

RESEARCH ARTICLE

Modelling and analysis of dengue disease transmission in urban Colombo: a wavelets and cross wavelets approach

W.P.T.M. Wickramaarachchi^{1*}, S.S.N. Perera¹ and S. Jayasinghe²

¹Research and Development Center for Mathematical Modelling, Department of Mathematics, Faculty of Science, University of Colombo, Colombo 03.

²Department of Clinical Medicine, Faculty of Medicine, Po Box 271, No 25, Kynsey Road, Colombo 08.

Revised: 02 April 2015; Accepted: 22 April 2015

Abstract: Transmission of dengue is a complex dynamic process. Dengue virus is transmitted to humans by the female *Aedes aegypti* mosquitoes. Dengue was first identified in Sri Lanka during 1960s and the risk due to dengue has increased rapidly during the past few decades mainly in the urban areas. There are several studies on dengue transmission based on mathematical and statistical models. However they are not capable of describing the complex dynamics of transmission since its transmission mechanism is highly dependent on various external factors such as climate, demography, geography and human mobility. Wavelet analysis is a powerful tool in mathematics, which can be used to reveal useful information of a highly non-stationary epidemiological time series. Wavelet theory can be used to obtain vital signals in a noisy epidemiological time series, which is useful to implement early warning systems to control the spread of epidemics. This particular study was aimed at investigating the applicability of wavelet theory to analyse dengue transmission in urban Colombo. The wavelet transformation was performed with data from weekly dengue cases time series from year 2006 to 2012 in the Colombo Municipal Council (CMC) area, and significant regions were observed in the spectrum that correspond to approximately 25 week cycles during the time period from year 2009 to 2012. The cross wavelet power spectrum showed that rainfall was leading the dengue cases in Colombo during the time period 2010 to 2012.

Keywords: Cross wavelets, dengue epidemics, dynamical system, wavelet coherence, wavelet power spectrum.

INTRODUCTION

Dengue is an arboviral disease, which is a major public health concern in the tropical regions. This disease has become a major public health problem in Sri Lanka

during the past five decades. Thousands of people have been vulnerable to dengue in Sri Lanka each year causing large number of deaths. During the first six months of the year 2014, 20074 dengue cases have been reported in Sri Lanka according to the Epidemiology Unit, Ministry of Health. Dynamics of the dengue disease is complex since its transmission process heavily depends on many biological, environmental, social and geographical factors. It is known to be highly influenced by climate variability. Increasing temperature reduces the length of the incubation period of mosquitoes, enabling a large number of infected female mosquitoes to transmit the disease (Watts *et al.*, 1987). Rainfall provides mosquitoes more breeding sites, and a favourable environment for breeding may increase the mosquito density (Huang *et al.*, 2013). Furthermore, unsystematic urban development and land use have changed the flow of natural rainwater in urban areas, resulting in the collection of rainwater in low wetlands, which are ideal for mosquito breeding. The rapid increase in the human population and migration to cities have led to overcrowding and high human mobility that facilitate dengue transmission. The interactions of all these variables make it extremely difficult to predict the transmission of dengue by using classical mathematical and statistical models. The classical SIR (S-susceptible, I-infected, R-recovered) models of dengue transmission are useful and reveal an effective framework for understanding the dynamics of the transmission. These SIR compartment models describe the dynamics of dengue transmission in terms of population dynamics. Further these models explain the interaction between susceptible, infected and recovered human populations together with susceptible and infected mosquito populations in a totally homogeneous environment (Pongsumpun, 2006). All these classical models mainly

*Corresponding author (wptharindu86uoc@gmail.com)

discuss the dynamics of the disease with fixed parameter spaces. Due to the high dependency on external factors, these models with fixed parameter values produce unrealistic results. Thus the complexity in biological systems should be addressed at the parameter estimation level so that suitable mathematical tools can be used to estimate the parameters under uncertainty.

Many studies based on time series approaches considering one parameter at one stage can be found. However, the inherent complexity of transmission and extremely non-stationary epidemiological datasets make time series models less effective. As a solution to overcome these problems, wavelet analysis has been applied by researchers to characterise and estimate dependencies among these non-stationary signals (Zang *et al.*, 2003; Cazelles *et al.*, 2007). Wavelet analysis is a powerful mathematical tool, which performs time-frequency decomposition of the signals, and estimate the spectral characteristics as a function of time. During the past three decades, signal processing has grown to be a major discipline in engineering, theoretical physics and computational mathematics. Since most finite energy signals, either natural or otherwise, are transient (or non-stationary) in nature, it is most natural and effective to represent these signals by localised finite energy bases. Because wavelets belong to this class of bases, the development of wavelet theory and its applications to signal processing in various engineering disciplines, mathematical epidemiology and computer science, have gained tremendous popularity in recent years.

