The Member Choice prediction scheme applied to a 2-strain host-vector SIR model for dengue with an SNA

Paul Krause
UFSC - Departamento de Matemática
Campus Trindade
88040-900, Florianópolis, SC
E-mail: p.krause@ufsc.br

Abstract: A non-statistical prediction scheme viable for high-dimensional nonlinear systems is presented. The scheme, called Member Choice (MC), consists of detecting the lowest-error state sample emerging from a data assimilation process and making a prediction from it. The lowest-error sample is detected by measuring the adherence of each sample to the process through the magnitude of state perturbations computed in every time step on each sample. When applied to a 2-strain host-vector SIR model for dengue with a strange nonchaotic attractor (SNA), MC shows a much higher predictive skill than straightforward Monte Carlo for any sample set size.

Keywords: *epidemiology, dengue, SIR, host-vector dynamics, SNA, determining variables, prediction*

1 Introduction

With strongly nonlinear dynamic systems, statistical methods for data assimilation and prediction have difficulty providing a reliable sample statistics, especially for high dimensions, and the average estimate tends to be a rare or inconsistent event (i.e. lie outside the sample space) [9]. For such systems, it is advisable to find an alternative to statistical prediction methods. This work presents a non-statistical prediction scheme viable for high-dimensional nonlinear systems. The scheme, called Member Choice (MC), consists of detecting the lowest-error state sample emerging from a data assimilation process and making a prediction from it. The lowest-error sample is detected by measuring the adherence of each sample to the process through the magnitude of state perturbations computed in every time step on each sample (but not applied to the samples unless dynamical stabilization is sought, which further requires that the computation of perturbations be iterated in each time step). To showcase its performance, MC is applied to a 2-strain host-vector SIR model for dengue (D2) that closely follows the model presented in [10], which stems from [1]. Model D2 is presented in section 2. MC is applied to D2 in section 4. In section 3, D2 is shown to have a strange nonchaotic attractor (SNA) [8, 2]. SNA are common in quasi-periodically forced nonlinear systems. They lie at the transition between order and chaos. Unlike chaos, they can occur in dimensions 1 and 2: SNA is the maximum complexity in these dimensions. When variables are controlled by forcing, the attractor is a smooth surface. In SNA, variables are not fully controlled by forcing, the attractor is non-smooth (which shows that strange attractors can arise before chaos) and the largest Lyapunov exponent is zero. SNA occurs when the maximum Lyapunov exponent is zero but intermittent bursts of the partial derivatives of the state variables with respect to the forcing phase show up in trajectories at periods of positive local Lyapunov exponents. The phase sensitivity function $\Gamma(t)$ provides a general characterization of this feature over all trajectories in the attractor and allows to distinguish smooth attractors from strange ones: $\Gamma(t)$ saturates for smooth attractors and grows unbounded for strange ones.

2 Model

In this section, the 2-strain host-vector SIR model for dengue (D2) is presented.

Variables:

S	susceptible individuals never exposed to infection
I_1,I_2	first-time infected individuals with strain type 1, or type 2
R_1,R_2	temporarily immune individuals recovered from first-infection with strain type 1, or type 2
S_1,S_2	susceptible individuals to strain type 2, or type 1, recovered from first-infection with strain type 1, or type 2
I_{12},I_{21}	cross-infected individuals with strain types 1 and 2
R	life-long immune individuals recovered from cross-infection
S_v	susceptible mosquitoes never exposed to any dengue infection
V_1,V_2	infected mosquitoes with strain type 1, or type 2

Parameters:

N	human population
$M(t)$	mosquito population
μ	human birth and death rates
$\xi(t)$	mosquito birth rate
ν	mosquito death rate
β	mosquito to human transmission rate
θ	human to mosquito transmission rate
ϕ	transmission amplification or attenuation factor
α	temporary cross-immunity rate
γ	recovery rate

