

Dengue Cases in Colombia: Mathematical Forecasts for 2018–2022

Luis Eduardo López-Montenegro MS PhD, Ana María Pulecio-Montoya MS PhD, Germán Arturo Marcillo-Hernández MS

ABSTRACT

INTRODUCTION Dengue is a disease caused by any one of five virus serotypes and transmitted to humans by the *Aedes aegypti* mosquito. Climate change and health conditions have combined to make dengue a global public health problem. The situation is especially serious in Colombia, where by week 36 of 2018, dengue incidence was 96 cases per 100,000 population, with a total of 111 deaths. Different mathematical and statistical models have been proposed to understand the dynamics of transmission and consequently to apply control strategies to reduce the number of dengue cases.

OBJECTIVE Forecast the number of dengue cases expected in Colombia from 2018 through 2022 with the stochastic Auto-Regressive Integrated Moving Average (ARIMA) model and use the results to adjust the parameters of an ordinary differential equations model in order to determine the disease's basic reproduction number in the year presenting the highest number of dengue cases.

METHODS An ecological time series study was conducted to forecast dengue incidence in Colombia from 2018 through 2022. The data were compiled from Colombia's National Health Institute series on dengue cases reported by epidemiological week from 2009 to 2017. The stochastic ARIMA time series model was applied. Forecasts were then analyzed, and the year with the highest number of predicted cases was used to adjust the parameters of an ordinary differential

equations model (ODE) through nonlinear least squares regression to calculate the vectorial capacity of the transmitting mosquito.

RESULTS Forecasts of the total number of dengue cases per year in Colombia for the following five years were: 32,411 (2018); 88,221 (2019); 56,392 (2020); 47,940 (2021); and 77,344 (2022). The highest number of cases was forecast for 2019. Values for the parameters affecting dengue transmission that year (by the year's four quarters), such as recovery rate (0.0992, 0.0838, 0.1177, and 0.1535, respectively), vectorial capacity of the transmitting mosquito (0.1720, 0.1705, 0.1204, and 0.2147, respectively) and the basic dengue reproduction number (1.73, 2.03, 1.02, and 1.40, respectively) were estimated, indicating that most cases would occur in the second quarter and, since the basic reproduction number values were >1 , the disease would persist in the country throughout the entire year.

CONCLUSIONS ARIMA model forecasts for 2018 through 2022 predicted the highest incidence of dengue cases in Colombia would occur in 2019. Comparison of ARIMA model forecasts and the ODE model allowed projections of possible variations in dengue cases reported, and the basic reproduction number predicted that dengue would persist throughout 2019.

KEYWORDS Arboviruses, climate, dengue, models, theoretical, basic reproduction number, prognosis, Colombia

INTRODUCTION

Arboviral diseases have become a global public health problem due to factors such as climate change, population growth, waste accumulation, pollution, inadequate recycling, and insufficient vector control.[1] Dengue is an acute viral disease transmitted to humans by its principal vector, the *Aedes aegypti* mosquito. Dengue virus (DENV) comprises five serotypes (DENV1, DENV2, DENV3, DENV4, DENV5);[2] the first four circulate simultaneously in Colombia.[3] The different serotypes do not confer cross immunity; an individual who recovers from one acquires permanent immunity against that serotype but only partial and temporary immunity against the other four types.[4]

Several studies have shown that one of dengue's greatest dangers is the increased severity of the infection when multiple serotypes are simultaneously present. That is, a secondary infection with another dengue serotype has a greater probability of causing severe acute infections than the primary disease, due to formation of autoimmune complexes that can attack the body.[5,6]

IMPORTANCE Forecasts of dengue cases in Colombia from 2018 through 2022 based on a time-series study design, combined with an ordinary differential equations model, can help health systems and institutions take more effective, precise preventive and control measures to reduce dengue infection, a serious health problem in the country.

Application of different theoretical and mathematical models has helped us understand the epidemiological dynamics of dengue. Nonlinear models based on systems of ordinary differential equations (ODE) describe variations in numbers of cases in a specific population. However, many factors and conditions are involved in dengue transmission dynamics, and any variation in one of them can lead to significant changes in the shape of the curves that describe growth of the infected population. For example, the model proposed by López[7] contains the following variables: average numbers of healthy, infected, and recovered individuals and average number of mosquitos in the environment; classified as larval (aquatic) or aerial (adult) phase, with adults also classified as female or male.

Through a series of calculations, this model was simplified to include only two variables (proportion of healthy individuals and proportion of infected individuals), thus making it easier to analyze while preserving the hypothetical viral transmission factors of the original version.

Amaku's model considers the virus's state of latency in the mosquito population, its vertical transmission, and how variations in climate factors can influence transmission dynamics.[8] Koiller proposes a population growth model for the *Aedes aegypti* mosquito that considers all phases and states in its life cycle.[9] The model then incorporates a biological control measure through use of *Wolbachia* bacterium. This mosquito-growth dynamic is included in a model to determine the dynamics of transmission to humans under these assumptions, considering the number of individuals exposed.

Each factor involved in the complex dynamics of the disease (presence of asymptomatic individuals and infected individuals, dispersion of the vector and infected individuals, application of one vector control or another, resistance to controls, reinfection by a different serotype, etc.) opens new modeling options. No single model can include all the real factors involved in dengue transmission, so selection of a forecasting model often depends on availability of reliable information on the variables and parameters involved, and on the specific objective of the projection.

