Forecasting Dengue Incidence Utilizing Geographic Information System and Autoregressive Integrated Moving Average Models

Markdy Y. Orong¹, Geraldine D. Durias¹, Rolysent K. Paredes¹, Jezreel Marc E. Pasay²

¹College of Computer Studies, Misamis University, Ozamiz City, Philippines
²College of Nursing and School of Midwifery, Misamis University, Ozamiz City, Philippines

Corresponding author: Markdy Y. Orong, email: markdy.orong@mu.edu.ph

Abstract

In the health discipline, forecasting is gaining importance due to its capability of anticipating the spread of diseases. The Philippine government launched the dengue surveillance map to help detect the spread of the disease. However, technology-based solutions to forecast dengue cases do not exist in the country. This study aimed to forecast the possible outbreak of dengue in Ozamiz City from 2016 to 2020 using a web-based technology system that generates maps and allows the input of new records of dengue cases to update the forecasting patterns. This study utilized the Geographic Information System (GIS) to map the spread of dengue outbreak and the Autoregressive Integrated Moving Average (ARIMA) to forecast the extent of dengue occurrence. The data included in this study were the dengue cases reported in seven hospitals in Ozamiz City on a monthly basis from January 2008 up to the second quarter of 2015. Based on the collected data, dengue incidence is lower in rural than in urban barangays. However, a rapid increase in dengue incidence in rural barangays is evident in the five-year forecast. This information can aid the community in designing and implementing preventive measures to address dengue epidemics during the forecasted year or season.

Keywords: epidemics, health, map, outbreak, technology
Introduction

Scientific research in our present time is more data extensive and collaborative than in the past (Tenopir et al., 2011). Data sharing then becomes an essential part of the scientific method to verify results and extend research from previous findings (Kim, 2013). Shared data have served as a benchmark that allows refinement of methods of collections and analysis (Vision, 2010). As the data increase, they are aggregated into collections and become available for reuse (Kim & Stanton, 2012). It is in this manner that data become an important end product to strengthen the process of discovery and decision.

Data sharing is gaining importance in the health sector (Roski et al., 2014; van Panhuis et al., 2014). Collecting data from disease outbreaks can generate valuable knowledge in determining strategies for disease control. Sharing of disease data to inform the community of their extent may allow designing preventive measures to stop the continued spread of risk. Disease data sharing is, therefore, an input for collaboration, response, and recovery stages of disease control (Gao et al., 2008). Since disease studies have strong spatial aspects including case disease location and diffusion (Koch, 2005), the goal of data sharing is to communicate the disease information directly to the community. In this context, the responsible units in the locality can warn the people and contain the possible spread of risk.

The use of technology can allow data sharing. The Local Area Network (LAN) enables users to share data within the network. The fast growing technology introduces the use of data sharing through Geographic Information System (GIS) which associates the concept of mapping to visualize the collated data. The system provides the procedure for managing, storing, analyzing, and mapping information (Gao et al., 2008). The technology is capable of analyzing spatial and non-spatial data. This system can also collate data to enhance substantially disease surveillance and render information along with other kinds of data such as demographic and environmental. Further, web-based GIS enables the generation of thematic maps dynamically and efficiently.

Since information has a significant role in data analysis, the study of Inoue et al. (2003) developed a database system that redistributes information via the Internet. The system automatically
generates graphic images based on the numerical data stored. The database system dynamically displays information on the map as well as numerical impression according to the real-time demand of the users. For example, Blanton et al. (2006) used a centralized database for diagnostic and demographic information collected by local, state, and federal agencies involved in rabies surveillance. The geo-referenced database remits data to internet-accessible mapping application that displays rabies surveillance data with environmental and geographic features. Furthermore, Qian et al. (2004) considered the web-based GIS approach to visualize global Severe Acute Respiratory Syndrome (SARS) information. These studies collected data and incorporated GIS to visualize information.

Aside from collecting and storing data, the extraction is also necessary to fully understand the meaning of these data. Buczak et al. (2014) utilized Fuzzy Association Rule Mining techniques to extract association rules from historical, epidemiological, environmental, and socio-economic data of disease incidence, as well as climate data indicating future weather patterns. A Prediction Model can then be generated by using selection criteria to choose a subset of rules for the classifier. The model can thereby predict the high or low incidence of a disease for a particular period.