Following [10], D2 is set as:

$$\dot{S} = -(\beta/M)S(V_1 + V_2) + \mu(N - S) \tag{1}$$
$$\dot{I}_1 = (\beta/M)SV_1 - (\gamma + \mu)I_1 \tag{2}$$
$$\dot{I}_2 = (\beta/M)SV_2 - (\gamma + \mu)I_2 \tag{3}$$
$$\dot{R}_1 = \gamma I_1 - (\alpha + \mu)R_1 \tag{4}$$
$$\dot{R}_2 = \gamma I_2 - (\alpha + \mu)R_2 \tag{5}$$
$$\dot{S}_1 = -(\beta/M)S_1V_2 + \alpha R_1 - \mu S_1 \tag{6}$$
$$\dot{S}_2 = -(\beta/M)S_2V_1 + \alpha R_2 - \mu S_2 \tag{7}$$
$$\dot{I}_{12} = (\beta/M)S_1V_2 - (\gamma + \mu)I_{12} \tag{8}$$
$$\dot{I}_{21} = (\beta/M)S_2V_1 - (\gamma + \mu)I_{21} \tag{9}$$
$$\dot{R} = \gamma(I_{12} + I_{21}) - \mu R \tag{10}$$
$$\dot{S}_v = -(\theta/N)S_v(I_1 + I_2 + \phi(I_{12} + I_{21})) + \xi M - \nu S_v \tag{11}$$
$$\dot{V}_1 = (\theta/N)S_v(I_1 + \phi I_{21}) - \nu V_1 \tag{12}$$
$$\dot{V}_2 = (\theta/N)S_v(I_2 + \phi I_{12}) - \nu V_2 \tag{13}$$

where

$$N = N(0) = S + I_1 + I_2 + R_1 + R_2 + S_1 + S_2 + I_{12} + I_{21} + R \tag{14}$$
$$M = M(t) = S_v + V_1 + V_2 \tag{15}$$

In this work, differently from [10], the mosquito birth rate ξ is set to

$$\xi = \xi(t) = \nu(1 + a\cos(\omega t)), \tag{16}$$

to account for seasonality, and parameter values to

$$\mu = 1/65 \text{ year}^{-1}, \nu = 36.5 \text{ year}^{-1}, \beta = 2\gamma \text{ year}^{-1}, \theta = 2\nu \text{ year}^{-1}, \tag{17}$$

$$\phi = 0.8, \alpha = 2 \text{ year}^{-1}, \gamma = 52 \text{ year}^{-1}, a = 0.4, \omega = 12\pi. \tag{18}$$

3 Strange nonchaotic attractor

In this section, D2 set with 17-18 is shown to have a strange nonchaotic attractor (SNA) [8, 2].

In SNA the largest Lyapunov exponent is zero and the phase sensitivity function $\Gamma(t)$ grows unbounded (in particular it scales with t^λ, where λ is the phase exponent).

Figure 1 shows the evolution of the time-averaged local Lyapunov exponents in a trajectory of D2 set with 17-18. One sees that the largest Lyapunov exponent of D2 is zero. Figure 2 shows that $\Gamma(t)$ grows unbounded and linearly ($\lambda = 1$) for S_v (purple), which is the most sensitive variable to perturbations in the forcing phase.

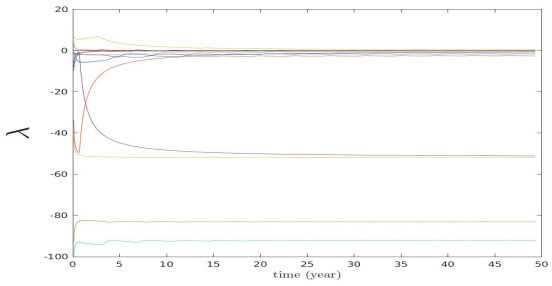

Figure 1: Evolution of the time-averaged local Lyapunov exponents in a trajectory of D2.

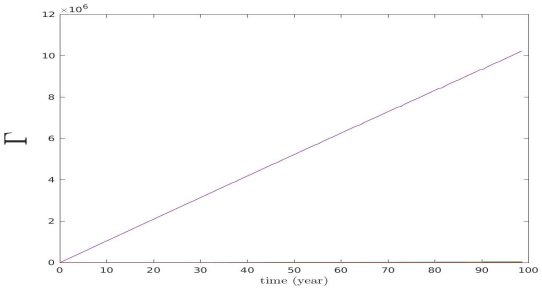

Figure 2: Phase sensitivity profile for each variable of D2.

4 Prediction scheme

In this section, the Member Choice (MC) prediction scheme is presented and applied to D2 set with 17-18.

Step 1: Among the least sensitive variables on prediction time scales, set a minimal group of determining variables for the evolution of relevant variables in the model, in the sense of Foias [3], endowed with a time series of data, to serve for data assimilation and against which to measure the adherence of the samples.

