

Developing a dengue early warning system using time series model: Case study in Tainan, Taiwan

Xiao-Wei Chen (1), Chyan-Deng Jan (2), and Ji-Shang Wang (3)

(1) International Master Program on Natural Hazards Mitigation and Management, National Cheng Kung University, Tainan 70101, Taiwan, R.O.C. (winnie5566b@gmail.com), (2) Department of Hydraulic and Ocean Engineering, National Cheng Kung University, Tainan 70101, Taiwan, R.O.C. (cdjan@mail.ncku.edu.tw), (3) Ecological Soil and Water Conservation Research Center, National Cheng Kung University, Tainan 70101, Taiwan, R.O.C. (rheo.js@gmail.com)

Dengue fever (DF) is a climate-sensitive disease that has been emerging in southern regions of Taiwan over the past few decades, causing a significant health burden to affected areas. This study aims to propose a predictive model to implement an early warning system so as to enhance dengue surveillance and control in Tainan, Taiwan. The Seasonal Autoregressive Integrated Moving Average (SARIMA) model was used herein to forecast dengue cases. Temporal correlation between dengue incidences and climate variables were examined by Pearson correlation analysis and Cross-correlation tests in order to identify key determinants to be included as predictors. The dengue surveillance data between 2000 and 2009, as well as their respective climate variables were then used as inputs for the model. We validated the model by forecasting the number of dengue cases expected to occur each week between January 1, 2010 and December 31, 2015. In addition, we analyzed historical dengue trends and found that 25 cases occurring in one week was a trigger point that often led to a dengue outbreak. This threshold point was combined with the season-based framework put forth by the World Health Organization to create a more accurate epidemic threshold for a Tainan-specific warning system.

A Seasonal ARIMA model with the general form: $(1,0,5)(1,1,1)_{52}$ is identified as the most appropriate model based on lowest AIC, and was proven significant in the prediction of observed dengue cases. Based on the correlation coefficient, Lag-11 maximum 1-hr rainfall ($r=0.319$, $P<0.05$) and Lag-11 minimum temperature ($r=0.416$, $P<0.05$) are found to be the most positively correlated climate variables. Comparing the four multivariate models (*i.e.* 1, 4, 9 and 13 weeks ahead), we found that including the climate variables improves the prediction RMSE as high as 3.24%, 10.39%, 17.96%, 21.81% respectively, in contrast to univariate models. Furthermore, the ability of the four multivariate models to determine whether the epidemic threshold would be exceeded in any given week during the forecasting period of 2010-2015 was analyzed using a contingency table. The 4 weeks-ahead approach was the most appropriate for an operational public health response with a 78.7% hit rate and 0.7% false alarm rate. Our findings indicate that SARIMA model is an ideal model for detecting outbreaks as it has high sensitivity and low risk of false alarms. Accurately forecasting future trends will provide valuable time to activate dengue surveillance and control in Tainan, Taiwan. We conclude that this timely dengue early warning system will enable public health services to allocate limited resources more effectively, and public health officials to adjust dengue emergency response plans to their maximum capabilities.