In mathematics, Fourier analysis has been the dominant tool for the representation of signals for past decades (Olkkonen, 1994). Although it is capable of quantifying periodic components in a time series, it fails when characterising signals whose frequency changes with respect to time (Cazelles *et al.*, 2007). Wavelets are finite energy functions with time-frequency localisation properties that represent the transient signals with small finite number of coefficients. The functions narrow when high frequency features are presented and widen when the low frequencies are present (Cazelles *et al.*, 2007; 2008). Generally, wavelets are small waves, so that the area under the graph of the wavelet function $\psi(t)$ is zero that is $\int_{-\infty}^{\infty} \psi(t) dt = 0$.

Due to these properties wavelet analysis has become an active and interesting area of research. Further mathematical development of wavelet theory may provide researchers in the fields of engineering, biology and epidemiology more research opportunities and a good understanding of complex sciences. Numerous studies

have shown that wavelet analysis is useful in analysing noisy time series data. Cazelles *et al.*, (2005) have analysed the *El Nino* effect on dengue cases in Thailand. They have found a strong association between monthly dengue cases and the dynamics of *El Nino* for a 2 – 3 year period. Santos *et al.* (2003) have presented some applications of wavelet theory to analyse the rainfall patterns in Japan and North Eastern Brazil. Various Sri Lankan researchers have applied time series and regression models to predict dengue outbreaks in Sri Lanka (Goto *et al.*, 2013; Kavinga *et al.*, 2013). However, the wavelet approach of analysing noisy epidemiological time series data has not been used so far in the Sri Lankan context.

The objective of this study was to focus on the wavelet approach to analyse dengue transmission in urban Colombo. First it discusses the wavelet theory and an application to wavelet power spectrum by using a classical dynamic model of dengue transmission. Next, it presents the wavelet power spectra obtained for the three time series simulated from the dynamic model and the wavelet power spectrum for the real dengue data time series reported in Colombo Municipal Council (CMC) area from year 2006 to 2012. Finally it presents the cross wavelet power spectra obtained for the number of weekly dengue cases with weekly averaged maximum temperature and rainfall in CMC area from year 2006 to 2012.

METHODS AND MATERIALS

Wavelet theory

For a given wavelet $\psi(t)$, a scaled and translated version is given by

$$\psi_{ba}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad \dots(1)$$

The parameter $a (\neq 0)$ corresponds to the scale while b represents the translation. The wavelet $\psi_{01}(t) = \psi(t)$ is called the basic wavelet or the mother wavelet. The discrete translations and dilations (b,a) form an orthonormal basis for L^2 (or finite energy) space (Chui, 1992).

Continuous wavelet transform (CWT)

The continuous wavelet transform of a signal (time series) is a linear transform defined by the integral

$$\begin{aligned} W_{\psi} x(b,a) &= \int_{-\infty}^{\infty} x(t) \overline{\psi_{ba}(t)} dt \\ &= \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \\ &= \langle x(t), \psi_{ba}(t) \rangle \end{aligned} \quad \dots(2)$$

where $a, b \in R$ and $a \neq 0$.

The term $\langle x(t), \psi_{ba}(t) \rangle$ represents the inner product of two functions, defined in the L^2 space, which is given by

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(t) \overline{g(t)} dt. \quad \dots(3)$$

The wavelet function in equation (2) should possess both the conditions given by

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad \dots(4)$$

$$\int_{-\infty}^{\infty} \|\psi(t)\|^2 dt = 1 \quad \dots(5)$$

The above equations (4) and (5) imply that the original signal can be recovered by means of the inverse wavelet transformation given by

$$x(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_\psi(x)(b, a) \psi_{ba}(t) \frac{db da}{a^2} \quad \dots(6)$$

where C_ψ is a finite constant given by the integral

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|}{|\omega|} d\omega < \infty \quad \dots(7)$$

where $\hat{\psi}$ denotes the Fourier transform of ψ .

