Constructing differential equations using only a scalar time-series about continuous time chaotic dynamics

Natsuki Tsutsumi,¹ Kengo Nakai,² and Yoshitaka Saiki³

¹Faculty of Commerce and Management, Hitotsubashi University, Tokyo 186-8601, Japan
²Faculty of Marine Technology, Tokyo University of Marine Science and Technology, Tokyo 135-8533, Japan
³Graduate School of Business Administration, Hitotsubashi University, Tokyo 186-8601, Japan

We propose a simple method of constructing a system of differential equations of chaotic behavior based on the regression only from a scalar observable time-series data. The estimated system enables us to reconstruct invariant sets and statistical properties as well as to infer short time-series. Our successful modeling relies on the introduction of a set of Gaussian radial basis functions to capture local structure. The proposed method is used to construct a system of ordinary differential equations whose orbit reconstructs a time-series of a variable of the well-known Lorenz system as a simple but typical example. A system for a macroscopic fluid variable is also constructed.

Keywords: data driven modeling | dynamical systems | linear regression | model discovery | fluid flow

Time series data governed by some deterministic law are everywhere around us. However, in most cases we do not know the governing law which enables us to predict behavior in the future. Hence, the identification of the law, a dynamical system model which governs the observable behavior is one of the primary issues in science and technology. In this paper we propose a simple but sophisticated method of modeling a dynamical system as differential equations only from a scalar time-series, when the dynamics is governed by a continuous time chaotic system. The method we propose is applicable even when the observable number of variables is limited, which is often the case in the real world application. Our method is expected to be used to model a complex behavior in various fields such as climate science, mechanical engineering and biology, especially when a governing law is unknown.

I. INTRODUCTION

Since the 1980s techniques of chaotic time-series analysis have been developed, and they are used to analyze experimental and observation data. When the number of observable time series data is limited or less than the dimension of the corresponding dynamics as is often the case in the real world application, the delay embedding method¹ can be used for the reconstruction of an attractor. The Lyapunov spectrum, the degree of stability, is also computed from observable time series data³. From a time series we can obtain the existence of nonlinearity behind the background dynamics using the method of surrogate data⁴. Using the Gaussian radial basis functions time series prediction has been attempted⁵.

We are often eager to identify a dynamical system model which generates an observable time series data. A dynamical model enables us to predict a behavior in the near future. Furthermore, by using the model we can recognize a rare event which has not been occurred before. However, it is often hard to derive a governing system of complex dynamics. In particular, derivation of a governing system for macroscopic dynamics has been one of the fundamental unsolved problems because of the existence of nonlinear interactions in microscopic dynamics.

Recently several approaches have been proposed concerning modeling dynamics from given time-series data using machine learning. These modern approaches employ high performance computations. Neural ODEs⁶ recently attracts much attention for using a neural network to model ODE. Reservoir computing is a recurrent neural network that is widely used for time-series inference⁷,⁸. The method employs relatively large number of nodes in compared to the dimension of the original dynamical system. It does not employ differential equations. Some studies⁹,¹⁰ estimate ODEs having essential structures using simple polynomial terms behind time series. Champion et al.¹¹ recently proposes a method which simultaneously learns the governing equations and the associated coordinate system by using deep neural networks and sparse identification of nonlinear dynamics. Their focuses are in finding the essential dynamical system with the fewest terms necessary to describe the dynamics behind a given data. They succeed in clarifying a mathematical skeleton behind the dynamics. Model variables in their approach are not physically understandable.

There are simple approaches which model a system of differential equations using an observable variable without using neural networks. Baake et al.¹² estimates a system of ODE by polynomial fittings. A higher order polynomial estimations of ODE system are attempted using a regression with Lasso regularization¹³,¹⁴. These methods can be applicable only under the condition that the time derivative of each variable is approximately equal to the low order polynomials of variables. When we have no knowledge of the choice of variables of the governing system, the above condition will not be satisfied.

