PARAMETER ESTIMATION AND REAL-TIME PREDICTION OF A DENGUE OUTBREAK USING MODEL AVERAGING: THE DENGUERT R PACKAGE

Carlos Rafael Sebrango Rodríguez*, Lizet Sánchez Valdés**, Ziv Shkedy***,Ewoud De Troyer***,Martin Otava***, Vivian Sistachs Vega****
*University of Sancti Spiritus "José Martí Pérez", Cuba (sebrango@uniss.edu.cu)
** Center of Molecular Immunology, Cuba (lsanchez@cim.sld.cu)
***Center for Statistics, Hasselt University, Belgium (ziv.shkedy@uhasselt.be, ewoud.detroyer@uhasselt.be, martin.otava@uhasselt.be)
****Havana University, Cuba (vivian@matcom.uh.cu)

ABSTRACT
Early estimation and prediction of the turning point and final size of any epidemic and in particular for dengue outbreaks can be useful for health authorities in order to explore the control measures and to plan the response to the outbreak. The Richards model is often been used to estimate epidemiological parameters for infectious diseases based on the reported cumulative cases. However, other nonlinear growth models can also fit the data well. Recently, we proposed the use of several nonlinear models for estimation and real time-prediction of epidemiological parameters via the method of model averaging. In order to implement this method, an R package was created. The ***DengueRT*** package uses the incidence data from a single dengue outbreak and gives estimates for the final size, the turning point of the epidemic and conduct a real-time prediction for these parameters using several nonlinear models via model averaging taking into account model uncertainty. The package also includes graphical tools for a visualization of the results. In this paper, we describe the ***DengueRT*** package and illustrate its use for a single dengue outbreak that occurred in two health areas of the Playa municipality in Havana City, Cuba during the 2001/2002 outbreak.

KEYWORDS: Dengue outbreak, R package, model averaging, parameter estimation, real-time prediction.

MSC: 62P10

RESUMEN
La estimación y predicción de parámetros epidemiológicos en brotes de dengue puede ser útil para las autoridades de salud con el fin de explorar las medidas de control y planear la respuesta al brote. El modelo de Richards se utiliza para estimar parámetros epidemiológicos en enfermedades infecciosas basadas en los casos acumulados notificados. Sin embargo, otros modelos de crecimiento pueden también ajustar bien los datos. Recientemente, propusimos el uso de varios modelos no lineales para la estimación y predicción en tiempo real de parámetros epidemiológicos a través del método de promedio de modelos. Para implementar este método, se creó un paquete de R. El paquete ***DengueR***T usa los datos de incidencia de un brote de dengue de una onda y brinda estimaciones para el acmé y tamaño final de la epidemia y conduce predicciones en tiempo real para estos parámetros utilizando varios modelos a través del promedio de modelo teniendo en cuenta la incertidumbre de modelos. El paquete también incluye herramientas gráficas para la visualización de los resultados. En este artículo se describe el paquete ***DengueRT*** y se ilustra su uso para los brotes de dengue que ocurrieron en dos áreas de salud de La Habana durante el brote 2001/2002.

PALABRAS CLAVES: Brote de dengue, paquete de R, promedio de modelos, estimación de parámetros, predicción en tiempo real.

1. INTRODUCTION

In recent years, there has been an increased interest in using statistical models for analysis of single dengue outbreaks [8] based on the reported cumulative cases. These models facilitate the estimation of primary epidemiological parameters, assess the impact of control interventions, and generate short and long-term forecasts, just to name a few [2]. During a single peak epidemic, the turning point, the time point at which the rate of accumulation changes from increasing to decreasing, and the final size of the epidemic are among the most important epidemiological parameters to be estimated [4] [5].
Phenomenological models emphasize the reproducibility of empirical observations using simple models [2]. Maximum likelihood fitting of phenomenological models remains important due to its simplicity, to the

difficulty of using modern methods in the context of limited data [7]. Among the most used phenomenological models is the Richards model [11]. In particular, Hsieh et al. [4] [5] proposed to use a specific nonlinear model, the Richards model, to estimate these two key parameters. In addition, the model proposed by Hsieh et al. [4] can be used for real-time prediction of these primary parameters. The Richards model considers only the cumulative infective population size with saturation in growth as the outbreak progresses [4].

A variety of nonlinear models has been considered to model growth data. Among them, we consider the three parameter logistic, five parameters logistic, Gompertz and Weibull models. All these models can be fitted to epidemic data well. The use of several models for fitting the same data raises the issue of model selection. Typically, one selects the best fitting model out of the set of fitted models and ignores the uncertainty due to model selection in estimation and inference. For these reasons, several authors (i.e., Burnham & Anderson [1], Claeskens & Hjort [3] and Lin [6]) advocate the use of model averaging techniques to perform multi-model for the estimated parameters. Model averaging is a method that takes into account all fitted models for the estimation of the parameters.

Recently, we proposed the use of several nonlinear models for estimation and real time-prediction of epidemiological parameters via the method of model averaging [12] [13]. In order to implement this method, an R package was created. In this paper, we describe the R package ***DengueRT*** and illustrate its use for a single dengue outbreak that occurred in two health areas of the Playa municipality in Havana City, Cuba during the 2001/2002 outbreak.