Forecasting then is gaining importance in health sector due to its capability of anticipating the spread of diseases. Much attention in the world is currently towards the mosquito-borne viral disease called dengue (World Health Organization [WHO], 2015a). The bites of infected Aedes mosquitoes, principally Aedes aegypti can transmit the virus to humans. More than 2.5 billion people or over 40% of the world’s population are at risk from dengue. Currently, there are about 50–100 million dengue infections every year worldwide. The number of incidences significantly varies in all countries.

In other countries, there were studies conducted predicting the possible spread of dengue and viral hemorrhagic fever. Dom (2013) developed a forecasting model for the incidence of dengue cases in Subang Jaya, Malaysia using time series analysis. The model utilized the Autoregressive Integrated Moving Average (ARIMA) as the statistical tool to analyze the collected data from 2005 to 2010 and generate a
temporal model using climate variables for the prediction of dengue cases in the area. The approach predicts dengue incidence by extrapolating dengue patterns. Karim et al. (2012) used the monthly reported dengue cases and climate data from the year 2000-2008 in Bangladesh and utilized multiple linear regressions to interpret the information. Log transformation of the dependent variable normalized the data for linear regression. The study of Liu et al. (2011) utilized ARIMA in forecasting incidence of hemorrhagic fever with renal syndrome in China. Luz et al. (2008) claimed that ARIMA models are useful tools for monitoring dengue incidence in Rio de Janeiro and for predicting trends in dengue incidence. Promprou et al. (2006) also used ARIMA model in forecasting dengue hemorrhagic fever cases in Southern Thailand with the Box-Jenkins approach analyzing which model was appropriate for a longer forecasting period.

Buczak et al. (2014) were also able to develop prediction models to predict high and low incidence of dengue in the Philippines using a largely automated method. The model can use the past and recent data to predict dengue incidence levels several weeks in advance of an outbreak with high accuracy for a specific period and a sub-national geographic region. The decision as to which specific model to use depends on how the predictions will be utilized in a particular situation.

Predicting dengue incidence in the Philippines is necessary being one of the four countries in the Western Pacific with the highest dengue incidence in recent years (Geronimo, 2015a). In 2014 alone, the National Epidemiology Center of the Department of Health reported a total of 59,943 dengue cases from January 1 to September 6, although the numbers are significantly lower from 2013 by 59.57% (Geronimo, 2015b). In 2015, WHO reported 21,962 cases of dengue including the 61 deaths from January to April 2015 (WHO, 2015b). The incidence is 7.28% higher compared in 2014. The Department of Health (DOH) has recorded 55,079 suspected dengue cases from January 1 to August 8, 2015 (Geronimo, 2015a). The record is 9.15% higher than the cases recorded during the same period in 2014. Most of the cases come from Luzon in Calabarzon (15.2%), Central Luzon (13.0%), National Capital Region (11.1%), Cagayan Valley (7.5%), and in Northern Mindanao (8.70%).
Dengue fever is endemic on many islands of the Philippines (ter Meulen et al., 2000). Compared to the other provinces in Northern Mindanao, the Misamis Occidental Province had the highest increase in dengue incidence in 2013 (Department of Health, 2014). Within the province, the most recent study of dengue cases was in Ozamiz City in 2013 with reported 1,908 cases from 2008 to 2013 (Pasay et al., 2013). Despite a drop in dengue cases by 13% in Northern Mindanao in the report of the DOH (Orias, 2015), the number of dengue cases is still alarming and requires control mechanism to reduce further the incidence. Bravo et al. (2014) conducted a systematic literature review of the epidemiology of dengue cases in the Philippines from 2000 to 2011. The study showed broad trends in dengue incidence, including age group predominance and disease severity.

The ongoing dengue surveillance in the Philippines is, therefore, a critical component of dengue prevention and control program as it provides the information necessary for risk assessment, epidemic response and program evaluation (World Health Organization et al., 2009). Despite the effort of the Philippine government for launching the dengue surveillance map to help detect the spread of dengue in Metro Manila (GlobalVoices, 2013), there is still need to use technology-based solutions to forecast dengue cases in other areas of the country. The current surveillance system makes use of Google Map in capturing the map details of Metro Manila as well as the identified points in every part that has some dengue cases. Activating the identified points will then display information about dengue cases including the action taken by the community.