We propose a simple method of constructing a system of differential equations (ODEs) of chaotic behavior based on the regression only from observable time-series data, which is applicable even when the time derivative of each variable is not approximated by the low order polynomials of variables. Our approach to a model construction has various advantages:

---

¹Electronic mail: yoshi.saiki@r.hit-u.ac.jp
Throughout the paper to simplify explanations the number of variable from given time-series data unless all the components of the \( D \) variables. Hence we can identify the invariant sets such as the fixed points and the periodic orbits. We are unable to reconstruct the original system Eq. (1) itself from given time-series data unless all the components of the variable \( x \) in Eq. (1) and their time-series are known. Hence, we assume that there exists a system of \( D \) dimensional ODEs:

\[
\frac{dX}{dt} = F(X),
\]

where the first \( I \) components of the variable \( X \) are \( X_i = w_i \) \((i = 1, \ldots, I)\) and the rest of \( X \) \((i = I + 1, \ldots, D)\) are created from the delay-coordinates of some \( w_i \). The aim of our proposed method is to model Eq. (3), which is a dynamical system that can describe the behavior of the observable variables \( w_i \) \((i = 1, \ldots, I)\) as components of the variable \( X \). Throughout the paper to simplify explanations the number of observable variables \( I \) is fixed as 1 which is less than the effective dimension of the dynamics to be modeled. In this paper, we focus on cases where the right-hand side of Eq. (3) cannot be described by low-order polynomials and present two examples. In the first example, we construct a system of differential equations which describes the dynamics of an \( x \) variable of the Lorenz system\(^{20}\). \( dx/dt = 10(y - x) \), \( dy/dt = 28x - y - xz \), \( dz/dt = xy - 8z/3 \), when the observable time-series is limited to that of the variable \( x \). We call the Lorenz system the original system corresponding to Eq. (1), and a trajectory created from it is called an actual trajectory. At least three variables are required to model a chaotic attractor of autonomous ODEs. Then we construct a system of ODEs of such \( D = 3 \) dimensional variables. Remark that we have no knowledge of the appropriate value of the dimension \( D \) and the degree of nonlinearity. In the second example, we construct a fluid dynamics model. We apply the same proposed method to construct an unknown system of ODEs that governs the time development of a macroscopic fluid variable\(^{20}\).

The rest of this paper is organized as follows. In Section II we present the proposed method for deriving a system of ODEs. In Section III we apply the proposed method to the first example, where the only observable is the first variable of the Lorenz system. In Section IV we analyse the basin of attraction of the constructed model in Section III In Section V we model a system of ODEs describing a macroscopic fluid flow for which the governing equations are unknown. Concluding remarks are given in Section VI.

II. METHOD

As shown in Fig. 1 our aim is to construct a system of \( D \) dimensional ODEs Eq. (3) based only on some observable deterministic time-series of discrete time. The steps of the proposed method are outlined below:

A. Choose the delay-coordinates: delay-time and dimension
B. Estimate the time derivative at sample points using the Taylor expansion
C. Choose the basis function\(^{21}\) used in Step D
D. Perform ridge regression at sample points
E. Evaluate the model quality according to the reproducibility of a delay structure

A. Case of limited variables

For practical purposes, as the number of observable variables is small, compared with the dimension of the background dynamical system. We can use the delay-coordinates \((w_1(t), w_1(t - \tau), w_1(t - 2\tau), \ldots, w_1(t - (D - 1)\tau))\) of a limited observable to obtain a system of ODEs. See Sec. III for the choice of \( \tau \).
B. Estimating the time derivative through smoothing the trajectory

We can use the sixth-order Taylor expansion to estimate the time derivative at each sample point $X(t)$ based only on discrete time points of a trajectory $X(t)$ ($t = \ldots, t - \Delta t, t, t + \Delta t, \ldots$):

$$\frac{dX(t)}{dt} \approx \{X(t + 3\Delta t) - 9X(t + 2\Delta t) + 45X(t + \Delta t) - 45X(t - \Delta t) + 9X(t - 2\Delta t) - X(t - 3\Delta t)\}/(60\Delta t),$$

where $\Delta t$ is the time step of sample points along a training trajectory. Lower- and higher-order Taylor expansions will also work well. When the observable data includes noise, we can use $l\Delta t$ for some positive integer $l$ instead of $\Delta t$ to estimate the time derivative, which enables to avoid high frequency oscillations.

