Efficacy of Super-models in Chaotic Systems

A. Stoken, J. Warshawer,Y. Liu, X. Tu, L. Mota

Department of Mathematics, University of Arizona

6 May 2019

THE UNIVERSITY OF ARIZONA
COLLEGE OF SCIENCE
Mathematics

Abstract

This paper explores the use of the 'super-modeling' approach to model chaotic systems, which calls for the creation of multiple imperfect parameter models which are then connected by connection coefficients. This vector of connection coefficients, **C**, is learned from historical data or the canonical model via the Conjugate Gradient Method. The purpose of this super-modeling approach is to get a more accurate prediction at future times compared to both the imperfect models that make up the super-model as well as compared to the unconnected ensemble techniques. In order to determine the validity of super-modeling, we have applied the approach on the Lorenz 63 system and evaluated the error, qualitative accurateness, and Lyapunov exponents of the model. The super-models we created synchronize with the Lorenz 63 better than both the individual and averaged systems. Through our calculation of the Lyapunov exponents, the super-models are also less chaotic than the Lorenz 63 system. Further applications of these methods may lead to more accurate predictions for chaotic systems such climate. Super-models may also be less sensitive to initial conditions when compared to other modeling techniques.

Contents

1 Introduction

Chaotic systems have proven extremely difficult to model accurately. Such systems are named chaotic for their sensitivity to initial conditions - if an initial condition is off by even a small amount, the model will produce dramatically different results. To combat this sensitivity, many modern climate models are produced via ensemble methods, which are combinations of models averaged after they have been individually constructed. While this has thus far been the most effective way of modeling chaotic systems, these averaged models still desynchronize quickly with the truth.

We believe we can improve on these ensemble techniques by using the super-model approach. We are testing whether this approach yields greater synchronization with the truth by using three averaged, imperfectly parameterized models which have been augmented through linear connection coefficients. These connection coefficients, which allow each of the three models to exchange information with one another at each time step, nudge the models toward each other in time. However, finding the optimal set of connection coefficients to reduce the model error is a computationally intensive task requiring gradient descent. Once the optimal coefficients are found, we will then be able to test the accuracy of our super-model.

Climate is a consequential example of a chaotic system that is notoriously difficult to model. Even the best models are known to predict changes in temperature that turn out to be two degrees off from the truth. [1] The super-modeling approach has previously shown some promise in improving climate models [1] [2] [3]. We seek to further explore this approach by applying it to a climate model, the Lorenz 63 system, and determining if it can outperform the traditional weighted average method. Should super-models prove to enhance synchronization, the approach may be used to improve modeling of more complex climate dynamics.

1.1 The Lorenz 63 System

The Lorenz 63 system is a system of non-linear ordinary linear differential equations, defined in Cartesian coordinates as follows:

$$
\begin{aligned}
\dot{x} &= \sigma(y - x) \\
\dot{y} &= x(\rho - z) - y \\
\dot{z} &= xy - \beta z
\end{aligned}
\tag{1}
$$

x: rate of atmospheric convection

y: horizontal temperature variation

z: vertical temperature variation

The standard parameters of the Lorenz 63 system are $\sigma_0 = 10$, $\rho_0 = 28$ and $\beta_0 = 8/3$. σ is known as the Prandtl number, and ρ the Rayleigh number. [4] Beyond climate, the Lorenz system has also found applications throughout physics in lasers, circuits, and chemical reactions.

1.2 Characteristics of Chaotic Systems

Chaotic systems are so named because of their high sensitivity to initial conditions. While a non-chaotic system will produce similar trajectories for an initial condition x_0 and a perturbed initial condition $x_0 + \epsilon$, a chaotic system will produce dramatically different trajectories given enough time. Such sensitivity has become known as the "butterfly effect". Only nonlinear or infinite-dimensional systems can exhibit chaotic behavior [5].

Another important aspect of chaotic systems is that they are deterministic. Given perfect initial conditions, one can predict future outcomes of the system.

However, in the real world, perfect initial conditions are not attainable. We can see these real life effects of chaos in weather. Despite the advances in technology, meteorologists still are unable to perfectly measure initial climate conditions, so weather is generally predictable only up to one week in advance [6].

Lorenz is known to have summed up these two defining properties of chaos as, "When the present determines the future, but the approximate present does not approximately determine the future." [4]

1.3 Measuring Chaos

It can be beneficial to quantitatively measure how chaotic a system is. In other words, it is useful to know how quickly two trajectories diverge when they have slightly different initial conditions. We will use the Lyapunov exponent to measure how chaotic the Lorenz 63 system and our super-models are. The Lyapunov exponent was originally posed as a measure of stability by Aleksandr Lyapunov in 1892, and has since become a classic stability measurement [7].

2 Methods

2.1 Data Collection

In order to generate data for the super-model connection coefficients to learn on, we modeled the canonical Lorenz 63 system. We implemented the Lorenz 63 system with the standard parameter values, and integrated the system for 20 time steps/number of iterations by using the 4th order Runge-Kutta Method. This gave us the $\mathbf{x_0}$ used in the cost function (2.3).