The paper is organized as follows: In the Methods section we present the set of the nonlinear models used and discuss the topics of model uncertainty, model selection and model averaging. In the next section, we introduce the ***DengueRT*** package, describing its main functions and illustrating its use with two examples, one for the retrospective parameter estimate and the other, for real time-prediction. Finally, we expose the conclusions.

2. Methods

2.1. Modeling dengue outbreak using nonlinear models

The Richards model belongs to the family of nonlinear models and its expression is given in the first line in Table 1. The parameter vector to be estimated is $\theta = (\alpha, \gamma, k, \eta)$ where α is the final size of the epidemic, γ is the per capita intrinsic growth rate of the infected population, k is the exponent of the deviation from the standard logistic curve and η is the turning point.

The cumulative number of reported cases in a dengue outbreak is an example of growth data. In addition to the Richards model, Table 1 presents five additional possible nonlinear models that can be used to model the outbreak data. The three parameter logistic model (3P logistic) is a special case of the Richards model, obtained when the exponent $k = 1$. The Gompertz model is another special case of the Richards function when $k \to 0$, and is frequently used in situations where growth is not symmetrical about the turning point. There are many variants of the Weibull model, the one we use in this paper is a modification of the Gompertz model when its independent variable time, is rescaled by logarithmic transformation. The Sigmoid Emax model and the five-parameter logistic (5P logistic) are commonly used in dose response modeling. The Sigmoid Emax model is obtained by mathematical transformation of 3P logistic model and rescaling the independent variable by a logarithmic transformation. The 5P logistic model is also obtained by rescaling the independent variable by logarithmic transformation and by doing a reparametrization, so that the model evaluated at the inflection point (η) reaches 50% of maximum response.

For all the models in Table 1, the turning point and the final size of the epidemic are parameters in the models. As in Hsieh et al. [4] [5], we assume that the cumulative number of reported cases at time t, Y_t, has asymptotic normal distribution, $Y_t \sim N(\mu(t,\theta), \sigma^2)$. Note that $\mu'(t,\theta) = \frac{\partial \mu(t,\theta)}{\partial t}$ is the incidence at time t.

Table 1: Nonlinear models considered to fit the cumulative cases of dengue outbreak.

Models	$\mu(t,\theta)$	$\mu'(t,\theta)$
Richards	$\frac{\alpha}{[1 + k.e^{-\gamma.k.(t-\eta)}]^{\frac{1}{k}}}$	$\gamma\mu(t)[1 - \left(\frac{\mu(t)}{\alpha}\right)^k]$
3P Logistic	$\frac{\alpha}{1 + e^{-\gamma(t-\eta)}}$	$\gamma\mu(t)\left[1 - \frac{\mu(t)}{\alpha}\right]$

5P Logistic	$\alpha + \frac{\alpha_0 - \alpha}{[1 + (2^{\frac{1}{k}} - 1)(\frac{t}{\eta})^{\gamma}]^k}$	$-\frac{k\gamma}{t}[\mu(t) - \alpha]\left[1 - \left(\frac{\mu(t) - \alpha}{\alpha_0 - \alpha}\right)^{\frac{1}{k}}\right]$
Sigmoid Emax	$\alpha_0 + \frac{t^n(\alpha - \alpha_0)}{t^n + \eta^n}$	$\frac{n}{t}[\mu(t) - \alpha_0]\left[1 - \frac{\mu(t) - \alpha_0}{\alpha - \alpha_0}\right]$
Gompertz	$\alpha_0 + (\alpha - \alpha_0)e^{-e^{-\gamma(t-\eta)}}$	$-\gamma[\mu(t) - \alpha_0]ln\left[\frac{\mu(t) - \alpha_0}{\alpha - \alpha_0}\right]$
Weibull	$\alpha + (\alpha_0 - \alpha)e^{-(\frac{t}{\eta})^{\gamma}}$	$\frac{\gamma}{t}[\mu(t) - \alpha]\ ln\left[\frac{\mu(t) - \alpha}{\alpha_0 - \alpha}\right]$

2.2. Model uncertainty, model selection and model averaging

In this section, we describe the model averaging (MA) technique [1] [3] [6], which is used to account for model uncertainty by combining the estimates from all the fitted models. It is based upon a weighted average of the parameter of primary interest obtained from different models, giving largest weights to those models that best fit the data [1]. Let us assume that the Akaike's Information Criterion (AIC) is used for model selection. For a given set of R candidate models $M_1, M_2, \dots, M_R$, Burnham & Anderson [1] proposed to rescale the AIC to

$$\Delta AIC_i = AIC_i - AIC_{min},\ \ i = 1, \dots, R$$

Here, AIC_{min} is the smallest AIC value across the set of R models. Burnham & Anderson [1] defined Akaike's weights as

$$w_i(AIC) = \frac{e^{-\frac{1}{2}\Delta AIC_i}}{\sum_{i=1}^{R} e^{-\frac{1}{2}\Delta AIC_i}}$$

