Data-driven Predictions of the Lorenz 96 System Using Echo State Networks

by

Nishanta Baral and Ariel J. Bonneau

Montclair State University

April 2022

Contents

1 Abstract

In this paper, we discuss a complex nonlinear dynamical system known as the Lorenz 96 model. This chaotic model is a system of ordinary differential equations (ODEs) which is computationally challenging to solve in real life applications. In order to make computation more practical, a helpful approach would be to implement machine learning (ML) and data driven modeling as means to simulate complex nonlinear systems like the Lorenz 96 model. Our objective is to use data driven processes to predict the spatiotemporal evolution of the chaotic Lorenz 96 model and discuss the techniques used to make future predictions of our model.

2 Introduction

2.1 The Lorenz 96 Model

Edward Norton Lorenz (1917 - 2008), an American mathematician and meteorologist, is highly regarded for his work on modern chaos theory. Lorenz focused on the behavior of dynamical systems and established a basis for meteorology, weather and climate predictability [1].

One of Lorenz's most well known contributions to chaos theory is the Lorenz 96 model. The Lorenz 96 model is a dynamical system of differential equations that describes a particle's movement in a geometrical space.This model is useful when dealing with complex nonlinear dynamical systems like weather, the flow of water, or the movement of animals throughout different seasons.

The Lorenz 96 model is defined as,

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F$$

such that $i = 1, 2, ..., N$, $x_{-1} = x_{N-1}$, $x_0 = x_N$, $x_{N+1} = x_1$, and $N \geq 4$ [2]. We assume x_i is the state of the system and the F is the forcing constant. In several extensions of this model, we assume $F = 8$ as it is as common value used to create chaotic behavior [2].

In figure 1, we generated an output that plots the first 5 variables of the simulation in order to see the evolution of x with respect to time. We can also see that when $N = 5$, the system is very chaotic. Since the Lorenz 96 model is commonly used as a base model which can be further extended to model more complex multi-scale systems, or to model systems when N gets larger, we will investigate and explore an extended Lorenz 96 model in the next section.

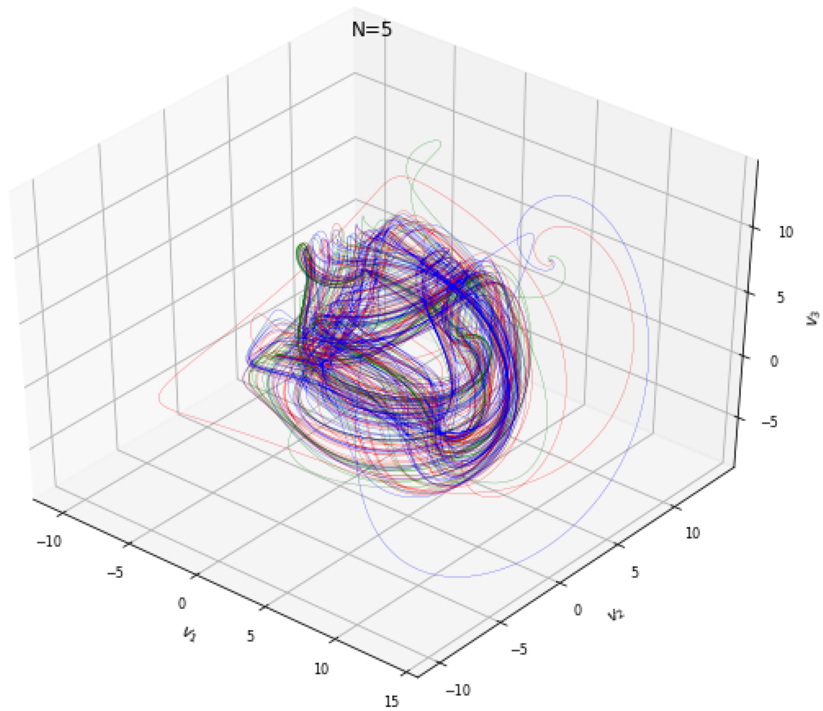

Figure 1: Lorenz 96 model where $i = 1, .., 5$, $N = 5$, and $F = 8$.

