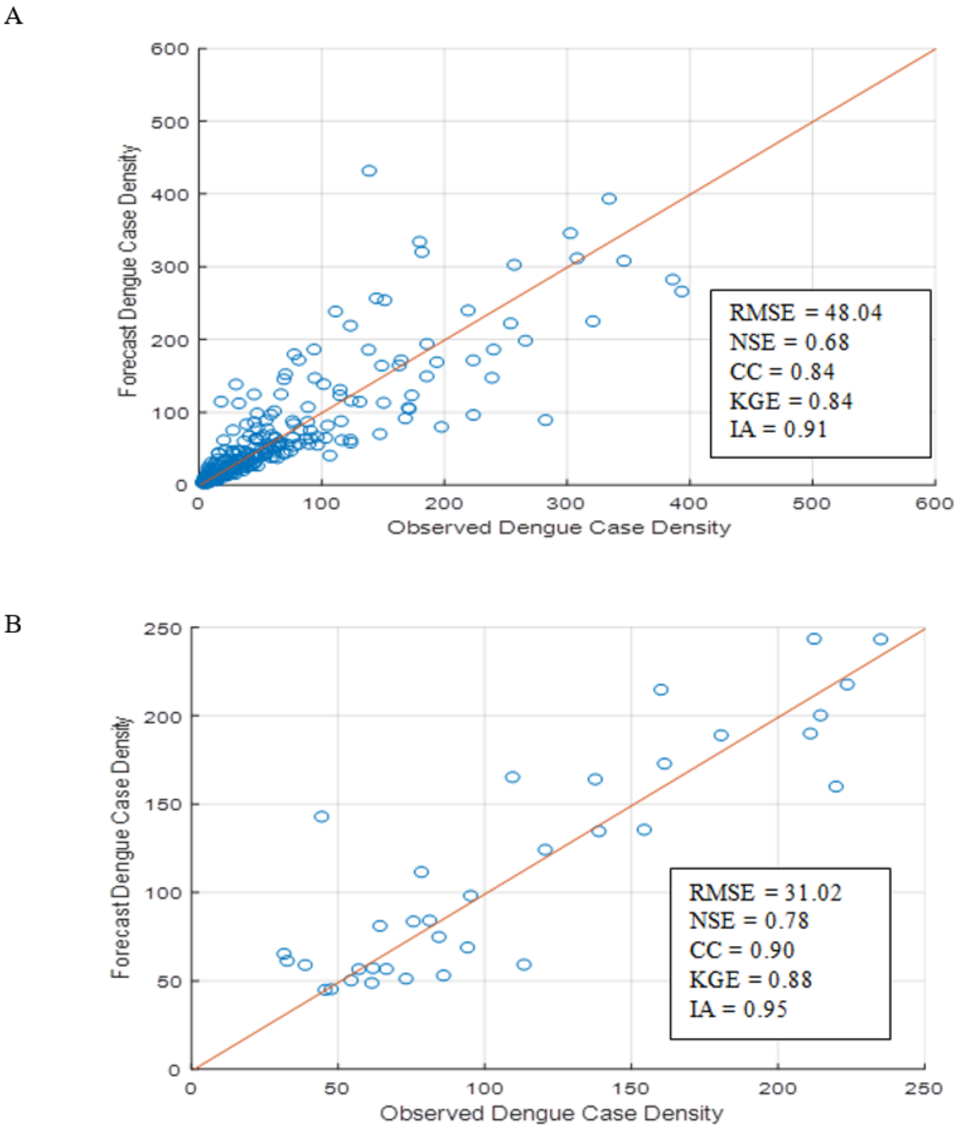

Fig 3 - Correlation of forecast and observed dengue case density

Panel B is an expansion of panel A in dengue case density range of 0-250. The LSTM model was applied to forecast dengue case density in the National Capital Region, the Philippines from January 1994 to December 2018 (a total of 300 observations).

