Chaotic System Prediction Using Data Assimilation and Machine Learning

Guo Yanan\(^1\), Cao Xiaoqun\(^{1,2*}\), Peng Kecheng\(^{1,2}\)

\(^1\) College of Computer, National University of Defense Technology, Changsha Hunan 410073, China
\(^2\) College of Meteorology and Oceanography, National University of Defense Technology, Changsha Hunan 410073, China

Abstract. Atmospheric systems are typically chaotic and their chaotic nature is an important limiting factor for weather forecasting and climate prediction. So far, there have been many studies on the simulation and prediction of chaotic systems using numerical simulation methods. However, there are many intractable problems in predicting chaotic systems using numerical simulation methods, such as initial value sensitivity, error accumulation, and unreasonable parameterization of physical processes, which often lead to forecast failure. With the continuous improvement of observational techniques, data assimilation has gradually become an effective method to improve the numerical simulation prediction. In addition, with the advent of big data and the enhancement of computing resources, machine learning has achieved great success. Studies have shown that deep neural networks are capable of mining and extracting the complex physical relationships behind large amounts of data to build very good forecasting models. Therefore, in this paper, we propose a prediction method for chaotic systems that combines deep neural networks and data assimilation. To test the effectiveness of the method, we use the model to perform forecasting experiments on the Lorenz96 model. The experimental results show that the prediction method that combines neural network and data assimilation is very effective in predicting the amount of state of Lorenz96. However, Lorenz96 is a relatively simple model, and our next step will be to continue the experiments on the complex system model to test the effectiveness of the proposed method in this paper and to further optimize and improve the proposed method.

1 Introduction

The prediction of chaotic systems is an important area of research [1-3]. Predicting chaotic systems is a very difficult task because the essential laws of chaotic systems are still unknown and are extremely sensitive to initial values. For complex systems, such as the atmosphere and oceans, even small perturbations can set off a chain of events with very different consequences. While it is not possible to make long-term predictions about chaotic systems, it is possible to make short-term predictions about chaotic systems. Over the years, many researchers have been modeling and predicting chaotic systems through known physical laws [4-7]. Currently, many numerical simulation methods have been used for the numerical simulation of chaotic systems. In the case of weather systems, for example, many numerical weather prediction models, as well as climate models, have now been used for weather forecasting and climate prediction, such as the Weather Research and Forecasting Model (WRF), the Global/Regional Assimilation and Prediction System model (GRAPES), and the Community Earth System Model (CESM) [8-11]. These numerical prediction models have been highly successful and have greatly improved weather and climate prediction. Besides, they have improved the understanding of atmospheric chaotic systems. However, they also face many problems, such as error accumulation, irrational parameterization of physical processes, and huge computational effort.

In recent years, with the advent of big data, data-driven approaches have received more and more attention and research [12-14]. Driven by big data, machine learning has achieved great success, especially deep neural networks have had a great impact on many fields [15-17]. Studies have shown that neural networks have powerful nonlinear fitting capabilities and can learn the complex physical relationships from massive amounts of data to build corresponding predictive models [18-20]. Since deep neural networks have performed well in a large number of prediction tasks, many researchers have started to try to use deep neural networks to predict chaotic time series, and a lot of work has been carried out. This study attempts to apply them to the prediction of chaotic atmospheric systems, using neural networks instead of traditional physics-based models, which can avoid the problems of discretization of control equations and make better use of observed

Corresponding author’s e-mail: caoxiaoqun@nudt.edu.cn
information, making the prediction process more flexible and efficient. However, the problem of initial value sensitivity still exists in the prediction of chaotic systems, so this study still adopts the data assimilation method [21-22] to make full use of the observed data, construct a reasonable initial field, and improve the prediction results.

2 Proposed Method

Artificial Neural Network (ANN) is a mathematical model that mimics biological neural networks [23-24]. ANN consists of a large number of artificial neurons linked together and is used to estimate or approximate functions. Like other machine learning methods, deep neural networks have been used as a nonlinear modeling tool to solve a wide variety of problems, such as image classification and speech recognition. Figure 1 shows a schematic diagram of an artificial neural network. As shown in Figure 1, an artificial neural network contains an input layer, one or more hidden layers, and an output layer. A deep neural network is capable of learning the complex physical relationships behind massive amounts of data. In this study, we train the neural network using historical observations of the chaotic system to obtain a predictive model of the neural network instead of traditional physics-based models.