2.2 System of Equations

For this project, we investigate the multi-scale Lorenz 96 system which is a three-tier extension of Lorenz's original model. The multi-scale system is widely known and frequently used as a prototype for multi-scale chaotic variability of the weather and climate systems [3]. This multi-scale Lorenz 96 system is defined as:

$$\frac{dX_k}{dt} = X_{k-1}(X_{k+1} - X_{k-2}) + F - \frac{hc}{b}\sum_j Y_{j,k}, \tag{1}$$

$$\frac{dY_{j,k}}{dt} = -cb(Y_{j+2,k} - Y_{j-1,k}) - cY_{j,k} + \frac{hc}{b}X_k - \frac{he}{d}\sum_i Z_{i,j,k}, \tag{2}$$

$$\frac{dZ_{i,j,k}}{dt} = edZ_{i-1,j,k}(Z_{i+1,j,k} - Z_{i-2,j,k}) - geZ_{i,j,k} + \frac{he}{d}Y_{j,k}, \tag{3}$$

such that i, j, k =1,2,. . . ,8, implying that X has 8 elements, Y has 64 elements, and Z has 512 elements. From [4], we assume $F = 20$ which is large enough to make the system very chaotic, $b = c = e = d = g = 10$, and $h = 1$ all of which have been tuned to ensure proper variability in X, Y, and Z. We note that b, c, d, e, g and h have all been fine tuned according to an existing multi-scale Lorenz 96 model sourced from [4].

In the context of weather, $X(t)$ is the the atmospheric circulation, $Y(t)$ is the atmospheric convection, and $Z(t)$ are the gravity waves [5]. The atmospheric circulation, $X(t)$, corresponds to the component of wind and air flow, the atmospheric convection, $Y(t)$, corresponds to the clouds which are responsible for precipitation, and the gravity waves, $Z(t)$ corresponds to ripples in space, generated by accelerated masses [6], [7].

2.3 Objective

Typically, to solve the Lorenz 96 system of equations we should know the initial conditions of $Y(t)$ and $Z(t)$. Through trial and error, this is computationally expensive and we do not have the computing power. Therefore, in order to lower the computational costs, the system of equations would need to be solvable without knowing $Y(t)$ and $Z(t)$ at any time, and the evolution of $X(t)$ would need to be predicted based on the previous observations of X.

Thus, our goal is the use the method of an echo state network (ESN) to analyze the effects of $Y(t)$ and $Z(t)$ on the evolution of $X(t)$ and to predict the time and space evolution of the multi-scale chaotic system. This is otherwise known as the spatiotemporal evolution of the Lorenz 96 model.

We use the true solution of $X(t)$ for our training data set in the Echo State Network algorithm applications. The true solution is shown in figure 2.

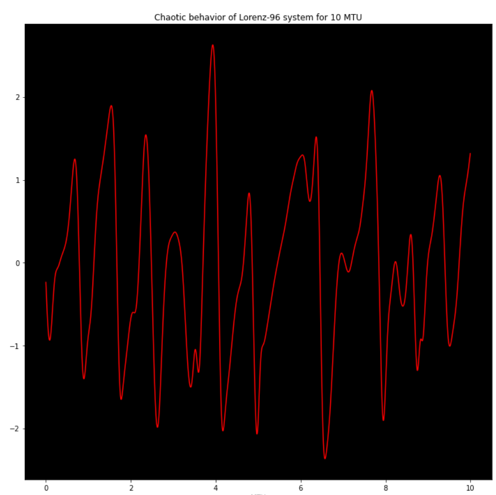

Figure 2: Chaotic behavior of the multi-scale Lorenz 96 model for 10MTU.

3 Echo State Network

The architecture of our ESN will look as follows:
a. First we create a random dynamical reservoir RNN, using any neuron model. Then we attach input units to the reservoir (input-to-reservoir layer) W_{in}. Finally, we create output units (reservoir-to-output layers) W_{out}.
b. Harvest reservoir states: The reservoir is driven by the input which results in a sequence of N-dimensional reservoir states.

c. Compute output weights which completes the training of the ESN.

The equations governing the RC-ESN training process are as follows:

$$r(t+\Delta t) = tanh((\widetilde{A}r(t)) + W_{in}X(t)), \tag{4}$$
$$W_{out} = \underset{W_{out}}{argmin} \parallel W_{out}\widetilde{r}(t) - X(t) \parallel + \alpha \parallel W_{out} \parallel . \tag{5}$$

Similarly, the prediction process is governed by:

$$v(t+\Delta t) = W_{out}\widetilde{r}(t+\Delta t) \tag{6}$$
$$X(t+\Delta t) = v(t+\Delta t) \tag{7}$$

$\widetilde{r}(t+\Delta t)$ in Equation (6) is computed by applying one of the T1, T2 or T3 algorithms on $r(t+\Delta t)$, which itself is calculated via Equation (4) from $X(t)$ that is either known from initial condition or has been previously predicted.

