

流体の一変数時系列データを用いた 機械学習によるモデリング

東京大学大学院 数理科学研究科
中井 拳吾 (Kengo NAKAI) *

1 導入.

機械学習は物理の様々な分野で注目されている. 近年, 機械学習の一種で, リカーレントネットワークをもったリザーバーコンピューティング [6, 9] が時系列やリャプノフ指数などの予測において有効であることが報告されている [1, 2, 3, 9, 11, 12, 13]. 我々はこの機械学習手法を用いて流体のエネルギー変数のモデルを構成し, 予想を成功させている [7].

リザーバーコンピューティングにおける学習は入力データから得られた中間層のベクトルと出力するベクトルを線形にフィッティングさせることに重点が置かれている. この学習では入力データから中間層のベクトルを得るために用いるニューラルネットワーク構造を学習しないため, フィッティングにかかる計算コストを減らしている.

そこで, 本誌ではリザーバーコンピューティングの手法を用いて流体のスカラーマクロ変数のモデルの構成をする. ただし, ここでは時間遅れ座標系を使うことで予想の精度を上げた. リザーバーコンピューティングの説明をしたのちに流体変数としてレイノルズ数についての構成したモデルについて述べる.

尚, 本稿で述べる研究内容は齊木 吉隆氏 (一橋大学) との共同研究 [7, 8] に基づく.

2 リザーバーコンピューティング.

$d\phi/dt = \mathbf{f}(\phi)$ で表される力学系の変数, $\mathbf{u} = \mathbf{h}_1(\phi) \in \mathbb{R}^M$ と $\mathbf{s} = \mathbf{h}_2(\phi) \in \mathbb{R}^P$ について考える. ただし, 時系列データ $\mathbf{u}$ は $\mathbf{s}$ の時系列を予測する時刻にデータはないとする. この予測する時刻に $\mathbf{u}$ の時系列データがない場合は Pathak らが行った考察 [11, 12] の特殊な場合に相当する.

リザーバーベクトル $\mathbf{r} \in \mathbb{R}^N$ ($N \gg M$) は次の式で支配されている.

$$\mathbf{r}(t + \Delta t) = (1 - \alpha)\mathbf{r}(t) + \alpha \tanh(\mathbf{A}\mathbf{r}(t) + \mathbf{W}_{\text{in}}\mathbf{u}(t)), \quad (1)$$

ただし, Δt は時間幅を表す. また, 行列 $\mathbf{A}$, $\mathbf{W}_{\text{in}}$ は $N \times N$, $N \times M$ 行列である. パラメータ α ($0 < \alpha \leq 1$) は $\mathbf{r}$ の力学系の非線形性を表し, 式 (1) の Δt に依存して決める. リザーバーベクトルは $(0, 1]$ の一様乱数から定めた $\mathbf{r}(-\tau)$ を初期値とする. ただし, $\tau/\Delta t (\gg 1)$ はトランジェント時間

* 本研究は科研費 (課題番号 :19J12482) の助成を受けたものである .

に相当する.

行列 $\mathbf{W}_{\text{in}}$ の各列はただ一つのために $[-\sigma, \sigma]$ の一様乱数から値を与え, それ以外は 0 とする. 行列 $\mathbf{A}$ は $D \times N$ 個の非 0 の要素を持つ疎行列とする. $D \times N$ 個の非 0 成分は $[-1, 1]$ の一様乱数から値を与える. さらに, 行列 $\mathbf{A}$ については最大固有値が ρ になるように行列のすべての要素を調整する.

リザーバーベクトル $\mathbf{r}$ は次のようにして得たいベクトルに変換する.

$$\hat{\mathbf{s}}(t) = \mathbf{W}_{\text{out}} \mathbf{r}(t) + \mathbf{c}. \quad (2)$$

これにより $L = T/\Delta t$ ステップのリザーバーベクトル $\{\mathbf{r}(l\Delta t)\}_{l=1}^L$ を得る.