2.2 Super-Model

To create a super-model, we first choose our system of interest. To begin, we will focus on the Lorenz 63 system. Three different imperfect models are created by perturbing the standard parameter values called Truth in Table 3. The super-model is created by adding linear connection terms into each equation for all three imperfect models. This transforms the traditional Lorenz 63 system of equations (1) into the *connected* Lorenz 63 system:

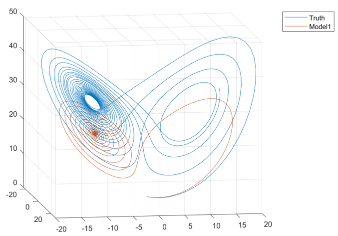
$$
\begin{aligned}
\dot{x_k} &= \sigma_k(y_k - x_k) + \sum_{j \neq k} C^x_{kj}(x_j - x_k) \\
\dot{y_k} &= x_k(\rho_k - z_k) - y_k + \sum_{j \neq k} C^y_{kj}(y_j - y_k) \\
\dot{z_k} &= x_k y_k - \beta_k z_k + \sum_{j \neq k} C^z_{kj}(z_j - z_k)
\end{aligned}
\tag{2}
$$

where k = 1,2,3 are the indices of the three imperfect models. σ_k, β_k and ρ_k are the perturbed parameter values. We begin our exploration of super-modeling by reproducing van den Berge, F. M. Selten, W. Wiegerinck, and G. S. Duane [1]. The perturbed parameters are given in Table 3. These perturbed parameter models produce the trajectories described in Figure 1.

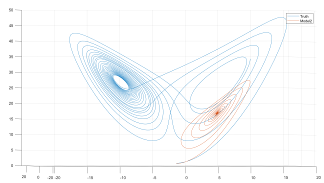
The $C^x_{kj}, C^y_{kj}, C^z_{kj}$ in the connected equations are referred to as connection coefficients. There are three imperfect models, and each model is three dimensional. Each dimension has two connection coefficients. This leads to 3 x 3 x 2 = 18 connection coefficients in total. We introduce $\mathbf{x}_s$ to denote the final solution of the supermodel, which is the the average of

	σ	ρ	β
Truth	10	28	$\frac{8}{3}$
Imperfect Model 1	13.25	19	3.5
Imperfect Model 2	7	18	3.7
Imperfect Model 3	6.5	38	1.7

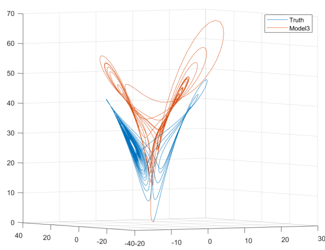
Table 1: Perturbed Parameters for Connected Lorenz System (2)



(a) Submodel 1



(b) Submodel 2



(c) Submodel 3

Figure 1: Sub-model Trajectories

$$
\begin{aligned}
x_s &= \frac{1}{3}(x_1 + x_2 + x_3) \\
y_s &= \frac{1}{3}(y_1 + y_2 + y_3) \\
z_s &= \frac{1}{3}(z_1 + z_2 + z_3)
\end{aligned} \tag{3}
$$

In order to learn the connection coefficients of the super-model, we will introduce a cost function. Minimizing this function leads to the optimal vector of connection coefficients, $\mathbf{C}_{opt}$ (Section 2.5).

2.3 Cost Function

We introduce the following cost function:

$$F(\boldsymbol{C}) = \frac{1}{K\Delta}\sum_{i=1}^{K}\int_{t_i}^{t_i+\Delta} \mid \boldsymbol{x}_s(\mathbf{C},t) - \boldsymbol{x}_0(t) \mid^2 \gamma^t dt \tag{4}$$

$\mathbf{x}_0$ denotes the trajectory of the standard Lorenz 63 system. $\mathbf{x}_s$ denotes the solution of the super-model for a given $\mathbf{C}$. The cost of a super-model solution with $\mathbf{C}$ is proportional to the mean square error between $\mathbf{x}_s$ and $\mathbf{x}_0$, integrated over a small time Δ. The γ^t term, where γ is between 0 and 1, is introduced inside the integral in order to give larger weight to error occurring early in each time-step.

We have K initilizations where we begin our integration procedure over a time Δ, yielding a total time of $K\Delta$. This is why we divide by $K\Delta$ in order to normalize the cost. F($\mathbf{C}$) outputs a numerical description of the error, which is the sum of the total area between our super-model and the truth.

Large values of F are indicative of a $\mathbf{C}$ that causes the super-model to stray from the truth, while low values of F show we have a good $\mathbf{C}$ causing the super-model to synchronize closer with the true system. For this reason, our goal is to pick a $\mathbf{C}$ that minimizes the cost function (see 2.5 for details).

2.4 Minimization

We next are going to determine what value of $\mathbf{C}$ that will minimize the cost function. Whichever $\mathbf{C}$ accomplishes this task is the one that will synchronize the super-model most closely with the truth.