Following Burnham & Anderson [1], we can calculate the model averaged estimator for turning point ($\hat{\eta}_{MA}$) and the final size of outbreak ($\hat{\alpha}_{MA}$) as follow:

$$\hat{\eta}_{MA} = \sum_{i=1}^{R} w_i(AIC)\hat{\eta}_i\ ,\ \ \hat{\alpha}_{MA} = \sum_{i=1}^{R} w_i(AIC)\hat{\alpha}_i$$

Here, $\hat{\eta}_i$ and $\hat{\alpha}_i$ are the parameter estimates for the turning point and final size of outbreak of *i*-th model, respectively. The estimators for variance for $\hat{\eta}_{MA}$ and $\hat{\alpha}_{MA}$ are given, respectively, by:

$$\hat{V}(\hat{\eta}_{MA}) = \left[\sum_{i=1}^{R} w_i(AIC)\sqrt{\hat{V}(\hat{\eta}_i|M_i) + (\hat{\eta}_i - \hat{\eta}_{MA})^2}\right]^2,$$

$$\hat{V}(\hat{\alpha}_{MA}) = \left[\sum_{i=1}^{R} w_i(AIC)\sqrt{\hat{V}(\hat{\alpha}_i|M_i) + (\hat{\alpha}_i - \hat{\alpha}_{MA})^2}\right]^2$$

Note that one can replace the AIC by other information criteria such as BIC, KIC [3] and calculate the model's weight based on these criteria.

3. INTRODUCTION TO *DENGUERT* PACKAGE

One of the strengths of R is the ability to share software as packages. Packages give users a reliable, convenient, and standardized way to access R functions, data, and documentation [9]. The nonlinear growth models and the model averaging method, discussed in the previous section, are implemented in the R package ***DengueRT***, which can be download from Comprehensive R Archive Network (CRAN) [10] in the Web site https://cran.r-project.org/web/packages/DengueRT/. The ***DengueRT*** package does require packages: ***nlme***, ***drc*** and ***ggplot2*** to work properly. Table 2 presents the main functions in the package and their descriptions. A S3 class named *dengue*, a list with the outputs of the *allmodels*, *allmodelpredict* and *changetimeFSTP* objects, was incorporated into the package. The last two main functions, the generic functions, dispatch the S3 methods *summary* and *plot*, which produce summaries and visualize the results of these objects of the S3 class *dengue*.

Table 2: The main ***DengueRT*** package functions.

Function	Description
allmodels()	Gives the parameter estimate for each built-in model and model-averaged estimate for final size and turning point of outbreak.
allmodelpredict()	Gives real-time predictions of these parameters for each built-in model and model-averaged.
changetimeFSTP()	Gives the changes over time of the final size and turning point estimates for each built-in model and model averaged.
summary()	Generic function which produces result summaries of the functions: *allmodels(), allmodelpredict()* and *changetimeFSTP()*.
plot()	Generic function which produces plots for a visualization of the results of the functions: *allmodels(), allmodelpredict()* y *changetimeFSTP()*.

Example 1: Parameter estimate of dengue outbreak.
In this section we illustrate the use of the package ***DengueRT*** on the data set related to a single dengue outbreak occurred in "26 de Julio" and "J.R. Ramirez" health areas in Playa municipality. The first step in R package ***DengueRT*** is typically to define the incidence and time vector. The data frame *dengueoutbreak1* and *dengueoutbreak2* contain the data for the two health areas, respectively.

```
> library(DengueRT)
# Data health area 1
> data(dengueoutbreak1)
# Data health area 2
> data(dengueoutbreak2)


> head(dengueoutbreak1)
  Incidence Time
1         1    1
2         0    2
3         2    3
4         5    4
5         2    5
6         2    6
> head(dengueoutbreak2)
  Incidence Time
1         5    1
2         4    2
3         4    3
4         3    4
5         7    5
6         1    6
```

In the second stage, we have to decide which nonlinear model from the built-in models to use to estimate the parameters or if all nonlinear models will be used. The function *allmodels()* provides a list of outputs including the parameter estimate for each built-in model and model averaged estimate for final size and turning point of outbreak and, when all the built-in models are used, gives the AIC of each model, the model averaged weights, the predicted incidence and cumulative cases. In order to produce result summary of it, the function *summary()*, S3 method from class *dengue*, is used. We show its use with a single model first and then with all built-in nonlinear models.

```
# Parameter estimate using Richards model
# (Health area 1)
> h1 <-allmodels(dengueoutbreak1$Incidence,dengueoutbreak1$Time, model="Richards")
> summary(h1)

# output 1
Richards model
AIC
214.0607
Parameter estimate
           Value  Std.Error    t-value      p-value
```

```
alpha 375.0832264 1.14702012 327.006667 5.761598e-58
k       2.2914315 0.23076559   9.929693 2.693529e-11
gamma   0.3284212 0.01383574  23.737170 6.951558e-22
eta    15.8038168 0.12280456 128.690793 5.126656e-45

Final size estimate:
   lower     est.    upper
372.7468 375.0832 377.4196

Turning point estimate:
   lower     est.    upper
15.55367 15.80382 16.05396
```

In case that a model average estimates for the parameters are of interest we use the *allmodels()*the following way:

```
# Parameter estimate using all nonlinear built-in model
#(Health area 2)
> h2 <-allmodels(dengueoutbreak2$Incidence,dengueoutbreak2$Time,model="all")
> summary(h2)
```