C. Choice of a basis function

When we model Eq. (3) by linear regression using limited computational resources, basis functions should be chosen appropriately. After we explain polynomial basis functions used in Wang et al., we introduce Gaussian radial basis functions. We combine them with polynomial basis functions for the basis of our regression.

1. Polynomial basis function

For ODE estimation through regression, one possible choice is the polynomial basis function:

$$F_k(X) \approx \beta^k_0 + \sum_{d=1,\ldots,D} \beta^k_d X_d + \sum_{1 \leq i \leq J} \beta^k_{i,j} X_i X_j + \cdots,$$

where $c(i,j) = iD + j - \frac{1}{2}(i - 1)i$ and $\beta^k$ is a vector whose components are $\beta^k_l$ ($l = 0, 1, 2, \ldots$), and $F_k(X)$ is the $k$th component of $F(X)$.

Wang et al. employ the polynomial basis function to estimate dynamical systems for the Hénon map, Rössler system and Lorenz system based on time-series data. They report that bifurcation diagrams are reconstructed well. They used Lasso regularization for the regression of the coefficients. In this paper we assume no knowledge of variables of a governing equation. Furthermore, the function $F$ in Eq. (3) generally does not have the form of low-order polynomials. We confirmed that the method is not applicable to our problem of ODE construction (see Appendix A). Although it tends to be successful at short-term time-series inference, it tends to fail at reconstructing statistical properties.

2. Gaussian radial basis function

To overcome the difficulty of modeling local structures in complex dynamics, we can approximate $F(X)$ by using the localized Gaussian radial basis function $\phi_j$:

$$F_k(X) \approx \beta^k_0 + \sum_{d=1,\ldots,D} \beta^k_d X_d + \sum_{j=1,\ldots,J} \beta^k_{D+j} \phi_j(X), \quad (5)$$

where $\beta^k$ is a set of estimated parameters and

$$\phi_j(X) = \exp\left(\frac{-||X - c_j||^2}{\sigma^2}\right), \quad (6)$$

where $||\cdot||$ denotes the $l^2$ norm. Here $c_j$ is the coordinate of the $j$th center point ($j = 1, \ldots, J$), and $\sigma^2$ is the parameter that determines the deviation of $\phi_j$. The parameter $c_j$ is distributed as lattice points with the grid size $\delta_{grid}$. In this paper, $\sigma^2$ is approximately equal to $1.7373\delta_{grid}$. Figure 2 shows the shape of $\phi_j$ in Eq. (6) for $D = 1$ or $2$ with $\delta_{grid} = 0.25$. See Kawano and Konishi for more details.

D. Ridge regression

To estimate a model of Eq. (3), we can apply least squares regression with an appropriate basis function using ridge regularization. Here we briefly review the regression method. We can estimate the coefficients $\beta$ of a linear model $y_j \approx a^T \cdot \beta$, where $y_j$ is a normalized target variable, $a_i$ is a set of normalized explanatory variables and $^T$ represents the transpose of the matrix. For this model, the ridge estimator is realized by minimizing the following function $L(b)$:

$$L(b) = \frac{1}{2n}||y - Ab||^2 + \frac{\lambda}{2}||b||^2, \quad (7)$$

where $n$ is the size of data, $y = [y_1, y_2, \ldots, y_n]^T$, $A = [a_1, a_2, \ldots, a_n]^T$, and $\lambda$ is a parameter determining the strength of regularization. The ridge estimator can be writ-
where $I$ is the identity matrix. When we estimate the system of ODEs Eq. (3), $y$ represents the time derivative, and $A$ corresponds to $\{x_{1,i}, \ldots, x_{D,i}, \phi_1(X_i), \ldots, \phi_J(X_i)\}_{i=1,\ldots,n}$ (see Fig. 5 in Sec. III).}

### E. Evaluation of the model quality

For the model to be appropriate we can confirm that the delay structure $(w_1(t), w_1(t-\tau), w_1(t-2\tau), \ldots, w_1(t-(D-1)\tau))$ is reconstructed in the model trajectory. See Figs. 6 and 7 in Section III and Fig. 10 in Section V. The hyper-parameters such as regularization parameter and grid size are chosen appropriately based on the degree of the reconstruction.