Wavelet power spectrum

The wavelet transform can be considered as a generalisation of the classical Fourier transform so that the spectral properties of the time series can be visualised. The wavelet power spectrum of the wavelet transform $W_\psi x(b, a)$ in equation (2) is computed as

$$S_\psi x(b, a) = \left\| W_\psi x(b, a) \right\|^2. \quad \dots(8)$$

The global wavelet power spectrum $\overline{S_\psi} x(a)$, which is comparable with the Fourier spectrum of a signal can be defined as the averaged energy of all wavelet coefficients of the same scale a .

$$\overline{S_\psi} x(a) = \frac{\sigma_x^2}{T} \int_0^T \left\| W_\psi x(b, a) \right\|^2 db, \quad \dots(9)$$

with σ_x is the standard deviation of the time series $x(t)$ and T the duration of the time series. Averaging the

scale components gives the mean variance of each time location defined as

$$\overline{S_\psi} x(b) = \frac{\sigma_x^2 \pi^{1/4} b^{1/2}}{C_\psi} \int_0^\infty a^{1/2} \|W_\psi x(b, a)\|^2 da, \quad \dots(10)$$

with C_ψ as in equation (7) and π is a constant in usual notations.

Wavelet coherency and phase difference

It is useful to quantify the statistical relationship between two signals if we have two non-stationary time series. The wavelet coherence function measures the correlation between two time series $x(y)$ and $y(t)$ (Cazelles *et al.*, 2007). The wavelet cross spectrum of the two time series $x(y)$ and $y(t)$ can be defined as

$$W_\psi x, y(b, a) = W_\psi x(b, a) \cdot W_\psi^* y(b, a), \quad \dots(11)$$

where ‘*’ denotes the complex conjugate. The cross spectrum normalised by the spectrum of each signal gives the wavelet coherence, which is defined as

$$R_\psi x, y(b, a) = \frac{\left\| \langle W_\psi x, y(b, a) \rangle \right\|}{\left\| \langle W_\psi x, x(b, a) \rangle \right\|^{1/2} \left\| \langle W_\psi y, y(b, a) \rangle \right\|^{1/2}}. \quad \dots(12)$$

The notation ‘ $\langle \rangle$ ’ stands for the smoothing operator in both time and scale parameters (more details on the smoothing can be found in Cazelles *et al.*, 2007). The wavelet coherency is similar to simple statistical correlation but $0 \leq R_\psi x, y(b, a) \leq 1$. This measure equals to 1 implies a perfect relationship between the two signals in both time scale, however this goes to 0 if the two time series are independent.

Choice of the mother wavelet function

The selection of the mother wavelet function is not done arbitrarily and will be dictated by the kind of real world application the researcher has in mind. It is required to provide quantitative information about the phase interactions between the two time series. In this regard continuous and complex valued analytic wavelets are the natural choices. A majority of the researchers prefer Morlet wavelet as the mother wavelet function due to its better time and frequency localisation capability (Torrence & Compo, 1998; Conraria & Soares, 2011).

The Morlet wavelet is defined as

$$\psi(t) = K \cdot \exp(-i 2 \omega_0 t) \cdot \exp(-t^2 / 2), \quad \dots(13)$$

where ω_0 represents the central angular frequency of the wavelet.

In order to $\psi(t)$ have unit energy, the normalisation constant K is selected such that

$$K = \pi^{-1/4} \text{ (Conraria \& Soares, 2011)}. \quad \dots(14)$$

The relationship between the frequencies and wavelet scales can be derived as $1/f = \frac{4\pi a}{\omega_0 + \sqrt{2 + \omega_0^2}}$. It can be easily shown that if $\omega_0 \approx 2\pi$ then $f \approx 1/a$. Thus the term, $1/f$ can be used instead of the scalar a in all the above equations. An example of the Morlet is given in Figure 1.

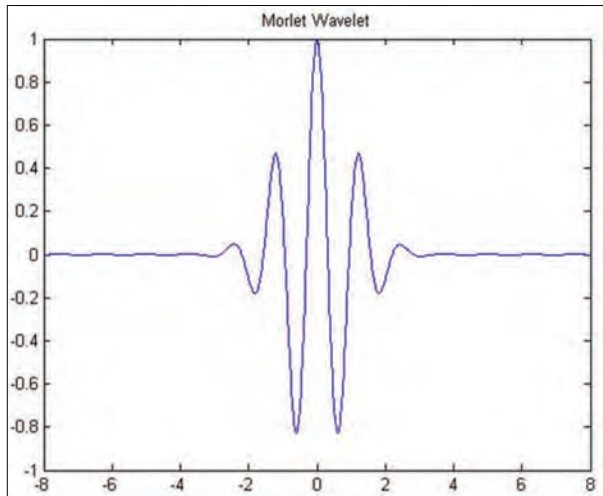


Figure 1: Morlet wavelet

Mathematical model of dengue transmission

The mathematical model of dengue transmission described by a normalised system of non-linear ordinary differential equations was considered (Pongsumpun, 2006). This model describes the interaction between susceptible (S) and infected (I) human populations with infected vector (I_v) population. In this model, it is assumed that only infected mosquitoes can transmit the dengue virus to the human population.

$$\frac{dS}{dt} = \lambda - \gamma'_h S I_v - \mu_h S$$

$$\frac{dI}{dt} = \gamma'_h S I_v - (\mu_h + r) I$$

$$\frac{dI_v}{dt} = \gamma_v (1 - I_v) I - \mu_v I_v, \quad \dots(15)$$

where $\gamma_v = b\beta'_v$ and $\gamma'_h = b\beta'_h n$ with $n = \frac{D}{N_T} \mu_v$. The left

hand side of the system represents the rate of change in the normalised susceptible human, infected human and infected mosquito populations, respectively.