The parameter estimates are shown in the panel below.

```
#Output 2
AIC
      Richards    3P Logistic Sigmoidal Emax       Gompertz        Weibull   5P Logistic
      234.7269       240.6719       222.4572       238.4828       250.1002    220.7851
Model weights
      Richards    3P Logistic Sigmoidal Emax       Gompertz        Weibull   5P Logistic
   6.54415e-04    3.34912e-05    3.02130e-01    1.00063e-04    3.00323e-07   6.97081e-01

Estimates of the final size
                 lower     est.    upper
Richards       339.8966 343.8305 347.7645
3P logistic    341.7937 346.0720 350.3503
SigmEmax       344.4335 348.0050 351.5764
Gompertz       343.5955 347.9745 352.3536
Weibull        347.2308 353.6570 360.0832
5P logistic    342.6305 346.2744 349.9184
Model averaged 341.8111 346.7956 351.7801

Estimates of the turning point
                 lower     est.    upper
Richards       13.29229 13.74728 14.20227
3P logistic    12.98689 13.16641 13.34593
SigmEmax       13.20638 13.35525 13.50413
Gompertz       12.14338 12.33392 12.52446
Weibull        12.15957 12.36931 12.57904
5P logistic    13.22684 13.37713 13.52742
Model averaged 13.17956 13.37064 13.56172
```

These results can be visualized using the *plot()* function, S3 method from class *dengue*. When the argument of this function is an allmodels object, using the graph options, *which*, this function plots the cumulative (*which*=1) or incidence (*which*=2) epidemic curves with the fitted models, the final size (*which*=3) or turning point (*which*=4) with respective 95% CI and point estimates. We show its use below

```
# Figure 1a, 1b
# Note that graph options which=3 and which=4 are not used
# because only one model was used
> h1 <-allmodels(dengueoutbreak1$Incidence,dengueoutbreak1$Time, model="Richards")
> plot(h1,which=c(1,2), xlab="Weeks")
```

Figure 1 is constructed using the default options in the command *plot*, which presents the outbreak data and fitted models when the Richards model is used.

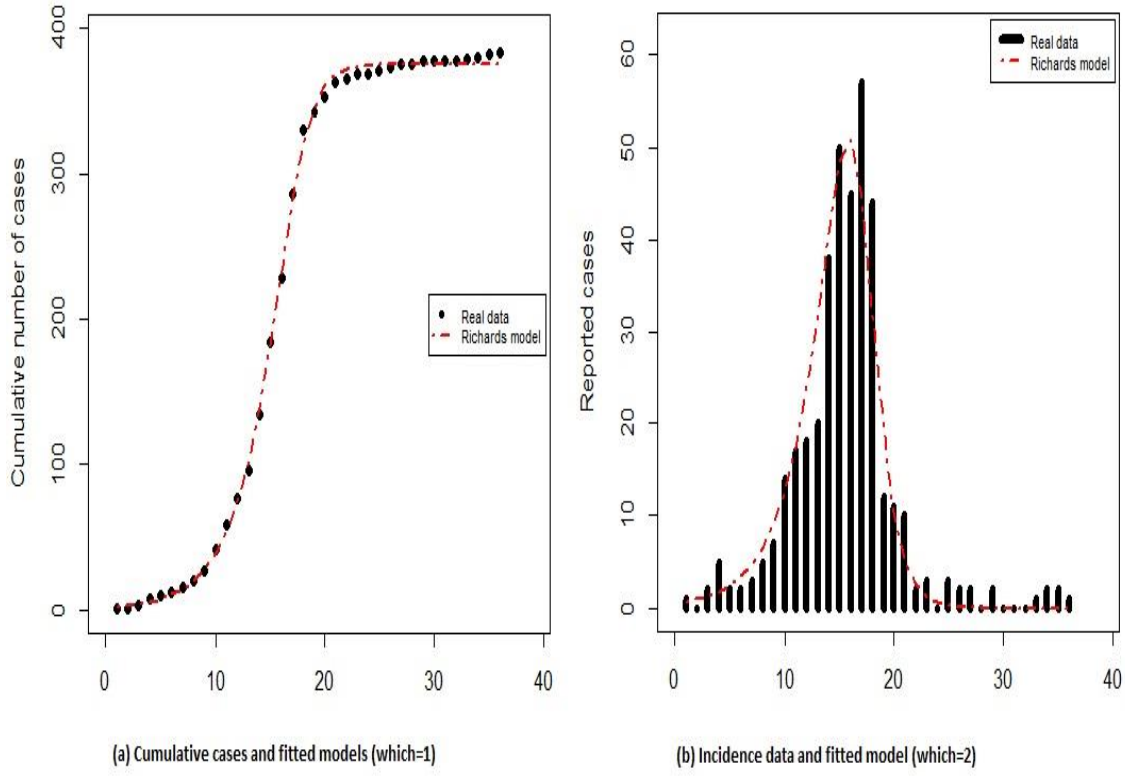

Figure 1: Example of outputs of function *plot()* when only one model is used.

