

# Model to Describe Fast Shutoff of CoVID-19 Pandemic Spread

Genghmun Eng

PhD Physics 1978, University of Illinois at Urbana-Champaign  
geng001@socal.rr.com

August 6, 2020

## Abstract

Early CoVID-19 growth obeys:  $N\{\hat{t}\} = N_I \exp[+K_o \hat{t}]$ , with  $K_o = [(\ln 2)/(t_{dbl})]$ , where  $t_{dbl}$  is the pandemic growth *doubling time*. Given  $N\{\hat{t}\}$ , the daily number of new CoVID-19 cases is  $\rho\{\hat{t}\} = dN\{\hat{t}\}/d\hat{t}$ . Implementing society-wide *Social Distancing* increases the  $t_{dbl}$  *doubling time*, and a linear function of time for  $t_{dbl}$  was used in our *Initial Model*:  $N_o[t] = \mathbf{1} \exp[+K_A t/(1 + \gamma_o t)] \equiv e^{+G_o} \exp(-Z_o[t])$ ,

to describe these changes, with  $G_o \equiv [K_A/\gamma_o]$ . However, this equation could not easily model some quickly decreasing  $\rho[t]$  cases, indicating that a second *Social Distancing* process was involved. This second process is most evident in the initial CoVID-19 data from *China*, *South Korea*, and *Italy*. The *Italy* data is analyzed here in detail as representative of this second process. Modifying  $Z_o[t]$  to allow exponential cutoffs:

$$Z_E[t] \equiv [+G_o / (1 + \gamma_o t)] [\exp(-\delta_o t - q_o t^2)] = Z_o[t] \exp(-\delta_o t - q_o t^2),$$

provides a new *Enhanced Initial Model (EIM)*, which significantly improves datafits, where  $N_E[t] = e^{+G_o} \exp(-Z_E[t])$ . Since large variations are present in  $\rho_{data}[t]$ , these models were generalized into an orthogonal function series, to provide additional data fitting parameters:

$$N(Z) = \sum_{m=0}^{m=M_F} g_m L_m(Z) \exp[-Z].$$

Its first term can give  $N_o[t]$  or  $N_E[t]$ , for  $Z[t] \rightarrow Z_o[t]$  or  $Z[t] \rightarrow Z_E[t]$ . The  $L_m(Z)$  are *Laguerre Polynomials*, with  $L_0(Z) = 1$ , and  $\{g_m; m = 0, M_F\}$  are constants derived from each dataset. When  $\rho[t] = dN[t]/dt$  gradually decreases, using  $Z_o[t]$  provided good datafits at small  $M_F$  values, but was inadequate if  $\rho[t]$  decreased faster. For those cases,  $Z_E[t]$  was used in the above  $N(Z)$  series to give the most general *Enhanced Orthogonal Function (EOF)* model developed here. Even with  $M_F = 0$ ,  $q_o = 0$ , this *EOF* model fit the *Italy* CoVID-19 data for  $\rho[t] \equiv dN[t]/dt$  fairly well.

When the  $\rho[t]$  post-peak behavior is not Gaussian, then  $Z_E[t]$  with  $\delta_o \neq 0, q_o = 0$ ; which we call  $Z_A[t]$ , is also likely to be a sufficient extension of the  $Z_o[t]$  model. The *EOF* model also can model a gradually decreasing  $\rho[t]$  tail using small  $\{\delta_o, q_o\}$  values [with 6 Figures].

# 1 Introduction

Let  $N\{\hat{t}\}$  be the total number of CoVID-19 cases in any given locality, with  $\rho\{\hat{t}\}$  being the predicted number of daily new CoVID-19 cases, so that:

$$N\{\hat{t}\} = \int_{t'=0}^{t'=\hat{t}} \rho\{t'\} dt', \quad [1.1a]$$

$$\rho\{\hat{t}\} = dN\{\hat{t}\} / d\hat{t}. \quad [1.1b]$$

Early CoVID-19 growth often obeys  $N\{\hat{t}\} \approx N_I \exp[+K_o \hat{t}]$ , with  $K_o = [(\ln 2)/t_{dbl}]$ , where  $t_{dbl}$  is the pandemic *doubling time*. The start of society-wide *Social Distancing* at  $\hat{t} = 0$  can gradually lengthen  $t_{dbl}$  for  $\hat{t} > 0$ . The  $\hat{t} < 0$  exponential growth phase is not applicable for estimating *Social Distancing* effects. For  $\hat{t} > 0$ , an *Initial Model* for CoVID-19 pandemic shutoff was first developed<sup>1</sup> using a linear function of time to describe the  $t_{dbl}$  changes:

$$N\{\hat{t}\} = N_I \exp[+K_o \hat{t} / (1 + \alpha_S \hat{t})]. \quad [1.2]$$

Given measured  $N_{data}\{\hat{t}\}$ , the data end-points  $\{N_I, N_F\}$  help to set  $\{K_o, \alpha_S\}$ . An *Orthogonal Function Model (OFM)* was developed next<sup>2</sup>, with Eq. [1.2] as the first term of the orthogonal function series. Each new *OFM* term provides another fitting parameter, to progressively better match  $N_{data}\{\hat{t}\}$  and  $\rho_{data}\{\hat{t}\} = (d/d\hat{t}) N_{data}\{\hat{t}\}$ .

The *OFM* improves on the *Initial Model*, and it works best with gradually decreasing  $\rho\{\hat{t}\}$  ["Slow Shutoff"]. In contrast, when  $\rho\{\hat{t}\}$  decreased quickly ["Fast Shutoff"], the *Initial Model* was not a good datafit, and a few-term *OFM* series only gave small improvements. This result indicates there is an inherent limit to what the gradually changing  $t_{dbl}$  *doubling time* of Eq. [1.2] can model.

For these cases, typified by CoVID-19 pandemic evolution in Italy, data often showed a stage where  $\rho_{data}\{\hat{t}\} \sim [\exp(-\delta_o \hat{t})]$  or  $\rho_{data}\{\hat{t}\} \sim [\exp(-q_o \hat{t}^2)]$ , which likely represents a second process, independent of the gradually changing  $t_{dbl}$  *doubling time*. An *Enhanced Initial Model (EIM)* is developed here to include this second process. The prior *OFM* methods can then be applied, giving an *Enhanced Orthogonal Function (EOF)* model for this more general case.

## 1.1 Review of Prior Models

The *Initial Model* of Eq. [1.2] is still needed as the first part of the *OFM*. The *Initial Model* starts with measured data end-points  $\{N_I, N_F\}$ , where  $\hat{t} = (t_F - t_I)$  is the largest data time interval so that  $N\{\hat{t} = (t_F - t_I)\} = N_F$ . Usually  $\alpha_S$  in Eq. [1.2] was chosen first, and  $K_o$  or  $t_{dbl}$  adjusted to match the  $N_F$  data end-point, using an *Excel<sup>TM</sup> – Goal-Seek* or its equivalent. The final  $\{K_o, \alpha_S\}$  values were the pair with the minimum *root-mean-square (rms)* error between the given data and the Eq. [1.2] model.

The above  $\{K_o, \alpha_S\}$  also provides a  $t = 0$  estimate for the pandemic start, and gives  $\{K_A, \gamma_o\}$  as new data fitting parameters:

$$N_o[t] = \mathbf{1} \exp[+K_A t / (1 + \gamma_o t)] \equiv \exp[+G_o] \exp(-Z_o[t]), \quad [1.3a]$$

$$G_o \equiv (K_A / \gamma_o), \quad [1.3b]$$

$$Z_o[t] = +[G_o / (1 + \gamma_o t)], \quad [1.3c]$$

$$\rho_o[t] \equiv dN_o[t] / dt. \quad [1.3d]$$

Using  $(t_F - t_I)$ , with  $N_o[t = t_I] = N_I$  and  $N_o[t = t_F] = N_F$ , sets:

$$t_I = \ln(N_I) / [K_o + \alpha_S \ln(N_I)]. \quad [1.4a]$$

$$t_F = \ln(N_I) / [K_o + \alpha_S \ln(N_I)] + (t_F - t_I), \quad [1.4b]$$

which determines  $\{K_A, \gamma_o\}$  in  $Z_o[t]$  for Eq. [1.3c]:

$$\gamma_o = \{ [\ln(N_I) / t_I] - [\ln(N_F) / t_F] \} / [\ln(N_F) - \ln(N_I)], \quad [1.5a]$$

$$K_A = [(1/t_I) - (1/t_F)] / \{ [1/\ln(N_I)] - [1/\ln(N_F)] \}. \quad [1.5b]$$

The Eq. [1.4a] value for  $t_I$  is what determines the new  $t = 0$  point, as an extrapolation for when  $N_o[t = 0] = 1$ . In addition:

$$N_o[t \rightarrow \infty, Z_o \rightarrow 0] \approx 1 \exp[+K_A / \gamma_o] = \exp[+G_o] \equiv N_{\max}^o, \quad [1.6a]$$

$$\rho_o[t] = dN_o[t] / dt \approx N_o[t] [G_o \gamma_o / (1 + \gamma_o t)^2] \rightarrow N_{\max}^o [G_o / (\gamma_o t^2)], \quad [1.6b]$$

provides an estimate for the total number of cases ( $N_{\max}^o$ ) at the pandemic end, and determines a function for the  $\rho_o[t]$  long-time tail. The  $\{0 < t < t_I\}$  period prior to the start of *Social Distancing*, extrapolates what pandemic progress would have looked like, if *Social Distancing* had begun at  $t = 0$ .

An *Orthogonal Function Model* [*OFM*] was then developed<sup>2</sup> to better model the different observed  $\rho_{data}[t]$  shapes, as an improvement of the *Initial Model*:

$$N(Z) = \sum_{m=0}^{m=M_F} g_m L_m(Z) \exp[-Z], \quad [1.7a]$$

$$R(Z) = \sum_{m=0}^{m=M_F} c_m L_m(Z) \exp[-Z], \quad [1.7b]$$

$$N(Z) \equiv \int_{Z'=Z}^{Z'=+\infty} R(Z') dZ', \quad [1.7c]$$

$$c_{M_F-k} = \sum_{m=0}^{m=k} g_m, \quad [1.7d]$$

with  $L_m(Z)$  being the *Laguerre Polynomials*, and  $L_m(Z = 0) = L_0(Z) \equiv 1$ . Using  $Z = Z_o[t]$  from Eq. [1.3c] gives  $N(Z) \rightarrow N(Z_o)$  and  $R(Z) \rightarrow R(Z_o)$ .

The  $\{g_m; m = (0, M_F)\}$  constants in Eq. [1.7a] can be arranged in a  $\vec{g}$ -vector form, with comparable constants for  $R(Z)$  from Eq. [1.7b] arranged in a  $\vec{C}$ -vector form. For  $M_F = 2$ , it allows Eq. [1.7d] to be written as:

$$\vec{C} = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \vec{g} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} g_0 \\ g_1 \\ g_2 \end{pmatrix}. \quad [1.8]$$

Once these  $c_m$  values are determined for  $R(Z)$  in Eqs. [1.7b]-[1.7c], an *OFM* feature is that  $c_0$ , by itself, becomes the *OFM* best estimate for the total number of CoVID-19 cases at the pandemic end.

For large enough  $M_F$  values, and a monotonic  $Z$ -function, this *OFM* can provide successively better approximations to almost any given set of  $N_{data}[Z]$ , with  $Z[t] \rightarrow Z_o[t]$  of Eq. [1.3c] being a specific case.

The *OFM* implicitly uses a *Linear Y-axis*, so its results differ from the *Initial Model* datafit on a *Logarithmic Y-axis*. As an example, compare the *Initial Model* result of  $N_o(Z_o) = G_0 \exp[-Z_o]$  with the Eq. [1.7a] *OFM* result of  $N(Z_o) = g_0 \exp[-Z_o]$  for  $M_F = 0$ . In the *Initial Model*,  $G_o$  is fixed so that  $N_o(Z_o)$  exactly matches  $\{N_I, N_F\}$  at the  $\{t_I, t_F\}$  boundaries. In the *OFM*,

$g_0 = G_o$  is no longer required, so that the OFM  $N(Z_o)$  best datafit is not constrained to exactly match  $\{N_I, N_F\}$  at  $\{t_I, t_F\}$ .