3.1 Numerical Solution

The multi-scale Lorenz 96 system, as described in equations 1 - 3, is solved in *Python* using a fourth-order Runge–Kutta solver with a time step of 0.005. Due to our limited computing power, the system of equations are integrated for 1 million time steps to generate a very large dataset that will be used for ESN training. The solutions are given with respect to model time units (MTUs). We note that 1 MTU corresponds to about 3.75 Earth days [4] and 1 time step is equivalent to 200 MTU. In terms of the Lyapunov timescale, 1 MTU in this system is $4.5/\lambda_{max}$ such that the maximum eigenvalue is 1 and 1 MTU is 4.5 Lyapunov$\vec{\Delta}t$. Figure 3 shows 8 plots which corresponds to one solution of $X(t)$ in R^8.

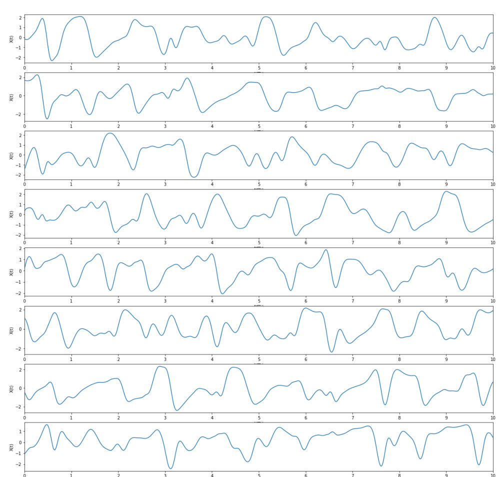

Figure 3: The 8 plots show one solution of $X(t)$ in R^8 for 10 MTU which corresponds to approximately 3.75 Earth days.

3.2 Echo State Network Predictions

We now investigate our model predictions (the predictive forecast) for different initial conditions, reservoir sizes, training lengths, and shifts. We then compare our ESN predictions to the true solution of $X(t)$. We note that achieving these results are computationally heavy and require high computing power. Thus, we used Amazon $EC2$ which is Amazon's computing platform to run our ESN codes faster. Although this computing cluster make computation manageable, one challenge is that computation time still took anywhere from 5 to 30 hours even for changing a single parameter.

Our first set of results, shown in figure 4, we use a reservoir size of 500, a prediction length of 10 MTU, a training length of 100,000 time steps which is 500 MTU (as per the unit conversion of 1 time step = 200 MTU), and a shift of 0 (indicating that we began predictions at the start of the data set). The red curve shows the predicted solution of $X(t)$ and the blue curve shows the true solution of $X(t)$. In figure 4, we can see that the predicted solution does not compare well to the true solution. After a lot of experimentation, we realized changing the reservoir size from 500 would produce more favorable results.

