AMSI VACATION RESEARCH SCHOLARSHIPS 2017-2018

Parameter Estimation for the New Two-term Fractional Order Nonlinear Dengue Fever Epidemic Model

Steven Kedda

Supervised by:

Prof. Fawang Liu & Dr. Tianzeng Li

Queensland University of Technology

Vacation Research Scholarships are funded jointly by the Department of Education and Training and the Australian Mathematical Sciences Institute.

Australian Government
Department of Education and Training

AMSI AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

Introduction

Fractional calculus techniques have proved abundantly useful in recent years. Fractional order models have been developed for a large variety of problems where integer order derivatives have insufficient capability to provide accurate agreement between simulated and real data. In this paper, we are concerned with a mathematical model to provide insight into an outbreak of dengue fever in the Cape Verde Islands, Africa, in 2009.

We will propose modelling techniques including the superposition of fractional order derivative terms, parameter estimation techniques to solve the inverse problem, and numerical methods to solve nonlinear systems. Previous studies have considered both fractional order models and parameter estimation techniques, but have not considered the two term fractional order model for an SIR dengue fever model.

Fractional Order Derivative Definitions

In this study, we propose the use of the Caputo definition of the fractional derivative. Another two commonly used fractional derivative definitions are the Grunwald-Letnikov and the Riemann-Liouville definitions.

Caputo derivative

$$ {}_a^C D_t^\alpha f(t) = {}_0D_t^{-(n-\alpha)} \frac{d^n}{dt^n} f(t) = \frac{1}{\Gamma(n-\alpha)} \int_a^t (t-\tau)^{n-\alpha-1} f^{(n)}(\tau) d\tau, $$

the Riemann-Liouville derivative

$$ {}_a^{RL} D_t^\alpha f(t) = \frac{d^n}{dt^n} {}_aD_t^{-(n-\alpha)} f(t) = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \int_a^t (t-\tau)^{n-\alpha-1} f(\tau) d\tau, $$

And the Grunwald-Letnikov derivative

$$ {}_a^{GL} D_t^\alpha f(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0) t^{k-\alpha}}{\Gamma(k+1-\alpha)} + \frac{1}{\Gamma(n-\alpha)} \cdot \int_a^t (t-\tau)^{n-\alpha-1} f^{(n)}(\tau) d\tau $$

We may define the Caputo derivative in terms of the Riemann-Liouville definition in the following way

$$ {}_{a}^{C}D_t^{\alpha} f(t) = {}_{a}^{RL}D_t^{\alpha} f(t) - \sum_{k=0}^{n-1} \frac{(t-a)^{k-\alpha} f^{(k)}(a)}{\Gamma(k-\alpha+1)}. $$

By letting

$$ h(t) = f(t) - \sum_{k=0}^{n-1} \frac{(t-a)^{k} f^{(k)}(a)}{k!}, $$

we may let

$$ {}_{a}^{C}D_t^{\alpha} h(t) = {}_{a}^{RL}D_t^{\alpha} h(t) = {}_{a}^{RL}D_t^{\alpha} \left(f(t) - \sum_{k=0}^{n-1} \frac{(t-a)^{k} f^{(k)}(a)}{k!} \right). $$

This may simplified down to the following expression

$$ {}_{a}^{C}D_t^{\alpha} h(t) = {}_{a}^{C}D_t^{\alpha} f(t). $$

We will consider the Caputo derivate, since it may be combined with classical initial conditions. The Riemann-Liouville derivative, for example, is not suitable to be combined with classical initial conditions.

Integer Order Model

The original model used to model this epidemic was the integer order SIR (susceptible, infected, recovered) model. This is a system of coupled differential equations in time. Of the 5 equations, there exists susceptible, infected, and recovered humans; and susceptible and infected mosquitos. The dynamical system is as described below

$$
\begin{aligned}
\frac{dS_h}{dt} &= \mu_h(N_h - S_h) - \frac{\beta_h b}{N_h + m} S_h I_m, \\
\frac{dI_h}{dt} &= \frac{\beta_h b}{N_h + m} S_h I_m - (\mu_h + \gamma) I_h, \\
\frac{dR_h}{dt} &= \gamma I_h - \mu_h R_h, \\
\frac{dS_m}{dt} &= \mu_m(N_m - S_m) - \frac{\beta_m b}{N_h} S_m I_h, \\
\frac{dI_m}{dt} &= \frac{\beta_m b}{N_h + m} S_m I_h - \mu_m I_m.
\end{aligned}
$$

Where $S_h, I_h,$ and R_h are susceptible, infected, and recovered humans, respectively. S_m and I_m are susceptible and infected mosquitos, respectively. μ_h is the mortality rate of humans, μ_m is the mortality rate for mosquitos, γ is the recovery rate of humans, b is the biting rate of mosquitos, β_m is the chance of transmission from a human to a mosquito, β_h is for mosquito to human, and lastly m represents the number of blood sources other than humans.