The above  $R(Z)$  and  $Z[t]$  gives  $N[t]$  and  $\rho[t]$  as an explicit functions of time:

$$N[t] \equiv \int_{\hat{z}=Z[t], t'=t}^{\hat{z}=+\infty, t'=(t_{\min})} R(Z[t']) \frac{dZ}{dt'} dt' \equiv \int_{t'=t}^{t'=t} \rho[t'] dt'. \quad [1.9]$$

Since Eq. [1.6b] gives  $\rho\{t\} \sim [1/t^2]$ , the aim here is to model faster decaying functions such as  $\rho\{t\} \sim [\exp(-\delta_o t)]$  or  $\rho\{t\} \sim [\exp(-q_o t^2)]$ .

## 1.2 Updated Initial Model Results for Italy

The  $Z_o[t]$  model of Eq. [1.3a] was applied to *bing.com* data<sup>9</sup> for Italy, starting with  $N_{data}^{2/23/2020} = 150$  CoVID-19 cases as an early pandemic point, up through June 15, 2020. Here,  $t_I$  is when mandatory *Social Distancing* was introduced at  $N_{data}^{3/10/2020} = 10,149$ ; with  $t_F$  being when  $N_{data}^{6/15/2020} = 237,290$ . Data prior to *Social Distancing* ( $t < t_I$ ) was excluding from this *Social Distancing* analysis.

**Figure 1** compares the  $\rho_{data}[t]$  results with the updated  $N_o[t]$  and  $\rho_o[t] = dN_o[t]/dt$  predictions using  $Z_o[t]$ . The  $N_{data}^{2/23/2020}$  (*Day 1*) to  $N_{data}^{3/10/2020}$  (*Day 17*) interval was examined for estimating a  $t = 0$  pandemic start where  $N_{data}[t = 0] \rightarrow 1$ . A best fit value of  $t_{offset} = 9.10055$  days was found, giving  $t_I = (17 - 9.10055) = 8.89945$  days for  $N_{data}^{3/10/2020}$  (*Day 17*), while  $N_{data}^{6/15/2020}$  (*Day 114*) gives  $t_F = (114 - 9.10055) = 104.89945$  days, so that:

$$N_{data}[t_I = 8.899] = 10,149; \quad [1.10a]$$

$$N_{data}[t_F = 104.899] = 237,290; \quad [1.10b]$$

$$N_{data}^{3/2/2020}[t = 0] = 1, \quad [1.10c]$$

$$N[t \rightarrow \infty] = 338,165; \quad [1.10d]$$

with  $\{K_A, \gamma_o\} \approx \{4.2405, 0.33078\}$  and  $(t_F - t_I) = 97$  days. In **Fig. 1**, the X-axis uses this  $t = 0$  point where  $N_o[t = 0] \rightarrow 1$ , and it shows what *Social Distancing* effects would have been, if it had been operating throughout the  $t > 0$  period. This  $\rho_o[t]$  prediction still has a much more gradual drop than the data. This discrepancy indicates that a second *Social Distancing* process is operating, besides just the gradual  $t_{dbl}$  lengthening of the *Initial Model*.

**Figure 2** compares the  $N_o[t]$  predictions for this model, to the measured  $N_{data}[t]$ . Systematic deviations are evident, with the net *rms* error on a *Logarithmic Y-axis* being  $rms_{error} = 0.097828$ . To cure these defects, an enhanced  $Z[t]$  model is developed next.

## 2 Developing Enhanced $Z[t]$ Models

To generalize  $Z[t]$  beyond Eq. [1.3c], it is convenient to use the  $t = \{0^+, \infty^-\}$  domain, and require convergence of  $Z[t] \rightarrow Z_o[t]$  in some limit, along with:

$$\lim_{t \rightarrow 0} \{Z[t]\} = G_o \equiv [K_A / \gamma_o], \quad [2.1a]$$

$$\lim_{t \rightarrow +\infty} \{Z[t]\} = 0, \quad [2.1b]$$

and that the  $M_F = 0$  case of Eq. [1.7a] remains as:

$$N(Z) = g_0 \exp[-Z]. \quad [2.2]$$

How to choose an appropriate  $Z[t]$ , as part of an *Enhanced Initial Model* (*EIM*), which also allows a  $\rho[t] \sim [\exp(-\delta_o t)]$  or  $\rho[t] \sim [\exp(-q_o t^2)]$  stage, is examined next. It can be motivated by studying a simple  $N_T[t]$  test-case, where  $\rho_T[t]$  itself is a pure exponential decay, as a function of time:

$$N_T[t] = [e^{+G_o}] [(1 + e^{-G_o}) - \exp(-\delta_o t)], \quad [2.3a]$$

$$\rho_T[t] = \{dN_T[t] / dt\} = [\delta_o e^{+G_o}] \exp(-\delta_o t), \quad [2.3b]$$

while also preserving  $N_T\{t \rightarrow 0\} = 1$ . Comparing Eq. [2.3a] to the Eq. [2.2], sets  $Z_T[t]$  for this test-case:

$$N_T[t] = \exp[+G_o] \exp(-Z_T[t]), \quad [2.4a]$$

$$Z_T[t] = (-1) \ln [(1 + e^{-G_o}) - \exp(-\delta_o t)]. \quad [2.4b]$$

At large times, Eq. [2.4b] gives:

$$\lim_{t \rightarrow \infty} \{Z_T[t]\} \approx \exp(-\delta_o t), \quad [2.5]$$

since  $(-\ln[1 - x] \approx x)$  for small  $x$ , which shows that if  $\rho_T[t]$  has an exponential tail, then  $Z_T[t]$  also has an exponential tail. A simple generalization for  $Z[t]$  in Eq. [1.7a]-[1.7d] would be either:

$$Z_E[t] \equiv +[G_o / (1 + \gamma_o t)] [\exp(-\delta_o t - q_o t^2)], \quad [2.6a]$$

$$Z_A[t] \equiv +[G_o / (1 + \gamma_o t)] [\exp(-\delta_o t)] = Z_E[t; q_o \equiv 0]. \quad [2.6b]$$

As with Eq. [1.3c], the original CoVID-19 exponential growth factor  $K_A$  remains only as part of the  $G_o$  scaling factor, while the  $\{\delta_o \rightarrow 0\}$  limit of Eq. [2.6b] converges back to the Eq. [1.3c] *Initial Model*.

Using Eq. [2.6a] for  $Z_E[t]$  in the Eq. [2.2]  $N(Z_E)$  example gives:

$$\rho[t] = \{dN[t] / dt\} = g_0 \frac{d}{dt} \exp[-Z_E] = -g_0 \exp[-Z_E] \frac{d}{dt} Z[t_E] = \quad [2.7]$$

$$+ N(Z_E) [\exp(-\delta_o t - q_o t^2)] \left\{ \frac{+G_o \gamma_o}{(1+\gamma_o t)^2} + \frac{G_o}{(1+\gamma_o t)} [+ \delta_o + 2q_o t] \right\},$$

which exhibits the following variety of long-time limits:

$$\rho[t; \delta_o = 0, q_o = 0] \rightarrow N(Z_E) \left\{ \frac{+G_o \gamma_o}{(1+\gamma_o t)^2} \right\}, \quad [2.8a]$$

$$\rho[t; q_o = 0] \rightarrow N(Z_E) [\exp(-\delta_o t)] \left\{ \frac{+G_o \delta_o}{(1+\gamma_o t)} \right\}, \quad [2.8b]$$

$$\rho[t; \delta_o = 0] \rightarrow N(Z_E) [\exp(-q_o t^2)] \left\{ \frac{2G_o q_o}{\gamma_o} \right\}, \quad [2.8c]$$

$$\rho[t] \rightarrow N(Z_E) \left\{ \exp[-q_o(t + \frac{\delta_o}{2q_o})^2] \right\} \left\{ \frac{2G_o q_o}{\gamma_o} \exp[+\frac{1}{4}\delta_o^2 / q_o] \right\}. \quad [2.8d]$$

Here, any  $q_o \neq 0$  Gaussian component in  $Z_E[t]$  gives a  $\rho[t]$  tail that is also a pure Gaussian. An exponential component ( $q_o \equiv 0$ ) in  $Z_E[t]$  gives a time-modified exponential  $\rho[t]$  tail, while having  $\{q \equiv 0, \delta_o \equiv 0\}$  in  $Z_E[t]$  gives the prior  $\rho[t] \sim (1/t^2)$  result of Eq. [1.6b].

### 3 Pandemic Fast vs Slow Shutoffs

Each  $Z_E[t]$  function modifies  $N[t]$  predictions for the pandemic start, pandemic end, and the mid-range where  $\rho[t]$  has its pandemic peak. The Eq. [2.6a]  $Z_E[t]$  function especially alters the calculated CoVID-19 pandemic tail. For either  $q_o \neq 0$  or  $\delta_o \neq 0$ , Eqs. [2.6a]-[2.6b] gives a pandemic *Fast Shutoff*, compared to the gradually decreasing  $Z_o[t]$  of Eq. [1.3c] in the *Initial Model*, which is a pandemic *Slow Shutoff*. However  $Z_o[t]$  from the *Initial Model*, and  $Z_A[t]$  from Eq. [2.6b] both gave long-term  $\rho[t]$  tails that decay much slower than the  $q_o \neq 0$  Eq. [2.6a] Gaussian.

If data does not show evidence of a Gaussian pandemic *Fast Shutoff*, assuming the post-peak  $\rho[t]$  data will be Gaussian is likely to provide optimistically inaccurate  $N[t]$  predictions for CoVID-19 pandemic evolution. Apparently, this is exactly what was done by the *University of Washington IHME (Institute of Health Metrics and Evaluation)* in their widely publicized initial preprint<sup>3</sup> of 27 March 2020, with this Gaussian model continuing throughout their subsequent updates<sup>4–6</sup> up through 29 April 2020.

*IHME* changed everything in their 4 May 2020<sup>7–8</sup> update. They no longer used  $\rho[t]$  Gaussian tails, and it doubled or tripled their predicted CoVID-19 pandemic death rates. Thus, unless the post-peak  $\rho[t]$  exhibits Gaussian behavior, the  $Z_A[t]$  with  $\delta_o \neq 0$  is likely the most important modification to  $Z_o[t]$ , which is the pandemic *Fast Shutoff* model used here.

Using  $Z_A[t]$ , the pandemic *Fast Shutoff* can be extrapolated to calculate a pandemic start point where  $N[t = 0] = 1$ . We then examine if this  $Z_A[t]$  must also carry over to the  $\rho[t]$  long-term tail,

Since the long-term low  $\rho_{data}[t]$  tail may differ among localities, and is not well known, the  $\delta_o \neq 0$  case of Eq. [2.6b] could end with a *Slow Shutoff*, giving:

$$Z_B[t] = +[G_o / (1 + \gamma_o t)] \{ \exp[-\delta_o t / (1 + \lambda_o t)] \}, \quad [3.1a]$$

$$\lim_{t \rightarrow \infty} \{Z_B[t]\} \approx +[G_o / (1 + \gamma_o t)] \{ \exp[-\delta_o / \lambda_o] \}, \quad [3.1b]$$

with  $G_o$  as in Eq. [2.1a]. This Eq. [3.1a]  $Z_B[t]$  function has the  $\{\delta_o, \lambda_o\}$  *Mitigation Measure* operating at the start of *Social Distancing*, but reverting to the *Initial Model* in the long-time limit. Combining Eq. [2.6b] and Eq. [3.1a] cases gives this *Enhanced Initial Model (EIM)* equation:

$$Z_B[t] = +[G_o / (1 + \gamma_o t)] \{ \exp[-\delta_o t / (1 + \kappa \gamma_o t)] \}, \quad [3.2]$$

where  $\kappa = 0$  is a pure exponential, and  $\kappa = 1$  has a modified tail that includes its own long-term shutoff. Comparing  $\kappa = \{0, 1\}$  in Eq. [3.2] provides a simple test for which model matches CoVID-19 data better in any locality. Any other  $\kappa > 0$  value then recovers the more general  $\lambda_o \equiv \kappa \gamma_o$  case. This Eq. [3.2]  $Z_B[t]$  replaces  $Z_o[t]$  of Eq. [1.3c], and its *EIM* companion  $N_B[t]$  is:

$$N_B[t] = [e^{+G_o}] \{ \exp(-Z_B[t]) \}, \quad [3.3]$$

while using  $Z[t] \rightarrow Z_B[t]$  in Eqs. [1.7a]-[1.7d] gives an *Enhanced Orthogonal Function (EOF)* model.