CC: correlation coefficient; KGE: Kling-Gupta efficiency; IA: index of agreement; NSE: Nash-Sutcliffe efficiency; RMSE: root mean square error

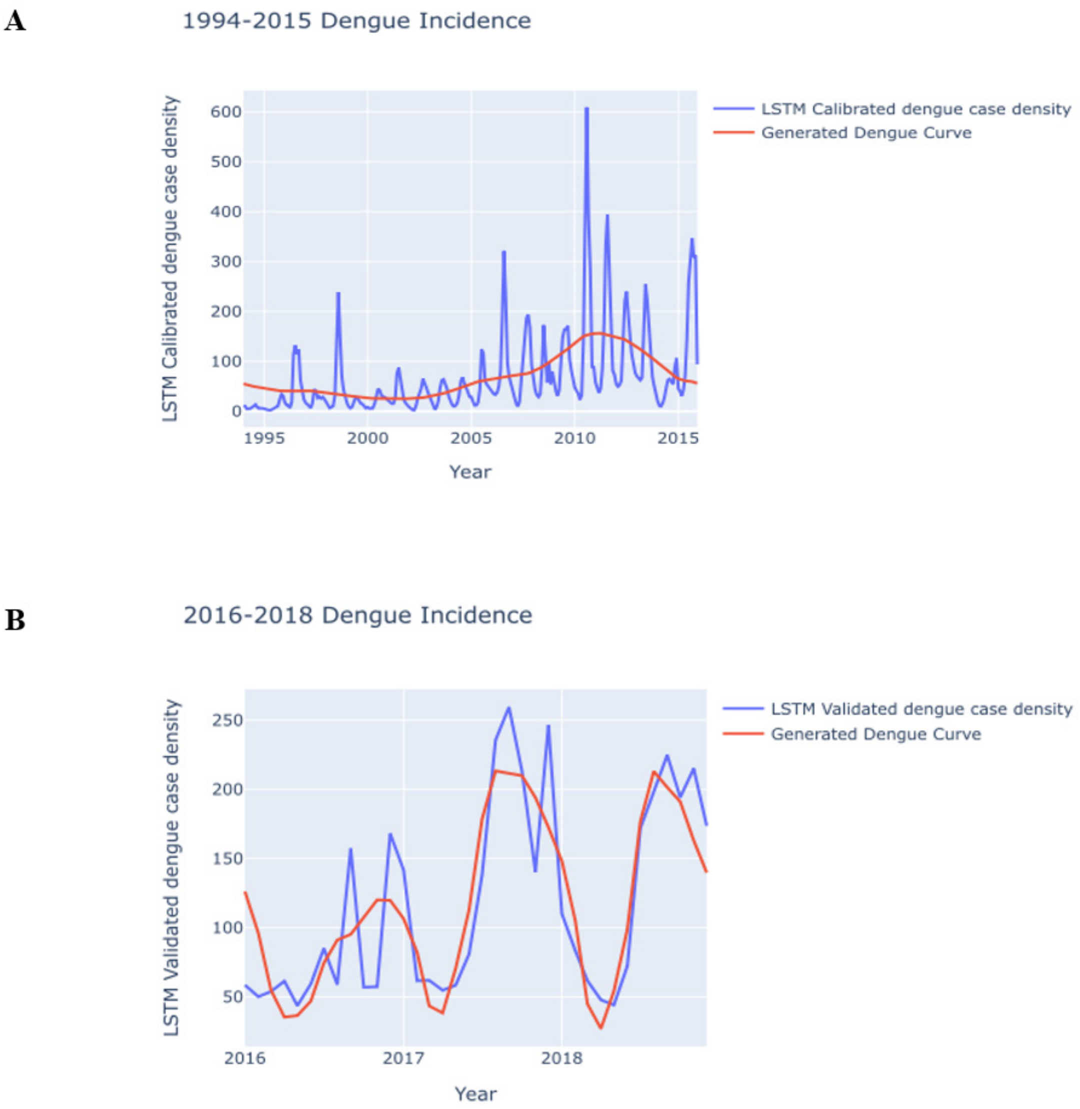

Fig 4 - Training dataset for generation of dengue density curve (A) and validation of generated dengue density curve (B).

Dengue density data were from Fig 1. The 0-month lagged model was chosen as the input signals to be reconstructed by a discrete wavelength transform.