With the alarming dengue incidence in Ozamiz City, the availability of a technology-based forecasting system is, therefore, necessary. Hence, this study aimed to forecast the possible outbreak of dengue in Ozamiz City from 2016 to 2020 using a web-based technology system that generates maps and allows the input of newly reported dengue cases to update the forecasting patterns. The information can aid the community in designing preventive measures during the forecasted month(s) or season when dengue is epidemic and in implementing prompt actions to address the current incidence.
Materials and Methods

The locale of the study

The study was carried out in Ozamiz City located in the southern part of Misamis Occidental Province in Northern Mindanao, Philippines. The area has geographic coordinates of $8^\circ \ 9' \ 0''$ N, $123^\circ \ 50' \ 59''$ E. Included in the study are the 51 barangays of the city, 23 of these are urban, and 28 are rural (Figure 1).

Data collection

The data included in this study were the dengue cases reported by Pasay et al. (2013) in seven hospitals in Ozamiz City in monthly basis from January 2008 to September 2013 and the data retrieved from the same hospitals from January 2014 up to the second quarter of 2015. Permission from the hospital administration was obtained to utilize the previous data in addition to the data collected for this present study.
Architectural design

The study used an architectural design (Figure 2) anchored on the Disease Mapping Architecture of Gao et al. (2008) with added component that includes data storage stage, ontology engine stage, standard health services, dengue forecasting stage and maps. Included in the design is the forecasting dengue incidence using ARIMA as the statistical tool. In data storage stage, a centralized database was created to store the data of dengue incidence in Ozamiz City from January 2008 to June 2015. The development of the database was anchored on the concept of Blanton et al. (2006). The study utilized MySQL, an open source relational database management system (MySQL AB, 1995).

Figure 2. The architectural design of the study.
In ontology engine stage, the heterogeneous data collected were updated using the Structured Query Language (SQL) command to produce the homogeneous data mostly preferred for the query process. The data gathered from the seven hospitals in the city were heterogeneous, and there was a need to do the process since these data could cause difficulty in retrieving exact information from the database.

In standard health services stage, the dengue cases were clustered as to the barangay address of the patient to provide the spatial autocorrelation indicators for detecting disease clusters. Barangays were grouped into rural and urban. The system provided disease information through dynamically generated maps online. Through this stage, hospitals have access to the system where they could key in the newly identified dengue cases reported to the hospital.

In dengue forecasting stage, the ARIMA was used as the statistical tool based on the approach used by Dom (2013). Shown below is the formula to forecast dengue incidence in all rural and urban barangays of Ozamiz City from 2016 to 2020.

\[ X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} \]

where \( \phi \) (phi) is the autoregressive parameter to be estimated, \( \theta \) (theta) is the moving average parameter to be estimated, \( X \) is the original series, and the \( a \) is a series of unknown random errors (or residuals) which are assumed to follow the normal probability distribution.

In this study, the dengue cases of each barangay from the year 2008-2015 were tabulated in a quarterly manner per year. Calculations were made for the four quarter moving average (4QMA) from the first quarter of the year 2008 until the second quarter of 2015 and for the centered average based on the results of the 4QMA respectively. The percentage of the average of the reported cases per quarter was calculated to obtain the adjacent factor needed for the calculation of the seasonal index per quarter for forecasting. The standard error of the estimate was determined by comparing the forecasted data from the first quarter to the second quarter of the year 2015 with the actual data gathered from the seven hospitals. The forecasted data on dengue incidence from 2016 to 2020 were calculated using the linear regression analysis.
In the maps and animation stage, the map of Ozamiz City showing the barangays with associated information about dengue cases was generated. The map identified the barangay with the highest incidence of dengue. Each hospital can update the map based on the currently reported dengue cases to generate new forecasting patterns.