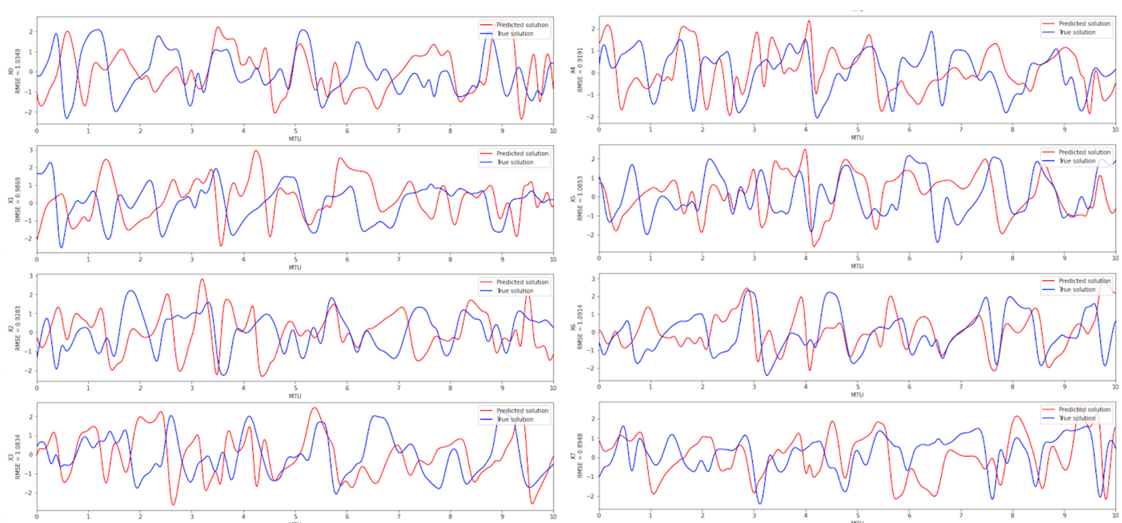

Figure 4: Predicted results of $X(t)$ are shown by the red curve and true results of $X(t)$ are shown by the blue curve for 10 MTU, and a reservoir size of 500. The results indicate poor prediction accuracy.

In figure, 5, we use a prediction length of 10 MTU, a reservoir size of 1000, a training length of 500,000 time steps which is 2500 MTU, and a shift of 0. At this stage, we noticed that a reservoir size of 1000 showed better prediction accuracy from 0 MTU to about 1.5 MTU. After that, the prediction solution diverges from the true solution.

Through continued trial and error, we found that a reservoir size of 5000 produced the most optimal results. Figure 6 shows predicted a true solutions of $X(t)$ for a prediction length of 50 MTU, reservoir size of 5000, a training length of 2500 MTU, and a shift of 0. Since we used a prediction length of 50 MTU, which is about

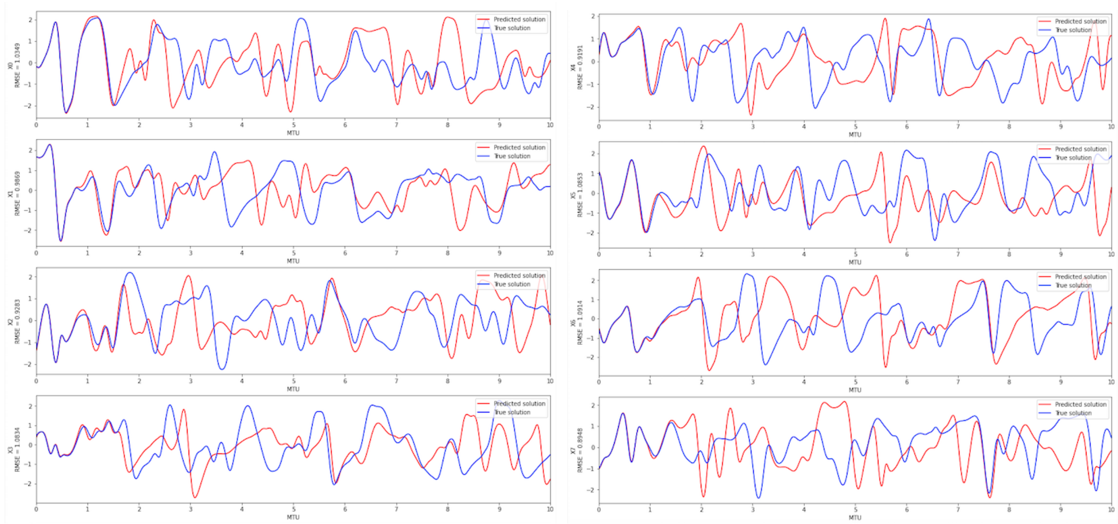

Figure 5: Predicted results of $X(t)$ are shown by the red curve and true results of $X(t)$ are shown by the blue curve for 10 MTU, and a reservoir size of 1000. The results indicate slightly better prediction accuracy as the predicted solution nearly matches the true solution from 0 to 1.5 MTU.

187.5 Earth days, there is a much larger data set to investigate. We notice that the prediction solution does well for 0 MTU to about 2 MTU and matches with the overall trend of the true solution very sporadically throughout the rest of the MTU's. This lead us to believe that ESN does not do well for predicting the long term evolution of $X(t)$, so we produced more short term results.