![Figure 1 Schematic diagram of an artificial neural network.](image)

In this study, the neural network is trained by assimilating the forecast state values to the data, while the data assimilation is trained by fusing the observed data to provide reasonable inputs to the forecast model. The training of the neural network is achieved by minimizing the following loss functions

\[
L(W) = \sum_{k=0}^{K-N-1} \sum_{i=1}^{N} \|N_W^{(i)}(x_k) - x_{k+i}\|_2^2
\]

where \(W\) is the target parameter to be continuously optimized. Through the above combination of neural networks and data assimilation, we predict the state changes of the chaotic system.

3 Experiments and Results

In order to analyze the performance of the method proposed in this study, the Lorenz-96 system [27] was selected for this experiment. Specifically, the Lorenz-96 system is controlled by a set of differential equations, which are mathematically formulated as follows:

\[
\text{\text{E3S Web of Conferences 185, 02025 (2020) ICEEB 2020}}
\]

http://doi.org/10.1051/e3sconf/202018502025

In this study, the neural network is trained by assimilating the forecast state values to the data, while the data assimilation is trained by fusing the observed data to provide reasonable inputs to the forecast model. The training of the neural network is achieved by minimizing the following loss functions

\[
L(W) = \sum_{k=0}^{K-N-1} \sum_{i=1}^{N} \|N_W^{(i)}(x_k) - x_{k+i}\|_2^2
\]

where \(W\) is the target parameter to be continuously optimized. Through the above combination of neural networks and data assimilation, we predict the state changes of the chaotic system.
\[
\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F \tag{4}
\]

where \( x_i \) (0 \(<\) i \(<\) m) is the value of the state variable. \( x_0 = x_m, x_1 = x_{m-1}, x_2 = x_{m-2} \), \( m = 40 \), and \( F = 8 \). The model is integrated using a fourth-order Runge-Kutta scheme with a time step \( dt = 0.001 \), and at the values of \( m \) and \( F \) given above, lorenz96 exhibits chaos. In this study, the number of sets used for data assimilation is \( N = 30 \). The model noise is taken as a normal distribution with a mean of zero and a standard deviation of 0.01. We constructed a neural network containing 6 hidden layers, each containing 80 neurons, and chose ReLU as the activation function. The input and output of the neural network are 40 variables. Specifically, we use the Tensorflow [28] framework to build and train the neural network. During the training process, we set the learning rate to 0.001 and selected Adam optimizer to optimize the loss function. Adam has been widely used as an effective stochastic optimization method. Therefore, by using Adam optimizer in the process of training the model, we can converge to good results faster. Moreover, to prevent the overfitting problem, we used dropout and set the dropout rate to 0.1.

Figure 2 shows the prediction results of the proposed method in this paper for the Lorenz96 system. The green solid line is the true state value, the red dashed line is the prediction result, and the blue scatter is the observed data. From Figure 2, it can be seen that the prediction model combining data assimilation and machine learning can predict the Lorenz96 system well. In this experiment, the prediction time can reach two thousand steps, and the prediction results are still very close to the real results. This is due to the strong fitting ability of the neural network and the data assimilation method can make full use of the observed information to provide high quality input data for the neural network prediction model. The experimental results show that the combination of machine learning and data assimilation has great potential in the prediction of chaotic systems and deserves further improvement and development.

![Figure 2](image_url)

**Figure 2** The true data and prediction results of the proposed method.

## 4 Conclusions

To improve the prediction of chaotic systems, a hybrid prediction model combining machine learning and data assimilation is proposed in this paper. The method uses neural networks to construct prediction models instead of physics-based models, in addition, data assimilation provides reasonable input for neural network, and the output of neural network model is used for the next round of data assimilation. We select the Lorenz-96 system to test the method proposed in this paper. The test results show that the method proposed in this paper can predict the amount of state of the Lorenz96 system well, and the prediction results are close to the true value over a certain time scale. However, it should be noted that the method in this paper needs to be validated on more chaotic systems. Moreover, there are many details to be further improved in this method, which has a lot of room for improvement. However, it is clear that the combination of machine learning and data assimilation methods is important to improve the prediction of chaotic systems and deserves continuous in-depth research.

## Acknowledgments

1. National key research and development program of China (Grant of No. 2018YFC1506704)
2. This research was funded by National Natural Science Foundation of China (Grant No.41475094) and the National Key R&D Program of China (Grant No.2018YFC1506704).