From Nishiura, H. (2006) and Diethelm, K. (2013), the parameters have been selected to be

$$\beta_h = 0.36, \beta = 0.36, b = 0.7, \gamma = \frac{1}{3}, \mu_m = \frac{1}{10}, \mu = \frac{1}{71 * 365}, m = 0.$$

With initial conditions

$$S_h(0) = 55784, I_h(0) = 216, R_h(0) = 0, S_m(0) = 168000, I_m(0) = 0$$

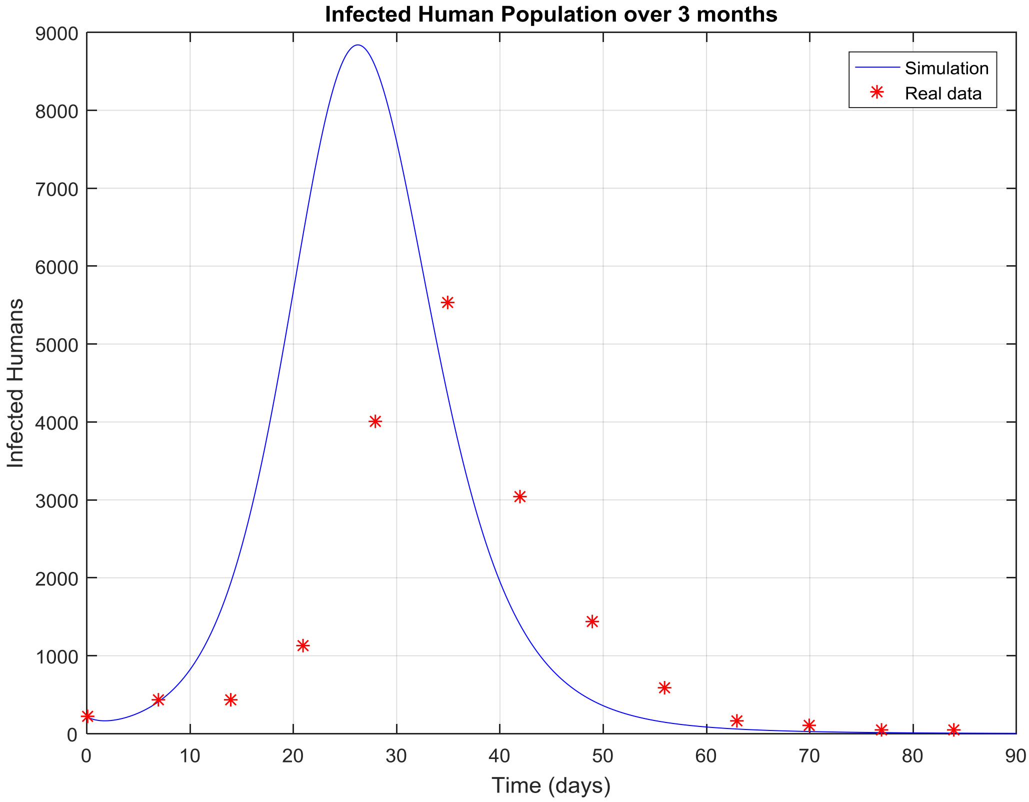

Figure 1: Integer Order Model. Simulation overlayed on real data.

This integer order simulated model provides a poor fit to the measured data. We now consider replacing the integer order time derivate terms with fractional order derivative terms.