## 4 Finding $\{K_A, \gamma_o, \delta_o\}$ for $Z_B[t]$ from Data

If  $\delta_o = 0$ , the prior Eqs. [1.4a]-[1.5b] for  $\{K_A, \gamma_o, \delta_o = 0\}$  and  $\{t_I, t_F\}$  can be used, with the initial best-fit  $\{K_o, \alpha_S\}$  values determined by minimizing the *rms* error between Eq. [1.2] and the measured data on a *Logarithmic Y-axis*. Unfortunately, Eqs. [1.4a]-[1.5b] cannot be used when  $\delta_o \neq 0$ , although finding a good  $\{K_A, \gamma_o, \delta_o\}$  starting point is still needed for the *EIM*:

$$Z_B[t] = +[G_o / (1 + \gamma_o t)] \{ \exp[-\delta_o t / (1 + \kappa \gamma_o t)] \}, \quad [4.1a]$$

prior to any *EOF* analysis. The  $\kappa = 1$  case also has this special symmetry:

$$[\delta_o t / (1 + \gamma_o t)] = (\delta_o / \gamma_o) [1 - \frac{1}{(1 + \gamma_o t)}], \quad [4.2a]$$

$$\exp[-\delta_o t / (1 + \gamma_o t)] = e^{-(\delta_o / \gamma_o)} \exp[+(\delta_o / \gamma_o) \frac{1}{(1 + \gamma_o t)}], \quad [4.2b]$$

$$Z_B[t] = +[\frac{G_o}{(1+\gamma_o t)}] e^{-(\delta_o / \gamma_o)} \exp[+(\delta_o / \gamma_o) \frac{1}{(1+\gamma_o t)}], \quad [4.2c]$$

$$Z_B[t] = +G_o [\frac{(\delta_o / \gamma_o)}{(1+\gamma_o t)}] \frac{1}{(\delta_o / \gamma_o) e^{+(\delta_o / \gamma_o)}} \exp[+(\delta_o / \gamma_o) \frac{1}{(1+\gamma_o t)}], \quad [4.2d]$$

which can re-written as:

$$W(X[t]) \equiv X[t] \exp(+X[t]), \quad [4.3a]$$

$$X[t] = [(\delta_o / \gamma_o) / (1 + \gamma_o t)], \quad [4.3b]$$

$$X[0] = [\delta_o / \gamma_o], \quad [4.3c]$$

$$Z_B[t]_{\kappa=1} = +G_o W(X[t]) / W(X[0]). \quad [4.3d]$$

$$N_B(Z_B) = [e^{+G_o}] \exp(-Z_B[t]_{\kappa=1}), \quad [4.3e]$$

while the  $\kappa = 0$  case is:

$$Z_A[t] = +[G_o / (1 + \gamma_o t)] \exp(-\delta_o t), \quad [4.4a]$$

$$N_A(Z_A) = [e^{+G_o}] \exp(-Z_B[t]_{\kappa=0}) = [e^{+G_o}] \exp(-Z_A[t]). \quad [4.4b]$$

For  $\kappa = \{0, 1\}$ , the  $t = 0$  point,  $G_o$  from Eq. [2.1a], and the  $N_I(t_I)$  and  $N_F(t_F)$  initial and final points, give these equations to help set  $\{K_A, \gamma_o, \delta_o\}$ :

$$N_I[t = 0] = 1 = G_o \exp(-Z_B[t = 0]), \quad [4.5a]$$

$$N_I[t_I] = N_I = G_o \exp(-Z_B[t_I]), \quad [4.5b]$$

$$N_F[t_F] = N_F = G_o \exp(-Z_B[t_F]), \quad [4.5c]$$

for the EIM. Minimizing the *rms* error between the Eq. [3.3]  $\{K_A, \gamma_o, \delta_o, t_I\}$  functions and measured data on a *Logarithmic Y-axis* can be done as follows.

Start with estimated values for  $\{\hat{K}_A, \hat{\gamma}_o, \hat{\delta}_o, \hat{t}_{offset}\}$  in:

$$\hat{Z}[t] = +[\hat{G}_o / (1 + \hat{\gamma}_o t)] \{\exp[-\hat{\delta}_o t / (1 + \kappa \hat{\gamma}_o t)]\}, \quad [4.6a]$$

$$\hat{N}[t] = [e^{+\hat{G}_o}] [\exp(-\hat{Z}[t])], \quad [4.6b]$$

$$\hat{G}_o \equiv (\hat{K}_A / \hat{\gamma}_o), \quad [4.6c]$$

$$t = (t_{data} - \hat{t}_{offset}), \quad [4.6d]$$

where  $t_{data}$  is the data start time. Set a preliminary value for  $\hat{t}_{offset}$  first, to fix the time scale for the  $N_{data}[t]$  measured values:

$$N_I \equiv N_{data}[t_{data}^I - \hat{t}_{offset}], \quad [4.7a]$$

$$N_F \equiv N_{data}[t_{data}^F - \hat{t}_{offset}], \quad [4.7b]$$

$$(t_F - t_I) \equiv (t_{data}^F - t_{data}^I). \quad [4.7c]$$

Next, pick values for  $\{\hat{\gamma}_o, \hat{\delta}_o\}$  for  $\hat{N}[t]$  in Eq. [4.6b], allowing direct comparison between  $N[t]$  and  $\hat{N}[t]$  at each data point:

$$\hat{N}[t] \equiv \hat{N}[t = (t_{data} - \hat{t}_{offset})], \quad [4.8a]$$

$$N_{data}[t] \equiv N_{data}[t = (t_{data} - \hat{t}_{offset})]. \quad [4.8b]$$

The resulting calculated values for both  $\{\hat{N}[t_I], \hat{N}[t_F]\}$  can often be much too high or low, compared to the  $\{N_I, N_F\}$  measured data, but those values can be renormalized to:

$$\bar{N}[t] = \hat{N}[t] (N_I / \hat{N}[t_I]) \equiv S_I \hat{N}[t], \quad [4.9a]$$

$$\bar{N}[t_F] = \hat{N}[t_F] (N_I / \hat{N}[t_I]) \equiv S_I \hat{N}[t_F]. \quad [4.9b]$$

Here,  $S_I$  is the renormalization coefficient, and  $\bar{N}[t]$  allows easy comparison to the measured  $N_{data}[t]$  since  $\bar{N}[t_I] \equiv N_I$ . Given  $\{\hat{\gamma}_o, \hat{\delta}_o\}$ , the  $\hat{K}_A(\hat{\gamma}_o)$  value that is needed to obey  $\bar{N}[t_F] \rightarrow N_F$  can be set by using *Excel<sup>TM</sup> Goal-Seek* or its equivalent, which also sets a particular  $S_I$  value. Next, the  $\hat{\gamma}_o$  value is adjusted to find the specific  $\{K_A, \gamma_o\}$  parameter pair that gives  $S_I = 1$ . This process is needed because these  $\hat{\delta}_o \neq 0$  cases do not allow easy determination of  $t_I$  as in Eq. [1.4a], or for  $\{K_A, \gamma_o\}$ , given  $t_I$ , as in Eqs. [1.5a]-[1.5b].

The *rms* error on a *Logarithmic Y-axis*, between this  $\bar{N}[t; K_A, \gamma_o]$  and the  $N_{data}[t]$  is one of many  $\{K_A, \gamma_o, \hat{\delta}_o, \hat{t}_{offset}; S_I = 1\}$  choices. The minimum *rms* error among all these  $S_I = 1$  cases and the  $N_{data}[t]$ , when varying  $\{\hat{\delta}_o, \hat{t}_{offset}\}$  gives the best  $\{K_A, \gamma_o, \delta_o, t_{offset}\}$  values for Eqs. [4.3a]-[4.3e].

## 5 Enhanced Initial Model [EIM] Results for Italy

The Eq.[1.3a] *Initial Model* results were shown in **Figs. 1-2**. Mandatory *Social Distancing* was introduced at  $N_{data}^{3/10/2020} = 10,149$  which is the  $t_I$  data point, with  $N_{data}^{6/15/2020} = 237,290$  being the  $t_F$  data point. The *EIM* was then applied to the same data, to highlight the improvements that can be obtained from using the *EIM* of  $Z_A[t]$  and  $N_A(Z_A)$ , in place of  $Z_o[t]$  and  $N_o(Z_o)$ .

**Figures 3-4** show the resulting *EIM* best-fits for  $Z_A[t]$  and  $N_A(Z_A)$ , along with using a  $\rho_A[t]$  tail that is a pure exponential decay.

For the *EIM*, a new best estimate of  $t_{offset} = 2.866$  days was found within the  $N_{data}^{2/23/2020}$  (*Day 1*) to  $N_{data}^{3/10/2020}$  (*Day 17*) data, setting the *EIM*  $t = 0$  point. Then  $t_I = (17 - 2.866) = 14.134$  days, while  $N_{data}^{6/15/2020}$  (*Day 114*) gives  $t_F = (114 - 2.866) = 111.134$  days, along with:

$$N_{data}^{3/10/2020}[t_I = 14.134] = 10,149; \quad [5.1a]$$

$$N_{data}^{6/15/2020}[t_F = 111.134] = 237,290; \quad [5.1b]$$

$$N_{data}^{2/25/2020}[t = 0] = 1, \quad [5.1c]$$

$$(t_F - t_I) = 97 \text{ days}. \quad [5.1d]$$

This X-axis  $t = 0$  point is a hypothetical *EIM* pandemic starting point, if *Social Distancing* had been operating throughout the initial CoVID-19 period.

**Figure 3** has a predicted CoVID-19 pandemic peak of  $\sim 5,217 / \text{day}$  at  $t = 29.134$  days on 3/25/2020, with  $\sim 243,100$  total cases at the pandemic end. This datafit has  $> 4X$  error reduction over the  $\delta_o = 0$  case, as summarized next:

$K_A$	$\gamma_o$	$t_{dbl}, \text{days}$	$N[t \rightarrow \infty]$	$\delta_o$	$t_I, \text{days}$	$rms_{ERR}$
1.285436	.101062	.553060	$\sim 243,100$	.0336	14.134	0.023755
4.240513	.33078	.163458	$\sim 338,165$	0	7.899	0.097828

[5.2]

The *EIM* with  $Z_A[t]$  and  $\delta_o = 0.0336$  gives a  $\rho_A[t]$  curve that is in excellent agreement with the **Fig. 3** measured  $\rho_{data}[t]$  data. Comparing  $N_A[t]$  and  $N_{data}[t]$  in **Fig. 4** also shows an excellent match over the whole *Logarithmic Y-axis* range, used in the *rms* error minimization.