Parameter definitions in the above system:

- N_T – the total number of human population
- λ – the birth rate of the human population
- b – the biting rate of the mosquitoes
- β_h – the transmission probability of the virus from mosquitoes to humans
- β_v – the transmission probability of the virus from humans to mosquitoes
- μ_h – the death rate of the human population
- r – the recovery rate of the human population
- D – the constant recruitment rate of the vector population
- μ_v – the death rate of the mosquito population

The seasonal variation of dengue transmission is modeled as a sinusoidal variation given by

$$\beta'_h = \beta_h (1 + \zeta \sin(\varphi t)). \quad \dots(16)$$

The parameter ζ in equation (16) measures the influence of seasonality towards the transmission of the dengue virus (Pongsumpun, 2006). Equation (15) was solved numerically using ode 45 solver in MatLab and obtained the three time series $S(t)$, $I(t)$ and $I_v(t)$ in Figure 2. This solver function implements a fourth order Runge-Kutta method with a variable time step for efficient computation (Isotani et al., 2012).

RESULTS AND DISCUSSION

These three time series were used to perform the wavelet analysis using the wavelet toolbox in MatLab. The Morlet wavelet was used as the mother wavelet function and the spectra was obtained using 5 % level of significance. The corresponding wavelet power spectra are presented in Figure 3.

The number of weekly dengue cases in the Colombo Municipal Council (CMC) area from year 2006 to 2012 was used. These data were recorded by the epidemiological unit in CMC from the reported dengue cases in the CMC area. The non-stationary time series obtained from the number of dengue cases database is given in Figure 4.

The non-stationary epidemiological time series was transformed into wavelet power spectrum using wavelet tool box in MatLab. The resulting wavelet power spectrum

is given in Figure 5. Again the Morlet wavelet was the mother wavelet function and 5 % level of significance was used.

Three time series namely, the number of weekly dengue cases in CMC area, weekly averaged maximum temperature, and rainfall from year 2006 to 2012 were used to obtain the cross wavelet power spectra presented in Figure 5. Climate data were obtained from the Department of Meteorology in Colombo.