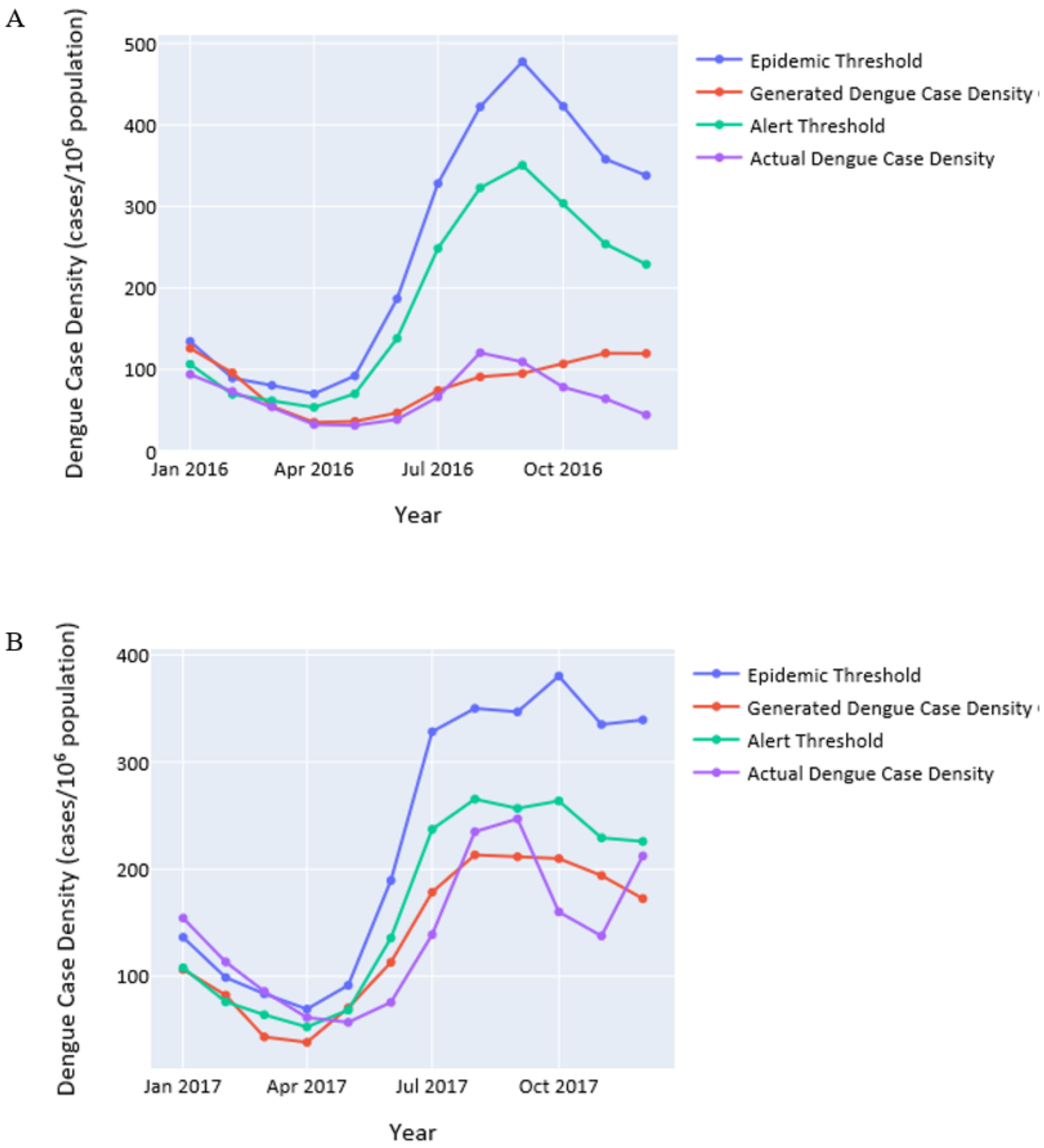

Fig 5 - Epidemic and alert thresholds, actual case density and generated dengue curve for 2016 (A), 2017 (B) and 2018 (C) in National Capital Region, the Philippines