### III. CONSTRUCTION OF DATA-DRIVEN ODES: LORENZ DYNAMICS

As the first example we construct a system of ODEs that describes the behavior of the Lorenz system. We assume the number of observable variables in $y$ is $I = 1$, and the variable is $w_1 = x$, where $x$ is a variable of the original Lorenz system. We set the model coordinate $X$ in Eq. (3) as $X(t) = (w_1(t), w_1(t-\tau), w_1(t-2\tau))$, where $\tau = 0.13$ ($D = 3$ in Eq. (3)). The time length $T$ of the training time-series is 5000 (time step $\Delta t = 0.005$), and 2% are chosen as the sample points for regression. The number of center points $c_j (j = 1, \ldots, J)$ of the Gaussian radial basis function is 1806 with a grid size of about 2 corresponding to the product of the standardization coefficient and normalized grid size ($\approx 7.9261 \cdot 0.25$). The regularization parameter $\lambda = 10^{-7}$.

We evaluate a model from various aspects. After confirming that the constructed model has a trajectory approximate to that created from the actual model, we investigate invariant sets such as the chaotic attractor and fixed points. We also confirm the reconstruction of the delay structure among components among a variable $X$.

**Short-term orbit.** We found that a time-series inference of $x$ can successfully be applied for a short time under many initial conditions. Figure 3 shows an example trajectory. The long-term time-series inference inevitably fails because of the chaotic property of the Lorenz system.

**Attractor and invariant density.** We find it difficult to represent long-term behavior accurately from a model constructed using only polynomial basis functions alone for the regression (see Appendix A). Here, we show that a model constructed using the proposed method can reconstruct a statistical quantity. Figure 4 shows long-term trajectories. Figure 5 shows an agreement of the density distributions computed from a model trajectory $\{X_1(t)\}$ and a trajectory $\{x(t)\}$ of the original Lorenz system.

The Lyapunov exponents are used to evaluate the degree of instability. We find that the positive and zero exponents of the model are $\lambda_1 = 0.998$ and $\lambda_2 = 0.000$, which agree well with those of the original Lorenz system ($\lambda_1 = 0.906$ and $\lambda_2 = 0.000$). The third negative exponent depends on the choice of coordinates; thus, the disagreement between that of the model ($\lambda_3 = -5.730$) and that of the original Lorenz system ($\lambda_3 = -14.572$) is reasonable.

**Delay structure.** As we employ the delay-coordinate with $D = 3$ and $\tau = 0.13$ for the variable $X$ of a model, the relation $X_1(t) \approx X_2(t+\tau) \approx X_3(t+2\tau)$ should hold for a model to be appropriate. Figure 6 shows time-series $X_1(t)$, $X_2(t+\tau)$ and $X_3(t+2\tau)$ that satisfy the relation. Figure 7 confirms that the density distributions of $X_1(t) - X_2(t+\tau)$ and $X_2(t) - X_3(t+\tau)$ are each localized around zero, which suggests the successful reconstruction of the delay structure. Note that the delay structure is not reconstructed well when a ridge regularization parameter for the regression is not chosen appropriately (see Appendix B).
The difference in area between that of the original Lorenz system and the model orbit deviates from the orbit of the original system. Even though the model orbit deviates from the orbit of the original system, the delay structure among the three variables is reconstructed by the model. The鬼固定点 in relation to the basin of attraction of the Lorenz system, the other two do not exist in the original system. The coordinates for the fixed points of the original system projected onto the delay-coordinate are shown to be close to each other, which confirms that the delay structure among the three variables is reconstructed by the model. Even though the model orbit deviates from the orbit of the original system, the delay structure among the three variables is reconstructed by the model. The Greek fixed points in relation to the basin of attraction of the Lorenz system, the other two do not exist in the original system. The coordinates for the fixed points of the original system projected onto the delay-coordinate are shown to be close to each other, which confirms that the delay structure among the three variables is reconstructed by the model.