Time-series modeling, particularly with ARIMA models[10]—which benefit from repeated autocorrelations to make extrapolations—allows reasonable estimates of future effects of an infectious outbreak on a specific population. In combination with other predictive nonlinear models, ARIMA models have been extremely useful in epidemiological investigations to prevent and control infectious diseases.[11–14]

In Colombia, temperature changes caused by climate phenomena and the growing population living in unhealthy conditions have caused intense propagation of dengue virus and the appearance of new viruses in different regions.[15–17] The community epidemiological bulletin published by the National Epidemiological Surveillance System (SIVIGILA) reported a total of 13,427 confirmed dengue cases at the national level by week 23 of 2018,[15] and by week 36, an incidence rate of 96 cases per 100,000 population, with 111 deaths.[16] The Ministry of Health and Social Protection advised health sector units to monitor and control virus transmission and its effects on the population, which reconfirms the importance of projections to support decision-making and financial resources for preventive actions in the face of such a serious health problem.

The purpose of this study was to apply mathematical models to forecast the average number of dengue cases per year from 2018 through 2022 in Colombia, and to determine dengue's basic reproduction number (BRN) in the year with the highest number of predicted cases in order to identify possible outbreaks.

METHODS

Type of study A time-series study was conducted using Colombia's National Institute of Health routine surveillance system database. The study was conducted from February 2017 through March 2018 by researchers in San Juan de Pasto, Nariño Department, Colombia.

Data sources The database from which the reports on incidence of dengue were taken as time series from 2009 to 2017 was obtained from the National Health Institute's SIVIGILA web portal.[15] Specifically, the data correspond to dengue cases reported by the country's departments (principal administrative divisions) and municipalities as "confirmed" cases of dengue or severe dengue. Data were compiled in XLS-formatted files and included records of dengue cases for each epidemiological week of the respective year.

Variables The study considered two variables: 1) number of dengue cases in Colombia and 2) number of epidemiological weeks.

Mathematical models

ARIMA model This model effectively combines three components: autoregressive (AR), integrated (I) and moving average (MA). It is used to analyze stochastic time series and takes into consideration the correlation between the data and the errors

corresponding to the preceding periods.[10] It is generally represented as $ARIMA(a, d, b)$, where a is the number of autoregressive terms, d is the number of differences, and b is the number of moving average terms. Mathematically it is expressed as:

$$y_t^d = c + \alpha_1 y_{t-1}^d + \dots + \alpha_a y_{t-a}^d + \vartheta_1 \varepsilon_{t-1}^d + \dots + \vartheta_b \varepsilon_{t-b}^d + \varepsilon_t^d$$

where $\alpha_i, i = 1, \dots, a$ and $\vartheta_j, j = 1, \dots, b$ are the parameters to estimate and ε_t^d is a white-noise process.

The mathematical processes and algorithms to analyze time series and their respective predictive process with ARIMA follow four steps:[18]

- 1) Definition of the model, which includes the assumption that forecast errors will be distributed normally around the mean and time-independent variance (white noise) and application of simple and partial autocorrelation functions (ACF and PACF, respectively).
- 2) Estimation of parameters with the orders of the fitting process, according to ACF and PACF functions.
- 3) Evaluation and validation of the tentative ARIMA model.
- 4) Forecasting of time series.

To carry out these four steps, researchers used R software,[19] a free programming language and statistical computing and graphics tool, together with a series of algorithms and packages related to the models under consideration.

Step 1: The time series was generated from the data to be analyzed. Trend and seasonality could be identified and removed through differentiation. A density histogram was also constructed to determine whether the series is stationary, which means forecast errors normally distributed, with zero mean and constant variance.

Step 2: ACF and PACF analyses were performed along with the application of several functions related to ARIMA in the R statistical package; respective graphics were plotted indicating the appropriate order of the lag in the linear $ARIMA(a, d, b)$ model, and respective parameters of the polynomial coefficients were estimated for the tentative model.

Step 3: The Ljung-Box[20] contrast was used to determine the dependency relation among residuals, enabling validation of the model's forecasting capacity.

Step 4: The previously validated $ARIMA(a, d, b)$ model and periodogram were used to forecast future incidence of dengue.

ODE model These equations relate a dependent variable with one or more of its derivatives with respect to an independent variable.[21] In infectious diseases such as dengue, ODEs are generally used to represent variations in time of average numbers of susceptible individuals, infected individuals and recovered individuals, providing a basis for fitting models with parameters that weight features associated with disease transmission. These models are used to estimate the disease prevalence or to predict its disappearance in a particular setting.[22,23]

This study is based on a system formulated by López[7] presented below and adjusted to fit the reported data on dengue in Colombia. The authors have shown that the new model, a simplified version of a more complex one that includes a larger number of variables and parameters, nevertheless preserves the properties of the original model. The rationale for the use of this simplified system

is that the only data available for Colombia from SIVIGILA and the National Statistical Administration Department (DANE)[24] are life expectancy and the number of dengue cases reported per epidemiological week. The transformed model is the simplest possible without affecting the close approximation of estimated case numbers.

The original model was rewritten as:

$$\begin{cases} \dot{p} = \mu(1 - p) - Bpq \\ \dot{q} = Bpq - (\mu + \theta)q \end{cases} \quad \text{MODEL 1}$$

Where

p and q are the proportion of susceptible and infected individuals, respectively;

\dot{p} is the variation in the proportion of susceptible individuals over time (measured in weeks);

\dot{q} is the variation in the proportion of infected individuals over time.

In this model, 1 (100%) represents the total human population under consideration, so the sum of p and q (proportions of susceptible and infected individuals) will never be >1 ; i.e., $p + q \leq 1$.

Given that μ represents the population growth minus death rate, it can be interpreted as the proportion of people entering weekly into the susceptible population, while μp and μq represent the proportion of susceptible individuals and infected individuals who die each week, respectively. Considering that B is the dengue recovery rate, θq represents the proportion of infected individuals who recover each week.