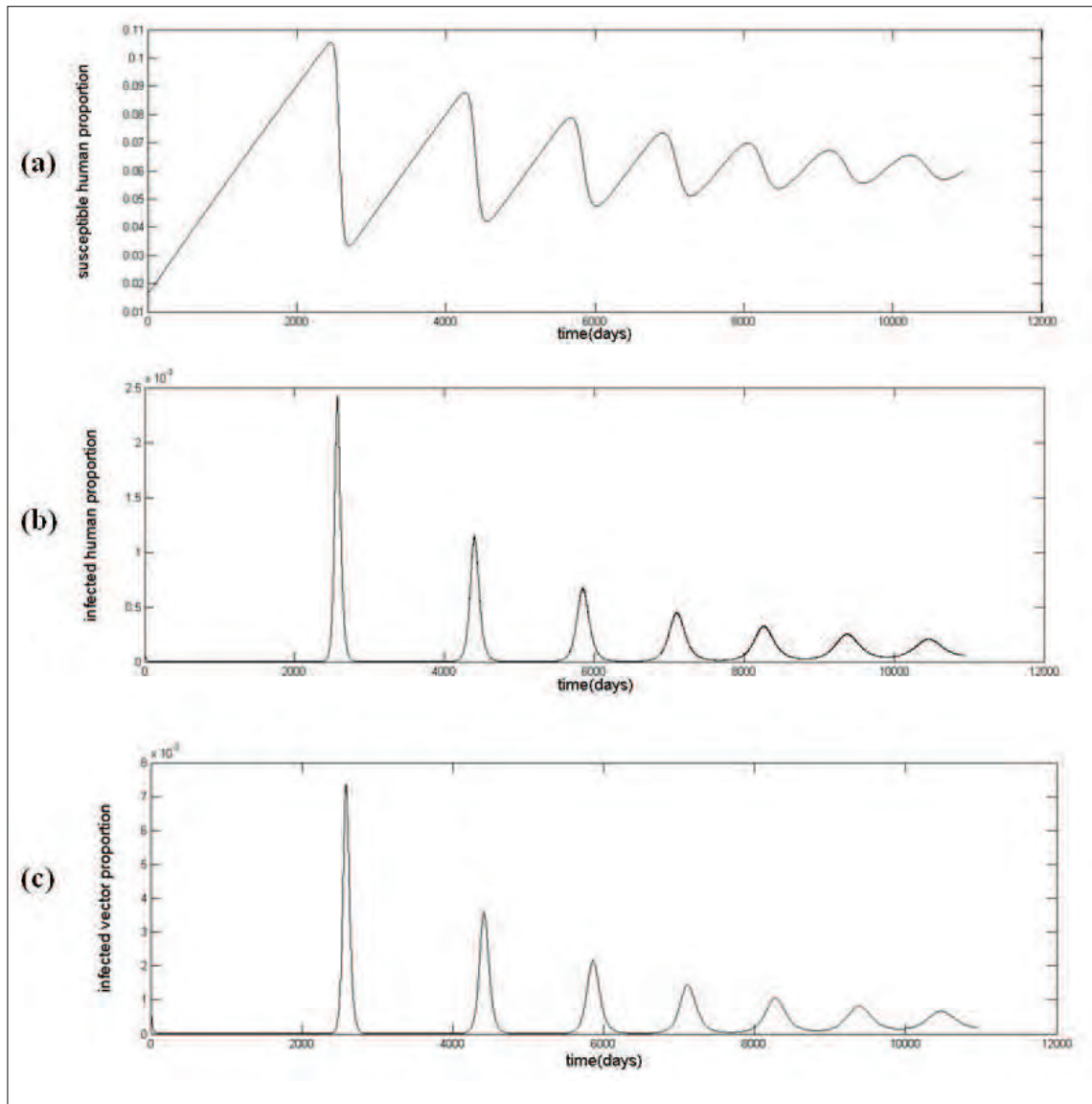


Figure 2: Time series of the susceptible human population (a); time series of the infected human population (b) and time series of the infected vector population (c)