(To be continued)

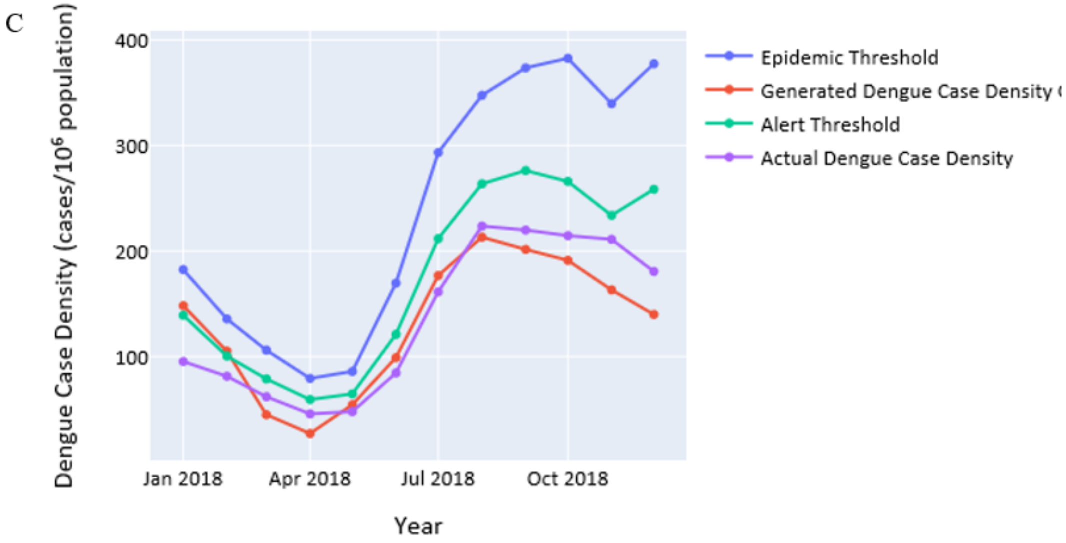

Fig 5 - Epidemic and alert thresholds, actual case density and generated dengue curve for 2016 (A), 2017 (B) and 2018 (C) in National Capital Region, the Philippines

Correlation coefficient, Kling-Gupta efficiency, index of agreement, Nash-Sutcliffe efficiency, and root mean square error value for actual dengue case density curve vs generated dengue case density curve for 2016 were 0.58, 0.50, 0.71, -0.31, and 32.14, respectively; for 2017: 0.82, 0.82, 0.90, 0.62, and 38.41, respectively; and for 2018: 0.92, 0.85, 0.95, 0.84, and 28.06, respectively.

DISCUSSION

A basic assumption of the LSTM method is that the state at the current time step depends on the previous time step (*ie* there is a connection between previous and future cases), which is advantageous over other methods which assume linearity in model parameters and independence of observations (Siami-Namini *et al*, 2018). Use of DWT in combination with LSTM transforms the time series to lessen singularities in order to produce a more consistent and better (Liu *et al*, 2021) dengue case density curve, *ie* there is certainty that y_2 is predictable because the value of y_1 is known. A good agreement between the elements of the datasets is an important factor in time series modeling, ensuring a correct model type will help in deciding the proper modeling technique.

The 0-month lag LSTM model outperformed the 1-month lag model. It is worth noting RMSE of the training dataset was higher than that of the validation dataset, indicating elimination of the possibility of overfitting. Other statistical parameters (CC, IA, KGE, and NSE), were in general consistent in demonstrating the superiority of 0-lagged over 1-lagged LSTM model. That CC (a subcomponent of KGE) value in the validation subset was slightly higher than the training subset implied the training data had slightly more crucial cases compared to the validation set, but despite CC value sensitivity to high values and insensitivity to additive and proportional differences between simulated and observed data (Legates and McCabe Jr, 1999), the performance result was still close to 1, indicates a positive linear association. Interestingly, when KGE was measured, values of the training and validation subset of the 0-lagged model were in close agreement; Gupta *et al* (2009) recommend use of KGE to lessen the tendency towards underestimation. On the other hand, IA, which is a more meaningful independent evaluation of performance (Wilmott *et al*, 2011), showed a lower value for the calibration compared to validation subsets in both models, implying both models gave equal best fit. NSE value (measure of how well the modelling technique reproduces the actual values) (Moriasi *et al*, 2007), showed the simulated value was a better predictor than the observed value.