Figure 2, in addition to the outbreak data and fitted models, shows the parameter estimate and CI for the turning point and the final size of the epidemic, for all built-in models to the R package.

```
# Figure 2a, 2b
> h2 <- allmodels(dengueoutbreak2$Incidence,dengueoutbreak2$Time, model="all")
> plot(h2,which=c(1:4),xlab="Weeks")
```

Real-Time predictions

Modeling based on model averaging, or on a single model, is particular useful for real-time prediction. We can use the R package ***DengueRT*** to forecast the eventual severity of the outbreak in real-time by estimating the carrying capacity. The two main functions of the R package ***DengueRT*** for providing real-time predictions are *allmodelpredict()* and *changetimeFSTP()*. In order to illustrate some of the features of the package for real-time prediction, the following example is shown.

Example 2: Real-time predictions of a dengue outbreak

Suppose that incidence data from the health area 1 are available through week 22 and is required to obtain the predictions of the final size and the turning point of at the end of the epidemic, as well as the incidence and cumulative number of cases at week 30.

```
# Real-time prediction using all nonlinear built-in models
# Incidence data available through week 22(Health area 1)
> hrp2 <-allmodelpredict(dengueoutbreak1$Incidence[1:22],dengueoutbreak1$Time[1:22],30,
+ model = "all")
> summary(hrp2)
```

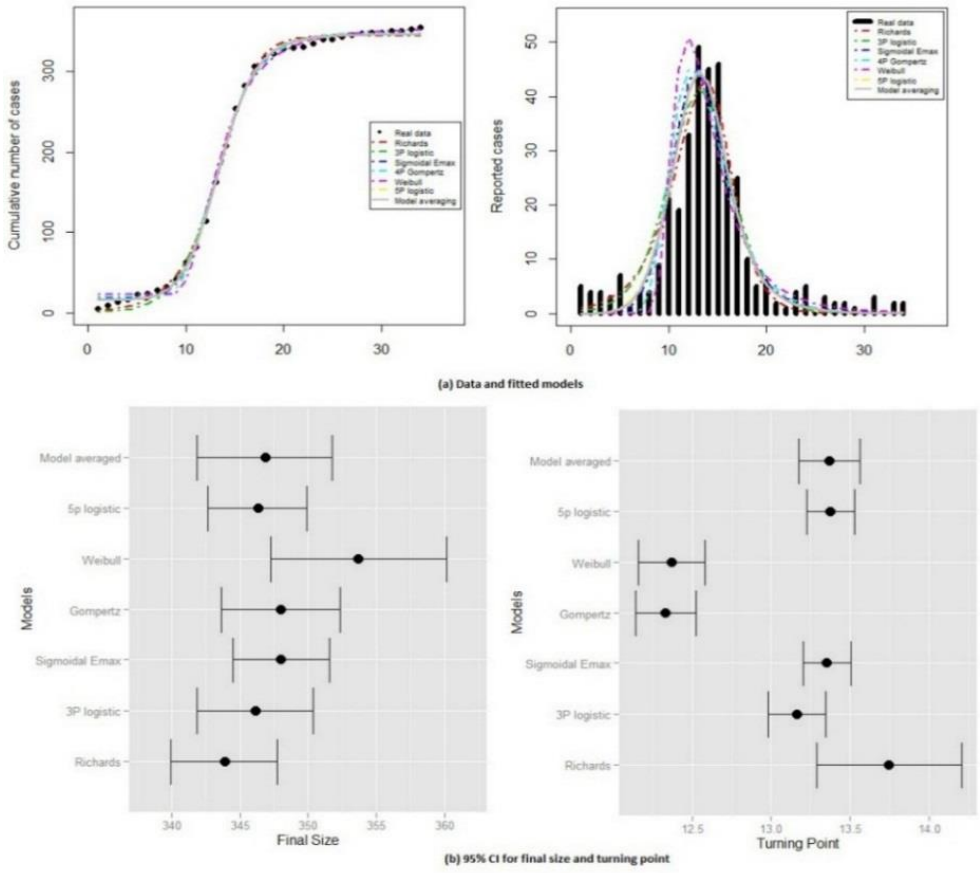

Figure 2: Example of outputs of function *plot()* when all built-in models are used.

AIC values and model weights are shown bellow.

```
#output 3
AIC
      Richards    3P Logistic Sigmoidal Emax       Gompertz        Weibull
      117.6249       160.8052       169.1075       179.6142       185.8068
   5P Logistic
      129.3018

Model weights
      Richards    3P Logistic Sigmoidal Emax       Gompertz        Weibull
   9.97095e-01    4.19025e-10    6.59831e-12    3.45087e-14    1.56036e-15
   5P Logistic
   2.90488e-03
```

Model specific and model average prediction of the final size at the end of epidemic are given by

```
                lower     est.    upper
Richards       358.7022 363.4467 368.1911
3P logistic    373.0097 389.2167 405.4237
SigmEmax       375.4903 402.4293 429.3684
Gompertz       378.5680 419.4317 460.2954
Weibull        384.1957 472.9414 561.6870
5P logistic    355.9536 362.1653 368.3770
Model averaged 357.6308 363.4429 369.2550
```