Figure 3 shows an initial value sensitivity analysis of a trajectory of D2 set with 17-18. The analysis was performed with 100 samples. On decadal scales one sees intermittent bursts of sensitivity which for I_1, I_2, I_{12}, I_{21} reach a maximum value of about 5 population units. These are the least sensitive variables of the system. Among them the cross-infection variables I_{12} and I_{21} may be endowed with data from hospitals.

Figure 4 shows a determining variables analysis of a trajectory of D2 set with 17-18. The analysis was performed with 100 state samples, including error in the human population N (up to 5%, uniformly drawn). One sees that the cross-infection variables determine the evolution of the first-infection and recovery variables in this trajectory (if N is perfectly known, they determine the evolution of all human variables). This property is supposed to apply to every trajectory

of D2 set with 17-18, in which case the cross-infection variables form a minimal group of determining variables for the evolution of the first-infection and recovery variables of D2 with that parameter setting, endowed with data.

Step 2: Among the minimal group of determining variables set in Step 1 and the trailing variables associated to them, and based upon the application of interest, set a decision variable(s). After a series of data insertions the sample choice made by the decision variable(s) stabilizes and a prediction can be launched from this sample (preferably after the data assimilation process has converged) using the model that originated it.

Following [4, 7], let the perturbation of a state sample x of D2 at time t_n, namely $x(t_n)$, be given by

$$dx(t_n) = (x(t_{n+1}) - x(t_n)) - \int_{t_n}^{t_{n+1}} f(\hat{x}(t_n), \tilde{x}(t), t)\, dt, \tag{19}$$

where $x(t) = (\hat{x}(t), \tilde{x}(t))$ is the trajectory of x under D2, $\hat{x}$ denotes the determining variables of D2 set in Step 1 and $\tilde{x}$ the remaining variables. Also, f is the o.d.e. field of D2 and $t_{n+1} = t_n + \Delta t$ where Δt is the time step of the o.d.e. solver of D2. In MC runs, state perturbations are not applied to samples unless dynamic stabilization is sought, which will be explored in a future publication with D2 set for more complex dynamical regimes.

Figure 5 shows that, with cross-infection data with up to 5% error (uniformly drawn once per insertion time) inserted into 4 state samples every once a year, the time-averaged normalized perturbation magnitude plots for the determining and associated trailing variables set in Step 1 for D2, namely the cross-infection, first-infection and recovery variables, eventually correlate with the error plots for these variables. The sample choice to be made by these variables at any time t is then set as the one whose time-averaged normalized perturbation magnitude at that time is minimum over all samples. Also seen in the figure is that the data assimilation process converges around year 30 in this run.

Figure 6 shows that, with cross-infection data with up to 5% error inserted into 4 state samples every once a year, the sample choice made by the decision variable, set to R, stabilizes after about 15 years at sample 2 in this run. Figure 8 shows that, with a similar run carried out with 400 samples, the sample choice made by R stabilizes after about 35 years at sample 99 in this run.

Figure 7 shows the error profiles obtained for the human population variables of D2 set with 17-18 in 150-year long Free Run (black), straightforward Monte Carlo (blue) and MC (red) predictions made after 40 years of cross-infection data insertions with up to 5% error into 4 samples (except for Free Run, which is a straightforward Monte Carlo run free of data insertions). The Free Run and Monte Carlo estimators are the sample average. Figure 9 shows the error profiles obtained for similar runs carried out with 400 samples. MC shows a much higher predictive skill than straightforward Monte Carlo in both cases.

5 Conclusion

A non-statistical prediction scheme viable for high-dimensional nonlinear systems was presented and applied to a 2-strain host-vector SIR model for dengue (D2).

The scheme, called Member Choice (MC), consists of detecting the lowest-error state sample emerging from a data assimilation process and making a prediction from it. The lowest-error sample is detected by measuring the adherence of each sample to the process through the magnitude of state perturbations given by 19 and computed in every time step on each sample

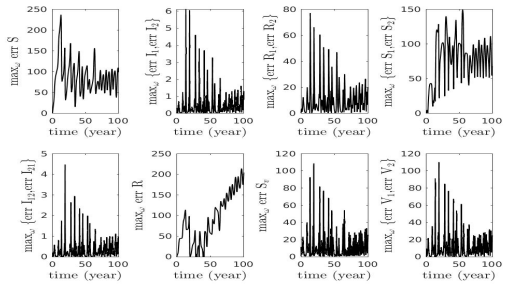

Figure 3: Initial value sensitivity analysis of a trajectory of D2.