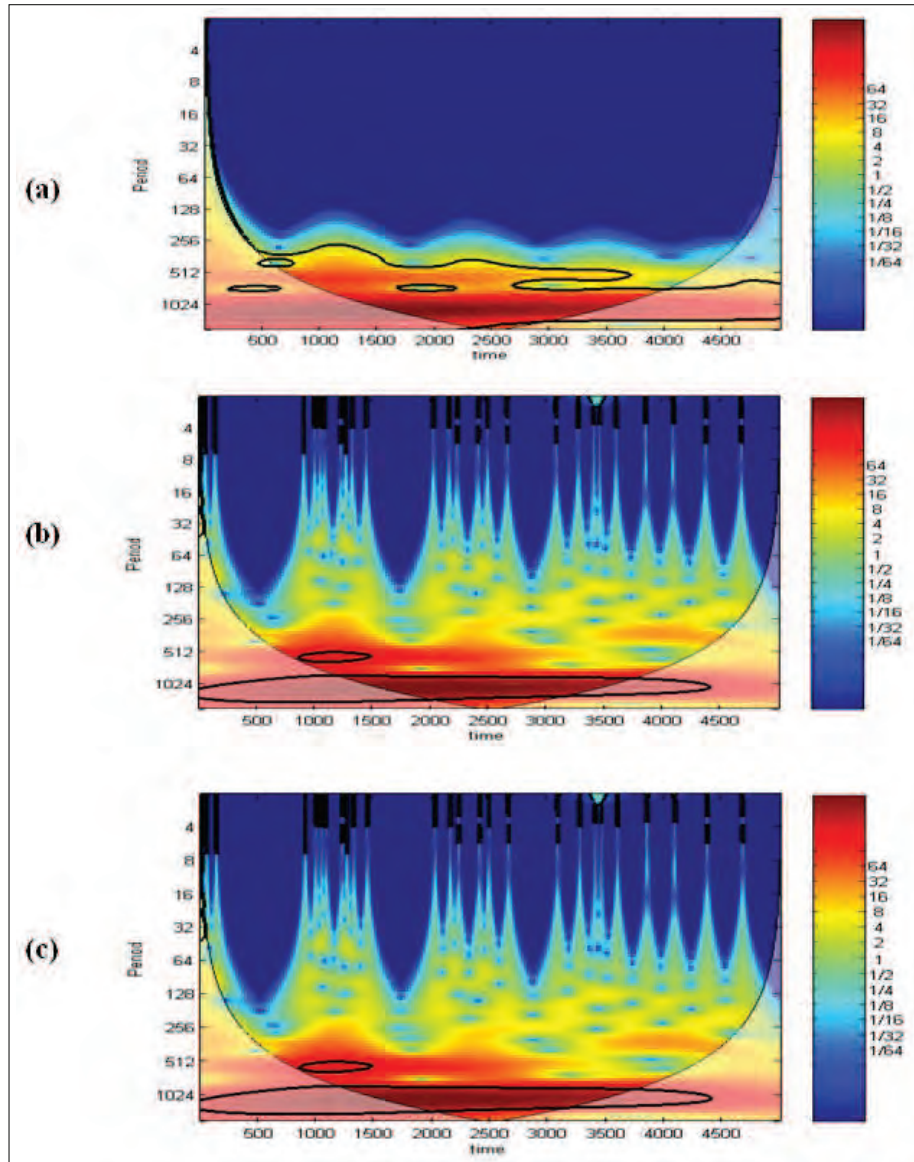


Figure 3: Wavelet power spectrum of the susceptible human population (a); infected human population (b) and the infected mosquito population (c) simulated from the dynamic model

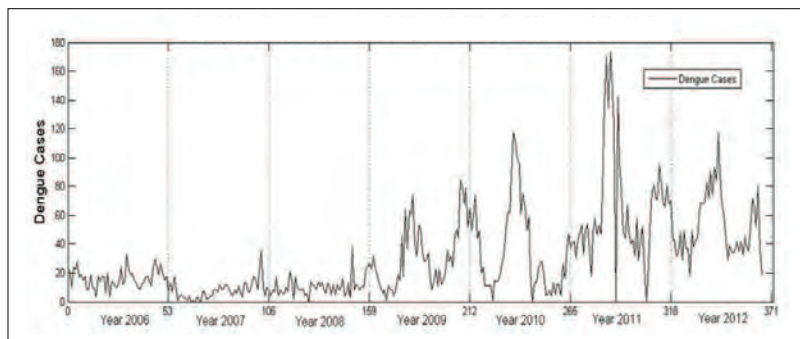


Figure 4: Weekly dengue cases time series for the reported dengue cases in CMC area from year 2006 to 2012

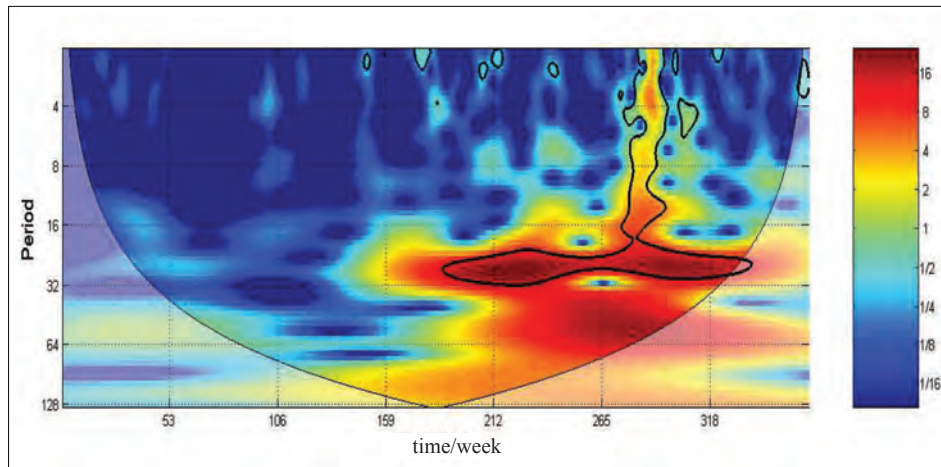


Figure 5: Wavelet power spectrum for the CMC dengue data series from year 2006 to 2012