Model specific and model average prediction of the turning point at the end of epidemic are given by

```
                lower     est.    upper
```

```
Richards       15.79110 15.98481 16.17853
3P logistic    14.79243 15.08987 15.38731
SigmEmax       14.88398 15.31167 15.73936
Gompertz       13.77158 14.26487 14.75817
Weibull        13.75640 14.68116 15.60592
5P logistic    14.96065 15.06854 15.17643
Model averaged 15.73889 15.98215 16.22541

Model averaged prediction of the incidence at the time point 30
[1] 0.0003387103

Model averaged prediction of the cumulative number of cases at the time point 30
[1] 363.4426
```

The observed final size of the outbreak in health area 1 is equal to 383 reported cases. Note that the model average estimate for the final size of the epidemic, obtained from the function *allmodelpredict()* (see output 3), is 363.44 (357.63, 369.26). This indicates that around 4 months before the end of the epidemic this valuable information can be available for health authorities. As with the function *allmodels()*, the results can be visualize using the function *plot()* with the same arguments as before. Figure 3 presents the outbreak data through week 22 and fitted models until week 30 when all built-in models to R package ***DengueRT*** are used.

```
# Figure 3a, 3b
# Note it is used only the graphs options which=c(1,2)
> hrp2 <-allmodelpredict(dengueoutbreak1$Incidence[1:22],dengueoutbreak1$Time[1:22],30,
+ model = "all")
> plot(hrp2,which=c(1,2))
```

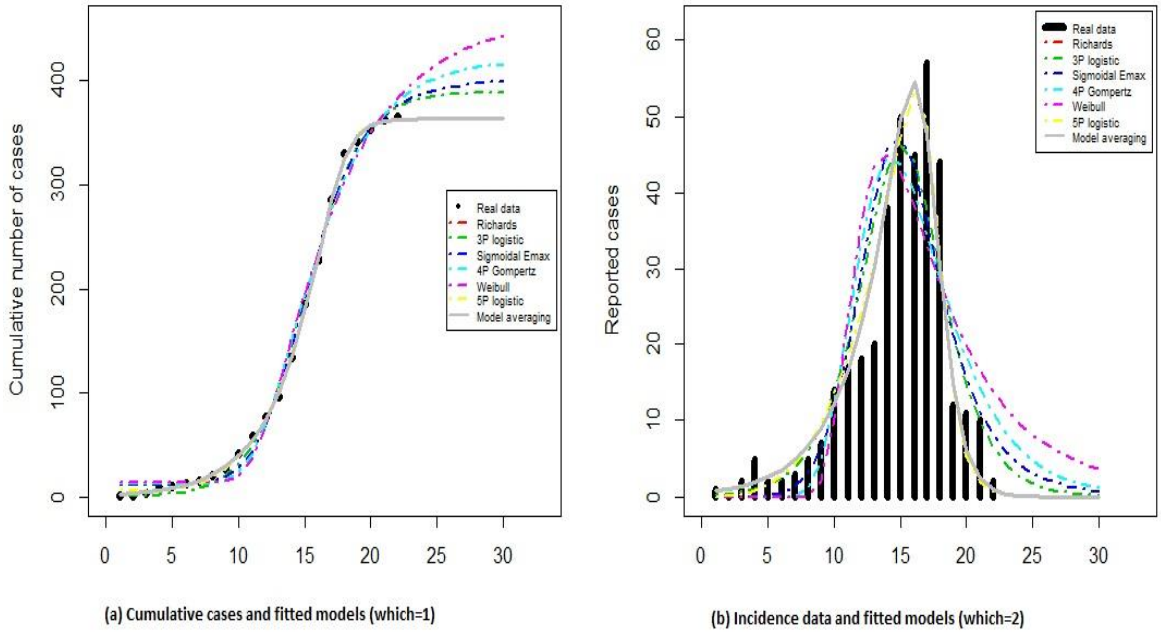

Figure 3: Example of outputs of function *plot()* when all built-in models are used to perform real-time predictions.

The last main function is *changetimeFSTP()*. This function provides the changes over time of the final size and turning point estimates (for specific models and model average) from the time point required until the last time point available. The output of the function are two tables, one for final size and the other for turning point. Suppose the incidence data in the health area 2 are available until week 34 and we want to analyze the change over time of both, final size and turning point, at the end of epidemic since week 19.

```
## (Health area 2)
## Changes over time, since time point 19 to time point 34, of the
## final size and turning point estimates for each built-in model
```

```
## and model averaged estimates
> ct1 <-changetimeFSTP(dengueoutbreak2$Incidence,dengueoutbreak2$Time,ini=19)
> summary(ct1)
```

The estimated final size is shown below (the rows represent the data used for the estimation of the model parameters, i.e., in the first row the first 19 weeks were used, in the second row, the first 20 weeks, etc.).