**IV. BASIN OF THE MODEL ATTRACTOR**

We compute the basin of attraction of the constructed data-driven model of the Lorenz system in order to capture the validity and the limitation of the model. A set of points that are attracted to the attractor of a model is called the basin. The basin of attraction of the original Lorenz system is known to be $\mathbb{R}^3$. The white region in Fig. 8 shows the basin of attraction together with the model attractor and the five fixed points of the data-driven model. Although three of the fixed points embedded in the attractor correspond to those of the original Lorenz system, the other two do not exist in the original system. We find that the two fixed points $L_{model}$ and $GR_{model}$ with two stable directions and one unstable direction in Table 1 are on the basin boundary formed by the stable manifold of each of the two fixed points (See Fig. 9). The computation of trajectories of the original system under $10^{10}$ initial conditions confirm that, in $(X_1, X_2, X_3)$-coordinates, at least one of the trajectories visit approximately 99.7% of the $80^3$ boxes in $[-20, 20]^3$. We confirm that a model for $D = 4$ with
For the second example, we construct a system of ODEs that governs macroscopic behavior of a fluid flow, which is known to be a difficult task. Macroscopic data of a fluid flow is obtained from a direct numerical simulation of the Navier–Stokes equation. We assume the number of observable variables in Eq. (2) is \( I = 1 \). The observable variable is \( w_1(t) = E(3, t) \) corresponding to the energy at a certain wavenumber range which is normalized in our analysis. See Nakai and Saiki. We set the model variables \( \mathbf{X} \) in Eq. (3) as \( \mathbf{X}(t) = (w_1(t), w_1(t - \tau), \ldots, w_1(t - 7\tau)) \), where \( \tau = 1.5 \) (\( D = 8 \) in Eq. (3)). Note that we do not have a knowledge of the Navier–Stokes equation during the construction procedure. The time length \( T \) of the training time-series is \( 10^5 \) (time step \( \Delta t = 0.05 \)), 2% of which are chosen as the sample points for regression. The number of center points \( c_j(j = 1, \ldots, J) \) of the Gaussian radial basis function is 922, 598 with the grid size of 0.25. The number of grids is restricted by our limited computational resources. The regularization parameter \( \lambda = 10^{-4} \).

**Short-term trajectories and the delay structure.** We find that time-series inference is successful for some time by integrating the obtained model. Two panels in Fig. (top) confirm that a single model could infer the time-series of \( E(3, t) \) under different initial conditions. As we employ the delay-coordinate as a model variable, we expect \( X_1(t) \approx X_2(t + \tau) \approx \cdots \approx X_8(t + 7\tau) \). The relation is confirmed for the short time-series of \( X_1(t), X_2(t + \tau), \cdots, X_8(t + 7\tau) \) as shown in Fig. (bottom).

**VI. CONCLUDING REMARKS**

We propose a simple method of constructing a system of differential equations by using observable time-series data. A model constructed using our method enables to reconstruct a long term dynamics, and the basin structures of the model are studied. We exemplify that the method can be applicable for modeling a macroscopic fluid flow as well as the chaotic property, but continues to satisfy the time delay structure among variables.
Lorenz system even when the number of observable variable is limited to one. As a future work, it is interesting to confirm the applicability of our method to various types of chaotic dynamics including an infinite dimensional one.

Appendix A: Failure of constructing ODEs by polynomial regression

![Diagrams](image1.png)

**FIG. 11. Short-term trajectories of two models.** The left and right panels show short-term trajectories of a model by using only polynomial terms and of the main model, each of which is plotted together with the corresponding trajectory of the original system.

![Diagrams](image2.png)

**FIG. 12. Density distributions of** $X_1$ **for two models together with that of the original system: using only polynomial terms (left) and the main model (right).** The errors in area from the distribution of the original system are 0.0235 (left) and 0.0042 (right).

For the estimation of ODE in Eq. (3) through regression, one possible choice is the regression using only polynomial basis functions Eq. (4).

Here, we present the difficulty of constructing ODEs by regression using only polynomial terms. More precisely, inferring a time-series in the short-term is relatively easy, but the density distribution is difficult to reconstruct because of the difficulty with modeling long-term behavior.