Now, if $B = \beta\psi \frac{\sigma}{\delta + \sigma}$

Where

β is the probability of mosquito-to-human viral transmission;

ψ is human-to-mosquito contagion rate;

σ is development rate of immature mosquitos into female adults; and

δ is development rate of immature mosquitos into male adults;

Then

Bq is infectious rate of susceptible individuals as a result of contact with a previously infected female adult mosquito (infected with the virus due to prior contact with an infected human).

Bpq is the proportion of susceptible individuals who become infected with the virus each week.

In this case, B represents the vectorial capacity of the transmitting mosquito (*Aedes aegypti*), i.e., the number of secondary infections per week following the entry of an infected individual into a susceptible setting.[25,26]

Nonlinear least squares regression This method is used to estimate the value of one or more parameters in a function in order to fit the function's graph more closely to the data distribution in the Cartesian plane.[27] Values of parameters B and in B Model 1 were estimated weekly to better fit the real-time data of dengue cases in Colombia.

According to the National Statistics Administrative Department,[24] average life expectancy in Colombia is 74 years (3848 weeks), so μ was defined as $\mu = \frac{1}{3848}$ per week. Estimated values of B and θ minimize the function:

$$z = \sum_{i=1}^n (q_i - \tilde{q})^2 \quad \text{MODEL 2}$$

Where

n is the number of data to be fitted;

q_i ($i = 1, \dots, n$) are the proportions of confirmed cases from week 1 to week n); and

\tilde{q}_i ($i = 1, \dots, n$) are estimated values of q for each week using Model 1.

Data collection and processing A time series was generated with SIVIGILA data using R statistical software to analyze the trend and seasonality of dengue cases. To generate the number of individuals infected with dengue for each week of 2009 through 2017 and determine the time series for the respective forecasting analysis, a filter was applied to the files reported annually, to select only dengue cases per week. This process was repeated for each year (2009 through 2017). The stochastic ARIMA model was also applied to generate annual forecasts of dengue cases for 2018 through 2022.

Then, totaling all cases reported in all epidemiological weeks of each year, data were taken from the 52 weeks of the year in which the most cases were reported. To determine the disease's BRN, mathematical software Matlab was used to fit a mathematical model based on a system of nonlinear ODEs.[28]

RESULTS

Figure 1a shows the time series of classic dengue cases reported by epidemiological week from 2009 through 2017 in Colombia. Week 1 corresponds to the first week of 2009 and week 469 to the last week of 2017. The time series shows the occurrence of several dengue outbreaks; for example, week 59 (7th week of 2010) when 1787 cases were reported, week 224 (16th week of 2013) with 3231 cases, and week 370 (5th week of 2017) with 3301 cases.

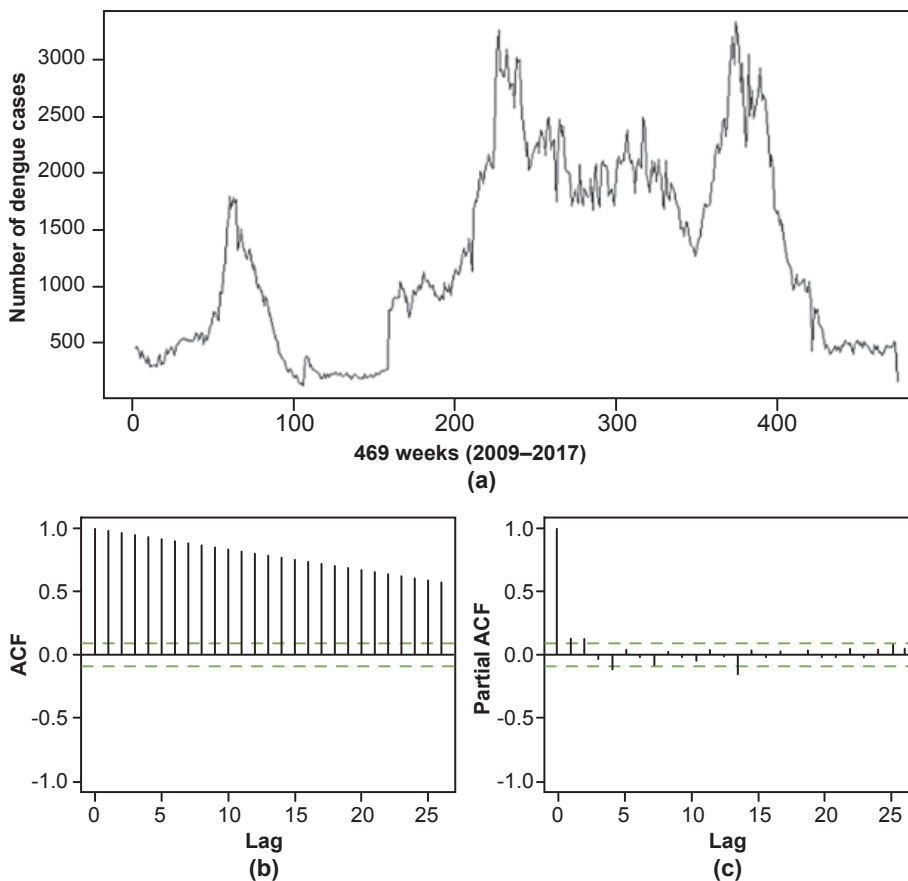
Figures 1b and 1c show the existence of simple and partial autocorrelations every 5 terms in the data, which suggests an autoregressive prediction of order 5, moving average of order 1; and order of integration 1. Thus, the appropriate method to project dengue cases in Colombia would be *ARIMA*(5,1,1), whose polynomial model is given by the equation:

$$y_t = -0.89y_{t-1} - 0.26y_{t-2} - 0.02y_{t-3} + 0.25y_{t-4} + 0.16y_{t-5} + 0.67\varepsilon_{t-1} + \varepsilon_t$$

Application of the Ljung-Box test in software R found the chi-square value to be $\chi^2 = 0.27607$ and p value $0.5993 > 0.05$. These differences showed that there was no autocorrelation in the residuals and the tentative model was valid for forecasting.