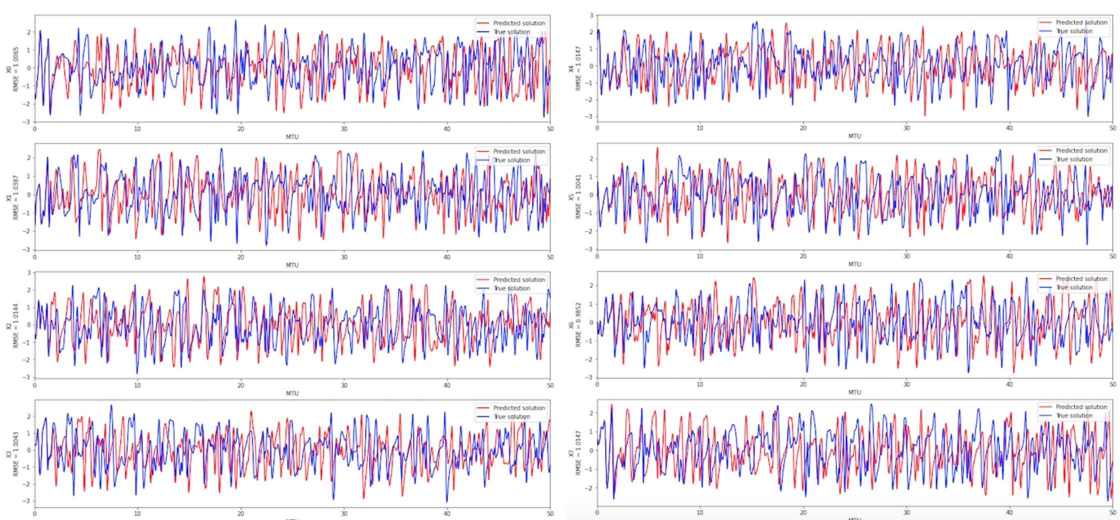

Figure 6: Predicted results of $X(t)$ are shown by the red curve and true results of $X(t)$ are shown by the blue curve for 50 MTU, and a reservoir size of 5000. The results indicate that ESN does not do well for long term predictions.

In figure 7, we use a prediction length of 10 MTU , a reservoir size of 5000, a training length of 500 MTU and a shift of 228750 which indicates that we begin training at 228750. We note that these predictions do fairly well from 0 MTU to about 4 MTU.

Lastly, in figure 8 we investigate the prediction results for a prediction length of 1 MTU a reservoir size of 5000, a training length of 25 MTU, and a shift of 228750. These results reconfirm that ESN does very well at predicting the short-term evolution of a multi-scale Lorenz 96 system.

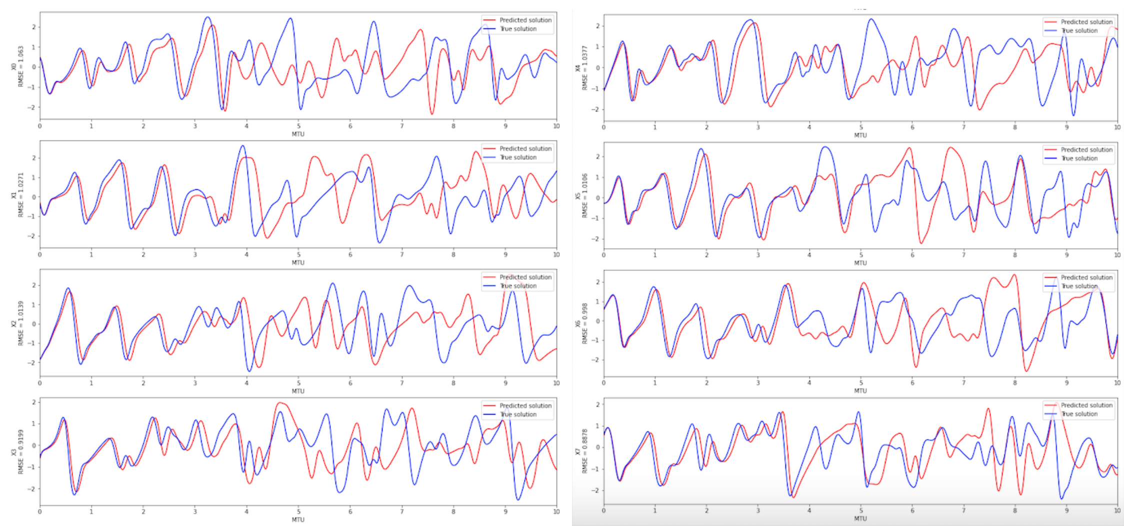

Figure 7: Predicted results of $X(t)$ are shown by the red curve and true results of $X(t)$ are shown by the blue curve for 10 MTU, and a reservoir size of 5000. The results show that the predictions compare to the true solution fairly well for 0 to 4 MTU.