Assume that the variable $X(t)$ follows Eq. (3) and $F$ is sufficiently smooth. We can estimate $F$ by regression using only polynomial basis functions Eq. (4). In fact, previous researchers have reported that simple systems such as the Lorenz system and Rössler system can be successfully estimated from time-series when $F$ is written using quadratic terms. However, our aim is to construct ODEs when $F$ is not necessarily written by low order polynomials. We apply 8th order polynomial regression to estimate a model with the parameter $\lambda = 10^{-6.2}$, where $\lambda$ is a coefficient of ridge regularization. Figures [11] and [12] confirm that the model can reproduce a short orbit but cannot approximate a density distribution created from a long orbit. Figure [13] shows that the relative regression error is larger for various $X$ than for the main model. This implies that localized structures cannot be captured well polynomial terms up to the eight order at least.

In theory the use of higher order polynomial terms will improve the model quality. However, in practice it is difficult to construct a model because of the computational resources and the difficulty in the choice of regularization parameter in ridge regression.

Appendix B: Choice of a regularization parameter from the reproducibility of the delay structure

We choose a regularization parameter in Eq. (7) based on the reproducibility of the delay structure. In constructing a ODE system we introduce delay-coordinate variables of an observable variable when the number of observable variables is not large relative to the dimension of the background dynamical system. We expect that the constructed model could recover delay relations among variables. For the first example of modeling the Lorenz dynamics, the relation $X_1(t) \approx X_2(t + \tau) \approx X_3(t + 2\tau)$ should hold for the model to be appropriate. This relation is confirmed by the short time-series of $X_1(t)$, $X_2(t + \tau)$, $X_3(t + 2\tau)$ shown in Fig. [6] and by the distributions of $X_1(t) - X_2(t + \tau)$, $X_2(t) - X_3(t + \tau)$ having sharp peaks around zeros as shown in Fig. [7]. Figures [14] and [15] show the attractor and density distributions of $X_1(t)$, $X_1(t) - X_2(t + \tau)$ and $X_2(t) - X_3(t + \tau)$ of the model when $\lambda = 10^{-3.9}$. These figures suggest that the model quality is worse with $\lambda = 10^{-3.9}$ than with $\lambda = 10^{-7}$, although the degrees of the model reconstruction are similar.
FIG. 14. Projection of a long-term trajectory to \((X_1, X_2, X_3)\). Approximation of an attractor of a constructed model with the regularization parameter \(\lambda = 10^{-3.9}\) (right) and that of a trajectory of the original Lorenz system (left) projected onto the delay coordinate.

FIG. 15. Density distributions for a model with the regularization parameter \(\lambda = 10^{-3.9}\). The discrepancy in density distributions of \(X_1\) along the trajectories of the model and the original system (left), and the distribution of \(X_2(t) - X_2(t + \tau) (\sigma = 0.0196)\) (center), and that of \(X_2(t) - X_2(t + \tau) (\sigma = 0.0128)\) (right) of the model. The discrepancy shown in the left panel is consistent with the fact that the delay relations are not satisfied well in comparison with Fig. 7 for a model with the regularization parameter \(\lambda = 10^{-3.9}\).

ACKNOWLEDGEMENTS.

YS was supported by JSPS KAKENHI (19KK0067, 21K18584). KN was supported by JSPS KAKENHI (22K17965) and the Project of President Discretionary Budget of TUMST. Part of the computation was supported by JP-PCN (jih210027, jh220007), HPCI (hp210072), and the Collaborative Research Program for Young-Women Scientists of ACCMS and IMIC, Kyoto University.