The normality test and the periodogram (shown in Figures 2a and 2b, respectively) helped determine that the adjusted residuals had a constant mean and variance, and enabled identification of periods with the highest peaks of dengue cases. Figure 2a shows that the difference between adjacent data of the series approaches a normal curve and Figure 2b shows the occurrence of three peaks; the first is presented in the logarithmic scale of 0.00625 which suggests selecting a cycle period of $\frac{1}{0.00625} = 160$ weeks to make forecasts. Considering the 160-week period and time-series data, a process to validate the model was undertaken for which

Figure 1: Number of dengue cases. (a) Number of cases. (b) Simple autocorrelations. (c) Partial autocorrelations. Colombia, 2009–2017



ACF: Autocorrelation functions

data from the 52 weeks of 2017 were used as real observations and the forecast was generated with the ARIMA(5,1,1) model of the same year. Figure 3 shows a good fit with a root mean squared error (RMSE) of 122.84 and Theil index (Theil's U) of 0.037. The model's fit for 2017 was validated and the stochastic model was used to generate a long forecast of potential dengue cases through 2022.

Weekly forecasts are shown in Figure 4a, which extends the series presented in Figure 1a to Week 729 (week 52 of 2022) with a margin of error of 0.5%. The forecast shows two other possible dengue outbreaks: one in 2019 and another in 2022. The first peak can be observed in Week 544 (week 23 of 2019) with an approximate total of 2624 cases. The second peak occurs in Week 690 (week 13 of 2022) with approximately 2336 cases. Figure 4b shows a bar chart with data on predicted cases per year. According to those projections, the highest number of cases will occur in 2019 (approximately 88,221) and in 2022 (approximately 77,344 cases).

Since the ARIMA forecasts of total dengue cases are highest for 2019, an adjustment was made to parameters B and θ of Model 1 through the nonlinear least squares regression given in Model 2 for data obtained for that year. Figure 5 shows data estimated by the ARIMA model projections and the solution for q of the ODE system presented in Table 1, showing the results of the two different analytic models (one stochastic and the other deterministic). Although the estimates differ, the deterministic model (the ODE system) approximates the stochastic model

(ARIMA) after calibration of parameters via the nonlinear least squares regression.

In Model 1, q is the proportion of infected individuals. To adjust the model without loss of generality, the data yielded by ARIMA were normalized per 10,000 population; for example, the initial number of 1099 dengue cases (Week 1) was given as: $q(1) = q_1 = 0.1099$. To better adjust q , the 52 weeks of 2019 were divided into four epidemiological quarters. The first quarter is composed of weeks 1–15, and the second, third, and fourth, of weeks 16–28, 29–43, and 44–52, respectively.

The initial conditions for the first quarter are given by $q(1) = 0.1099$ (1099 infected individuals). The nonlinear least-squares regression estimates were $B \cong 0.1720$ per week and $\theta \cong 0.0992$ per week. For the second quarter, the initial condition was set as the final number of cases in the first quarter, i.e., $q(16) = 0.1802$ and after the adjustment, the following estimates were obtained: $B \cong 0.1705$ and $\theta \cong 0.0838$ per week. Similarly, for the third quarter, the initial condition was: $q(29) = 0.2433$, and $B \cong 0.1204$ and $\theta \cong 0.1177$ per week. Finally, for the fourth quarter, the initial condition was: $q(44) = 0.1342$, and $B \cong 0.2147$ and $\theta \cong 0.1535$ per week.

Using the values of parameters estimated for each quarter in 2019, dengue's BRN values were calculated. Vectorial capacity was recalculated as $B = R_0(\mu + \theta)$ where R_0 represents the BRN, defined by López[7] as the number of secondary infections produced after introduction of an infected individual into a susceptible setting. Unlike vectorial capacity, BRN is not measured in units of time, since it measures secondary infections produced during an individual's entire infectious period, while vectorial capacity represents the mosquito's capacity per unit of time to infect humans due to the original infected individual. When the BRN value is >1 , the disease will persist in the setting. But when BRN is <1 , dengue will disappear.[7]

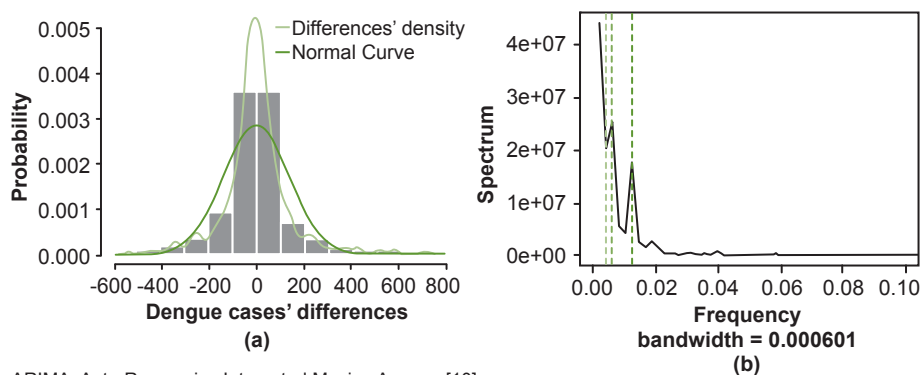
R_0 values estimated for each period in 2019 were:

- 1st quarter: $R_0 = 1.73$
- 2nd quarter: $R_0 = 2.03$
- 3rd quarter: $R_0 = 1.02$
- 4th quarter: $R_0 = 1.4$

During all four quarters of 2019, R_0 value was >1 , indicating that the disease will persist in Colombia throughout the entire year. Since the second quarter had the highest value, the number of cases will be higher during those weeks, as observed in Figure 5.