Fractional Order Model

$$
\begin{aligned}
{}_0^C D_t^{\alpha_1} S_h &= \mu_h(N_h - S_h) - \frac{\beta_h b}{N_h + m} S_h I_m,\\
{}_0^C D_t^{\alpha_2} I_h &= \frac{\beta_h b}{N_h + m} S_h I_m - (\mu_h + \gamma) I_h,\\
{}_0^C D_t^{\alpha_3} R_h &= \gamma I_h - \mu_h R_h,\\
{}_0^C D_t^{\alpha_4} S_m &= \mu_m(N_m - S_m) - \frac{\beta_m b}{N_h} S_m I_h,\\
{}_0^C D_t^{\alpha_5} I_m &= \frac{\beta_m b}{N_h + m} S_m I_h - \mu_m I_m,
\end{aligned}
$$

This model now allows for extra flexibility, where the fractional order of each term may now be adjusted to a non-integer value. Diethelm (2013) proposes the following choices for the fractional order $\alpha_h = \alpha_1 = \alpha_2 = \alpha_3$ and $\alpha_m = \alpha_4 = \alpha_5$, where $\alpha_h = 1$ and $\alpha_m = 0.77$. The result for this specific choice of the alphas reduces the root-mean-square value down to $MSE = 762$. This exhibits a major improvement from the integer model, however noticeable error in figure 2 encourages further exploration into the choice of the fractional order terms and the 6 other parameters in the model. Let's discuss the numerical solver and propose some parameter estimation techniques.

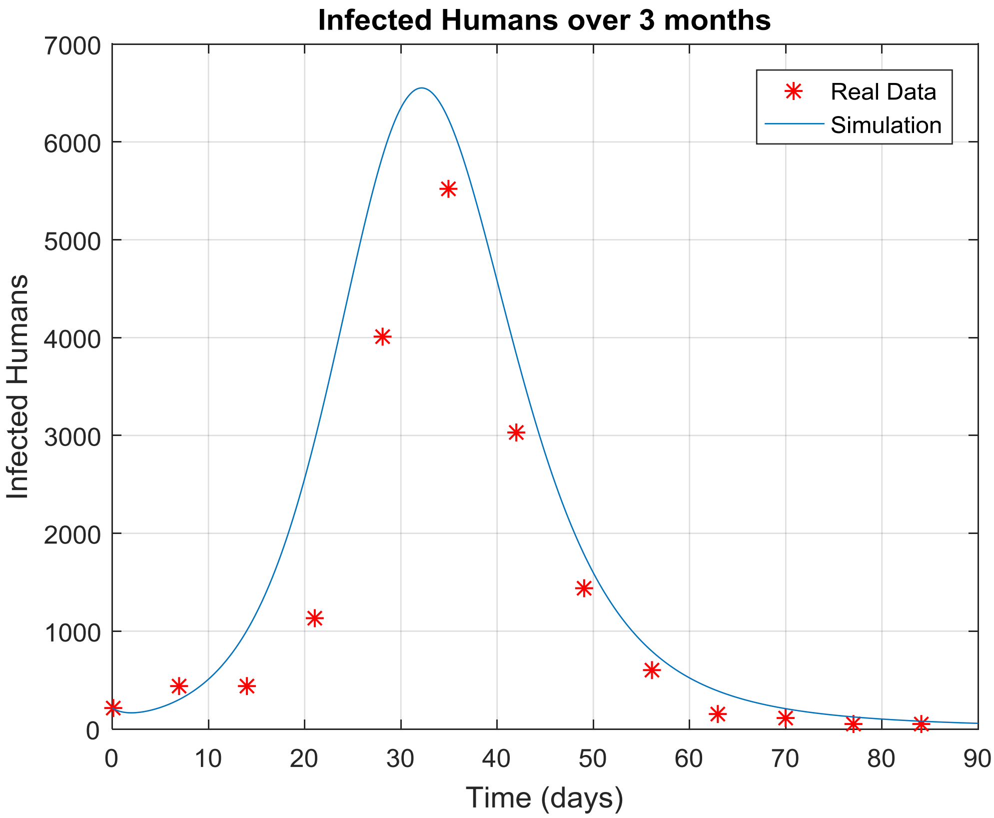

Figure 2: Fractional Order Model

Numerical solver for the Fraction Order Model

Since the system in question is nonlinear, Newton-GMRES has been chosen to solve the system numerically. Compared to other conventional numerical solvers, using the GMMP (Goronflo Mainardi Moretti Paradisi) scheme and Newton method is very time efficient.

We generate a uniform grid $t_j = a + jh$, $j = 0,1,2, \dots, N$, $Nh = t - a$ and $\alpha > 0$.