#output 4
Changes over time of the parameter estimates for the final size

	Richards	logistic3P	SigmEmax	Gompertz	Weibull	logistic5P	Model averaged
1-19	322.4869	354.2292	356.3453	370.3399	421.0725	326.5975	326.1941
1-20	325.5440	347.5214	350.1577	358.4578	392.2822	329.0374	328.8119
1-21	327.3609	343.4819	346.0693	351.2540	375.4671	330.2713	330.1456
1-22	328.4894	340.8797	343.2368	346.6224	364.8610	330.9176	330.8455
1-23	330.0702	339.9997	342.2648	344.6123	359.4146	332.3353	332.2914
1-24	332.1078	340.3032	342.5960	344.2955	357.0755	334.4296	334.3963
1-25	333.5184	340.4077	342.6088	343.8483	354.8899	335.7741	335.7534
1-26	335.0099	340.8982	343.0567	343.9746	353.8106	337.2756	337.2631
1-27	336.4161	341.5110	343.6344	344.3154	353.2253	338.7061	338.7043
1-28	337.7641	342.2137	344.3130	344.8135	353.0033	340.0937	340.1121
1-29	338.9533	342.8721	344.9358	345.2966	352.8694	341.3028	341.3550
1-30	339.9168	343.3990	345.4026	345.6574	352.6572	342.2436	342.3315
1-31	341.0061	344.1077	346.0944	346.2576	352.8560	343.3696	343.5455
1-32	341.9181	344.6992	346.6498	346.7415	352.9534	344.2836	344.5462
1-33	342.8614	345.3585	347.2924	347.3205	353.2335	345.2549	345.6415
1-34	343.8305	346.0720	348.0050	347.9745	353.6570	346.2744	346.7956

Changes over time of the parameter estimates for the turning point is shown below

	Richards	logistic3P	SigmEmax	Gompertz	Weibull	logistic5P	Model averaged
1-19	14.02646	13.28732	13.46911	12.53888	12.94928	13.22999	13.31139
1-20	14.01093	13.18380	13.38643	12.43053	12.67944	13.25221	13.30232
1-21	13.99784	13.12017	13.33196	12.36903	12.53247	13.26291	13.29506
1-22	13.98775	13.07874	13.29457	12.33237	12.44737	13.26822	13.28977
1-23	13.97135	13.06464	13.28186	12.31760	12.40741	13.27937	13.29288
1-24	13.94743	13.06954	13.28617	12.31541	12.39158	13.29529	13.30502
1-25	13.92910	13.07124	13.28634	12.31251	12.37788	13.30523	13.31152
1-26	13.90808	13.07927	13.29210	12.31328	12.37164	13.31606	13.32045
1-27	13.88672	13.08937	13.29952	12.31529	12.36850	13.32618	13.32928
1-28	13.86482	13.10104	13.30822	12.31809	12.36740	13.33583	13.33797
1-29	13.84430	13.11203	13.31618	12.32071	12.36678	13.34412	13.34530
1-30	13.82680	13.12088	13.32214	12.32259	12.36587	13.35049	13.35075
1-31	13.80606	13.13285	13.33095	12.32564	12.36666	13.35802	13.35701
1-32	13.78789	13.14289	13.33802	12.32803	12.36702	13.36408	13.36165
1-33	13.76829	13.15415	13.34619	12.33083	12.36797	13.37047	13.36615
1-34	13.74728	13.16641	13.35525	12.33392	12.36931	13.37713	13.37064

As can be seen, the model average estimate for the final size stabilizes slowly to the estimate of 346.80 at week 34. The real final size of this outbreak is 353. Thus, an accurate estimate for the final size could be reported to the health authority already in the middle of the outbreak. The model average estimate for the turning point is stabilized since week 19. Note that the point estimate for the turning point in week 19, 13.31, is very similar to the estimate if all data are used for estimation, 13.37.
As before, the function *plot()* can be used for visualization of the results (with the options *which=5* for final size and *which=6* for turning point). Figure 4 presents the change over time of the final size and turning point estimates from week 19 to the end of outbreak (week 34) for each built-in model and model averaged.

```
# Figure 4a, 4b
# Note it is used only the graphs options which=c(5,6)
# Incidence data available through week 34
# It is required to analyze the change over time from week 19
> ct1 <-changetimeFSTP(dengueoutbreak2$Incidence,dengueoutbreak2$Time,ini=19)
> plot(ct1,which=c(5,6))
```

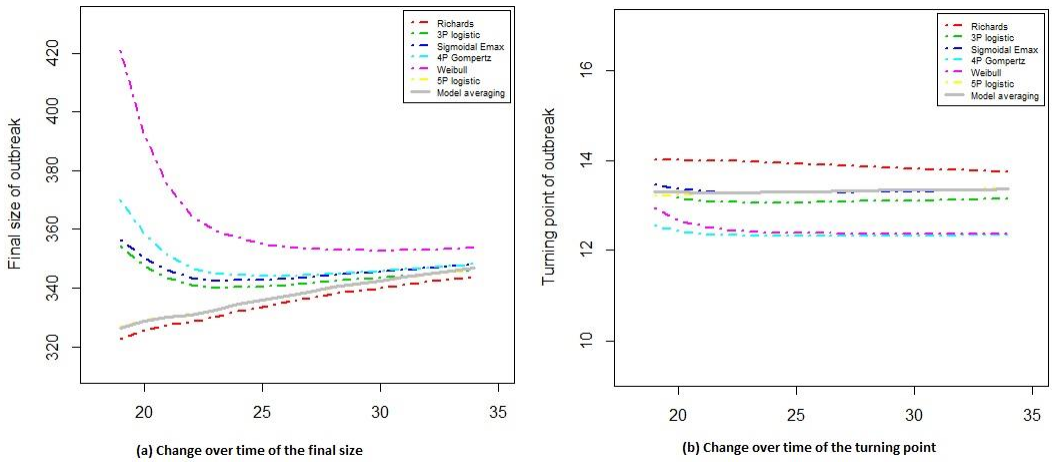

Figure 4: Example of outputs of function *plot()* when the arguments is a changetimeFSTP object with the graph options which=c(5,6).