## References

1. Samanta B. Prediction of chaotic time series using computational intelligence[J]. Expert Systems with Applications, 2011, 38(9): 11406-11411.
2. Jian-Cheng S, Ya-Tong Z, Jian-Guo L. Prediction of chaotic systems with multidimensional recurrent least squares support vector machines[J]. Chinese Physics, 2006, 15(6): 1208.
3. Pathak J, Hunt B, Girvan M, et al. Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach[J]. Physical review letters, 2018, 120(2): 024102.
4. Tél T, Gruiz M. Chaotic dynamics: an introduction based on classical mechanics[M]. Cambridge University Press, 2006.
5. Laws of chaos: invariant measures and dynamical systems in one dimension[M]. Springer Science & Business Media, 2012.
6. Quantum chaos: between order and disorder[M]. Cambridge University Press, 2006.
7. Peitgen H O, Jürgens H, Saupe D. Chaos and fractals: new frontiers of science[M]. Springer Science & Business Media, 2006.
8. Powers J G, Klepp J B, Skamarock W C, et al. The weather research and forecasting model: Overview, system efforts, and future directions[J]. Bulletin of the American Meteorological Society, 2017, 98(8): 1717-1737.
9. Zhang R H, Shen X S. On the development of the GRAPES—A new generation of the national operational NWP system in China[J]. Chinese Science Bulletin, 2008, 53(22): 3429-3432.
10. Chen D H, Xue J S, Yang X S, et al. New generation of multi-scale NWP system (GRAPES): general scientific design[J]. Chinese Science Bulletin, 2008, 53(22): 3433-3445.
11. Kay J E, Deser C, Phillips A, et al. The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability[J]. Bulletin of the American Meteorological Society, 2015, 96(8): 1333-1349.
12. Kusiak A, Zhang Z. Short-horizon prediction of wind power: A data-driven approach[J]. IEEE Transactions on Energy Conversion, 2010, 25(4): 1112-1122.
13. Zio E, Di Maio F, Stasi M. A data-driven approach for predicting failure scenarios in nuclear systems[J]. Annals of Nuclear Energy, 2010, 37(4): 482-491.
14. Formentin S, Karimi A. A data-driven approach to mixed-sensitivity control with application to an active suspension system[J]. IEEE Transactions on Industrial Informatics, 2012, 9(4): 2293-2300.
15. Szegedy C, Toshev A, Erhan D. Deep neural networks for object detection[C]//Advances in neural information processing systems. 2013: 2553-2561.
16. Yegnanarayana B. Artificial neural networks[M]. PHI Learning Pvt. Ltd., 2009.
17. Nielsen M A. Neural networks and deep learning[M]. San Francisco, CA: Determination press, 2015.
18. Nelles O. Nonlinear system identification: from classical approaches to neural networks and fuzzy models[M]. Springer Science & Business Media, 2013.
19. Saxe A M, McClelland J L, Ganguli S. Exact solutions to the nonlinear dynamics of learning in deep linear neural networks[J]. arXiv preprint arXiv:1312.6120, 2013.
20. Pan Y, Wang J. Model predictive control of unknown nonlinear dynamical systems based on recurrent neural networks[J]. IEEE Transactions on Industrial Electronics, 2011, 59(8): 3089-3101.
21. Evensen G. Data assimilation: the ensemble Kalman filter[M]. Springer Science & Business Media, 2009.
22. Reichle R H. Data assimilation methods in the Earth sciences[J]. Advances in water resources, 2008, 31(11): 1411-1418.
23. Goodfellow I, Bengio Y, Courville A. Deep learning[M]. MIT press, 2016.
24. LeCun Y, Bengio Y, Hinton G. Deep learning[J]. nature, 2015, 521(7553): 436-444.
25. Bocquet M. Ensemble Kalman filtering without the intrinsic need for inflation[J]. 2011.
26. Bocquet M, Raanes P N, Hannart A. Expanding the validity of the ensemble Kalman filter without the intrinsic need for inflation[J]. Nonlinear Processes in Geophysics, 2015, 22(6): 645.
27. Fertig E J, Harlim J, Hunt B R. A comparative study of 4D-VAR and a 4D ensemble Kalman filter: Perfect model simulations with Lorenz-96[J]. Tellus A: Dynamic Meteorology and Oceanography, 2007, 59(1): 96-100.
28. Ramosundar B, Zadeh R B. TensorFlow for deep learning: from linear regression to reinforcement learning[M]. " O'Reilly Media, Inc.", 2018.