The Caputo derivative is approximated as

$$ {}_a^C D_t^\alpha f(t) \approx \frac{1}{h^\alpha} \sum_{k=0}^{N} c_k^\alpha \left(f(t_{N-k}) - \sum_{j=0}^{n-1} \frac{(t-a)^j f^{(j)}(a)}{j!} \right). $$

Where $c_k^\alpha = (-1)^k \binom{\alpha}{j}$.

Consider the general case for our fractional order nonlinear system

$$ {}_{a}^{C}D_{t}^{\alpha}x(t) = f(t, x(t)) $$

Where the right hand side is the right hand side in one of the 5 equations in the dynamical system. The Caputo derivative is now submitted into this expression to receive

$$ \sum_{k=0}^{N} c_k^{\alpha} \left(x(t_{N-k}) - \sum_{j=0}^{n-1} \frac{(t-a)^j x^{(j)}(a)}{j!} \right) = h^{\alpha} f(t_N, x(t_N)) $$

Now by bringing $x(t_N)$ to the left hand side

$$ x(t_N) = h^{\alpha} f(t_N, x(t_N)) + \sum_{j=0}^{n-1} \frac{(t-a)^j x^{(j)}(a)}{j!} - \sum_{k=1}^{N} c_k^{\alpha} \left(x(t_{N-k}) - \sum_{j=0}^{n-1} \frac{(t-a)^j x^{(j)}(a)}{j!} \right). $$

However when $0 < \alpha < 1$, the expression simplifies to

$$ x(t_N) = h^{\alpha} f(t_N, x(t_N)) + x(a) - \sum_{k=1}^{N} c_k^{\alpha} \left(x(t_{N-k}) - x(a) \right). $$

When we plough through the algebra for each equation, we generate the solution equations

$$ S_h(t_N) = h^{\alpha_1} f\big(t_N, S_h(t_N)\big) + S_h(a) - \sum_{k=1}^{N} c_k^{\alpha_1} \left[S_h(t_{N-k}) - S_h(a) \right] $$

$$ I_h(t_N) = h^{\alpha_2} f\big(t_N, I_h(t_N)\big) + I_h(a) - \sum_{k=1}^{N} c_k^{\alpha_2} \left[I_h(t_{N-k}) - I_h(a) \right] $$

$$ R_h(t_N) = h^{\alpha_3} f\big(t_N, R_h(t_N)\big) + R_h(a) - \sum_{k=1}^{N} c_k^{\alpha_3} \left[R_h(t_{N-k}) - R_h(a) \right] $$

$$ S_m(t_N) = h^{\alpha_4} f\big(t_N, S_m(t_N)\big) + S_m(a) - \sum_{k=1}^{N} c_k^{\alpha_4} \left[S_m(t_{N-k}) - S_m(a) \right] $$

$$ I_m(t_N) = h^{\alpha_5} f\big(t_N, I_m(t_N)\big) + I_m(a) - \sum_{k=1}^{N} c_k^{\alpha_5} \left[I_m(t_{N-k}) - I_m(a) \right] $$

The Newton-GMRES method is

$$ x_{n+1} = x_n - J_F(x_n)^{-1} F(x_n), n = 0, 1, 2, \cdots , $$

$J_F(x_n)^{-1}$ is the Jacobian matrix. This method was iterated over 1001 nodes across the 90 day domain, to provide a sufficiently smooth and accurate solution.

Inverse Problem Parameter Estimation Techniques Single Term Model

To solve the inverse problem for this study, both the Nelder-Mead simplex search (NMSS) and the particle swarm optimisation (PSO) were used. Both of these techniques have different traits when it comes to estimating parameters. For example in the NMSS method, the initial points are predefined and the method moves the parameter away from points with poorer function values. The PSO method has a set of randomly chosen points and moves towards points with better function values. The Modified Hybrid Nelder-Mead simplex search and particle swarm optimisation (MH-NMSS-PSO) was adapted from these methods. The steps of this method are as follows

Define the function MSE as g

$$g(P^*) = \min_{P \in D} g(P) = \min_{p \in D} \sqrt{\frac{\sum_{j=0}^{N}(x(t_j) - x_j)^2}{N+1}}$$

1. Create a population of size 3m + 1. Form a simplex of m dimensions
   $$P_i = (p_{1,i}, p_{2,i}, \cdots, p_{m,i}) \in D, i = 1, 2, \cdots, m+1$$