The **Figs. 3-4**  $\kappa = 0$  results were then compared to the  $\kappa = 1$  case, using  $Z_B[t]$  and  $N_B(Z_B)$  of Eqs. [3.2]-[3.3]. The  $\kappa = 1$  case has a *Social Distancing* factor that gradually turns off the *EIM* exponential decay. The resulting *rms* error best fits converged to  $\delta_o \rightarrow 0$ , as follows:

$\gamma_o$	$t_{dbl}$	$rms_{ERR}$	$\delta_o$	$t_{offset}^{days}$	$t_I^{days} :: t_F^{days}$
.07146	.775863	.2223783	.2000	10.0	7 :: 104
.18623	.295372	.1751635	.2000	11.0	6 :: 103
.30360	.181417	.2102839	.2000	12.0	5 :: 102

.39748	.137925	.1430275	.0500	11.0	6 :: 103
.43604	.125722	.1425737	.0135	11.0	6 :: 103
.44265	.123844	.1425524	.0070	11.0	6 :: 103
.44466	.123284	.1425480	.0050	11.0	6 :: 103
.44868	.122179	.1425445	.0010	11.0	6 :: 103
.44968	.121907	.1425442	.0000	11.0	6 :: 103 . [5.3]

This  $\kappa = 1$  result shows that the process giving rise to exponential tails in the  $\rho_{data}[t]$  for a pandemic *Fast Shutoff* is independent of the *Social Distancing* process that gradually lengthens the  $t_{dbl}$  *doubling time*, as measured by  $\gamma_o$ . Comparing the  $\kappa = \{0, 1\}$  cases shows that the pure exponential tail with  $Z_A[t] = Z_B[t]_{\kappa=0}$  matches the  $\rho_{data}[t]$  data best.

## 6 Enhanced Orthogonal Functions for Italy

Any monotonic  $Z[t]$  can convert measured  $N_{data}[t]$  data into  $N_{data}(Z)$ . Using the Eq. [2.6b]  $Z_A[t] \rightarrow Z[t]$  in Eqs. [1.7a]-[1.7d] extends the *EIM* into an *EOF* model, where Eq. [1.9] gives:

$$\rho[t] = R(Z_A[t]) \frac{dZ_A}{dt} = Z_A[t] R(Z_A[t]) \{ \delta_o + \frac{\gamma_o}{(1+\gamma_o t)} \}, \quad [6.1a]$$

$$Z_A[t] = +[G_o / (1 + \gamma_o t)] \exp(-\delta_o t). \quad [6.1b]$$

The Eq. [5.2]  $\delta_o \neq 0$  entries and Eq. [5.1a]-[5.1d] boundary conditions give:

$$Z_A^{\min}[t_F = 111.134] = 0.024226510, \quad [6.2a]$$

$$Z_A^{\max}[t_I = 14.134] = 3.176125728. \quad [6.2b]$$

These data-driven  $Z_A[t]$  limits are used next, along with  $Z[t] \rightarrow Z_A[t]$  in:

$$N(Z) = \sum_{m=0}^{m=M_F} g_m L_m(Z) \exp[-Z], \quad [6.3a]$$

$$g_n \equiv \sum_{m=0}^{m=M_F} g_m \int_{Z=0}^{Z=+\infty} L_n(Z) L_m(Z) \exp[-Z] dZ = \\ \int_{Z=Z_A^{\max}}^{Z=Z_A^{\min}} L_n(Z) N_{data}(Z) dZ + \\ \int_{Z=Z_A^{\max}}^{Z=+\infty} L_n(Z) N(Z) dZ + \int_{Z=0}^{Z=Z_A^{\min}} L_n(Z) N(Z) dZ, \quad [6.3b]$$

where  $m = \{0, M_F\}$  sets how many terms are in the Eq. [6.3a] series. Generally  $M_F = 2$  is used here. The  $L_m(Z)$  are the *Laguerre Polynomials*, with the first few  $L_m(Z)$  being:

$$L_{-1}(Z) \equiv 0, \quad [6.4a]$$

$$L_0(Z) \equiv 1 = L_m(Z = 0), \quad [6.4b]$$

$$L_1(Z) \equiv (1 - Z), \quad [6.4c]$$

$$L_2(Z) \equiv (1 - 2Z + \frac{1}{2}Z^2), \quad [6.4d]$$

$$L_3(Z) \equiv (1 - 3Z + \frac{3}{2}Z^2 - \frac{1}{6}Z^3), \quad [6.4e]$$

$$L_4(Z) \equiv (1 - 4Z + 3Z^2 - \frac{2}{3}Z^3 + \frac{1}{24}Z^4). \quad [6.4f]$$

Some important properties of the *Laguerre Polynomials* are:

$$\int_{Z=0}^{Z=+\infty} L_m(Z) L_n(Z) \exp(-Z) dZ = \mathbf{1} \delta_{m,n}, \quad [6.5a]$$

It is made available under a CC-BY 4.0 International license .

$$\delta_{m,n} = \begin{cases} 1 & \text{for } m=n \\ 0 & \text{otherwise} \end{cases}, \quad [6.5b]$$

$$\int_{Z'=Z}^{Z'=\infty} L_m(Z') \exp(-Z') dZ' = [L_m(Z) - L_{m-1}(Z)] \exp(-Z), \quad [6.5c]$$

$$L_m(Z) \exp(-Z) = \frac{1}{m!} \frac{d^m}{dZ^m} [Z^m e^{-Z}] = e^{-Z} \sum_{k=0}^{k=m} (-1)^k \frac{m!}{k!(m-k)!} \left[ \frac{Z^k}{k!} \right], \quad [6.5d]$$

$$L_{m \geq 2}(Z) = [2 - \frac{(Z+1)}{m}] L_{m-1}(Z) - [1 - \frac{1}{m}] L_{m-2}(Z). \quad [6.5e]$$

Here Eq. [6.5a] defines an *orthogonal function set*. The "**n!**" (**n-factorial**) in Eq. [6.5d], for **n** an integer, is defined as the product:

$$\mathbf{n}! \equiv (\mathbf{n})(\mathbf{n}-1)(\mathbf{n}-2)(\mathbf{n}-3)\dots(3)(2)(1), \quad [6.6a]$$

$$\mathbf{1}! \equiv \mathbf{0}! \equiv 1, \quad [6.6b]$$

where *factorials* with negative integers are not allowed. For  $M_F > 2$ , the following equations developed by Watson<sup>10</sup>, and improved by Gillis and Weiss<sup>11</sup>, helps in evaluating Eq. [6.3b]:

$$L_r(Z) L_s(Z) = \sum_{t=|r-s|}^{t=(r+s)} C_{rst} L_t(Z), \quad [6.7a]$$

$$C_{rst} = \int_{X=0}^{X=\infty} L_r(X) L_s(X) L_t(X) \exp(-X) dX, \quad [6.7b]$$

$$C_{rst} \equiv \frac{(-1)^p}{2^p} \sum_{n=0}^{n=(r+s)} (2^{2n}) \frac{(r+s-n)!}{(r-n)!(s-n)!(2n-p)!(p-n)!}, \quad [6.7c]$$

$$p \equiv (r + s - t), \quad [6.7d]$$

where ALL terms in the Eq. [6.7c] sum for  $n = \{0, (r+s)\}$  have an implicit requirement that all negative *factorial* arguments are excluded<sup>10-11</sup>.

Since the  $N(Z)$  of Eq. [6.3a] has  $\{g_m; m = (0, M_F)\}$ , and  $N(Z)$  also appears in each  $g_n$ -equation of Eq. [6.3b], how to determine each  $g_m$  by itself, can be done as follows. First define:

$$Q_n \equiv \int_{Z=Z_A^{\min}}^{Z=Z_A^{\max}} L_n(Z) N_{data}(Z) dZ, \quad [6.8a]$$

$$K_{m,n} \equiv \int_{Z=Z_A^{\min}}^{Z=Z_A^{\max}} L_m(Z) L_n(Z) \exp(-Z) dZ = K_{n,m}. \quad [6.8b]$$

When the  $N_{data}(Z)$  is comprised of  $j = \{1, 2, \dots, J\}$  discrete values between  $\{Z_A^{\min}, Z_A^{\max}\}$ , with each  $Z_j$  having an  $N_{data}^{(j)}(Z_j)$  value, the Eq. [6.8a] integral needs to be replaced by a sum. Let  $Z_0 = Z_1$  and  $Z_{J+1} = Z_J$ , the  $Q_n$  replacement for Eq. [6.8a] is then:

$$Q_n \equiv \sum_{j=1}^{j=J} L_n(Z_j) N_{data}^{(j)}(Z_j) \Delta_j, \quad [6.9a]$$

$$\Delta_j \equiv \frac{1}{2}|Z_{j+1} - Z_{j-1}|. \quad [6.9b]$$

Eq. [6.3b] can then be re-written as a  $3 \times 3$  matrix  $\underline{\mathbf{M}}_3$ , which relates a data-driven  $\vec{Q}_3$ -vector to a resultant  $\vec{g}_3$ -vector:

$$\vec{Q}_3 = \underline{\mathbf{M}}_3 \vec{g}_3, \quad [6.10a]$$

$$\begin{pmatrix} Q_o \\ Q_1 \\ Q_2 \end{pmatrix} = \begin{pmatrix} K_{0,0} & K_{0,1} & K_{0,2} \\ K_{1,0} & K_{1,1} & K_{1,2} \\ K_{2,0} & K_{2,1} & K_{2,2} \end{pmatrix} \begin{pmatrix} g_0 \\ g_1 \\ g_2 \end{pmatrix}. \quad [6.10b]$$

It is made available under a CC-BY 4.0 International license .

$$(\underline{\mathbf{M}}_3)^{-1} \vec{Q}_3 \equiv \vec{g}_3, \quad [6.10c]$$

where  $(\underline{\mathbf{M}}_3)^{-1}$  is the matrix inverse of  $\underline{\mathbf{M}}_3$ . When  $\{Z_A^{\min}, Z_A^{\max}\} \rightarrow \{0, +\infty\}$ , this  $\underline{\mathbf{M}}_3$  becomes the Identity Matrix. The following  $k_{m,n}(Z)$  integrals set  $K_{m,n}$ :

$$k_{m,n}(Z) = \int_{Z'=Z}^{Z'=+\infty} L_m(Z') L_n(Z') \exp(-Z') dZ' = k_{n,m}(Z), \quad [6.11a]$$

$$K_{m,n} \equiv k_{m,n}(Z_A^{\min}) - k_{m,n}(Z_A^{\max}) = K_{n,m}. \quad [6.11b]$$

The  $k_{m,n}(Z)$  integrals can be determined using Eq. [6.5c], which gives:

$$k_{0,0}(Z) = 1 \exp(-Z), \quad [6.12a]$$

$$k_{1,1}(Z) = \{1 + Z^2\} \exp(-Z), \quad [6.12b]$$

$$k_{2,2}(Z) = \{1 + 2Z^2 - Z^3 + \frac{1}{4}Z^4\} \exp(-Z), \quad [6.12c]$$

$$k_{0,1}(Z) = (-Z) \exp(-Z), \quad [6.12d]$$

$$k_{0,2}(Z) = (-Z) \{1 - \frac{1}{2}Z\} \exp(-Z), \quad [6.12e]$$

$$k_{1,2}(Z) = (-Z) \{1 - Z + \frac{1}{2}Z^2\} \exp(-Z). \quad [6.12f]$$

To extract  $\{g_0, g_1, g_2\}$ , the  $3 \times 3$  symmetric  $\underline{\mathbf{M}}_3$  matrix needs inversion:

$$\underline{\mathbf{M}} = \begin{pmatrix} a & d & f \\ d & b & e \\ f & e & c \end{pmatrix}, \quad [6.13a]$$

$$\det[\underline{\mathbf{M}}] \equiv [abc - ae^2 - bf^2 - cd^2 + 2def], \quad [6.13b]$$

$$\det[\underline{\mathbf{M}}] (\underline{\mathbf{M}})^{-1} \equiv \begin{pmatrix} [bc - e^2] & -(cd - ef) & -(bf - de) \\ -(cd - ef) & [ac - f^2] & -(ae - df) \\ -(bf - de) & -(ae - df) & [ab - d^2] \end{pmatrix}, \quad [6.13c]$$

which determines  $\{g_0, g_1, g_2\}$  from the  $\{Q_0, Q_1, Q_2\}$  data. A best-fit  $N(Z)$  for  $Z = \{0^+, \infty^-\}$  results, along with an equivalent fit for  $R(Z)$  using Eq. [1.7d].