In order to minimize the 19-dimensional (18 inputs and 1 output) cost function, we will utilize the Fletcher-Reeves-Polak-Ribiere Conjugate Gradient method. [8]

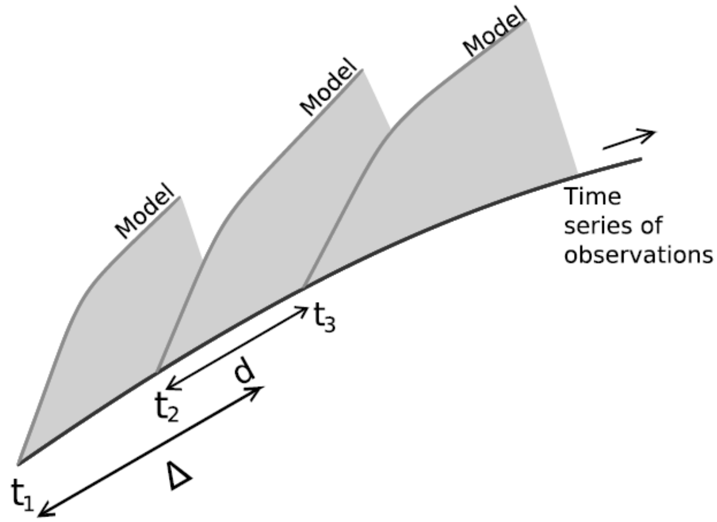


Figure 2: Graphical representation of the Cost Function [1]

For our purposes, we select a **C** to be the initial value in our minimization procedure. The Conjugate Gradient method will then change **C** one component at a time, testing whether this new **C** yields a smaller F than it did previously. If F becomes smaller, it will continue to change that specific parameter in the same way. If the new **C** did not yield an improvement in the cost function, the CG method will try another perturbation of **C**. This process moves through 18 dimensional space until a local minima is found. The **C** that produces this local minima will then serve as our final value of **C** for the super-model. This method has disadvantages that will be analyzed in the discussion section.

2.5 Lyapunov Exponents

The distance u_t between two trajectories at time t with initial separation u_0 can be estimated by the following equation:

$$|u_t| = e^{\lambda t}|u_0| \tag{5}$$

Where λ is the Lyapunov exponent. If the Lyapunov exponent is negative, the two trajectories come closer together over time, which implies the system is not chaotic. On the other hand, if the Lyapunov exponent is greater than zero, it demonstrates two trajectories come apart over time, meaning the system is chaotic. The greater the Lyapunov exponent, the faster the two trajectories diverge from one another, and the more chaotic the system is.

In order to calculate it, it helps to put the above equation into a different form by taking the natural logarithm of the both sides:

$$\ln|u_t| = \lambda t + \ln|u_0| \tag{6}$$

We graphed the above equation and found it's slope, giving us the Lyapunov exponent of the system.

2.6 Parameter Estimation

We now propose a novel method of parameter estimation via super-modeling. Take a fully-defined and optimized super-model of the form in (2) from three sub-models of imperfect parameters with unknown original parameters $\rho_0, \beta_0, \sigma_0$. Then, we take the resulting (x, y, z) predictions from this super-model at some time t. These coordinates (and the derivative, approximated by $(x, y, z)|_{t+1} - (x, y, z)|_{t-1}$ can be used in the standard Lorenz 63 system (1) to calculate $\rho_0, \beta_0, \sigma_0$.

Then, rearranging system (1) to solve for parameters gives

$$
\begin{aligned}
\sigma_0 &= \frac{\dot{x}}{y - x} \\
\rho_0 &= \frac{\dot{y} + y}{x} + z \\
\beta_0 &= \frac{xy - \dot{z}}{z}
\end{aligned} \tag{7}
$$

We will use this technique to approximate the true parameter values of the Lorenz 63 system from our super-model output.

3 Results

The following sections describe the result from super-modeling the Lorenz 63 system.