Thus, the performance indicators demonstrated 0-lagged LSTM with univariate variable (dengue case density) had the capability to predict dengue disease trends using only dengue monthly case numbers and without requiring additional environmental parameters as input. In combination with DWT reconstruction technique, we identified probable points of noises or surges, namely, June to October 1996 and July to November 1998, with the highest residual dengue density at 115 cases/km^2. High frequency noises were also found in August to November 2010 that peaked at 610 cases/km^2. We noted several data points in the dry months (December to May) from the DWT curve that exceeded LSTM estimation (Fig 4A), signifying at those particular points in the series (dry season), no abnormal increase of dengue incidences was detected. However, the opposite scenarios (occurrences or surges in dry months) and noises during wet months were also present (Fig 4B). The model-generated dengue case density curve of 2018 was the best predictor of the actual dengue case density in NCR.

The study had two major limitations. Firstly, the modeling approach was limited to a single variable. Secondly, the study was conducted with the assumption that the available data were limited only to monthly dengue case density and that there were no interventions (*eg* vector control) and influence from climatic factors was negligible, which was not valid as external environmental factors and interventions can interrupt seasonal patterns of the disease, invalidating the assumption of dependence and consequently affecting the accuracy of generated predictions.

In summary, the study shows LSTM combined with DWT could be used to model disease data containing surges or noises associated with an oscillating system or a seasonal pattern, such as dengue. This method may outperform other models as it can learn patterns of outspread and predict/estimate numbers of probable cases even in the absence of knowledge of other factors correlated with disease spread and transmission. In order to improve the accuracy of the predictions generated by the model, we recommend incorporating meteorological variables, using daily/weekly dengue disease datasets and geographical data with finer granularity.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2017R1A2B3005695).

REFERENCES

Ali MN, El-Dahshan ESA, Yahia AH. Denoising of heart sound signals using discrete wavelet transform. *Circuits Syst Signal Process* 2017; 36: 4482-97.

Alimohamadi Y, Zahraei SM, Karami M, Yaseri M, Lotfizad M, Holakouie-Naieni K. Aberration detection of pertussis from the Mazandaran province, Iran, from 2012 to 2018: application of discrete wavelet transform. *J Acute Dis* 2020; 9: 114-20.

AnggraeniW, Nurmasari R, Riksakomara E, *et al*. Modified regression approach for predicting number of dengue incidents in Malang Indonesia. *Procedia Comput Sci* 2017; 124: 142-50.

Barbazan P, Guiserix M, Boonyuan W, Tuntaprasart W, Pontier D, Gonzalez JP. Modelling the effect of temperature on transmission of dengue. *Med Vet Entomol* 2010; 24: 66-73.

Bhatt S, Gething PW, Brady OJ, *et al*. The global distribution and burden of dengue. *Nature* 2013; 496: 504-7.

Buczak AL, Baugher B, Babin SM, *et al*. Prediction of high incidence of dengue in the Philippines. *PLoS Negl Trop Dis* 2014; 8: e2771.

Centers for Disease Control and Prevention (CDC). Mosquito-borne transmission, 2018 [cited 2021 Jan 03]. Available from: URL: https://www.cdc.gov/dengue/training/cme/ccm/page45915. html

Coberly JS, Fink CR, Elbert E, *et al*. Tweeting fever: are tweet extracts a valid surrogate data source for dengue fever? *Online J Public Health Inform* 2013; 5: e64.

Cheong YL, Burkart K, Leitão PJ, Lakes T. Assessing weather effects on dengue disease in Malaysia. *Int J Environ Res Public Health* 2013; 10: 6319-34.

Dela Cruz AC, Lubrca JA, Punzalan BD, Martin MC. Time series analysis of dengue incidence: forcasting model using climate variables as predictors. *Acta Manilana* 2012; 60: 19-26.