### Results and Discussion

Table 1 presents the forecasted dengue cases for the first quarter and second quarter of 2015 using the ARIMA model. As shown, there were 276 reported dengue cases in rural barangays and 2,293 cases in urban barangays from the year 2008 until the second quarter of 2015. The calculated standard error of the estimate is substantial since it yields a lower value compared to the highest actual or forecasted value. The result is similar to the study of Li et al. (2012) which found that the actual data did not match the predicted data of the ARIMA model entirely, but they fell within the predicted 95% confidence interval. There are many reasons for the discrepancies between the actual and predicted data. The forecasted dengue cases in rural and urban barangays differed on the average percentage of increase from 2016 to 2020 (Table 2).

<table>
<thead>
<tr>
<th>Barangay</th>
<th>Reported Dengue Cases</th>
<th>Average Adjacent Factor</th>
<th>Average Seasonal Index</th>
<th>Actual Dengue Cases</th>
<th>Forecasted Dengue Cases</th>
<th>Standard Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>276</td>
<td>2.12</td>
<td>100</td>
<td>28</td>
<td>18.81</td>
<td>0.69</td>
</tr>
<tr>
<td>Urban</td>
<td>2,293</td>
<td>1.30</td>
<td>100</td>
<td>204</td>
<td>155.91</td>
<td>3.31</td>
</tr>
</tbody>
</table>

Table 1. Forecasted dengue cases for the first and second quarter of 2015 using the ARIMA model.
Table 2. Forecasted dengue cases in Ozamiz City from 2016 to 2020.

<table>
<thead>
<tr>
<th>Barangay</th>
<th>Forecasted Dengue Cases</th>
<th>Average increase (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>48.49</td>
<td>51.67</td>
</tr>
<tr>
<td>Total</td>
<td>349.31</td>
<td>352.66</td>
</tr>
</tbody>
</table>

Based on the results, the overall forecasted dengue incidence in Ozamiz City showed an increasing pattern (Figure 3). The result of this dengue forecasting could provide an early warning of dengue outbreak and allow the Department of Health as well as the communities to implement control mechanisms that are timely and effective for the next five years. The development of dengue forecasting model in other Asian countries such as Malaysia and China shows promising results in helping public health officers plan and prepare their limited resources to prevent dengue outbreak and in providing evidence-based interventions for disease control and prevention (Ho & Ting, 2015; Sang et al., 2015).

Several modeling studies differ in their approach and the disease data they use (Messina et al., 2015; Wilson & Chen; 2015). In the forecasting model of this present study, only the dengue incidence and disease distribution in the 51 barangays of Ozamiz City are anticipated for the next five years based on the past events or data from 2008 to 2015 without accounting the causes of the pattern derived. Nevertheless, the forecasted dengue incidence shown in this study gives information in identifying potential local endemic areas in Ozamiz City considering that dengue is endemic in the Philippines (Sayson et al., 2015). The data pattern generated is also useful in predicting the possible spread of infection through distribution mapping using the GIS since dengue is a contagious disease and can become endemic (Munir & Winarko, 2015). The forecasting model thereby provides risk mapping needed to identify vulnerable communities to dengue and prevent local transmission. Moreover, the dengue forecasting model generated will show any possible shift in the incidence and distribution over time with updates of disease data that can prompt health authorities in Ozamiz City to take serious attention.
The increasing pattern of forecasted dengue incidence in Ozamiz City may remain the same over time if the high population density at present continues to escalate. Currently, about 724.55 people by average live in one square kilometer (Philippine Islands, 2015). Population growth is one of the putative drivers for the rise of dengue in Singapore (Struchiner et al., 2015). An outbreak can become explosive with increasing human-vector contact in the densely-populated area. Climate parameters also have impacts on dengue incidence rate and vector distribution (Campbell et al., 2015; Liao et al., 2015; Struchiner et al., 2015). However, in Ozamiz City, the study of Pasay et al. (2013) did not show a significant correlation between climate factors and dengue incidence.