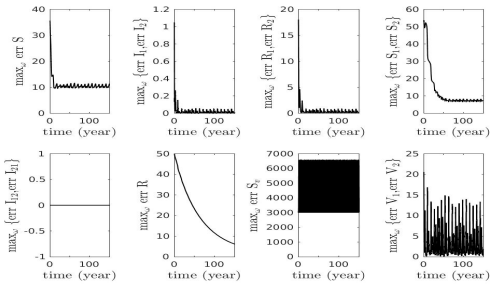

Figure 4: Determining variables analysis of a trajectory of D2.

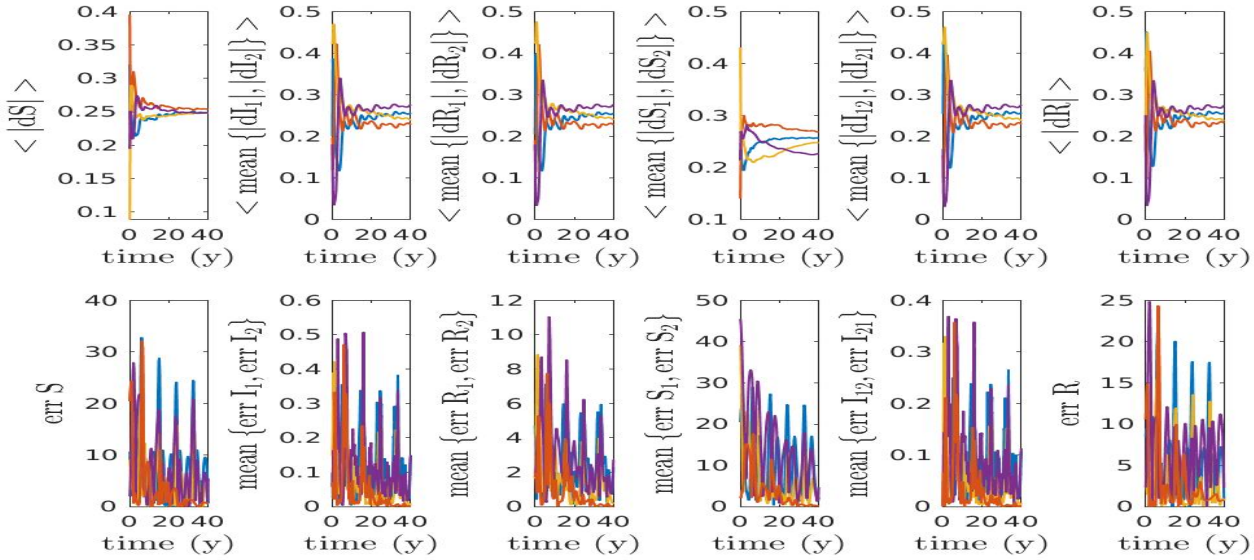

Figure 5: Evolution of error (bottom) and time-averaged normalized perturbation magnitude (top) for the human population variables of D2 in a MC data assimilation run with 4 samples.

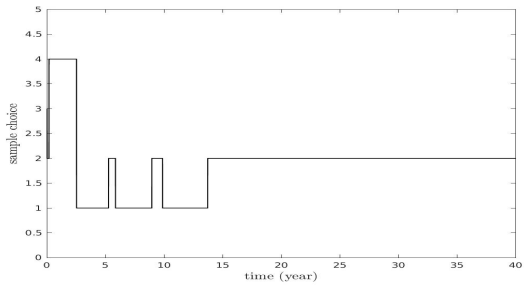

Figure 6: Evolution of the sample choice made by R in a MC data assimilation run with 4 samples.

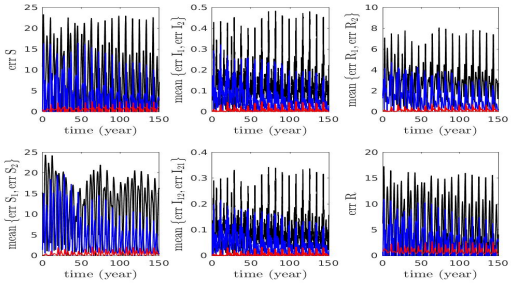

Figure 7: Evolution of error for the human population variables of D2 in 150-year long Free Run (black), straightforward Monte Carlo (blue) and MC (red) prediction runs made after 40 years of cross-infection data insertions with up to 5% error into 4 samples.