Department of Health (DOH). The 2018 Philippine Health Statistics, 2018 [cited 2019 Dec 06]. Available from: URL: https://doh.gov.ph/sites/default/files/publications/2018%20Philippine%20Health%20Statistics.pdf

Espino F, Marco J, Salazar NP, Salazar F, Mendoza Y, Velazco A. Community-based dengue vector control: experiences in behavior change in Metropolitan Manila, Philippines. *Pathog Glob Health* 2012; 106: 455-61.

Focks DA, Daniels E, Haile DG, Keesling JE. A simulation model of the epidemiology of urban dengue fever: literature analysis, model development, preliminary validation, and samples of simulation results. *Am J Trop Med Hyg* 1995; 53: 489-506.

Gbadamosi B, Ojo MM, Oke SI, Matadi MB. Qualitative analysis of a dengue fever model. *Math Comput Appl* 2018; 23: 33.

Gubler DJ. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. *Trends Microbiol* 2002; 10: 100-3.

Gubler DJ. Dengue and dengue hemorrhagic fever: its history and resurgence as a global public health problem. In: Gubler DJ, Kuno G, editors. Dengue and dengue hemorrhagic fever. Oxford: CAB International; 1997. p. 1-22.

Gupta HV, Kling H, Yilmazi KK, Martinez GF. Decomposition of the mean squared error and NSE performance criteria: implications for imroving hydrological modelling. *J Hydrol* 2009; 377: 80-91.

Halide H, Ridd P. A predictive model for dengue hemorrhagic fever epidemics. *Int J Environ Health Res* 2008; 18: 253-65.

Halstead SB. Controversies in dengue pathogenesis. *Paediatr Int Child Health* 2012; 32 Suppl 1: 5-9.

Hii YL, Zaki RA, Aghamohammadi N, Rocklöv J. Research on climate and dengue in Malaysia: a systematic review. *Curr Environ Health Rep* 2016; 3: 81-90.

Hii YL, Zhu H, Ng N, Ng LC, Rocklöv J. Forecast of dengue incidence using temperature and rainfall. *PLoS Negl Trop Dis* 2012; 6: e1908.

Ho CC, Ting CY. Time series analysis and forecasting of dengue using open data. *Lect Notes Comput Sc* 2015; 9429: 51-63.

Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput* 1997; 9: 1735-80.

Hsia CH, Chiang JS, Guo JM. Multiple moving objects detection and tracking using discrete wavelet transform, discrete wavelet transforms, 2011 [cited 2020 Jun 20]. Available from: URL: https://www.intechopen.com/books/discrete-wavelet-transforms-biomedical-applications/multiple-moving-objects-detection-and-tracking-using-discrete-wavelet-transform

Iguchi JA, Seposo XT, Honda Y. Meteorological factors affecting dengue incidence in Davao, Philippines. *BMC Public Health* 2018; 18: 629.

Lee HS, Nguyen-Viet H, Nam VS, *et al*. Seasonal patterns of dengue fever and associated climate factors in 4 provinces in Vietnam from 1994 to 2013. *BMC Infect Dis* 2017; 17: 218.

Legates DR, McCabe Jr GJ. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimaticmodel validation. *Water Resour Res* 1999; 35: 233-41.

Li CF, Lim TW, Han LL, Fang R. Rainfall, abundance of *Aedes aegypti* and dengue infection in Selangor, Malaysia. *Southeast Asian J Trop Med Public Health* 1985; 16: 560-8.

Li Z, Liu T, Zhu G, *et al*. Dengue Baidu Search Index data can improve the prediction of local dengue epidemic: a case study in Guangzhou, China. *Plos Negl Trop Dis* 2017; 11: e0005354.

Liu B, Zhang L, Wang Q, Chen J. A novel method for regional No2 concentration prediction using Discrete Wavelet Transform and an LSTM network. *Comput Intell Neurosci* 2021; 2021: 6631614.

McLennan-Smith TA, Mercer GN. Complex behaviour in a dengue model with seasonally varying vector population. *Math Biosci* 2014; 248: 22-30.