Vector environment also plays a crucial role in dengue transmission. Recently, *Aedes aegypti* becomes more prevalent in urban and in suburban tropical areas (Murugan et al., 2015). Several places in Ozamiz City are flood-prone (Enguito et al., 2013) and are potential breeding grounds for dengue mosquitoes. The forecasted dengue incidence in Ozamiz City may only shift to a non-alarming level if the health authorities and the residents carry out seriously the activities of
the national dengue prevention and control program of the Department of Health. The measures include proper environmental sanitation to eliminate mosquito breeding sites. With the forecasted dengue incidence in Ozamiz City for the next five years, the primary drivers of dengue may be given attention by the responsible sectors of the community so that interventions can blunt the impact of possible dengue outbreak.

Figure 4 presents the forecasted pattern of dengue incidence in rural and urban barangays of Ozamiz City from the year 2016 to 2020. Dengue cases are lesser in rural barangays than in urban with higher population. The estimated total population in urban barangays is 112,150 while the population in the rural areas is only 10,987 (Philippine Islands, 2015). However, there is a rapid increase in forecasted dengue incidence from the year 2016 to 2020 in rural barangays than in the urban areas. The pattern is similar to the study of Liu et al. (2011) that also used the ARIMA model to forecast the incidence of hemorrhagic fever with renal syndrome.

In Ozamiz City, dengue mosquitoes are actively present not only in urban but rural areas as well. In rural settings, most of the residents are having the bamboo fence of their backyards that are potential
breeding areas for mosquitoes. As shown in the study of Sunahara et al. (2002), small containers that hold water such as tree holes, bamboo stumps, plant pitchers, edges of lake and ponds common in the rural barangays are typical mosquito larval habitats. Derraik et al. (2005) added that mosquitoes could spread and vertically stratified in the native forest. While some mosquito species are active throughout the vertical forest strata from ground level to the top of the canopy, others may display a marked preference for forest canopy or close to the ground (Braack et al., 1994).

Relatively, there is a very slight increase in the pattern of forecasted dengue incidence from the year 2016 to 2020 in the urban barangays. The pattern is not similar to the study of Martinez et al. (2011) in Campinas State of São Paulo, Brazil. The effectiveness of the intervention programs implemented by the Local Government Unit (LGU) of Ozamiz City has contributed to the projected dengue incidence. The government in cooperation with the academic sectors regularly conducts a clean-up activity in esteros to reduce the possibility of dengue mosquito spread. Stagnant water of esteros in urban areas resulting from the high discharge of organic waste (Enguito et al., 2013) may provide the *Aedes aegypti* mosquitoes suitable habitat for oviposition, thereby increasing the vector population. Hence, the clean-up drive of academic institutions may help mitigate the possible outbreak in the next five years. The intensified dissemination campaign and information availability in urban barangays may also alleviate the risk of spread compared to rural barangays. Gürtler et al. (2009) conducted a five-year intervention program in Clorinda, northeastern Argentina that reduced significantly infestations of mosquito larva compared to pre-intervention levels. Further, the program most likely prevented or limited new local dengue outbreaks.

Despite the very slight increase in forecasted dengue incidence in urban barangays in Ozamiz City, the result of GIS mapping showed a higher number of dengue cases compared to the rural barangays. Estero clean-up may not suffice to reduce the current incidence of dengue with the rapid population growth in urban places. The incidence of dengue in Singapore for over the past 40 years has been escalating as a result of increasing population density (Struchiner et al., 2015).
Immediate attention is needed not only in rural barangays with increasing pattern of forecasted dengue cases for the next five years but urban barangays as well, to reduce dengue incidence at present and through the succeeding years.

Figures 5a to 5e display the maps showing the forecasted dengue cases in every barangay of Ozamiz City from the year 2016 to 2020. The map was divided into rural and urban areas showing the barangays covered. Among the rural barangays, Labinay has the highest number of forecasted dengue cases from the year 2016 to 2020 having an average value of 11.90. Among the urban barangays, Aguada has the highest forecasted cases in the year 2016 to 2020 having an average value of 38.16. Being the most populated barangay in the city with some establishments such as schools and fast food chains and an increased solid waste generation rate (Matunog & Awa, 2013) could be the apparent cause of the increasing forecasted rate of dengue cases in Aguada. Moreover, de Mattos et al. (2007) stated that information campaigns related to basic sanitation in the urban area are not adequate to discriminate levels of susceptibility to dengue. They also stressed that population density and concentration of establishments could increase vulnerability to vector infestation.