1F. Takens, “Detecting strange attractors in turbulence.” (Springer Berlin Heidelberg, Berlin, Heidelberg, 1981) pp. 366–381.
2T. Sauer, J. A. Yorke, and M. Casdagli, “Embedology,” J. Stat. Phys. 65, 579–616 (1991).
3M. Sano and Y. Sawada, “Measurement of the lyapunov spectrum from a chaotic time series,” Physical review letters 55, 1082 (1985).
4A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, “Determining lyapunov exponents from a time series,” Physica D: nonlinear phenomena 16, 285–317 (1985).
5J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, “Testing for nonlinearity in time series: the method of surrogate data,” Physica D: Nonlinear Phenomena 58, 77–94 (1992).
6M. J. Powell, “Radial basis functions for multivariable interpolation: a review,” Algorithms for approximation (1987).
7M. Casdagli, “Nonlinear prediction of chaotic time series,” Physica D: Nonlinear Phenomena 35, 335–356 (1989).
8R. T. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, “Neural ordinary differential equations,” Advances in Neural Information Processing Systems , 6571–6583 (2018).
9J. Pathak, Z. Lu, B. R. Hunt, M. Girvan, and E. Ott, “Using machine learning to replicate chaotic attractors and calculate lyapunov exponents from data,” Chaos 27, 121102 (2017).
10K. Nakai and Y. Saiki, “Machine-learning inference of fluid variables from data using reservoir computing.” Phys. Rev. E 98, 023111 (2018).
11K. Nakai and Y. Saiki, “Machine-learning construction of a model for a macroscopic fluid variable using the delay-coordinate of a scalar observable,” Discrete Contin. Dyn. Syst. S 14, 1079–1092 (2021).
12M. U. Kobayashi, K. Nakai, Y. Saiki, and N. Tsutsumi, “Dynamical system analysis of a data-driven model constructed by reservoir computing,” Phys. Rev. E 104, 044215 (2021).
13S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, “Chaos as an intermittently forced linear system,” Nature communications 8, 1–9 (2017).
14Y. Yuan, X. Tang, W. Zhou, W. Pan, X. Li, H.-T. Zhang, H. Ding, and J. Goncalves, “Data driven discovery of cyber physical systems,” Nature communications 10, 1–9 (2019).
15K. Champion, B. Lusch, J. N. Kutz, and S. L. Brunton, “Data-driven discovery of coordinates and governing equations,” Proceedings of the National Academy of Sciences 116, 22445–22451 (2019).
16E. Baake, M. Baake, H. G. Bock, and K. M. Briggs, “Fitting ordinary differential equations to chaotic data,” Phys. Rev. A 45, 5524 (1992).
17W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and C. Grebogi, “Predicting catastrophes in nonlinear dynamical systems by compressive sensing,” Phys. Rev. Lett. 106, 154101 (2011).
18S. L. Brunton, J. L. Proctor, and J. N. Kutz, “Discovering governing equations from data by sparse identification of nonlinear dynamical systems,” Proceedings of the National Academy of Sciences 113, 3932–3937 (2016).
19Almost all the deterministic chaotic time series of continuous time are governed by a differential equation. Such a differential equation can be autonomous differential equation, non-autonomous differential equation, partial differential equation, delay differential equation. They can be approximated well by ordinary differential equations if they have finite dimensional attractors.
20E. N. Lorenz, “Deterministic nonperiodic flow,” J. Atmos. Sci. 20, 130–141 (1963).
21We do not assume some particular vector space in advance as is usual in the literature for machine learning.
22The choice of \(\sigma^2\) is obtained from
\[
\sigma^2 := \frac{(m - 1) \delta_{\text{pred}}}{-\log_p p},
\]
where \(m\) is the degree of the corresponding B-spline basis function and \(p (>0)\) is a small value. We consider \(c_j\) only if there exists a data point in the \((m - 1) \delta_{\text{pred}}\)-neighborhood. We set \(m = 3\), and \(p = 0.1\), then \((m - 1)^2 / (-\log_p p) \approx 1.7372\).
23S. Kawano and S. Konishi, “Nonlinear regression modeling via regularized gaussian basis function,” Bulletin of Informatics and Cybernetics 39, 83–96 (2007).
24The auto-correlation is referred to for choosing the delay-coordinate. In this case the correlation coefficient of \(w_1(t)\) and \(w_1(t - \tau) = 0.79\), and that of \(w_1(t)\) and \(w_1(t - 2\tau) = 0.46\).
25R. Temam, Infinite-dimensional dynamical systems in mechanics and physics, Vol. 68 (Springer Science & Business Media, 2012)