3.1 Trajectories

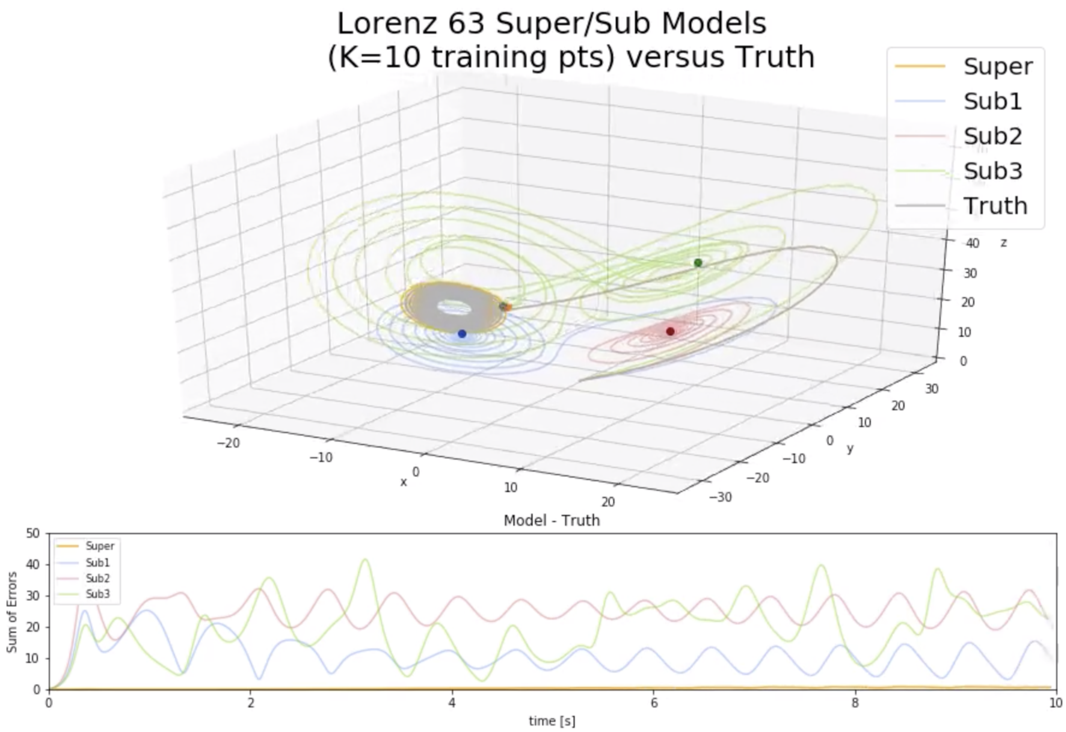


Figure 3: Lorenz 63 Super and Sub Models with Error

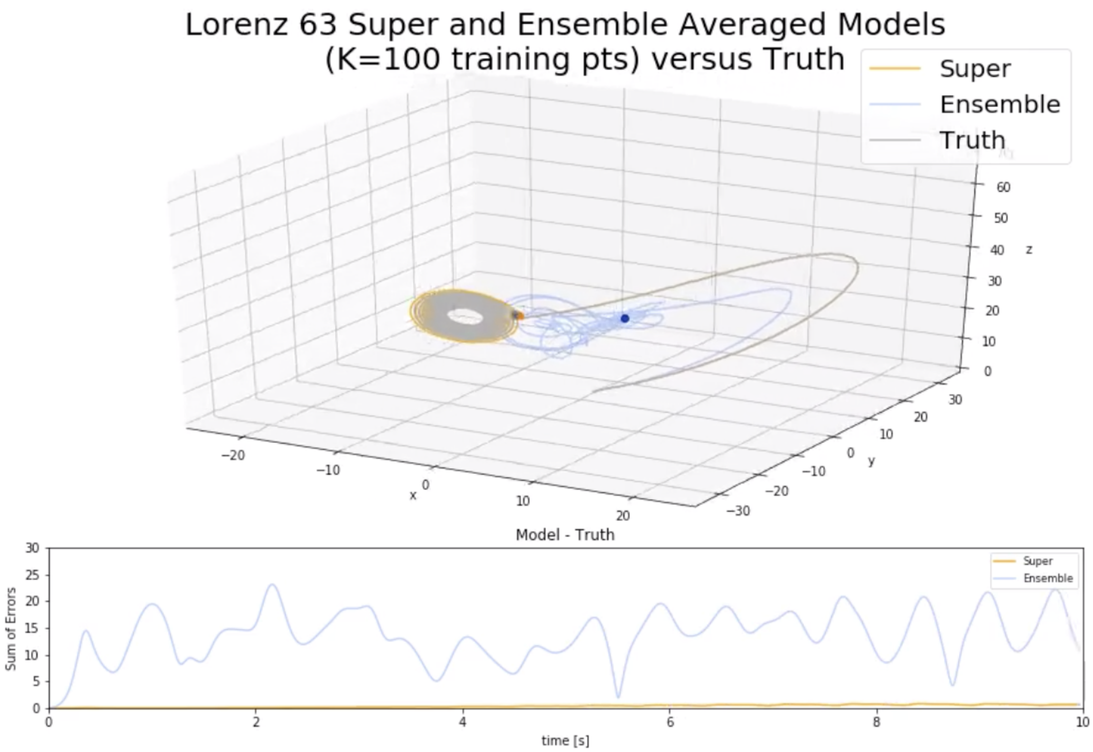


Figure 4: Lorenz 63 Super and Averaged-Sub Models with Error

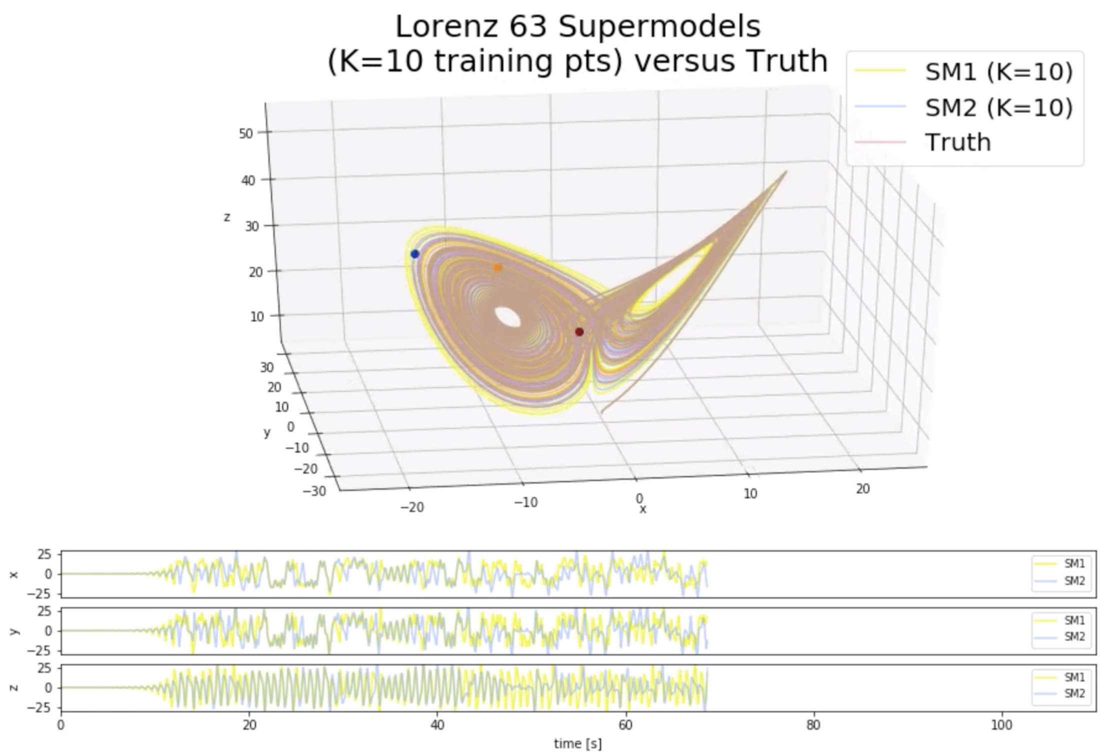


Figure 5: Lorenz 63 Unique Super Models

3.2 Conjugate Gradient Minimization

The minimization procedure produced the minimum **C** vectors in Table 2. These are plotted against the cost function in Figure 6.

	Supermodel 1	Supermodel 2
C_{12}^x	18.43	17.19
C_{13}^x	-11.25	-11.20
C_{21}^x	13.27	13.81
C_{23}^x	10.99	9.91
C_{31}^x	50.96	49.47
C_{32}^x	26.93	26.88
C_{12}^y	-46.95	-45.25