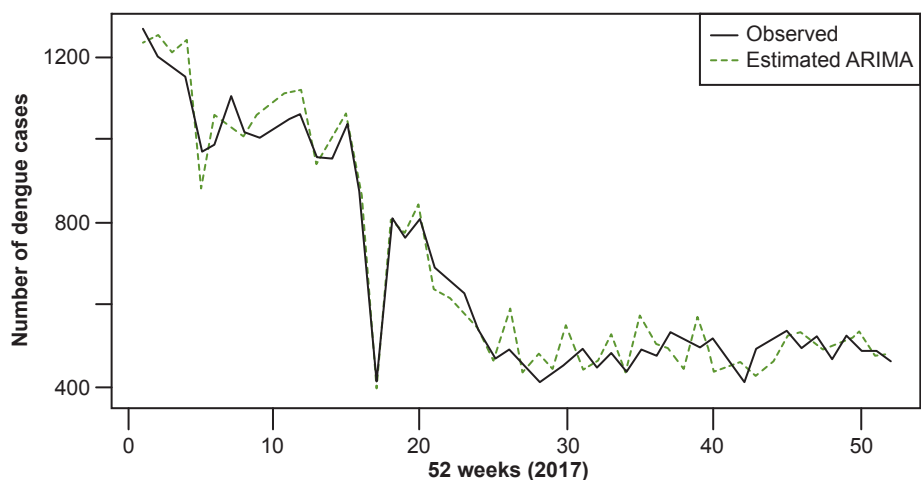
Table 1 summarizes the forecasts for 2019 by epidemiological week, with both the ARIMA and ODE models. The two models' forecasts presented only minor differences, except in the second

Figure 2: ARIMA model. (a) Normality test of residuals of adjusted model. (b) Periodogram of frequencies of peaks in which time-series cycles are presented



ARIMA: Auto-Regressive Integrated Moving Average[10]

Figure 3: Real values (observed) and estimated values (obtained via ARIMA model), Colombia, 2017



ARIMA: Auto-Regressive Integrated Moving Average[10]

and third weeks of May, when the numbers of predicted cases differed by more than 350.

DISCUSSION

No prior studies for Colombia have used adjusted numbers for forecasting dengue cases via time series using the ARIMA model, so there are no comparable results to examine. However, time-series analysis to forecast incidence of dengue has been the subject of research in other countries. In Brazil, for example, Cortes[29] applied and adjusted an ARIMA model, which he used to detect peak incidence and distribution patterns of the disease in two different areas of the country. In Sri Lanka, in the Gampaha District, Withanage[30] developed a forecasting model for dengue using multiple time-series regressions and successfully modeled the effects of climatic factors on imminent outbreaks. Both studies predicted the disease's distribution and incidence in a population, but unlike the present study, they did not determine other factors or measurements of the disease such as BRN and vectorial capacity (via integration with a determinist model such as ODE). Although this study's forecasts do not exactly mirror reality, application of a time-series model and a dynamic system to forecast disease incidence can produce a mathematical modeling trajectory that can be further applied and improved.

Several ODE systems-based mathematical models describe dengue transmission in the population and in gener-

al make it possible to determine the BRN and vectorial capacity based on parameters included in the model. Sardar,[23] for example, obtained the BRN based on an ODE system, and Liu-Helmersson[26] generated both the BRN and vectorial capacity. However, neither study identified a specific value for parameters in a given region. Other studies have taken hypothetical parameters or parameters obtained in laboratory studies in countries with very different climatic and health conditions from the country of application. Sepúlveda-Salcedo,[31] for example, adjusted the mathematical model parameters to the data reported in one year, and based on that, obtained forecasts of dengue cases in subsequent periods, the forecasts preserving the properties and values of the previous period's parameters.

In this study, parameter values were estimated for 2019 based on Colombian data and on forecasts for that same year. Though similar to other studies in expressing vectorial capacity and BRN under the parameters of the ODE system, the ARIMA model forecasts and subsequent adjustment of ODE parameters through nonlinear least-squares regression applied here enabled projection of specific BRN values and vectorial capacity specific to Colombia in 2019. Thus, ODE forecasts do not depend on parameters from other countries or time periods. The predicted numbers of dengue cases and BRN provide information on the disease's behavior in upcoming years and specifically predict the weeks that could present more severe outbreaks. Since the ARIMA-model forecasts extend through

2022, adjustments can be made over time to ODE parameters for subsequent periods to find the corresponding $\overline{R_0}$.

To the best of our knowledge, no previous studies compared the application of ARIMA and ODE methods for viral diseases, or specifically for dengue. In addition to forecasting possible variations in the number of cases, the importance of this approach is that the ARIMA method allows forecasting of potential cases based solely on data already reported; with this forecast in mind, the ODE parameters can then be adjusted to provide information on which qualities (ODE parameters) are affecting the rise or fall in the number of cases. While it is important to predict occurrence of high numbers of cases, it is also important to identify the factors causing the increase, such as BRN, vectorial capacity, and other relevant parameters.