Where

$$p_{j,i} = p_j^{(min)} + (i-1) \times (p_j^{(max)} - p_j^{(min)})/(m+1), j = 1, 2, \cdots, m, i = 1, 2, \cdots, m+1.$$

A pair of particles are created from the PSO method

$$P_i = (p_{1,i}, p_{2,i}, \cdots, p_{m,i}) \in D, i = m+2, \cdots, 3m+1$$

Where

$$p_{j,i} = p_j^{(min)} + Rand \times (p_j^{(max)} - p_j^{(min)}), j = 1, 2, \cdots, m, i = m+2, \cdots, 3m+1.$$

Lastly their velocities are calculated.

$$V_{j,i} = (V_j^{(max)} - V_j^{(min)})/L_j, j = 1, 2, \cdots, m, i = m+2, \cdots, 3m+1,$$

2. Evaluate g at each particle P. Order them from smallest to largest

$$g(P_1) \leq g(P_2) \leq \cdots \leq g(P_{3m+1}).$$

3. Incorporate the NMSS, calculate centre of gravity

$$P_o = (p_{1,o}, p_{2,o}, \cdots, p_{m,o}) \in D,$$

Where

$$p_{j,O} = \frac{\sum_{i=1}^{m} p_{j,i}}{m}, j = 1, 2, \cdots, m.$$

Calculate

$$P_r = (1+\alpha)P_o - \alpha P_{m+1},$$

Where $\alpha > 0$, recommended that $\alpha = 1$.

Case 1: If $g(P_1) \leq g(P_r) \leq g(P_m)$ then $P_{m+1} = P_r$

Case 2: If $g(P_r) \leq g(P_1)$ then compute

$$P_e = \gamma P_r + (1-\gamma)P_o.$$

Where $\gamma = 2$. If $g(P_e) \leq g(P_1)$, $P_{m+1} = P_e$. Else $P_{m+1} = P_r$

Case 3: If $g(P_r) \geq g(P_m)$ and $g(P_r) \leq g(P_{m+1})$, $P_{m+1} = P_r$ Compute

$$P_c = \beta P_{m+1} + (1-\beta)P_c.$$

If $g(P_c) \leq g(P_{m+1})$, $P_{m+1} = P_r$ else let

$$P_i = \sigma P_i + (1-\sigma)P_1, i = 1, 2, \cdots, m+1.$$

Choose $\beta = 0.5$ and $\sigma = 0.5$

4. Incorporate PSO. Update $2m$ particles with the poorest MSE function value.
5. If $S_c < \varepsilon$, stop. Otherwise return to Step 2.

$$S_c = \left[\sum_{i=1}^{m+1} \frac{(\overline{g} - \sqrt{g_i})^2}{m+1}\right]^{\frac{1}{2}}$$

Where $\overline{g} = \sum_{i=1}^{m+1} \frac{g_i^*}{m+1}$ and $g_i^* = \sqrt{g_i} = \sqrt{g_i(p_1, p_2, \cdots, P_m)}$.

From this parameter estimation method, the model now fits much better, as seen in figure 3. The MSE of this new model is $MSE = 249$, which is a great improvement from the model without parameter estimation, showing that solving the inverse problem can provide a much better model.

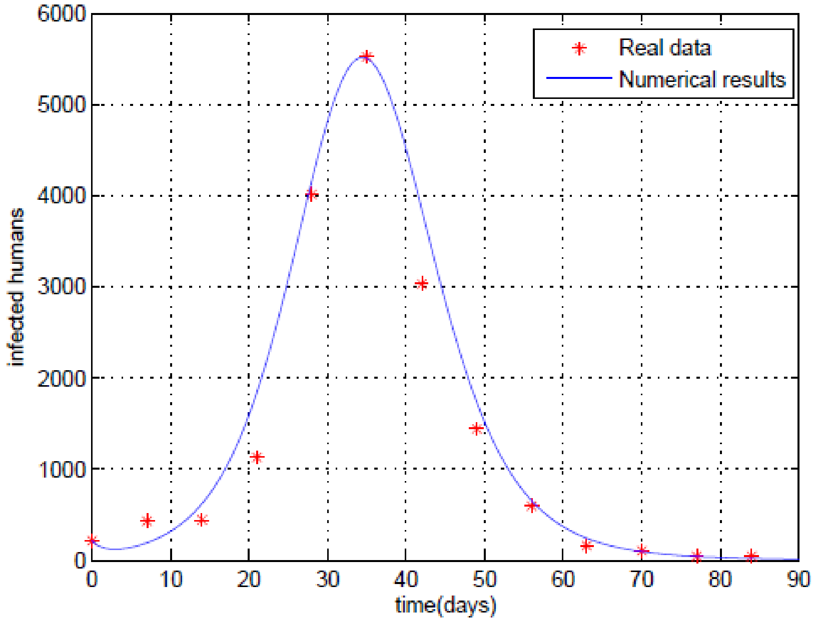

Figure 3: Fractional order model with estimated parameters.