4 Conclusion and Remarks

In this paper, we discuss the multi-scale Lorenz 96 model and its applications to weather predictions. This has proven to be difficult since weather is very chaotic and high computational power is needed to predict large weather data sets. Due to the computational complexity of this project, we study the evolution of atmospheric circulation, $X(t)$, to make future predictions about the weather based on the previous observations of X without knowing $Y(t)$ and $Z(t)$ using a machine learning technique known as Echo State Networks.

Our results show that predictive accuracy is not significantly dependent on size of training length. When we look back at our ESN prediction figures, it is the reservoir size that impacts our predictive accuracy. We found that the optimal reservoir size is 5000.

Based on our knowledge of non-linear dynamical systems and chaotic systems, we know that these systems are extremely hard, if not impossible to perfectly predict. That being said, we are confident that the ESN does very well at predicting the

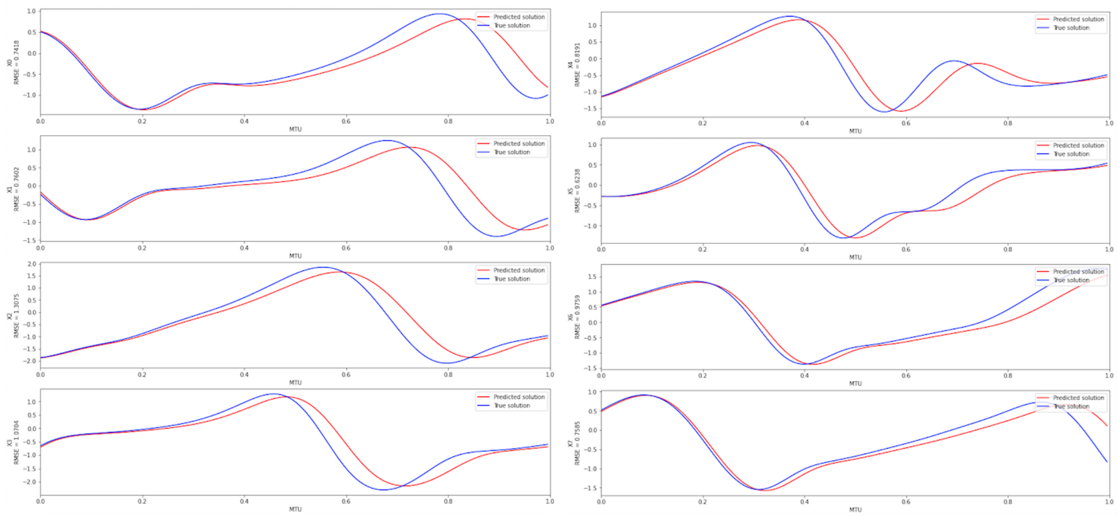

Figure 8: Predicted results of $X(t)$ are shown by the red curve and true results of $X(t)$ are shown by the blue curve for 1 MTU, and a reservoir size of 5000. These results show that ESN does very well at predicting short-term evolution.

short-term evolution of a multi-scale Lorenz 96 system because the mean squared error between the predicted solution and the true solution is the smallest when we make predictions for our system for a shorter number of days. On the contrary, we know that this means that ESN does poorly at long term prediction. We lastly want to note that all code and data files are available upon request.

References

[1] V. Krishnamurthy, "Edward norton lorenz," *Resonance*, vol. 20, no. 3, pp. 191–197, 2015.

[2] Wikipedia, "Lorenz 96 model." `https://en.wikipedia.org/wiki/Lorenz_96_model`, 2022.

[3] T. P. T. Thornes, P. Düben, "On the use of scale-dependent precision in earth system modelling," *Quarterly Journal of the Royal Meteorological Society*, vol. 143, no. 703, pp. 897 – 908, 2017.

[4] D. S. Ashesh Chattopadhyay, Pedram Hassanzadeh, "Data-driven predictions of a multiscale lorenz 96 chaotic system using machine-learning methods: reservoir computing, artificial neural network, and long short-term memory network." `https://npg.copernicus.org/articles/27/373/2020/#section3`,, 2020.

[5] D. S. Ashesh Chattopadhyay, Pedram Hassanzadeh, "Data-driven superparameterization using deep learning: Experimentation with multiscale lorenz 96 systems and transfer learning." `https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020MS002084`,, 2020.

[6] Brittanica, "Atmospheric circulation." `https://www.britannica.com/science/atmospheric-circulationl`, 2022.

[7] K. Erickson, "What is a gravitational wave?." `https://spaceplace.nasa.gov/gravitational-waves/en/`, 2022.