One shortcoming of the study is that it was based on data for past and present conditions, and results could change significantly if, among other things, vector controls are applied or climatic and environmental conditions change; in the transmission of infectious diseases, a small change in conditions can lead to totally different behaviors than those previously forecast. Another limitation is that since the ARIMA model is not recommended for long-term forecasting, a four-year projection is likely to yield less accurate estimates. However, the 2019 forecasts in this study are in the range of close approximation. It is recommended that followup to the proposed forecasts be considered in order to establish them

Figure 4: Number of dengue cases forecast (ARIMA), Colombia, 2018–2022.
 (a) Forecasts with time series. (b) Forecasts with bar chart

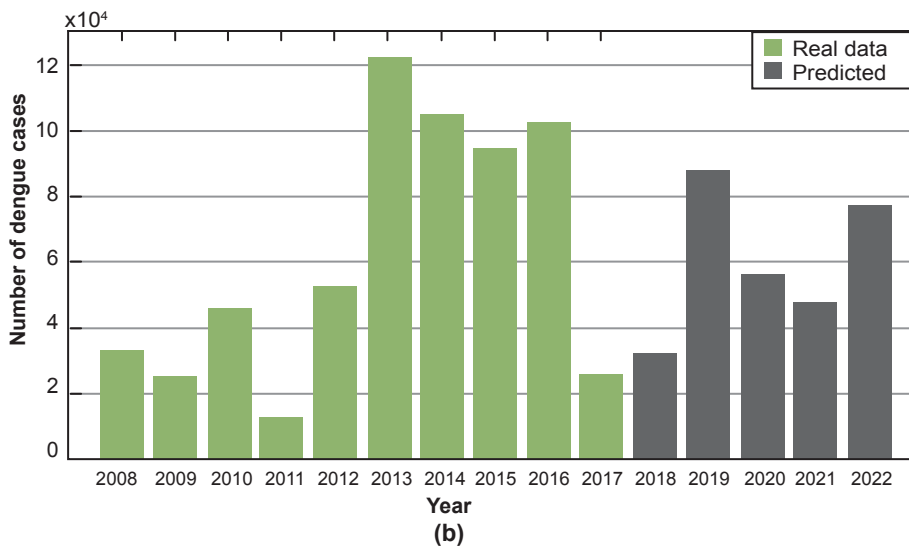
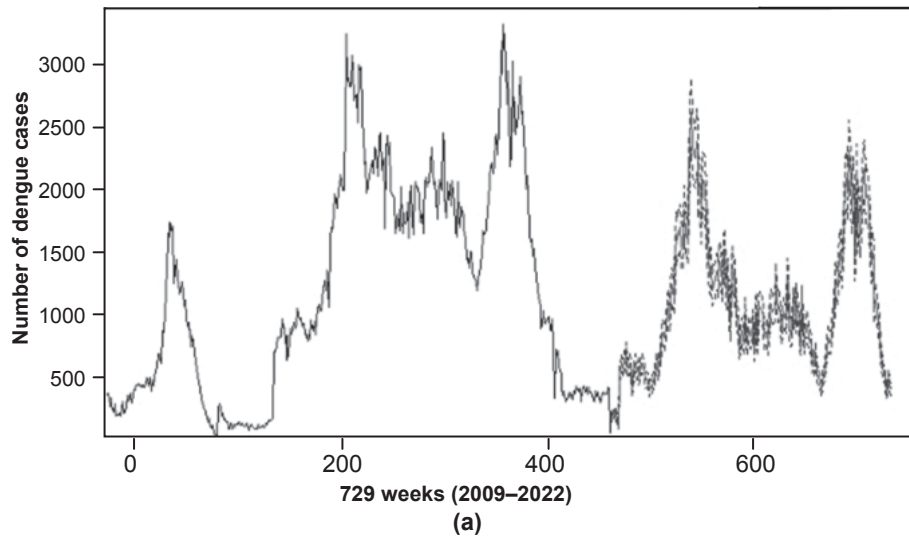
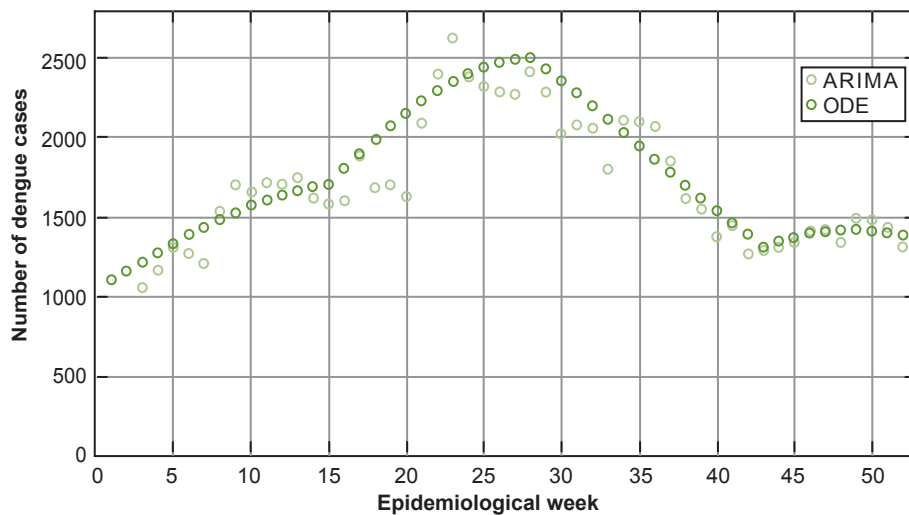


Figure 5: Number of dengue cases forecast, Colombia, 2019



ARIMA: Auto-Regressive Integrated Moving Average[10]
 ODE: Ordinary differential equations[21]

Table 1: Number of dengue cases forecast, Colombia by week for 2019 based on two mathematical models (ARIMA and ODE)