Dengue cases continue to climb despite intensive and numerous interventions, but the disease is still significantly underreported (Haider & Turner, 2015). Before the conduct of this research, the only report of dengue incidence in Ozamiz City with results that are available to the community was the study of Pasay et al. in 2013. Results of that study showed that dengue incidence is escalating and is a health issue that requires effective action. Orias (2015) noted a drop in dengue cases by 13% in Northern Mindanao but the possibility of underreported cases posts a concern. Although other researchers conducted dengue studies in various localities in the country (Bravo et al., 2014; Geronimo, 2015a), dengue cases in remote areas such as in the rural barangays of Ozamiz City may not be accounted.
Forecasting Dengue Incidence Utilizing Geographic Information System and Autoregressive Integrated Moving Average Models

M. Y. Orong, G. D. Durias, R. K. Paredes, J. M. E. Pasay

Figure 5a. The forecasted dengue incidence in Ozamiz City in the year 2016.
Figure 5b. The forecasted dengue incidence in Ozamiz City in the year 2017.
Figure 5c. The forecasted dengue incidence in Ozamiz City in the year 2018.
Figure 5d. The forecasted dengue incidence in Ozamiz City in the year 2019.
Figure 5e. The forecasted dengue incidence in Ozamiz City in the year 2020.
With the exponential increase in dengue incidence at present, dengue surveillance and forecasting are critical approaches that can help mitigate the occurrence. Hence, the use of GIS in this study generates maps that do not only serve as visual displays of dengue incidence from the year 2008 up to 2015 but as a forecasting system that can anticipate future incidence of dengue. Moreover, Gao et al. (2008) consider web-based GIS as an approach that enables the generation of thematic maps dynamically and efficiently. Hence, this strategy can improve the reporting system of dengue cases in Ozamiz City.

The use of ARIMA in this study as a tool for dengue forecasting allows the local government in Ozamiz City as well as the hospitals to input new records of dengue cases and update the forecasting pattern in the succeeding years. The transfer to stakeholders and other local agencies of this map-based information about the forecasted dengue incidence in Ozamiz City for full utilization can benefit the community. In this study, ARIMA model offers immediate result of the spatial patterns of dengue incidence which is an important component of disease data sharing (Eisen & Eisen, 2011). The findings of Tran et al. (2004) and Sharma et al. (2015) demonstrate the relevance and potential use of GIS for refining the surveillance strategy for dengue based on the spatiotemporal patterns of dengue. Significantly, the use of ARIMA in this study provides a web-based technology for use in a resource-poor environment such as in the rural barangays of Ozamiz City. On this context, this study shows that ARIMA is a useful tool for dengue surveillance and anticipating trends in dengue incidence. Several studies yield a similar outcome with the use of ARIMA (Promprou et al., 2006; Luz et al., 2008; Karim et al., 2012; Dom, 2013).

Conclusion and Recommendations

The use of GIS and ARIMA offers immediate results on visualization and patterns, respectively, which allow interpretation and analysis of dengue data in Ozamiz City. The generated maps showing dengue cases in each barangay and the patterns of forecasted dengue incidence from 2016-2020 has provided the community in Ozamiz City relevant information in designing measures that prevent the possible
outbreak and in implementing prompt actions to address the current incidence. The ability of the forecasting system to allow the input of newly reported dengue cases can update the forecasting patterns of the incidence. The higher incidence of dengue in urban barangays and the increasing pattern of forecasted incidence in rural areas require appropriate and immediate attention. Thus, this web-based technology serves as the dengue-monitoring system in Ozamiz City to improve the preparedness and resilience of the people to the disease.

The LGU and health agencies in Ozamiz City must intensify awareness campaign of the intervention programs for dengue reduction and prevention. Other sectors of the community, mainly the academic institutions, shall formulate mechanisms for dengue mitigation giving attention also to the rural barangays. There is also a need for the responsible sectors to ensure that the community can become aware of the existence of this technology and can fully access the system.

Acknowledgment

Authors are grateful to Misamis University for funding this study. The Department of Health in Ozamiz City and the respective hospital administration are also acknowledged for the full cooperation given.

Literature Cited


166