Instead of having to find the best  $\{g_0, g_1, g_2\}$  triplet, one could find the best  $\{g'_0, g'_1\}$  by just using using  $\{Q_0, Q_1\}$  and an  $\underline{\mathbf{M}}_2$  sub-matrix; or one could find the best  $\{g_o^+\}$  by itself by just using  $\{Q_0\}$  and an  $\underline{\mathbf{M}}_1$  sub-matrix:

$$\vec{Q}_2 = \underline{\mathbf{M}}_2 \vec{g}_2, \quad [6.14a]$$

$$\begin{pmatrix} Q_o \\ Q_1 \end{pmatrix} = \begin{pmatrix} K_{0,0} & K_{0,1} \\ K_{1,0} & K_{1,1} \end{pmatrix} \begin{pmatrix} g'_0 \\ g'_1 \end{pmatrix}, \quad [6.14b]$$

$$\begin{pmatrix} g'_0 \\ g'_1 \end{pmatrix} = \begin{pmatrix} K_{0,0} & K_{0,1} \\ K_{1,0} & K_{1,1} \end{pmatrix}^{-1} \begin{pmatrix} Q_o \\ Q_1 \end{pmatrix} = \frac{1}{[K_{0,0} K_{1,1} - K_{0,1} K_{1,0}]} \begin{pmatrix} K_{1,1} & K_{1,0} \\ K_{0,1} & K_{0,0} \end{pmatrix} \begin{pmatrix} Q_o \\ Q_1 \end{pmatrix}, \quad [6.14c]$$

$$\vec{Q}_1 = \underline{\mathbf{M}}_1 \vec{g}_1, \quad [6.14d]$$

$$\begin{pmatrix} Q_o \end{pmatrix} = \begin{pmatrix} K_{0,0} \end{pmatrix} \begin{pmatrix} g_o^+ \end{pmatrix}, \quad [6.14e]$$

$$\begin{pmatrix} g_o^+ \end{pmatrix} = \begin{pmatrix} K_{0,0} \end{pmatrix}^{-1} \begin{pmatrix} Q_o \end{pmatrix}. \quad [6.14f]$$

Once the  $\{g_m; m = (0, M_F)\}$  constants are found and used in Eq. [1.8], its  $c_0$  value provides the new EOF estimate for the predicted total number of CoVID-19 cases at the pandemic end, refining the initial Eq. [5.2]  $N[t \rightarrow \infty]$  EIM value.

## 7 EOF Model Results for Italy

The *EOF* model starts with the *EIM* of Eq. [2.6b] using  $Z_A[t]$ , and the *bing.com* Italy data<sup>9</sup>, which gives:

$$N_{data}[t_I = 14.134] = 10,149; \quad [7.1a]$$

$$N_{data}[t_F = 111.134] = 237,290; \quad [7.1b]$$

$$N_A[t \rightarrow \infty] = 243,109; \quad [7.1c]$$

$$Z_A^{\min}[t_F = 111.134] = 0.024226510; \quad [7.1d]$$

$$Z_A^{\max}[t_I = 14.134] = 3.176125728; \quad [7.1e]$$

via Eqs. [5.1a]-[5.1d], [5.2], and [6.2a]-[6.2b]. For these  $\{Z_A^{\min}, Z_A^{\max}\}$  values, with  $(t_F - t_I) = 97$  days, the  $\underline{\mathbf{M}}_3$  matrix of  $K_{m,n}$  entries, via Eq. [6.8b], is:

$$\underline{\mathbf{M}}_3 = \begin{pmatrix} K_{0,0} & K_{0,1} & K_{0,2} \\ K_{1,0} & K_{1,1} & K_{1,2} \\ K_{2,0} & K_{2,1} & K_{2,2} \end{pmatrix} = \begin{pmatrix} 0.93432 & 0.10895 & -0.10133 \\ 0.10895 & 0.51376 & 0.35717 \\ -0.10133 & 0.35717 & 0.36868 \end{pmatrix} \quad [7.2]$$

It has a rather small  $\det[\underline{\mathbf{M}}_3] = 0.04024218$  value, with an inverse of:

$$(\underline{\mathbf{M}}_3)^{-1} \equiv \begin{pmatrix} 1.536\,745 & -1.897\,502 & 2.260\,643 \\ -1.897\,502 & 8.304\,570 & -8.566\,823 \\ 2.260\,643 & -8.566\,823 & 11.633\,092 \end{pmatrix}. \quad [7.3]$$

A convolution of  $L_m(Z_A)$  functions with the measured  $\vec{Q}_3$  dataset vector of Eqs. [6.9a]-[6.9b], along with the above  $(\underline{\mathbf{M}}_3)^{-1}$ , gives this final  $\vec{g}$ -vector<sup>12</sup>:

$$(\underline{\mathbf{M}}_3)^{-1} \vec{Q}_3 \equiv (\underline{\mathbf{M}}_3)^{-1} \begin{pmatrix} Q_o \\ Q_1 \\ Q_2 \end{pmatrix} = (\underline{\mathbf{M}}_3)^{-1} \begin{pmatrix} +226,767 \\ +26,978 \\ -22,399 \end{pmatrix} \equiv$$

$$\vec{g}_3 = \begin{pmatrix} g_o \\ g_1 \\ g_2 \end{pmatrix} = \begin{pmatrix} +246,656 \\ -14,362 \\ +20,954 \end{pmatrix}, \quad [7.4]$$

determining the constants for  $N(Z_A)$  in Eq. [1.7a]. The coefficients for  $R(Z_A)$ , which sets the predicted number of daily new CoVID-19 cases, are:

$$\vec{C}_3 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \vec{g}_3 = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} +253,248 \\ +6,592 \\ +20,954 \end{pmatrix}, \quad [7.5]$$

determining the constants needed for  $R(Z_A)$  in Eq. [1.7b]. Using these  $\{g_0, g_1, g_2\}$  values along with Eq. [1.8] gives:

$$N(Z_A \rightarrow 0) \equiv N[t \rightarrow \infty] = c_0 \equiv \{253,248\}, \quad [7.6]$$

as a new predicted total number of CoVID-19 cases at the pandemic end for the *EOF* model, which is a ~4.17% or 10,139 increase in the number of cases, compared to the *EIM* value of Eq. [7.1c].

Using Eq. [6.1b] for  $Z_A[t]$ , and substituting the Eq. [7.3]  $\vec{C}_3$  values into Eq. [1.7b] gives  $R(Z_A)$ . The  $\rho[t]$  in Eq. [6.1a] is derived from  $R(Z_A)$  using Eq. [1.7b], with the resulting *EOF*  $\rho[t]$  plotted in **Figure 5**, along with the  $t > t_I$  raw data for the daily new CoVID-19 cases.

The **Figure 5** *EOF* model also gives a  $t < t_I$  extrapolation, which shows what the combination of processes would look like, if they all had been operating continuously from the CoVID-19 pandemic start. The companion  $N[t]$  analytic

result, along with the  $t > t_I$  raw data for the total number of CoVID-19 cases is show in **Figure 6**.

Comparing the size and timing of the  $\rho[t]$  pandemic peak, and its Day 200 value, between the *EIM* (**Figs. 3-4**) and *EOF* model (**Figs. 5-6**), gives:

$$\begin{pmatrix} \text{Parameter} & \text{EIM} & \text{EOF Model} \\ N[t \rightarrow \infty] & 243, 109 & 253, 248 \\ \max\{\rho[t_p]\} & 5, 217 / \text{day} & 5, 142 / \text{day} \\ [t_p] \text{ Date} & 3/25/2020 & 3/27/2020 \\ \frac{200}{\text{Day}} \rho[t] & 6.67 / \text{day} & 7.59 / \text{day} \end{pmatrix}, \quad [7.7]$$

showing the *EOF* model predicts more cases total and more daily new CoVID-19 cases at Day 200, as well as modifying the pandemic peak predictions.

While the above analysis used  $M_F = 2$ , with the Eq. [7.4]  $\vec{g}_3$  setting the best  $\{g_0, g_1, g_2\}$  values, this *EOF* model also provides estimates for the simpler  $M_F = \{0, 1\}$  cases, as outlined by Eqs. [6.14a]-[6.14f]. For  $M_F = 1$ , the best two  $\{g'_0, g'_1\}$  values were gotten by only using  $\{Q_0, Q_1\}$  and an  $M_2$  sub-matrix of  $M_3$ . For  $M_F = 0$ , the best  $\{g_0^+\}$  by itself is derived by using  $\{Q_0\}$  and the  $M_1$  sub-matrix. These alternative estimates give:

$$\begin{aligned} \vec{g}'_2 &= \begin{pmatrix} g'_0 \\ g'_1 \end{pmatrix} = \begin{pmatrix} .93432 & .10895 \\ .10895 & .51376 \end{pmatrix}^{-1} \begin{pmatrix} 226, 767 \\ 26, 978 \end{pmatrix} = \begin{pmatrix} 242, 584 \\ 1, 069 \end{pmatrix} \\ \vec{g}'_1 &= (g_0^+) = (0.93432)^{-1} (+226, 767) = (242, 709). \quad [7.8a]-[7.8b] \end{aligned}$$

These additional calculations give the following progression of estimates for  $N[t \rightarrow \infty]$ , which is the final number of CoVID-19 cases at the pandemic end:

$$\begin{pmatrix} N_A[t \rightarrow \infty; \text{EIM}] \\ g_o^+ \\ c'_o = g'_0 + g'_1 \\ c_o = g_0 + g_1 + g_2 \end{pmatrix} = \begin{pmatrix} 243, 109 \\ 242, 709 \\ 243, 653 \\ 253, 248 \end{pmatrix}, \quad [7.9]$$

based on increasing the number of data fitting parameters used with the original data. This summary shows the  $N[t \rightarrow \infty]$  projections are fairly stable, with an average and  $1\sigma$  standard deviation:

$$\langle N[t \rightarrow \infty] \rangle = 245, 680 \pm 5, 060; \quad [7.10]$$

among these different calculations, where  $1\sigma$  is  $\sim 2.06\%$  of the overall average.

Comparing the results among **Figs. 1-6** also highlights these items:

(a) All  $\rho[t]$  functions have a sharp rise, and slower decreasing tail. The fastest changing  $\rho_{\text{data}}[t]$  tail, as in the Italy CoVID-19 (*Fast Shutoff*) data, was successfully modeled by adding in an exponential term, as in Eq. [6.1b].

(b) The datafits in **Fig. 4** and **Fig. 6** show that the extra parameters in the *EIM* and *EOF* model fits the  $\rho_{\text{data}}[t]$  shape progressively better.

(c) The *EOF* model shows only relatively small changes of  $\sim 2.06\%$  in the  $N[t \rightarrow \infty]$  limits (Eq. [7.10]), as an estimate of uncertainty in the *EIM*.

(d) The *Enhanced Initial Model (EIM)* function captures much of the progression to a pandemic *Fast Shutoff*, as seen in the Italy data.

The  $\rho[t]$  tail may still differ from these predictions, due to factors such as:

(i) The CoVID-19 dynamics may change in the long-term low  $\rho[t]$  regime;

(ii) A "second wave" or multiple waves of  $\rho[t]$  resurgence may occur, which are beyond the scope of this CoVID-19 pandemic modeling.

## 8 Summary and Conclusions

The early stages of the CoVID-19 coronavirus pandemic starts off with a nearly exponential rise in the number of infections with time. Defining  $N[t]$  as the expected total number of CoVID-19 cases vs time, this basic function:

$$N_o[t] = \mathbf{1} \exp[+K_A t / (1 + \gamma_o t)] = \exp[+G_o] \exp[-Z_o], \quad [8.1a]$$

$$Z_o[t] \equiv +[G_o / (1 + \gamma_o t)], \quad G_o \equiv [K_A / \gamma_o], \quad [8.1b]$$

$$\rho_o[t] \equiv dN_o[t] / dt, \quad [8.1c]$$

models *Social Distancing* effects as a gradual lengthening of the pandemic growth *doubling time*, which enables pandemic shutoff with only a small population of infected persons. The Eq. [8.1b]  $Z_o[t]$  was our *Initial Model*<sup>1</sup>, and gives a CoVID-19 *Slow Shutoff* with a long-term  $\rho_o[t] \sim [1/t^2]$  tail. Previously we showed<sup>1-2</sup> that this  $Z_o[t]$  model fits many  $N_{data}[t]$  and  $\rho_{data}[t]$  cases.