Month	Week	ARIMA	ODE	Month	Week	ARIMA	ODE
January	W1	1099	1099	February	W1	1306	1329
	W2	1159	1158		W2	1265	1383
	W3	1053	1216		W3	1211	1435
	W4	1166	1273		W4	1528	1483
	TOTAL	4476	4746		TOTAL	5308	5630
Month	Week	ARIMA	ODE	Month	Week	ARIMA	ODE
March	W1	1698	1528	April	W1	1618	1688
	W2	1655	1569		W2	1577	1705
	W3	1713	1606		W3	1606	1802
	W4	1703	1638		W4	1884	1896
	W5	1745	1665				
	TOTAL	8515	8006		TOTAL	6684	7090
Month	Week	ARIMA	ODE	Month	Week	ARIMA	ODE
May	W1	1683	1987	June	W1	2624	2352
	W2	1697	2073		W2	2391	2401
	W3	1626	2154		W3	2320	2442
	W4	2086	2228		W4	2287	2472
	W5	2398	2294				
	TOTAL	9491	10736		TOTAL	9622	9667
Month	Week	ARIMA	ODE	Month	Week	ARIMA	ODE
July	W1	2273	2493	August	W1	2081	2279
	W2	2408	2505		W2	2059	2198
	W3	2283	2433		W3	1801	2115
	W4	2023	2358		W4	2114	2031
	TOTAL	8988	9789		W5	2100	1947
			TOTAL	10156	10570		
Month	Week	ARIMA	ODE	Month	Week	ARIMA	ODE
September	W1	2070	1863	October	W1	1373	1539
	W2	1848	1780		W2	1448	1462
	W3	1618	1698		W3	1267	1388
	W4	1550	1618		W4	1291	1307
	TOTAL	7087	6960		TOTAL	5378	5697
Month	Week	ARIMA	ODE	Month	Week	ARIMA	ODE
November	W1	1310	1342	December	W1	1493	1418
	W2	1341	1370		W2	1479	1413
	W3	1401	1392		W3	1431	1402
	W4	1414	1408		W4	1310	1385
	W5	1338	1416				
	TOTAL	6803	6928		TOTAL	5713	5619

ARIMA: Auto-Regressive Integrated Moving Average[10] ODE: Ordinary differential equations[21]


as a viable methodology that could be applied to other infectious diseases.

Despite its limitations, the study shows the usefulness of synergy between two models which, although with different rationales, complement one another to make forecasts useful for timely preventive actions.

CONCLUSIONS

Our forecasts for dengue in Colombia for 2018 through 2022 using the ARIMA model show that most cases will likely occur in 2019.

The ARIMA forecasts enabled adjustment to ODE parameters and subsequent BRN estimates for dengue for that year, including possible variations in different quarters, and revealed the disease's persistence.

Despite the BRN's importance (since it predicts the disappearance or persistence of a viral disease in a population), there are no reports on BRN for dengue. These estimations are essential for health institutions particularly before seasons with high predicted BRN values. These methods can be applied to other vector-transmitted diseases such as Zika and chikungunya. 