Messer WB, Gubler DJ, Harris E, Sivananthan K, de Silva AM. Emergence and global spread of a dengue serotype 3, subtype III virus. *Emerg Infect Dis* 2003; 9: 800-9.

Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans ASABE* 2007; 50: 885-900.

Naik Singru R, Lindfield M. Republic of the Philippines National Urban Assessment, 2014 [cited 2019 Dec 06]. Available from: URL: https://think-asia.org/bitstream/handle/11540/754/philippines-national-urban-assessment.pdf?sequence=1

Olah C. Understanding LSTM networks, 2015 [cited 2020 Jun 20]. Available from: URL: http://colah.github.io/posts/2015-08-Understanding-LSTMs/

Pandey A, Mubayi A, Medlock J. Comapring vector-host SIR models for dengue transmission. *Math Biosci* 2013; 246: 252-9.

Patil GM, Rao KS, Satyanarayana K. Heart disease classification using discrete wavelet transform coefficients of isolated beats. In: Lim CT, Goh JCH, editors. Proceedings of the 13th International Conference on Biomedical Engineering; 2008 Dec 3-6; Singapore. Berlin, Germany: Springer; 2009.

Percival D, Walden A. Wavelet methods for time-series analysis. Cambridge, UK: Cambridge University Press; 2000.

Pham HV, Doan HT, Phan TT, Minh NN. Ecological factors associated with dengue fever in a central highlands province, Vietnam. *BMC Infect Dis* 2011; 11: 172.

Potts JA, Gibbons RV, Rothman AL, *et al*. Prediction of dengue disease severity among pediatric Thai patients using early clinical laboratory indicators. *Plos Negl Trop Dis* 2010; 4: e769.

Promprou S, Jaroensutasinee M, Jaroensuta K. Climatic factors affecting dengue haemorrhagic fever incidence in southern Thailand. *Dengue Bull* 2005; 29: 41-8.

Sang S, Yin W, Bi P, *et al*. Predicting local dengue transmission in Guangzhou, China, through the influence of imported cases, mosquito density and climate variability. *PLoS One* 2014; 9: e102755.

Shang CS, Fang CT, Liu CM, Wen TH, Tsai KH, King CC. The role of imported cases and favorable meteorological conditions in the onset of dengue epidemics. *PLoS Negl Trop Dis* 2010; 4: e775.

Shmueli G. Wavelet-based monitoring for biosurveillance. *Axioms* 2013; 2: 345-70.

Siami-Namini S, Tavakoli N, Siami Namin A. A Comparison of ARIMA and LSTM in Forecasting Time Series, 2018 [cited 2020 Jun 20]. Available from: URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8614252&casa_token=I1H_-oT040cAAAAA:FQsm3kckvCElOvirMWeU8dwJ4uVzY-LY60BxPSHC2EWCM4QGH-oczqkZe-2Bk3IxQP3thEbFE&tag=1

Su GL. Correlation of climatic factors and dengue incidence in Metro Manila, Philippines. *Ambio* 2008; 37: 292-4.

Tsui F. Time series prediction using a multi-resolution dynamic predictor [dissertation]. Pittsburgh (PA): University of Pittsburgh; 1996.

van Panhuis WG, Choisy M, Xiong X, *et al*. Region-wide synchrony and traveling waves of dengue across eight countries in Southeast Asia. *Proc Natl Acad Sci USA* 2015; 112: 13069-74.

Willmott CJ, Robeson SM, Matsuura K. A refined index of model performances. *Int J Climatol* 2011; 32: 2088-94.

Wongkoon S, Jaroensutasinee M, Jaroensutasinee K, Preechaporn W, Chumkiew S. Larval occurrence and climatic factors affecting DHF incidence in Samui Islands, Thailand, 2007 [cited 2020 Jun 20]. Available from: URL: https://publications.waset.org/389/larval-occurrence-and-climatic-factors-affecting-dhf-incidence-in-samui-islands-thailand

World Health Organization (WHO). Handbook for integrated vector management, 2012 [cited 2020 Jun 20]. Available from: URL: http://apps.who.int/iris/bitstream/handle/10665/44768/9789241502801_eng.pdf?sequence=1