However, some data had a CoVID-19 *Fast Shutoff*, with a  $\rho_{data}[t] \sim [\exp(-\delta_o t)]$  exponential tail, such as in Italy<sup>9</sup>, where a Gaussian tail  $\rho[t] \sim [\exp(-q_o t^2)]$  would have decreased too quickly. An *Enhanced Initial Model (EIM)* was developed here, using this  $Z_A[t]$  function:

$$N_A[t] \approx \mathbf{1} \exp[+G_o] \exp(-Z_A[t]), \quad [8.2a]$$

$$Z_A[t] \equiv +[G_o / (1 + \gamma_o t)] \exp(-\delta_o t), \quad [8.2b]$$

$$G_o \equiv [K_A / \gamma_o], \quad [8.2c]$$

$$\rho_A[t] \equiv dN_A[t] / dt. \quad [8.2d]$$

We also examined if the  $\exp(-\delta_o t)$  exponential decay could also be subject to a *Slow Shutoff*, giving  $\exp[-\delta_o t / (1 + \gamma_o t)]$  instead of  $\exp(-\delta_o t)$ , but that did not match the Italy data. To allow more data fitting parameters beyond just  $\{K_A, \gamma_o, \delta_o\}$ , an orthogonal function method was developed<sup>2</sup>:

$$N(Z) = \sum_{m=0}^{m=M_F} g_m L_m(Z) \exp[-Z], \quad [8.3a]$$

$$R(Z) = \sum_{m=0}^{m=M_F} c_m L_m(Z) \exp[-Z], \quad [8.3b]$$

$$N(Z) \equiv \int_{Z'=+\infty}^{Z=Z} R(Z') dZ', \quad [8.3c]$$

$$c_{M_F-k} = \sum_{m=0}^{m=k} g_m, \quad [8.3d]$$

which is applicable to a generic  $Z[t]$  function, with  $N[t] = N(Z[t])$ , where  $Z[t] \rightarrow Z_o[t]$  and  $Z[t] \rightarrow Z_A[t]$  are special cases. Larger  $M_F$  with more  $\{L_m(Z); m = (0, +M_F)\}$  terms can match almost any *arbitrary function*, enabling fits to a variety of  $N[t]$  and  $\rho[t]$  shapes. The  $\{g_m; m = (0, +M_F)\}$  are constants determined from each dataset. The  $L_m(Z)$  are the *Laguerre Polynomials*, with several important properties given in Eqs. [6.4a]-[6.5e].

Using  $Z_A[t]$  in Eqs. [8.3a]-[8.3d] results in this *Enhanced Orthogonal Function [EOF]* model, which is applicable to both *Slow* or *Fast Shutoff* CoVID-19 pandemic data. The  $\rho[t]$  expected number of daily new CoVID-19 cases is:

$$N[t] \equiv \int_{t'=(-1/\gamma_o)}^{t'=t} \rho[t'] dt', \quad [8.4a]$$

$$\rho[t] = R(Z_A[t]) \frac{dZ_A}{dt} = R(Z_A[t]) \left[ \frac{G_o}{(1+\gamma_o t)} \right] \left\{ \delta_o + \frac{\gamma_o}{(1+\gamma_o t)} \right\} [\exp(-\delta_o t)]. \quad [8.4b]$$

Methods were developed to derive the  $\{K_A, \gamma_o, \delta_o\}$  values, and to determine the  $\{g_m; m = (0, +M_F)\}$  and  $\{c_m; m = (0, +M_F)\}$  constants from data. Whereas our *Initial Model* and *EIM* were  $M_F = 0$  cases, the  $M_F = 2$  case was used here to examine the Italy CoVID-19 data, as an *EOF* model example.

The *bing.com* data for Italy up to  $\sim 6/15/2020$  was then analyzed, with **Figures 3-6** giving the new Italy results. Both the *EIM* and the *EOF* model provided good datafits, giving similar  $N[t \rightarrow \infty]$  results for the final number of CoVID-19 pandemic cases, differing by only  $\sim 2\%$  at the  $1\sigma$  level.

The  $\rho[t]$  post-peak behavior best indicates if a  $\delta_o \neq 0$  model (CoVID-19 pandemic *Fast Shutoff*) is applicable. The  $\delta_o \neq 0$  case likely is a second *Social Distancing* process, that operates along with, but is independent of the gradual pandemic *doubling time* changes. That *doubling time* change gives rise to a CoVID-19 pandemic *Slow Shutoff* ( $\gamma_o \neq 0$ ), and that process still operates concurrently with the  $\delta_o \neq 0$  CoVID-19 pandemic *Fast Shutoff*.

This analysis shows a wide variety of CoVID-19 data can be modeled using  $\{K_A, \gamma_o, \delta_o, t_{offset}\}$  as parameters, covering: (I) an exponential rise at CoVID-19 pandemic start; (II) a gradual lengthening of *doubling times* for a pandemic *Slow Shutoff*; and (III) an exponential decay for pandemic *Fast Shutoffs*.

## 9 List of Figures

**Figure 1:** Updated Initial Model for ITALY, CoVID-19 data to 6/15/2020. Number of daily CoVID-19 cases calculated as if *Social Distancing* started at 3/2/2020, but only data from 3/10/20 actual *Social Distancing* start, with  $N = 10,149$  cases, was used in calculations.

**Figure 2:** Comparison of Initial Model Best-Fit to Measured Data. Best-fit done on *Logarithmic Y-axis*, using data from 3/10/20 actual *Social Distancing* start with  $N = 10,149$  cases, through 6/15/20 with  $N = 237,290$ .

**Figure 3:** Enhanced Initial Model (EIM) for ITALY CoVID-19 data to 6/15/20. EIM best fit with  $N_A[t] \sim \exp(-Z_A[t])$  using enhanced  $Z_A[t]$  function having an exponential decay. Adding in exponential decay term gives significantly improved fit, compared to prior OFM.

**Figure 4:** Enhanced Initial Model (EIM) datafit for ITALY CoVID-19 data to 6/15/20. EIM best fit with  $N_A[t] \sim \exp(-Z_A[t])$  using enhanced  $Z_A[t]$  function having an exponential decay. Datafit minimizing *rms* error on *Logarithmic Y-axis* gives significant improvement vs OFM.

**Figure 5:** Enhanced Orthogonal Functions (EOF) for ITALY CoVID-19 data to 6/15/20. Orthogonal functions give additional parameters for further datafit improvement, using a 3-term series ( $M_F = 2$ ). Result shows EIM, by itself, provides most of the improvement.

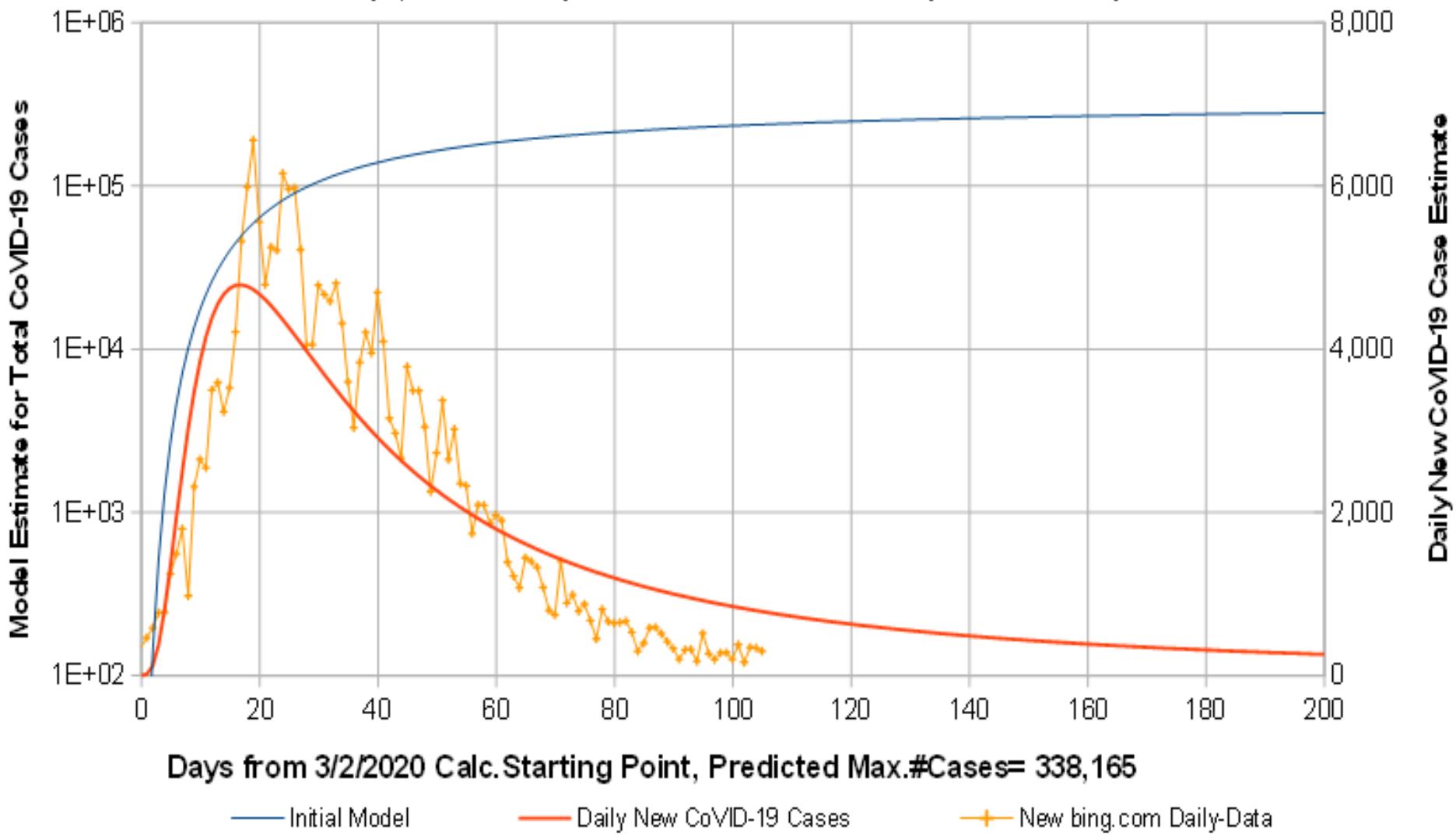
**Figure 6:** Enhanced Orthogonal Functions (EOF) for ITALY CoVID-19 data to 6/15/20. Data for total number of CoVID-19 cases versus time, compared with 3-term EOF model for  $N(Z_A[t])$  shows excellent match after *Social Distancing* start at  $N = 10,149$ .