REFERENCES

- Rossati A. Global warming and its health impact. *Int J Occupational Environ Med.* 2017 Jan;8(1):7–20.
- Mustafa M, Rasotgi V, Jain S, Gupta V. Discovery of fifth serotype of dengue virus (DENV-5): A new public health dilemma in dengue control. *Med J Armed Forces India.* 2015 Jan;71(1):67–70.
- Pérez-Castro R, Castellanos JE, Olano VA, Matiz MI, Jaramillo JF, Vargas SL, et al. Detection of all four dengue serotypes in *Aedes aegypti* female mosquitoes collected in a rural area in Colombia. *Mems Inst Oswaldo Cruz.* 2016 Apr;111(4):233–40.
- Muller DA, Depelsenaire AC, Young PR. Clinical and laboratory diagnosis of dengue virus infection. *J Infect Dis.* 2017 Mar 1;215(Suppl 2):S89–95.
- Soo KM, Khalid B, Ching SM, Chee HY. Meta-analysis of dengue severity during infection by different dengue virus serotypes in primary and secondary infections. *PLoS One* [Internet]. 2016 May 23 [cited 2017 Aug 2];11(5):e0154760. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4877104/pdf/pone.0154760.pdf>
- Anantapreecha S, Chanama S, A-nuegoonpipat A, Naemkhunthot S, Sa-Ngasang A, Sawanpanyalert P, et al. Serological and virological features of dengue fever and dengue haemorrhagic fever in Thailand from 1999 to 2002. *Epidemiol Infect.* 2005 Jun;133(3):503–7.
- López LE, Loaiza Muñoz A, Olivar Tost G. A mathematical model for transmission of dengue. *App Mathematical Sci.* 2016;10(7):345–55.
- Amaku M, Coutinho F, Raimundo S, Lopez L, Burattini M, Massad E. A comparative analysis of the relative efficacy of vector-control strategies against dengue fever. *Bull Math Biol.* 2014 Mar;76(3):697–717.
- Koiller J, Da Silva MA, Souza MO, Codeço C, Iggidr A, Sallet G. *Aedes. wolbachia* and dengue. [Research Report] RR-8462 [Internet]. Rocquencourt (FR): National Institute for Research in Computer Science and Automation (INRIA). 2014 Jan [cited 2017 Oct 3]. 44 p. Available from: <https://hal.inria.fr/hal-00939411/document>
- Box GEP, Jenkins GM, Reinsel GC, Ljung GM. *Time series analysis: forecasting and control.* 5th ed. New Jersey: John Wiley & Sons Inc; 2015 Jun. 712 p.
- Wang KW, Deng C, Li JP, ZhangYY, Li XY, Wu MC. Hybrid methodology for tuberculosis incidence time-series forecasting based on ARIMA and a NAR neural network. *Epidemiol Infect.* 2017 Apr;145(6):1118–29.
- Sierra W, Argoty C, Franco H. Varicella Incidence Rate Forecasting in Bogotá DC (Colombia) by Stochastic Time Series Analysis. In: *Workshop on Engineering Applications.* Cham (CH): Springer; 2017 Aug 29. p. 647–58.
- Pan Y, Zhang M, Chen Z, Zhou M, Zhang Z. An ARIMA based model for forecasting the patient number of epidemic disease. In: *Service Systems and Service Management (ICSSSM), 2016* [Internet]. 13th International Conference on IEEE; 2016 Jun 24–26; San Francisco, California, USA. California: IEEE; 2016 Aug 11 [cited 2017 Nov 12]. p.1–4. Available from: <https://ieeexplore.ieee.org/abstract/document/7538560>
- Dinh TQ, Le HV, Cao TH, Luong QC, Diep HT. Forecasting the magnitude of dengue in South-east Vietnam. In: *Asian Conference on Intelligent Information and Database Systems.* Berlin: Springer; 2016. p. 554–63.
- SIVIGILA. Boletín Epidemiológico semana 23 del 2018 [Internet]. Bogotá: National Health Institute (CO); 2018 Jun [cited 2018 Jul 25]. 34 p. Available from: <http://www.ins.gov.co/buscador-eventos/Paginas/Vista-Boletin-Epidemiologico.aspx>. Spanish.
- SIVIGILA. Boletín Epidemiológico semana 36 del 2018 [Internet]. Bogotá: Instituto Nacional de Salud (CO); 2018 [cited 2018 Sep 20]. Available from: <http://www.ins.gov.co/buscador-eventos/Paginas/Vista-Boletin-Epidemiologico.aspx>. Spanish.
- Padilla JC, Lizarazo FE, Murillo OL, Mendigaña FA, Pachón E, Vera MJ. Epidemiología de las principales enfermedades transmitidas por vectores en Colombia, 1990–2016. *Biomédica.* 2017;37(Suppl 2):27–40. Spanish.
- Cryer JD, Chan KS. *Time Series Analysis with Applications in R.* 2nd ed. New York: Springer; 2008 Jan 1.
- Baayen RH. *Analyzing Linguistic Data: A practical introduction to statistics using R.* 1st ed. Cambridge: Cambridge University Press [Internet]. 2008 Mar 17 [cited 2018 Sep 12]. 390 p. Available from: <http://www.sfs.uni-tuebingen.de/~hbaayen/publications/baayenCUPstats.pdf>
- García Díaz JC. *Predicción en el dominio del tiempo. Análisis de series temporales para ingenieros.* Valencia: Editorial Universitat Politècnica de València; 2016 Oct 19. 270 p. Spanish.
- Zill D, Cullen M. *Ecuaciones diferenciales.* Mexico, DF: McGraw-Hill Interamericana; 2013. Spanish.
- Bustamam A, Aldia D, Yuwanda A. Understanding dengue control for short-and long-term intervention with a mathematical model approach. *J App Math.* 2018 Jan 1;2018.
- Sardar T, Rana S, Chattopadhyay J. A mathematical model of dengue transmission with memory. *Comm Nonlinear Sci Numer Simulat.* 2015;22(1–3):511–25.
- Departamento Administrativo Nacional de Estadísticas (DANE) (CO). *Proyecciones nacionales y departamentales de población 2005–2020. Estudios post-censales* [Internet]. Bogotá: DANE; 2010 [cited 2017 May 23]. Available from: http://www.dane.gov.co/files/investigaciones/poblacion/proyepobla06_2017Proyecciones_poblacion.pdf. Spanish.
- Basáñez MG, Rodríguez DJ. *Dinámica de transmisión y modelos matemáticos en enfermedades transmitidas por vectores.* Entomotrópica. 2004 Dec;19(3):113–34. Spanish.
- Liu-Helmersson J, Stenlund H, Wilder-Smith A, Rocklöv J. Vectorial capacity of *Aedes aegypti*: Effects of temperature and implications for global dengue epidemic potential. *PLoS One.* 2014 Mar 6;9(3):e89783.
- Brauer F, Castillo-Chávez C, editors. *Mathematical models in population biology and epidemiology.* 2nd ed. New York: Springer; 2012.
- MathWorks, Inc. *MATLAB: Application program interface guide.* Massachusetts: MathWorks; 1996. 357 p.
- Cortes F, Turchi Martelli CM, Arraes de Alencar Ximenes R, Montarrouos UR, Siqueira Junior JB, Gonçalves Cruz O, et al. Time series analysis of dengue surveillance data in two Brazilian cities. *Acta Trop.* 2018 Jun;182:190–7.
- Withanage GP, Viswakula SD, Nilmini Silva Gunawardena YI, Hapugoda MD. A forecasting model for dengue incidence in the District of Gampaha, Sri Lanka. *Parasit Vectors.* 2018 Apr 24;11(1):262.
- Sepúlveda-Salcedo LS, Vasileva O, Martínez-Romero HJ, Arias-Castro JH, Ross McDonald: Un modelo para la dinámica del dengue en Cali, Colombia. *Rev Salud Pública.* 2016 Feb;17(5):749–61. Spanish.

THE AUTHORS

Luis Eduardo López-Montenegro (Corresponding author: lelopez@ucesmag.edu.co), mathematician with a master's degree in applied mathematics and a doctorate in engineering. Member of the Scientific Implementation of General Mathematical Applications (SIGMA) research group at the Maria Goretti Center for Advanced Studies (CESMAG), San Juan de Pasto, Colombia.

Ana María Pulecio-Montoya, mathematician with a master's degree in applied mathematics and a doctorate in engineering. Member of the SIGMA research group at CESMAG, San Juan de Pasto, Colombia.

Germán Arturo Marcillo-Hernández, mathematician with a master's degree in applied mathematics. Member of the FORMA Research Group, Mariana University, San Juan de Pasto, Colombia.

Submitted: July 12, 2018

Approved for publication: April 11, 2019

Disclosures: None