Estimating Parameters for a Dynamical Dengue Model Using Genetic Algorithms

Joshua Uyheng
Ateneo de Manila University
Quezon City, Metro Manila, Philippines
joshuauyheng@gmail.com

John Clifford Rosales
Ateneo de Manila University
Quezon City, Metro Manila, Philippines
johnclifford.rosales@gmail.com

Kennedy Espina
Ateneo de Manila University
Quezon City, Metro Manila, Philippines
kespina@ateneo.edu

Ma. Regina Justina Estuar
Ateneo de Manila University
Quezon City, Metro Manila, Philippines
restuar@ateneo.edu

ABSTRACT

Dynamical models are a mathematical framework for understanding the spread of a disease using various epidemiological parameters. However, in data-scarce regions like the Philippines, local estimates of epidemiological parameters are difficult to obtain because methods to obtain these values are costly or inaccessible. In this paper, we employ genetic algorithms trained with novel fitness functions as a low-cost, data-driven method to estimate parameters for dengue incidence in the Western Visayas Region of the Philippines (2011-2016). Initial results show good fit between monthly historical values and model outputs using parameter estimates, with a best Pearson correlation of 0.86 and normalized error of 0.65 over the selected 72-month period. Furthermore, we demonstrate a quality assessment procedure for selecting biologically feasible and numerically stable parameter estimates. Implications of our findings are discussed in both epidemiological and computational contexts, highlighting their application in FASSSTER, an integrated syndromic surveillance system for infectious diseases in the Philippines.

CCS CONCEPTS

• **Applied computing** → **Health informatics**; • **Mathematics of computing** → *Combinatorial optimization*; Ordinary differential equations;

KEYWORDS

Disease modeling, dynamical systems, genetic algorithms, parameter estimation

ACM Reference Format:
Joshua Uyheng, John Clifford Rosales, Kennedy Espina, and Ma. Regina Justina Estuar. 2018. Estimating Parameters for a Dynamical Dengue Model Using Genetic Algorithms. In *GECCO '18 Companion: Genetic and Evolutionary Computation Conference Companion, July 15–19, 2018, Kyoto, Japan.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3205651.3205716

GECCO '18 Companion, July 15–19, 2018, Kyoto, Japan

ACM ISBN 978-1-4503-5764-7/18/07.
https://doi.org/10.1145/3205651.3205716

1 INTRODUCTION

Dynamical models of disease describe the spread of an epidemic throughout a population. The population is subdivided into several classes, often including a combination of the following: the susceptible class (S), the infected class (I), and the recovered class (R). Vector-borne diseases such as dengue fever require a special class of dynamical models as they consider the disease dynamics between and within more than one population. Dengue, for instance, is carried by the mosquito *Aedes egypti*, which spreads the disease by biting humans. In this paper, the following system of ordinary differential equations (ODEs) is used to model dengue spread.

$$\frac{dH_S}{dt} = \lambda_S H_S - aV_I H_S - \mu_H H_S + rH_I \quad (1)$$

$$\frac{dH_I}{dt} = aV_I H_S - \psi H_I - rH_I \quad (2)$$

$$\frac{dV_S}{dt} = \lambda_S V_S - bH_I V_S - \mu_V V_S \quad (3)$$

$$\frac{dV_I}{dt} = bH_I V_S - \mu_V V_I \quad (4)$$

In the above system of ODEs, H_S refers to the susceptible fraction of the human population, H_I refers to infected humans, V_S refers to susceptible vectors, and V_I refers to infected vectors. The epidemiological parameters used in the model include a or the biting-transmission rate from vectors to humans, b the biting-transmission rate from humans to vectors, r the daily recovery rate of humans, ψ the dengue mortality rate of humans, λ_H the natural human birth rate, μ_H the natural human mortality rate, λ_V the natural vector birth rate, and μ_V the natural vector mortality rate.

By describing how infections spread across human and vector populations, parameter values characterize regional disease dynamics and calibrate appropriate public health responses [2]. However, regional values for these parameters often involve significant costs and specialized procedures to obtain [3]. In addition, results do not necessarily hold outside the laboratory setting, as parameter values tend to vary across time and specific geographical location [1]. To address this issue, we employ genetic algorithms [4] requiring only cumulative daily dengue incidence, thus offering a low-cost, data-driven method for regional parameter estimation. The procedures in this project are conducted toward implementation in FASSSTER, a web service that enables users to geospatially model and visualize dengue spread.

2 METHOD

Daily dengue incidence counts from the Western Visayas Region were harvested from the Philippine Integrated Disease Surveillance and Response System (PIDSR) as raw data. The available counts spanned the 72-month period from January 1st, 2011 to December 31st, 2016.

Using the PIDSR data, we implemented genetic algorithms through the open-source R package GA [**?**] to estimate parameters for a dynamical dengue model on a monthly basis. Using the observed data in each month, parameter values were generated with parameter populations of size 100, elitism 0.5, mutation probability 0.3, and 100 iterations per algorithm. Three genetic algorithms were used, each with different fitness functions: Pearson correlation, normalized root mean square error, and the harmonic mean of the first two.