Inverse Problem Parameter Estimation Techniques Two Term Model

We propose the two term fractional order model as follows

$$\lambda_1 {}_a^C D_t^{\alpha_1} x(t) + \lambda_2 {}_a^C D_t^{\alpha_2} x(t) = f\left(t_N, x(t_N)\right)$$

Apply our approximated definition for the Caputo derivative

$$\frac{\lambda_1}{h^{\alpha_1}} \sum_{k=0}^{N} c_k^{\alpha_1} \left[x(t_{N-k}) - x(a)\right] + \frac{\lambda_2}{h^{\alpha_2}} \sum_{k=0}^{N} c_k^{\alpha_2} \left[x(t_{N-k}) - x(a)\right] = f\left(t_N, x(t_N)\right)$$

Emit a $k = 0$ term and solve for $x(t_N)$:

$$\left[\frac{\lambda_1}{h^{\alpha_1}} + \frac{\lambda_2}{h^{\alpha_2}}\right] \left[x(t_N) - x(a)\right] + \frac{\lambda_1}{h^{\alpha_1}} \sum_{k=1}^{N} c_k^{\alpha_1} \left[x(t_{N-k}) - x(a)\right] + \frac{\lambda_2}{h^{\alpha_2}} \sum_{k=1}^{N} c_k^{\alpha_2} \left[x(t_{N-k}) - x(a)\right] = f\left(t_N, x(t_N)\right)$$

Simplify by letting

$$\kappa_1 = \left[\frac{\lambda_1}{h^{\alpha_1}} + \frac{\lambda_2}{h^{\alpha_2}}\right]$$

$$B_1 = \frac{\lambda_1}{h^{\alpha_1}} \sum_{k=1}^{N} c_k^{\alpha_1} \left[x(t_{N-k}) - x(a)\right]$$

$$B_2 = \frac{\lambda_2}{h^{\alpha_2}} \sum_{k=1}^{N} c_k^{\alpha_2} \left[x(t_{N-k}) - x(a)\right]$$

Thus, the simplified and rearranged expression is given as

$$x(t_N) = x(0) + \frac{1}{\kappa}\Big[f\big(t_N, x(t_N)\big) - (B_1 + B_2)\Big].$$

From here, the GMMP scheme and Newton-GMRES method can be applied to obtain a numerical solution for x. So the new two term fractional order dynamical system is

$$\begin{aligned}
\lambda_1{}_a^C D_t^{\alpha_1} S_h + \lambda_2{}_a^C D_t^{\alpha_2} S_h &= \mu_h(N_h - S_h) - \frac{\beta_h b}{N_h + m} S_h I_m \\
\lambda_3{}_a^C D_t^{\alpha_3} I_h + \lambda_4{}_a^C D_t^{\alpha_4} I_h &= \frac{\beta_h b}{N_h + m} S_h I_m - (\mu_h + \gamma) I_h \\
\lambda_5{}_a^C D_t^{\alpha_5} R_h + \lambda_6{}_a^C D_t^{\alpha_6} R_h &= \gamma I_h - \mu_h R_h \\
\lambda_7{}_a^C D_t^{\alpha_7} S_m + \lambda_8{}_a^C D_t^{\alpha_8} S_m &= \mu_m(N_m - S_m) - \frac{\beta_m b}{N_h} S_m I_h \\
\lambda_9{}_a^C D_t^{\alpha_9} I_m + \lambda_{10}{}_a^C D_t^{\alpha_{10}} I_m &= \frac{\beta_m b}{N_h + m} S_m I_h - \mu_m I_m
\end{aligned}$$

Following the same methodology from the single term counterpart, we can numerically solve this two term fractional order nonlinear system and its 26 parameters (10 alphas, 10 lambdas, 6 others) to produce the following model with $MSE = 128$.