## 10 References

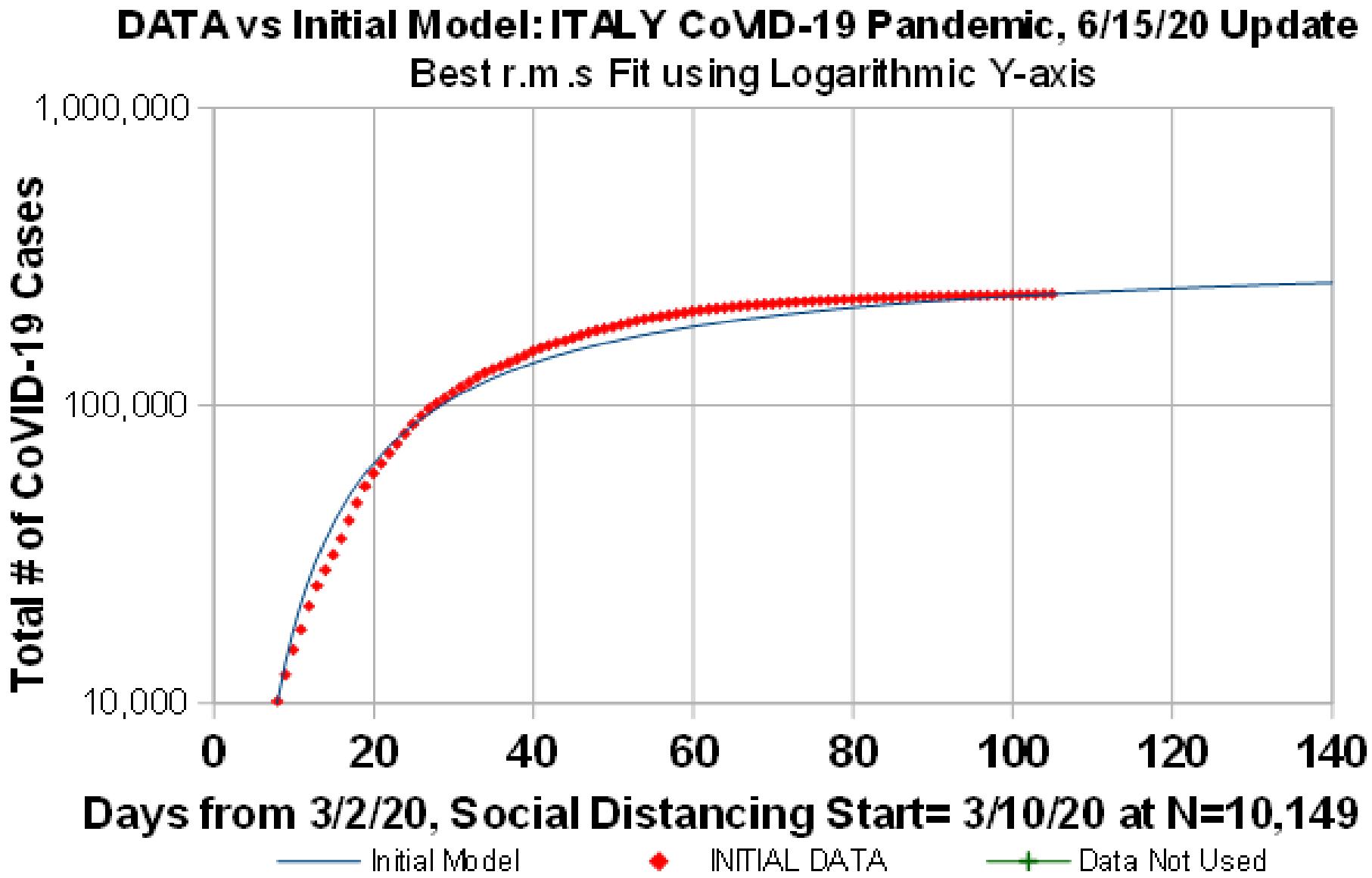
1. <https://www.MedRxiv.org/content/10.1101/2020.05.04.20091207v1>,  
<https://doi.org/10.1101/2020.05.04.20091207>, "Initial Model for the Impact of Social Distancing on CoVID-19 Spread", Gengmun Eng.
2. <https://www.MedRxiv.org/content/10.1101/2020.06.30.20143149v1>,  
<https://doi.org/10.1101/2020.06.30.20143149>, "Orthogonal Functions for Evaluating Social Distancing Impact on Covid-19 Spread", Gengmun Eng
3. <https://www.medrxiv.org/content/10.1101/2020.03.27.20043752v1>,  
"Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months", IHME COVID-19 Health Service Utilization Forecasting Team.
4. <https://www.geekwire.com/2020/univ-washington-epidemiologists-predict-80000-covid-19-deaths-u-s-july>  
"Univ. of Washington researchers predict 80,000 COVID-19 deaths in U.S. by July", Alan Boyle, GeekWire, March 26, 2020.
5. <https://www.yahoo.com/finance/news/coronavirus-modelers-raise-projected-u-041641553.html>, "Coronavirus modelers raise projected U.S. death toll and lengthen state-by-state recovery timeline", Alan Boyle, GeekWire, April 27, 2020.
6. <https://covid19.healthdata.org>, update of 29 April 2020.
7. <http://www.healthdata.org/covid/updates>  
"COVID-19: What's New for May 4, 2020: Updated IHME COVID-19 projections: Predicting the Next Phase of the Epidemic", IHME COVID-19 Health Service Utilization Forecasting Team.
8. <https://finance.yahoo.com/news/pandemic-projection-puts-u-death-220824741.html>,  
"New pandemic projection puts U.S. death toll at nearly 135,000, due to less social distancing", Alan Boyle, GeekWire, May 4, 2020.
9. [www.bing.com/covid](http://www.bing.com/covid): 'Bing COVID-19 Tracker', and  
<https://www.bing.com/covid?form=CPVD07>.
10. G. N. Watson, "A Note of the Polynomials of Hermite and Laguerre", Journal of the London Mathematical Society, **13**(1938), pp. 29-32.
11. J. Gillis and G. Weiss, "Products of Laguerre Polynomials", Math. Comput., **14**(69), Jan. 1960, pp. 60-63
12. For correct results, calculations done by *swp55.exe* requires **no** commas in the integer matrix or vector entries, when using: *ScientificWorkplace<sup>TM</sup>\_Compute\_EvaluateNumerically*. It is due to *swp55.exe* also accepting European notation where  $10,149 \Leftrightarrow 10.149$ .

## Updated Initial Model for ITALY CoVID-19 Pandemic, 6/15/20 Update; Best Fit on Logarithmic Y-axis

$K/A=4.2405/\text{day}$ ,  $\gamma/\delta=.33078/\text{day}$ ,  $\delta/\alpha=0.0$ ; Pred.Max.= 4,787/day at 16.89945 days on 3/19/2020

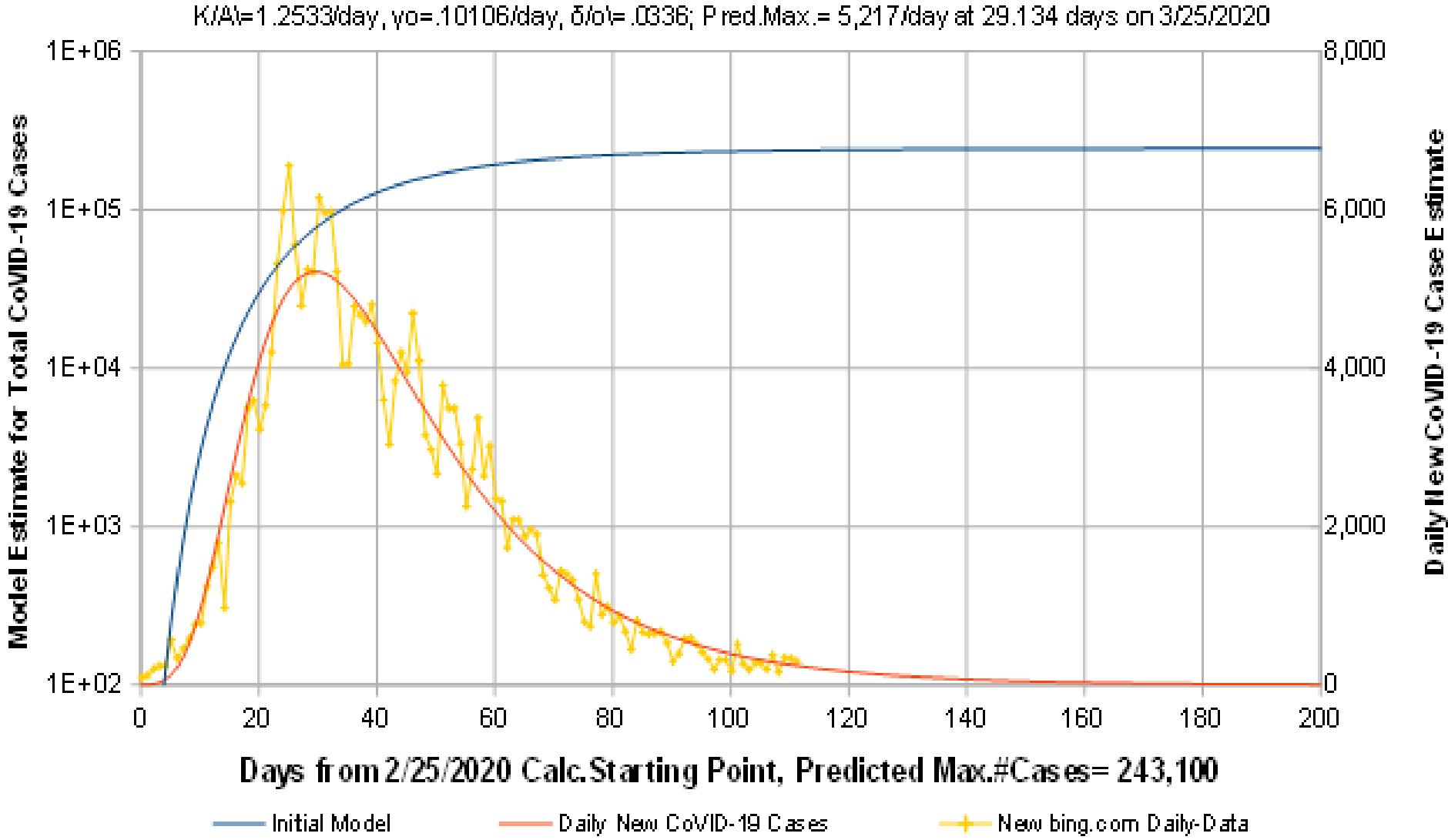


**Figure 1: Updated *Initial Model* for ITALY, CoVID-19 data to 6/15/2020.** Number of daily CoVID-19 cases calculated as if Social Distancing started at 3/2/2020, but only data from 3/10/20 actual *Social Distancing* start, with **N=10,149** cases, was used in calculations.



**Figure 2: Comparison of *Initial Model* Best-Fit to Measured Data.** Best-fit done on Logarithmic Y-axis, using data from 3/10/20 actual Social Distancing start with **N=10,149** cases, through 6/15/20 with **N=237,290**.