For the given dynamical model, the parameters to be estimated included a, b, r, ψ, λ_V, and μ_V. Both λ_H and μ_H, which represent human birth and mortality rates, were fixed values based on estimates by the Philippine Statistical Authority. The initial value of H_I used was the fraction of infected humans on the first day of the month, while H_S was taken as $1 - H_I$. Furthermore, an additional parameter k was estimated as the initial fraction of infected vectors. Thus, $1 - k$ represented the initial fraction of susceptible vectors.

3 RESULTS

In Table 1, we summarize the median parameter estimates for each type of genetic algorithm. The median was used instead of the mean to avoid the effects of outlier values. All genetic algorithms produced comparable estimates for each parameter, except for μ_V and k, for which the correlation-based genetic algorithm deviated considerably from the estimates of the latter two.

Parameter	Cor-GA	Err-GA	Harm-GA
a (biting-transmission V-H)	0.282	0.217	0.231
b (biting-transmission H-V)	0.324	0.285	0.254
r (recovery)	0.222	0.253	0.256
ψ (dengue mortality)	0.215	0.246	0.238
λ_V (vector birth)	0.300	0.248	0.233
μ_V (vector mortality)	0.146	0.225	0.223
k (initially infected vectors)	0.141	0.000	0.001

Table 1. Median values of parameter estimates.

3.1 Quality Assessment

To assess the biological feasibility of produced parameters, the proportion of estimates falling within benchmarks in the literature [2, 3] was calculated for each genetic algorithm. Furthermore, to quantify numerical stability, the variance of parameter estimates was taken and normalized over the median value of each parameter estimate. These results are summarized in Tables 2 and 3.

Parameter	Cor-GA	Err-GA	Harm-GA
$a\,[0.01, 0.5]$	1.000	1.000	1.000
$b\,[0.01, 0.5]$	1.000	1.000	1.000
$r\,[0.15, 0.5]$	0.694	0.875	0.917
$\psi\,[0, 0.01]$	0.014	0.000	0.000
$\lambda_V\,[0.05, 0.5]$	0.958	0.958	0.972
$\mu_V\,[0.05, 0.5]$	0.847	0.958	0.931

Table 2. Proportion of biologically feasible estimates.

Parameter	Cor-GA	Err-GA	Harm-GA
a	0.050	0.047	0.060
b	0.035	0.036	0.033
r	0.065	0.035	0.022
ψ	0.091	0.038	0.037
λ_V	0.045	0.029	0.030
μ_V	0.094	0.031	0.040
k	0.209	0.004	0.097

Table 3. Normalized variance of parameter estimates.

4 DISCUSSION

Parameter values for both biting-transmission rates a and b across all genetic algorithms are in the upper-middle region of the biologically feasible range. These moderately high rates suggest that in the Western Visayas Region, more serious measures could be taken in reducing contact between humans and vectors to avoid widespread infectious biting.

Meanwhile, in terms of biological feasibility, we note that most values fall within the literature-documented ranges, especially a, b, and the vector birth and death rates. However, the dengue mortality rate of humans across all fitness functions is severely overestimated, with nearly none of the estimates falling within the biologically feasible ranges. We hypothesize that this is due to the lack of an immune recovered human class in the dynamical model. Humans who have survived dengue and experience immunity to the specific serotype they had been exposed to. Because this is not accounted for in the dynamical model, the genetic algorithm does not return recently infected humans to the susceptible class, but assumes they have died from the disease.

Finally, in terms of numerical stability, all parameter estimates show very small normalized variance. Our data show that overall, parameter estimates tend to display variances of less than a tenth of their median value. This indicates reliable estimates from the genetic algorithms, and suggests that the parameter values obtained are not simply local optima.

5 CONCLUSION

Genetic algorithms provide a low-cost and data-driven method for estimating parameters of dynamical disease models even in data-scarce regions like the Philippines. Best parameter estimates suggest the critical importance of reducing human-vector interactions in addressing dengue outbreaks in the area. Integration of genetic algorithms with comprehensive databases such as FASSSTER will enable public health practitioners to understand and respond to disease outbreaks in a timely and strategic manner.

REFERENCES

[1] Ernest Tambo et al. 2016. Outwitting dengue threat and epidemics resurgence in Asia-Pacific countries: Strengthening integrated dengue surveillance, monitoring and response systems. *Infect Dis Poverty* 5, 56 (2016), 5. https://doi.org/10.1186/s40249-016-0148-3

[2] Eduardo Undurraga et al. 2017. Disease burden of dengue in the Philippines: Adjusting for underreporting by comparing active and passive dengue surveillance in Punta Princesa, Cebu City. *Am J Trop Med Hyg* 96, 4 (2017), 12.

[3] Sandip Mandal et al. 2011. Mathematical models of malaria - a review. *Malar J* 10, 202 (2011), 19. https://doi.org/10.1186/1475-2875-10-202

[4] L. Scrucca. 2013. GA: A package for genetic algorithms in R. *J. Stat. Softw.* 53, 4 (2013), 37.