C_{13}^y	62.78	59.23
C_{21}^y	24.77	25.31
C_{23}^y	33.55	32.61
C_{31}^y	-10.40	-11.03
C_{32}^y	51.09	49.93
C_{12}^z	11.47	12.26
C_{13}^z	3.35	3.54
C_{21}^z	17.54	16.15
C_{23}^z	17.45	16.40
C_{31}^z	-0.01	0.11
C_{32}^z	11.74	11.21
$F(C)$	0.011	0.076

Table 2: **C** values minimizing the two super-models

For Super-model 1, the dependence of the cost on each component of **C** is illustrated below.

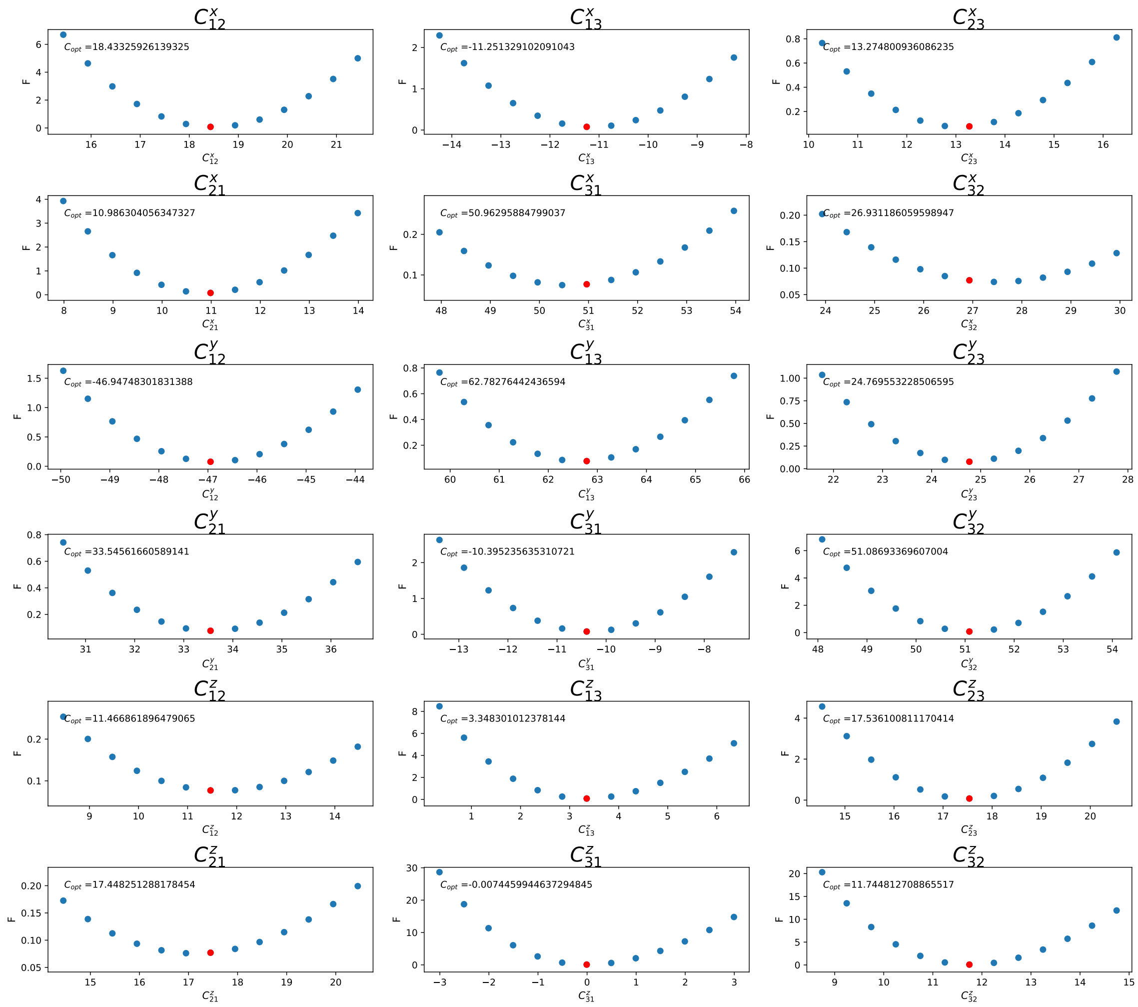


Figure 6: F vs $C^i_{\mu\nu}$

3.3 Lyapunov Exponents

The calculated Lyapunov exponents are showin below in Table 3

	Lyapunov Exponent
Lorenz 63	0.922
Averaged-Model	0.267
Super-Model 1	0.690
Super-Model 2	0.686