4. CONCLUSIONS

The modeling approach based on model averaging provides an attractive framework for early estimation and prediction of the turning point and final size of any epidemic and in particular for dengue outbreaks since it takes into account a set of nonlinear models and the real-time prediction is dominated by the model(s) with the best goodness-of-fit to the data.

The nonlinear growth models: Richards, 3PL, 5PL, Weibull, Emax and 4P Gompertz models, which have the turning point and final size as model parameters, and the model averaging method are implemented in the R package ***DengueRT***. The ***DengueRT*** package uses the incidence data from a single dengue outbreak, gives estimates for the final size, the turning point of the epidemic and conducts a real-time prediction for these parameters using the model averaging method.

The package includes the main functions *allmodels()*, *allmodelpredict()* and *changetimeFSTP()* for computing the results, the *summary()* to produce result summaries of the functions and *plot()* to visualize results. The main features and the functionalities of the package ***DengueRT*** have been illustrated using two dataset corresponding to single dengue outbreak occurred in one municipalities in Havana, Cuba.

A limitation of the package is that it is not possible to estimate the key epidemiological parameter, the basic reproductive number R_0, since not all built-in models have the growth rate as model parameter.

As the functions provided by this package requires only the incidence data, the use of this package can be extended to other infectious diseases such as SARS, H1N1, Zika, Ebola, etc. This modeling approach and its implementation in the R package ***DengueRT*** could be a valuable tool to public health policymakers for responding to future disease outbreaks in order to explore the control measures and plan the response to the outbreak. This package was developed for a single-wave outbreak, in future research we will extend it to a multi-wave outbreaks setting as well.

RECEIVED: JULY, 2018
REVISED: OCTOBER, 2018

REFERENCES

[1] BURNHAM, K. and ANDERSON, D. (2004): Multimodel inference: Understanding AIC and BIC in model selection, **Sociological Methods Research** 33, 261-304.

[2] CHOWELL, G., HINCAPIA-PALACIO, D., OSPINA, J., PELL, B., TARIQ, A., DAHAL S., MOGHADAS S., SMIRNOVA A., SIMONEN L. and VIBOUD C. (2016). Using phenomenological models to characterize transmissibility and forecast patterns and final Burden of Zika epidemics. **PLOS Currents Outbreaks**. http://dx.doi.org/10.1371/currents.outbreaks.f14b2217c902f453d9320a43a35b9583

[3] CLAESKENS, G. and HJORT, N. (2008): **Model selection and model averaging**, Cambridge Univ. Press.

[4] HSIEH, Y.H. and MA, S. (2009): Intervention measures, turning point, and reproduction number for dengue, Singapore, 2005, **American Journal of Tropical Medicine and Hygiene** 80, 66-71.
[5] HSIEH, Y.H., ARAZOZA, H. and LOUNES, R. (2013): Temporal trends and regional variability of 2001-2002 multiwave denv-3 epidemic in Havana City: did hurricane Michelle contribute to its severity?, **Tropical Medicine and International Health** 18 (7), 830-838.
[6] LIN, D., SHKEDY, Z., YEKUTIELI, D., AMARATUNGA, D. and BIJNENS, L. (2012): **Modeling Dose-response Microarray Data in Early Drug Development Experiments Using R**, Springer-Verlag.
[7] MA, J., DUSHOFF, J., BOLKER, B. M. and EARN, D.J.D. (2014): Estimating initial epidemic growth rate, **Bull Math Biol** 76:245-260. DOI 10.1007/s11538-013-9918-2.
[8] NISHIURA, H. (2006) Mathematical and statistical analyses of the spread of dengue. **Dengue Bulletin** 30: 51–67.
[9] PACE, L. (2012): **Beginning R: An Introduction to Statistical Programming**. Springer, New York.
[10] R DEVELOPMENT CORE TEAM (2017). R: a language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. http://www.R-project.org/..
[11] RICHARDS, F.J. (1959): A flexible growth function for empirical use. **Journal of Experimental Botany** 10, 290–300.
[12] SEBRANGO-RODRIGUEZ, C.R., MARTÍNEZ-BELLO, D.A., SÁNCHEZ, L., THILAKARATHNE P.J., DEL FAVA, E., VAND DER STUYFT, P., LÓPEZ-QUÍLEZ, A. and SHKEDY, Z. (2017): Real Time Parameter Estimation of Zika Outbreaks using Model Averaging, **Epidemiology and Infection** 145(11), 2313-2323.
[13] SEBRANGO RODRIGUEZ, C. R (2018): **Modeling dengue outbreak data, estimation and prediction of epidemiological parameters**, doctoral thesis, Havana University, Havana City, Cuba.