### Enhanced Initial Model (EIM): ITALY CoVID-19 Pandemic, 6/15/20 Update; Best Fit on Logarithmic Y-axis

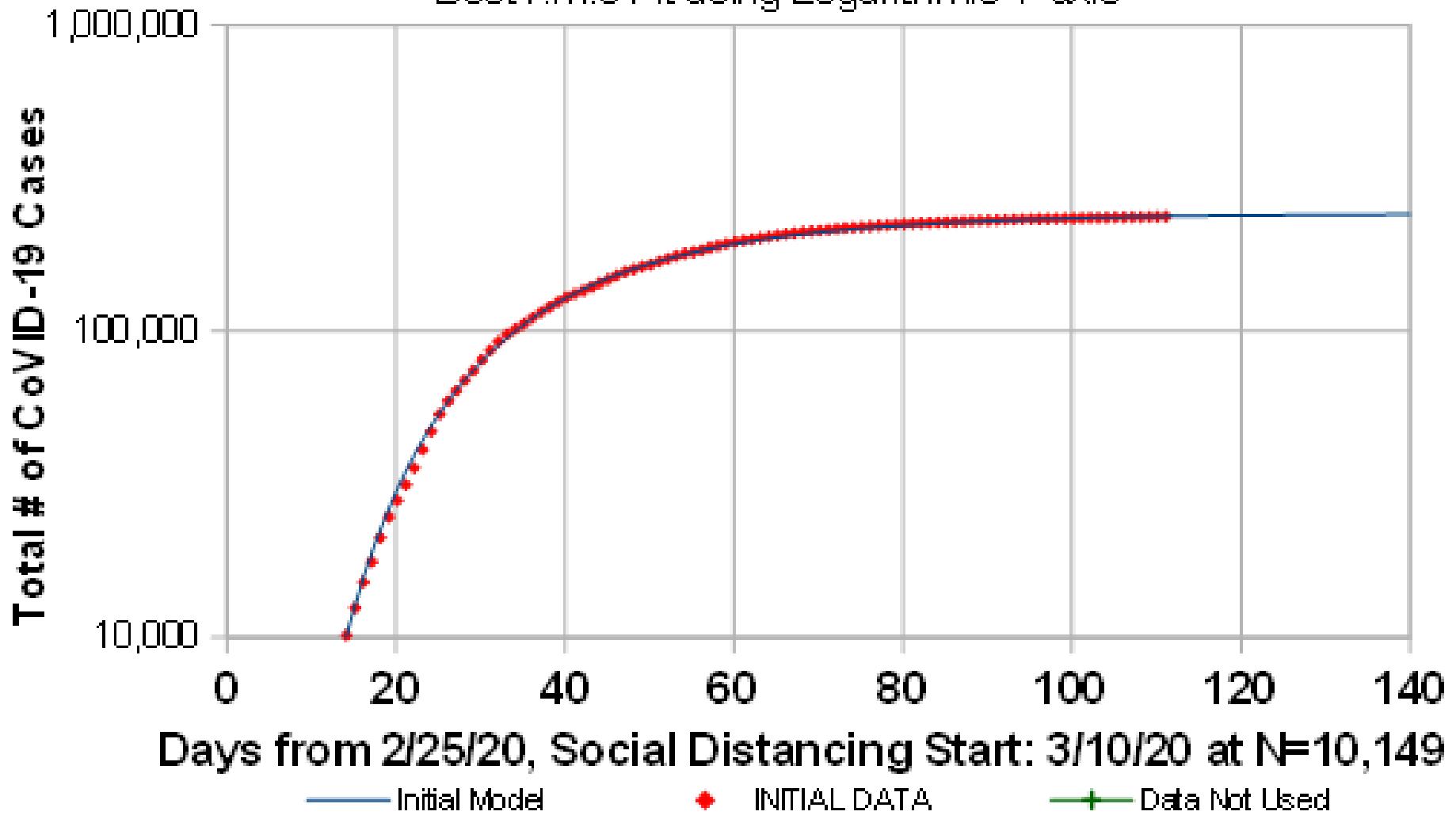


**Figure 3: Enhanced Initial Model (EIM) for ITALY CoVID-19 data to 6/15/20.**

EIM best fit with  $NA[t] \sim \exp(-ZA[t])$  using enhanced  $ZA[t]$  function having an exponential decay. Adding in exponential decay term gives significantly improved fit, compared to prior OFM.

## DATA vs Initial Model: ITALY CoVID-19 Pandemic, 6/15/20 Update

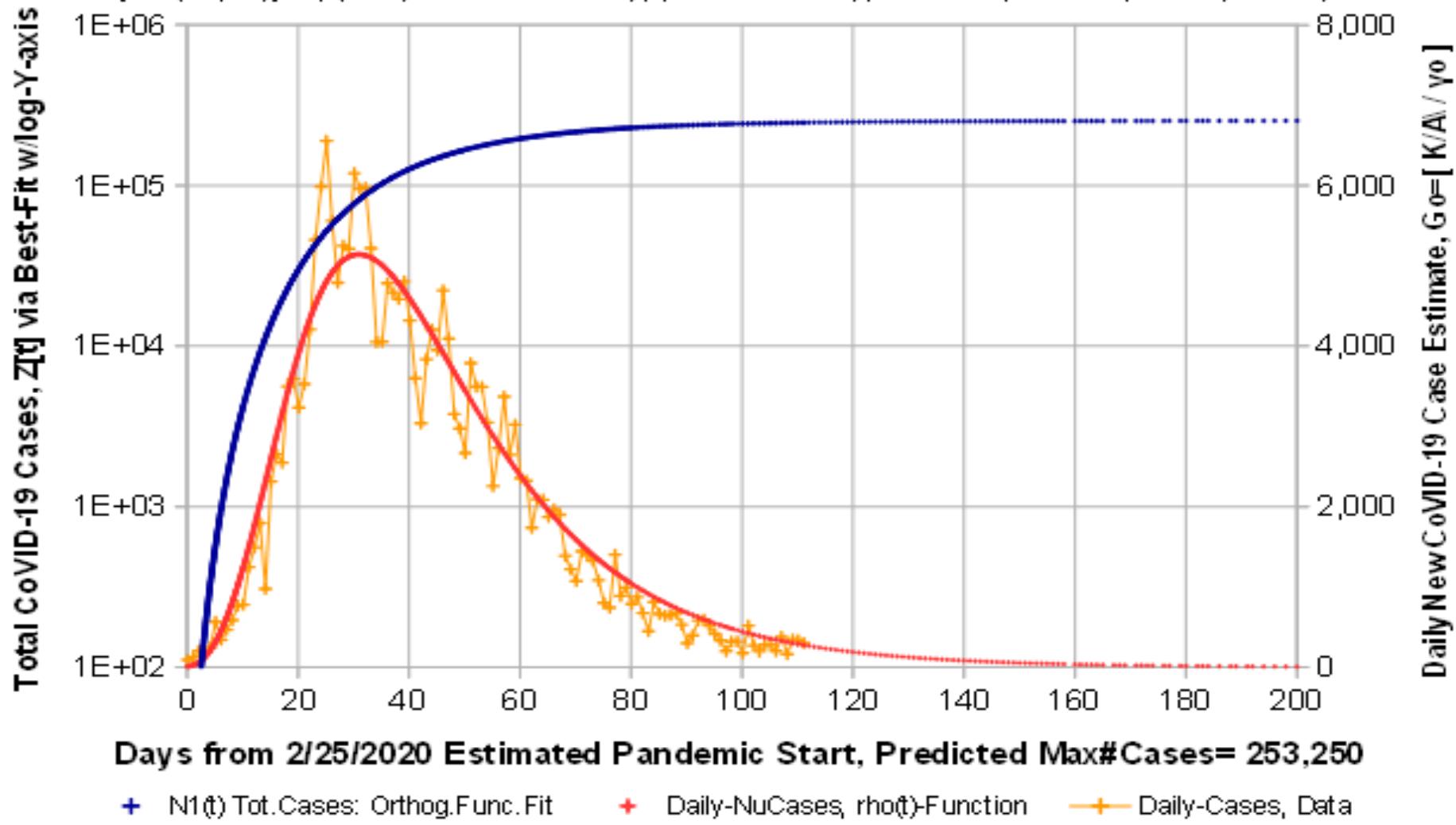
Best r.m.s Fit using Logarithmic Y-axis



**Figure 4: Enhanced Initial Model (EIM) datafit for ITALY CoVID-19 data to 6/15/20.**  
EIM best fit with  $NA[t] \sim \exp(-ZA[t])$  using enhanced  $ZA[t]$  function having an exponential decay.  
Datafit minimizing *rms* error on *Logarithmic* Y-axis gives significant improvement vs OFM.

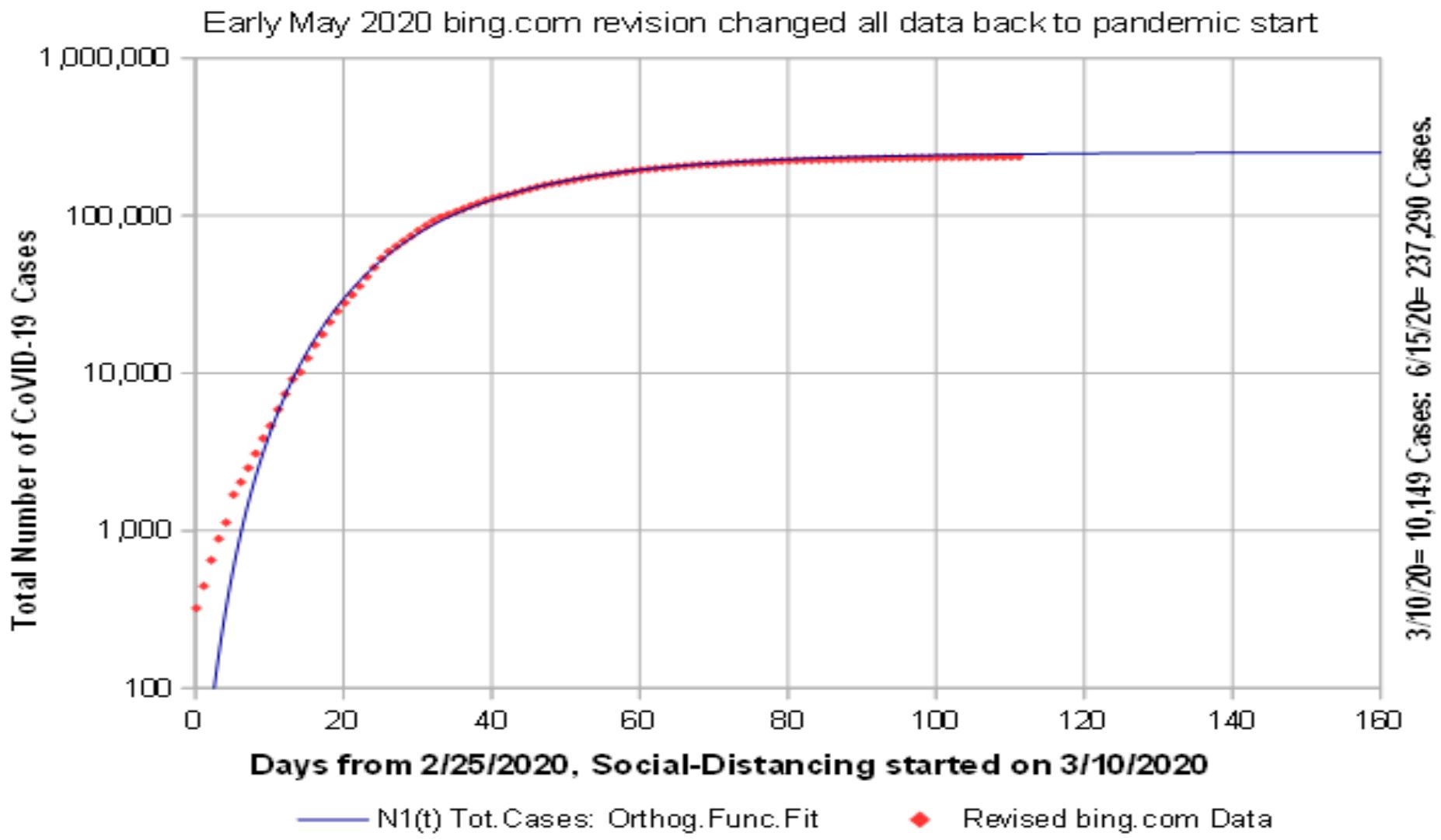
**Enhanced Orthog.Func. [EOF] Model: ITALY CoVID-19 Pandemic, 6/15/20 Update: Calc.Peak= 3/27/2020**

$Z = [G_0 / (1 + \gamma_0 * t)] \exp(-\delta_0 t)$ :  $K/A = 1.2533/\text{day}$ ,  $\gamma_0 = 0.10106/\text{day}$ ,  $\delta_0 = .0336$ , Peak= 5,142/day on Day 30.95



**Figure 5: Enhanced Orthogonal Functions (EOF) for ITALY CoVID-19 data to 6/15/20.**  
Orthogonal functions give additional parameters for datafit improvement, using a 3-term series (**MF=2**). Result shows *EIM*, by itself, provides most of the improvement.

## Enhanced Orthog.Func.[EOF] Model: ITALY CoVID-19 Pandemic, 6/15/20 Update



**Figure 6: Enhanced Orthogonal Functions (EOF) for ITALY CoVID-19 data to 6/15/20.**  
Data for total number of CoVID-19 cases versus time, compared with 3-term EOF model for  $N(ZA[t])$  shows excellent match after Social Distancing start at  **$N=10,149$** .