Table 3: Calculated Lyapunov Exponents for all systems studied

The graphs used to calculate each Lyapunov exponent are shown below.

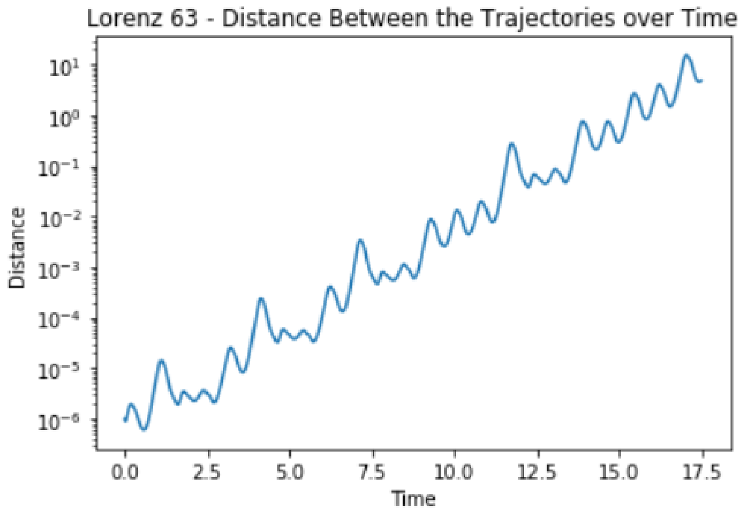


Figure 7: Calculating the Lyapunov Exponent for the Lorenz 63 System

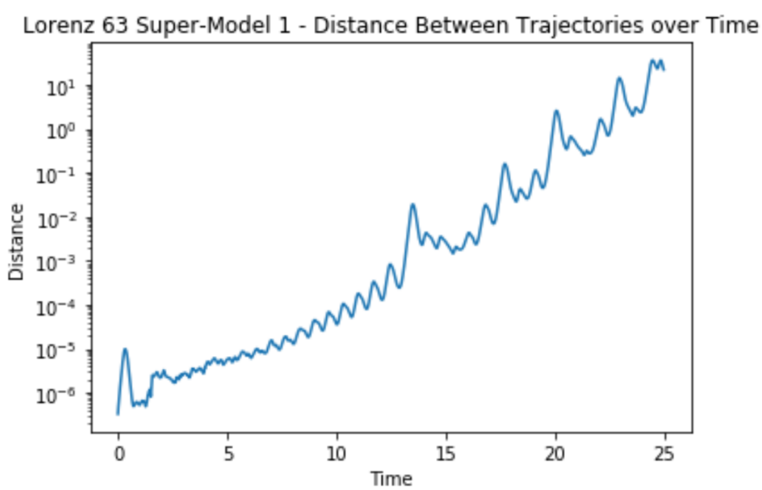


Figure 8: Calculating the Lyapunov Exponent for the super-model 1

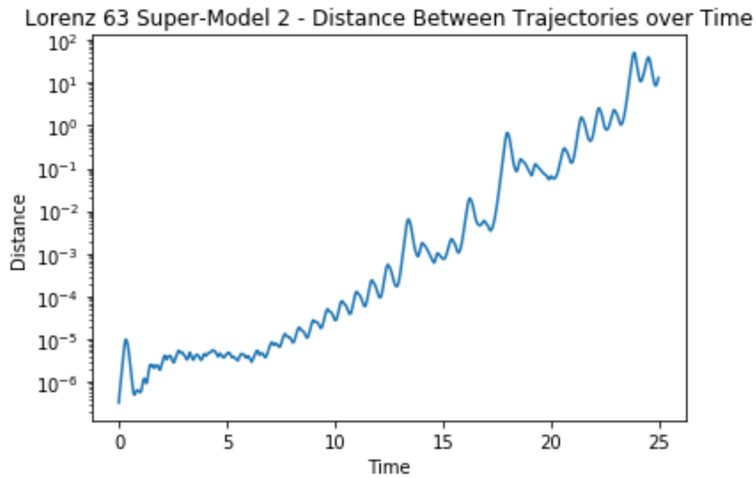


Figure 9: Calculating the Lyapunov Exponent for the super-model 2

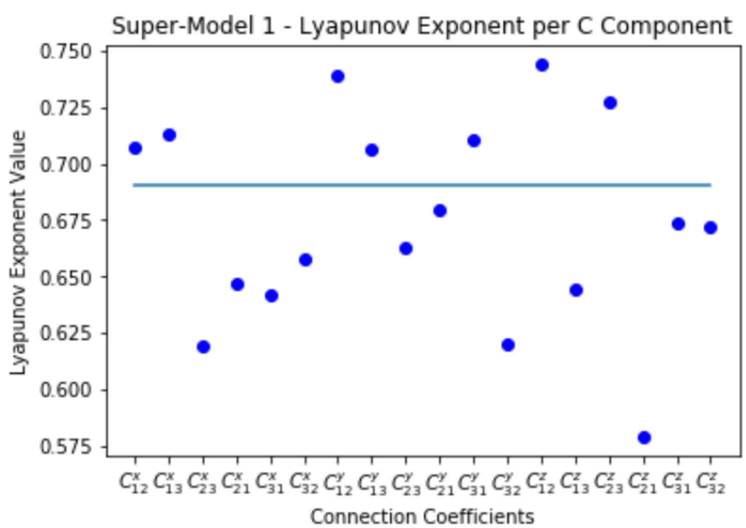


Figure 10: Calculated Lyapunov Exponents for reach C component perturbed

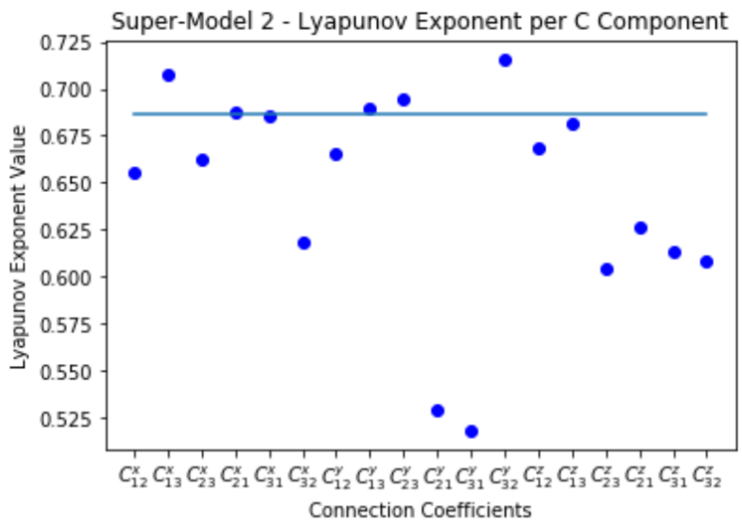


Figure 11: Calculated Lyapunov Exponents for reach C component perturbed

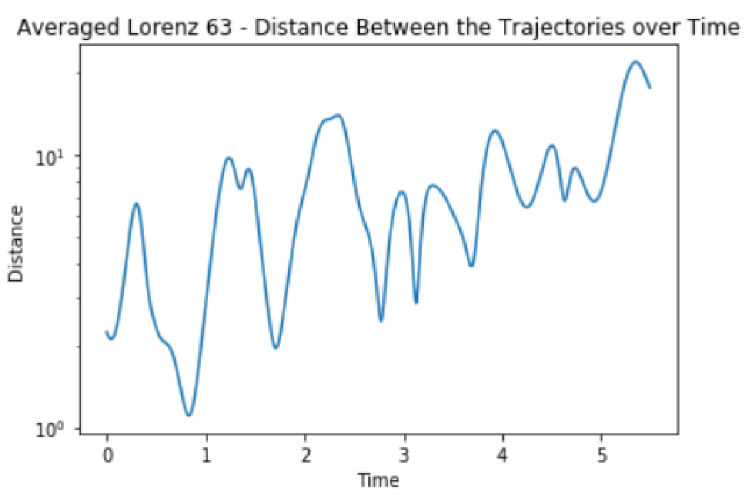


Figure 12: Calculating Lyapunov Exponent for the Averaged System

3.4 Parameter Estimation

The parameter estimates for the Lorenz 63 system generated using the process described in Section 2.6 are in Table 2.6.

Parameter	Estimate	Percent Error
σ	10.05	0.50
ρ	27.96	0.14
β	2.62	0.18

Table 4: Paramter estimates via Super-modeling

4 Discussion

4.1 Conjugate Gradient Method

The conjugate gradient method produced multiple solutions for optimal **C** vectors, depending on the initial random start point and which local minima was encountered. While these local minima both produce qualitatively correct solutions, some local minima for example, SM1) are deeper than others and produce 'better' solutions. See Figure 5 for the differences in trajectory from the different local minima. These differences, especially compared to the ensemble model and other modeling techniques, are negligible.

Once optimized, the conjugate gradient method was relatively computationally efficient. The method finished within 100 function calls each time, and the gradient estimation was quick for K = 10. Adding higher K values increased the optimization time significantly but did not show large increases in performance.

4.2 Accuracy of Super-Model

We have created a super-model that outperforms the sub-models from which it is made and a model averaging those sub-models as shown in Figure 3 and Figure 4. The error of the super-model is close to zero while the other models have completely desychnronized. The averaged model in Figure 4 does not sychncronize with the Lorenz 63 system either.

4.3 Connection Coefficients

Table 2 shows our connection coefficients for our two super-models. These super-models were created using the same process of Section 2.5, but with different initial conditions. It is important to note the connection coefficients are different. It is possible the connection coefficients are approaching one single local minima, it is also possible there are mulitple minima that give similar results.

Figure 6 shows what happens as you change one connection coefficient as you keep all others constant. These graphs were made once we had already found the minimum vector of connection coefficients. However, it is valuable to do this type of analysis to see how sensitive certain connection coefficients are.