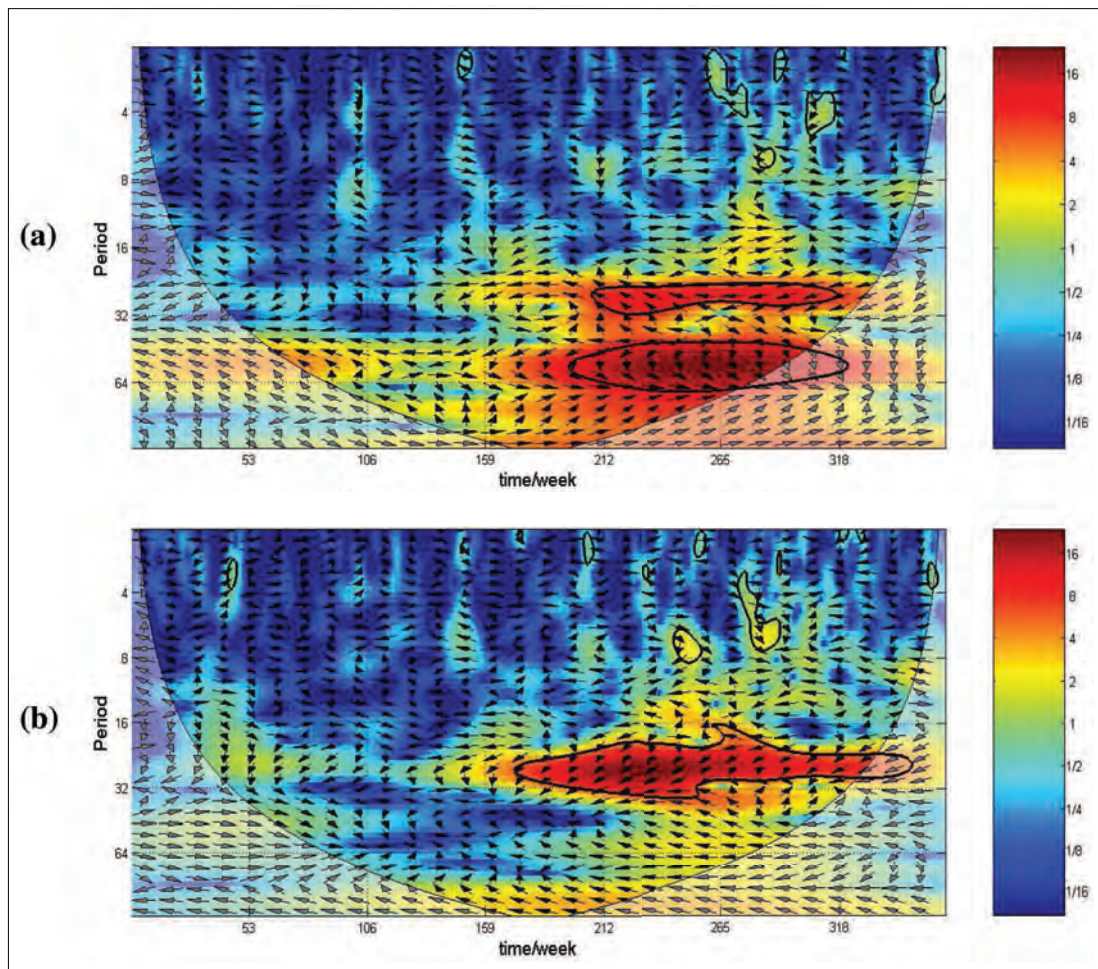


Figure 6: Cross wavelet power spectrum for weekly dengue cases with maximum temperature (a) and the cross wavelet power spectrum for weekly dengue cases with rainfall from year 2006 to 2012 (b)

In this study the simulated three time series from the dynamic model of dengue transmission was first analysed. It is clearly seen from Figure 2 that all three time series have regular peaks at the beginning but the system reaches to its equilibrium as $t \rightarrow \infty$. The simulated time series for infected humans and infected mosquitoes seem to be identical but they have different initial conditions and different populations in each time. Therefore the corresponding wavelet power spectra in Figure 3 seem to be identical with approximately 1024 day cycles, which are significant at 5 % level.

The real dengue data time series in Figure 4 is highly non-stationary and the corresponding wavelet power spectrum in Figure 5 reflects an approximately 25 week cycle from year 2009 to 2012, which can be explained as the semiannual cycle of dengue. The cross wavelet power spectrum of the number of dengue cases and the maximum temperature in Figure 6 (a) shows two significant regions after year 2010. According to Figure 6 (b), arrows pointing down in the significant region at 5 % level indicate that the rainfall follow the number of dengue cases during the period from year 2010 to 2012.

CONCLUSION

Dengue is one of the major diseases in the world having a complex transmission mechanism. The existing classical mathematical models of dengue transmission with fixed parameter spaces have not been fully capable of describing the complex transmission dynamics of dengue. Previous studies can be found based on statistical models of dengue such as time series. However time series models applied to analyse epidemiological data do not reveal all the information because they are extremely non-stationary.

In this paper, the wavelet power spectrum is proposed to identify significant cycles of the epidemic disease transmission, and hence it can further analyse potential external factors, which influence those cycles. The power spectra corresponding to the simulated time series showed significant periodic cycles due to equation (16), which was used to theoretically address the periodicity in the dynamic model. Based on this study in real number of dengue cases, an approximately 25 week cycle of dengue transmission from year 2009 to 2012 was found, which was significant at 5 % level. The cross wavelet power spectrum and the coherence can be useful to analyse the influence of climate factors such as rainfall, temperature and humidity. Using the cross wavelet power spectrum, it was observed that the number of dengue cases from year 2010 to 2012 is determined by rainfall.