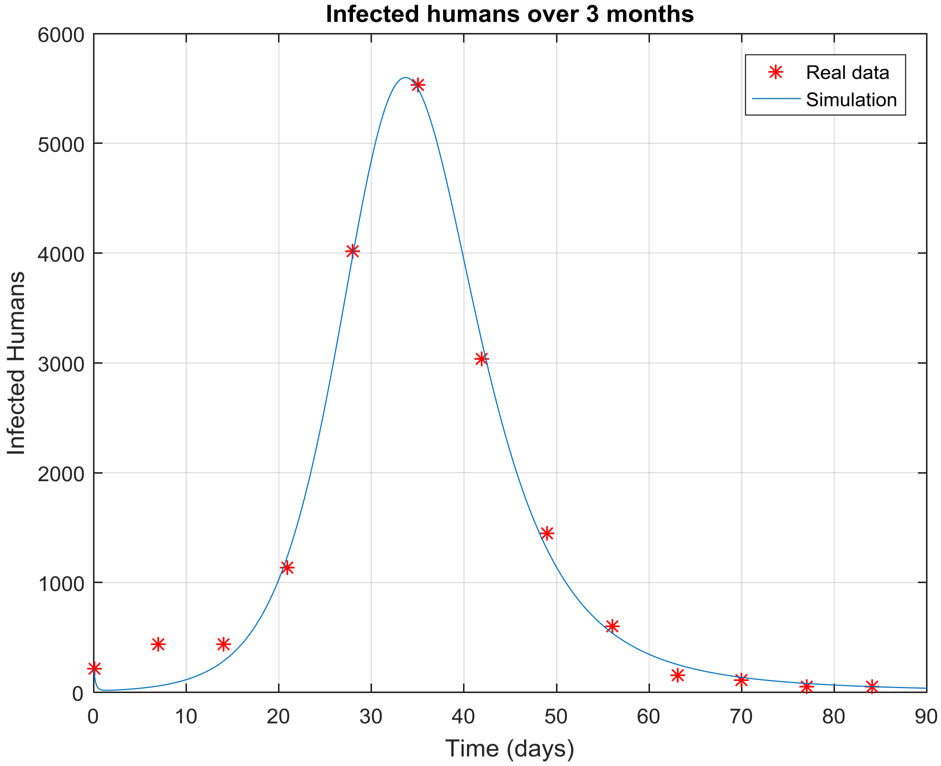

Figure 4: Plot of the two term fractional order model with parameter estimation.

Conclusion

The final result of the two term technique shows a drastic drop in the MSE test, demonstrating great effectiveness in modelling this epidemic using the two term fractional order model with parameter estimation. However, there is noticeable error in the first 2 weeks of this model. This is the main contribution towards the error exhibited in the MSE. However, despite this poor fitting in the section, the rest of the model agrees very well with the real data, showing this model's effectiveness. Perturbations in each parameter was attempted, whilst keeping all other parameters fixed, but the only significant change to the model was in the middle region, between days 20 and 60. So this contribution of error could not be reduced in this two term model.

For the purpose of using this model to make predictions on the behaviour of this disease, the single term model in figure 3 would suffice. Despite its greater error than the two term model, it models the first 2 weeks of the data better than the two term model in figure 4, and also sufficiently models the entire span of the real data.

This paper has proposed a new two-term method to achieve more accurate model simulations for a case of dengue fever epidemic data from Cape Verde Islands in 2009.

Pre-existing methods have been used and built upon to further develop the effectiveness and efficiency for mathematical modelling of epidemics. The comparison of the computed model to the real data from Cape Verde 2009 is evidence to suggest that the two term fractional order model, with parameter estimation, can be used to study the effects of different underlying mechanisms and make better predictions about the behaviour of large scale spread of dengue fever. In addition, this method may be used to model epidemics beyond dengue fever, including other dynamical systems that have the SIR structure. Future work in this field could see the two term model having more than two fractional order terms, however computationally expensive that may be. Though it could not be well predicted if a 3 term model would provide better results than a 2 term model. As for most modelling exercises, acquiring more data and modelling other epidemics would prove the versatility of this two term approach to modelling.

References

1. Nishiura, H. (2006). Mathematical and statistical analyses of the spread of dengue.
2. Diethelm, K. (2013). A fractional calculus based model for the simulation of an outbreak of dengue fever. Nonlinear Dynamics, 71(4), 613-619.