(but not applied to the samples unless dynamical stabilization is sought, which further requires that the computation of perturbations be iterated in each time step). Dynamical stabilization will be explored in a future publication. MC can be described in 2 steps: 1) Among the least

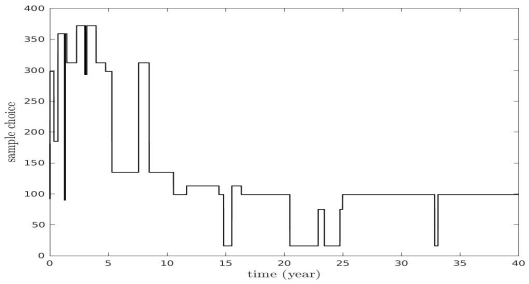

Figure 8: Evolution of the sample choice made by R in a MC data assimilation run with 400 samples.

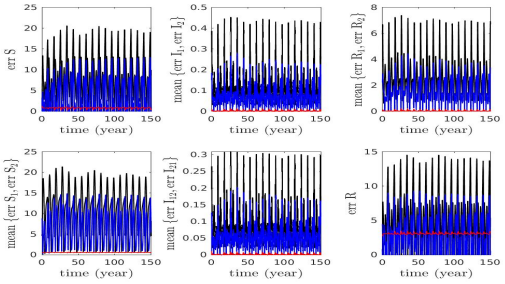

Figure 9: Evolution of error for the human population variables of D2 in 150-year long Free Run (black), straightforward Monte Carlo (blue) and MC (red) prediction runs made after 40 years of cross-infection data insertions with up to 5% error into 400 samples.

sensitive variables on prediction time scales set a minimal group of determining variables for the evolution of relevant variables in the model, in the sense of Foias [3], endowed with a time series of data; 2) Among the minimal group of determining variables set in the previous step and the trailing variables associated to them, and based upon the application of interest, set a decision variable(s). After a series of data insertions the sample choice made by the decision variable(s) stabilizes and a prediction can be launched from this sample (preferably after the data assimilation process has converged) using the model that originated it.

Model D2 set with 17-18 was shown to have a strange nonchaotic attractor with phase exponent $\lambda = 1$ for S_v, which is the most sensitive variable to perturbations in the forcing phase, wherein the cross-infection variables form a minimal group of determining variables for the evolution of the first-infection and recovery variables, endowed with data. When applied to D2 with this setting, MC showed a much higher predictive skill than straightforward Monte Carlo for any sample set size.

Acknowledgements

The author thanks Maíra Aguiar (BCAM) for helpful discussions on epidemiological modeling of vector-borne diseases and acknowledges IMPA's support in providing computational resources to this work.

References

[1] M. Aguiar, S. Ballesteros, B.W. Kooi and N. Stollenwerk, The role of seasonality and import in a minimalistic multi-strain dengue model capturing differences between primary and secondary infections: Complex dynamics and its implications for data analysis, *Journal of Theoretical Biology*, 289 (2011) 181-196.

[2] U. Feudel, S. Kuznetsov and A.S. Pikovsky, "Strange Nonchaotic Attractors", World Scientific, 2006.

[3] C. Foias and G. Prodi, Sur le comportement global des solutions non-stationnaires des équations de Navier-Stokes en dimension 2, *Rend. Sem. Mat. Univ. Padova*, 39 (1967) 1-34.

[4] P. Krause, Influence sampling of trailing variables of dynamical systems, *Math. Clim. Weather Forecasting*, 3 (2017) 51-63.

[5] P. Krause, Dyson's split action formula for transport operators, *Matemática Contemporânea*, 56:8 (2023) 135-142.

[6] P. Krause, Chaos control in high-resolution atmospheric predictions with a cloud-resolving model, In: "Proceeding Series of the Brazilian Society of Computational and Applied Mathematics" pp. 11:1, 2025.

[7] P. Krause and J. Tribbia, Initialization and dynamical stabilization of a cloud-resolving model, *Physics Open*, 12 (2022) 100117.

[8] A.S. Pikovsky and U. Feudel, Characterizing strange nonchaotic attractors, *Chaos*, 5:1 (1995) 253-260.

[9] G. Rubino and B. Tuffin (eds.), "Rare Event Simulation Using Monte Carlo Methods", John Wiley & Sons, 2009.

[10] A.K. Srivastav, V. Steindorf, N. Stollenwerk and M. Aguiar, The effects of public health measures on severe dengue cases: An optimal control approach, *Chaos, Solitons and Fractals*, 172 (2023) 113577.