Cross wavelet power spectrum is an important application in wavelet analysis. This may be used to further analyse the influence of other external factors such as geography, demography and human mobility towards the transmission of dengue disease.

Acknowledgement

The HETC Window 3 grant under the Ministry of Higher Education is acknowledged for funding the corresponding author. The Colombo Municipal Council (CMC) is also acknowledged for providing dengue cases data for this study.

REFERENCES

1. Cazelles B., Chavez M. & Constantin de Magny G. (2007). Time-dependent spectral analysis of epidemiological time-series with wavelets. *Journal of the Royal Society* **4**: 625 – 636.
DOI: <http://dx.doi.org/10.1098/rsif.2007.0212>
2. Cazelles B., Chavez M., Berteaux D., Menard F., Vik J.O., Jenouvrier S. & Stenseth N.C. (2008). Wavelet analysis of ecological time series. *Oecologia* **156**: 287 – 304.
DOI: <http://dx.doi.org/10.1007/s00442-008-0993-2>
3. Cazelles B., Chavez M., McMichael A.J. & Hales S. (2005). Nonstationary influence of El Niño on the synchronous dengue epidemics in Thailand. *PLOS Medicine* **2**(4): 313 – 318.
DOI: <http://dx.doi.org/10.1371/journal.pmed.0020106>
4. Chui C.K. (1992). *Wavelet Analysis and its Applications: An Introduction to Wavelets*, volume 1, pp. 49 – 80. Academic Press, San Diego, USA.
DOI: <http://dx.doi.org/10.1016/b978-0-12-174590-5.50029-0>
5. Conraria L.A. & Soares M.J. (2011). The continuous wavelet transform: a primer. *NIPE Working Paper Series* WP 23/2010: 1 – 42.
6. Goto K., Kumarendran B., Mettananda S. & Gunasekera D. (2013). Analysis of effects of meteorological factors on dengue incidence in Sri Lanka using time series data. *PLOS ONE* **8**(5): 1 – 8.
DOI: <http://dx.doi.org/10.1371/journal.pone.0063717>
7. Huang X., Clements C.A., William G., Milinovich G. & Hu W. (2013). A threshold analysis of dengue transmission in terms of weather variables and imported dengue cases in Australia. *Emerging Microbes and Infections* **2**(e87): 1 – 7.
DOI: <http://dx.doi.org/10.1038/emi.2013.85>
8. Isotani S., Pontuschka W.M. & Isotani S. (2012). An algorithm to optimize the calculation of the fourth order Runge-Kutta method applied to the numerical integration of kinetics coupled differential equations. *Applied*

- Mathematics* **3**: 1583 – 1592.
DOI: <http://dx.doi.org/10.4236/am.2012.311218>
9. Kavinga H.W.B., Jayasundara D.D.M. & Jayakody D.N.K. (2013). A new dengue outbreak statistical model using the time series analysis. *European International Journal of Science and Technology* **2**(10): 35 – 52.
 10. Olkkonen H. (1994). Running discrete Fourier transform for time-frequency analysis of biomedical signals. *Medical Engineering and Physics* **17**(6): 455 – 458.
 11. Pongsumpun P. (2006). Transmission model for dengue disease with and without the effect of extrinsic incubation period. *KMITL Science and Technology Journal* **6**(2): 74 – 82.
 12. Santos C.A.G., Galvao O.L. & Trigo R.M. (2003). Rainfall data analysis using wavelet transform. *Proceedings of the International Conference: Hydrology of the Mediterranean and Semiarid Regions*, 1 - 4 April. Montpellier, France, pp. 195 – 201.
 13. Torrence C. & Compo G.P. (1998). A practical guide to wavelet analysis; programme in atmospheric and oceanic sciences. University of Colorado, Boulder, Colorado. *Bulletin of the American Meteorological Society* **79**(1): 61 – 78.
 14. Watts D.M., Burke D.S., Harrison B.A., Whitmore R.E. & Nisalak A. (1987). Effect of temperature on the vector efficiency of *Aedes aegypti* for dengue 2 virus. *American Journal of Tropical Medicine and Hygiene* **36**(1): 143 – 152.
 15. Zang J., Tsui F., Wagner M.M. & Hogan W.R. (2003). Detection of outbreaks from time series data using wavelet transform. *Proceedings of the AMIA 2003 Symposium*, pp. 748 – 752.