For example, reducing proper C^z_{31} by three increases the error from approximately 0.01 to 30. On the other hand, decreasing C^z_{21} by the same amount only increases the total error from 0.01 to 0.018.

4.4 Lyapunov Exponents

We found the Lyapunov exponents in the super-models smaller than the traditional systems, yet still positive. This demonstrates our super-models are less chaotic than the systems they were derived from. In other words, they are less sensitive to initial conditions.

Section 4.1 and 4.2 discussed how well our super-models synchronizes with the the original models while starting at the same initial conditions. The study of how chaotic these super-models through Lyapunov exponents shows something else: they are less sensitive to initial conditions. This may mean that upon utilizing these models in more practical situations where the initial conditions are not completely known, super-models have the potential to handle that uncertainty better than traditional models.

In Section 3.2, we determined which connection coefficient was more sensitive to change. Some connection coefficients could be changed with little increase in overall error, while others were incredibly sensitive, with little change causing large changes in overall model error.

Similarly, we tested the chaotic nature of each component of $\mathbf{C}$. One at a time, we calculated the Lyapunov exponent after perturbing the a singular connection coefficient by 1. Figures 7-12 shows the results of this process. The horizontal line represents the Lyapunov exponent of the unperturbed super-model, while each dot represents the Lyapunov exponent for its corresponding connection coefficient being perturbed. For dots above the line, the super-model became more chaotic as a result of the connection coefficient perturbation, while the reverse is true for dots below the line.

5 Conclusion

We have demonstrated the super-modeling shows promise to better synchronizes with chaotic systems. After generating data by modeling a standard Lorenz 63 system, we created a super-model with three sub-models. These three sub-models have perturbed parameters, but they also are interconnected with connection coefficients. Using a cost function based on mean square error and a minimization technique suitable for the 18 dimensional vector of connection coefficients, the Fletcher-Reeves-Polak-Ribiere Conjugate Gradient method, we found an optimal **C**. Our results show our super-models yield less cumulative error, better synchronizing with the truth when compared to individual or averaged systems.

5.1 Further Analysis

Further analysis would be needed to show if and why super-models generally are less chaotic than their individual components. This could take the form of additional work applying the super-modeling method to different chaotic systems and calculating the Lyapunov exponents. It could also take the form of theoretical work to explain why super-models may or my not be generally less chaotic than their traditional system counterparts.

Another possible area of further analysis would be to determine if connection coefficients relate in some way to the real-world. For example, why are some connection coefficients small and some large? Why are some connection coefficients extremely sensitive to how accurate the model is and some are not? Why are some connection coefficients sensitive to how chaotic the overall super-model is and some are not? It is possible these coefficients have some connection to the real-world contexts in which they model, or perhaps not.

One additional area of exploration would be to attempt the super-modeling approach with differing amounts of sub-models. Perhaps a different amount of sub-models can increase model accuracy without significantly increasing computational time.

References

[1] L. A. van den Berge et al. “A multi-model ensemble method that combines imperfect models through learning”. In: *Earth System Dynamics* 2.1 (2011), pp. 161–177. DOI: 10.5194/esd-2-161-2011. URL: https://www.earth-syst-dynam.net/2/161/2011/.

[2] Mao-Lin Shen et al. “Role of atmosphere-ocean interactions in supermodeling the tropical Pacific climate”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.12 (2017), p. 126704. DOI: 10.1063/1.4990713. eprint: https://doi.org/10.1063/1.4990713. URL: https://doi.org/10.1063/1.4990713.

[3] Frank M. Selten, Francine J. Schevenhoven, and Gregory S. Duane. “Simulating climate with a synchronization-based supermodel”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.12 (2017), p. 126903. DOI: 10.1063/1.4990721. eprint: https://doi.org/10.1063/1.4990721. URL: https://doi.org/10.1063/1.4990721.

[4] URL: http://mpe.dimacs.rutgers.edu/2013/03/17/chaos-in-an-atmosphere-hanging-on-a-wall/.

[5] Ivar Bendixson. “Sur les courbes définies par des équations différentielles”. In: *Acta Math.* 24 (1901), pp. 1–88. DOI: 10.1007/BF02403068. URL: https://doi.org/10.1007/BF02403068.

[6] Robert G Watts. “Global Warming and the Future of the Earth”. In: *The computer journal* (2007), p. 17.

[7] Antonio Politi. *Lyapunov exponent*. URL: http://www.scholarpedia.org/article/Lyapunov_exponent.

[8] Reeves Fletcher and Colin M Reeves. “Function minimization by conjugate gradients”. In: *The computer journal* 7.2 (1964), pp. 149–154.

6 Report Components

Alex wrote the code for the Lorenz 63, super-modeling systems, model the trajectories, and to estimate the parameters. Alex and Jacob wrote the bulk of the paper. Jacob wrote much of the code dealing with Lyapunov exponents. Lourdes wrote the code that made the Lorenz 63 data. Yuanjie and Xiaokun wrote code to expand to other systems, however the code was not used in the final paper. Xiaokun, Lourdes, and Yuanjie wrote some components of the paper.

By percentage: Alex: 27.5 Jacob: 27.5 Yuanjie: 15 Xiaokun: 15 Lourdes: 15