$0 < t \leq T$ (学習時間と呼ぶ) での 時系列 $\hat{\mathbf{s}}$ が既知の時系列 $\mathbf{s}$ の近似になるように $\mathbf{W}_{\text{out}}$ と $\mathbf{c}$ を決定する. この決定がリザーバーコンピューティングの学習に相当する. これは次の二次形式が最小化するように $\mathbf{W}_{\text{out}}$ と $\mathbf{c}$ を決定することに対応する.

$$\sum_{l=1}^L \|(\mathbf{W}_{\text{out}} \mathbf{r}(l\Delta t) + \mathbf{c}) - \mathbf{s}(l\Delta t)\|^2 + \beta [\text{Tr}(\mathbf{W}_{\text{out}} \mathbf{W}_{\text{out}}^T)], \quad (3)$$

ただし $\mathbf{q}$ に対して $\|\mathbf{q}\|^2 = \mathbf{q}^T \mathbf{q}$ とし, $\beta (\geq 0)$ を含む項は $\mathbf{W}_{\text{out}}$ の過剰適合を避けるために導入している. この決定が妥当であれば, $t > T$ (予測時間と呼ぶ) での $\hat{\mathbf{s}}(t)$ は $\mathbf{s}(t)$ を再現できるはずである. 二次形式 (3) を最小化する解を $\mathbf{W}_{\text{out}}^*$ と $\mathbf{c}^*$ とすると予測時間では

$$\hat{\mathbf{s}}(t) = \mathbf{W}_{\text{out}}^* \mathbf{r}(t) + \mathbf{c}^*, \quad (4)$$

として時系列 $\hat{\mathbf{s}}$ を得る. $\mathbf{W}_{\text{out}}^*$ と $\mathbf{c}^*$ は次のように書くことができることが知られている (see [10] P.140 for details):

$$\begin{aligned} \mathbf{W}_{\text{out}}^* &= \delta \mathbf{S} \delta \mathbf{R}^T (\delta \mathbf{R} \delta \mathbf{R}^T + \beta \mathbf{I})^{-1}, \\ \mathbf{c}^* &= -[\mathbf{W}_{\text{out}}^* \bar{\mathbf{r}} - \bar{\mathbf{s}}], \end{aligned}$$

ただし, $\bar{\mathbf{r}} = \sum_{l=1}^L \mathbf{r}(l\Delta t)/L$, $\bar{\mathbf{s}} = \sum_{l=1}^L \mathbf{s}(l\Delta t)/L$ とし, $\mathbf{I}$ は $N \times N$ の単位行列, $\delta \mathbf{R}$ (rsp, $\delta \mathbf{S}$) は l 行が $\mathbf{r}(l\Delta t) - \bar{\mathbf{r}}$ (rsp, $\mathbf{s}(l\Delta t) - \bar{\mathbf{s}}$) となるような行列とする.

ただし, 予測時間 $t > T$ での時系列データ $\mathbf{u}$ は利用できないため, $\mathbf{u}(t)$ として 式 (1) で得た $\hat{\mathbf{s}}(t) = \hat{E}(\cdot, t)$ を用いる:

$$\mathbf{r}(t + \Delta t) = (1 - \alpha) \mathbf{r}(t) + \alpha \tanh(\mathbf{A} \mathbf{r}(t) + \mathbf{W}_{\text{in}} \hat{\mathbf{s}}(t)).$$

学習に用いる変数 $X(t)$ を対等に扱うために正規化した $\tilde{X}(t)$ を使うことにする:

$$\tilde{X}(t) = [X(t) - X_1]/X_2.$$

ただし, X_1 は平均値, X_2 は偏差値を表す. 予測時間の時系列データ $X(t)$ を $\tilde{X}(t)$ から再現するとき, 学習時間の正規化に用いた X_1 と X_2 を用いる. この正規化より σ の調整を避けることができる.

3 流体.

周期境界条件の下で 3 次元非圧縮 Navier–Stokes 方程式の直接数値計算によって得られた時系列データをリザーバーコンピューティングの学習データとして用いる:

$$\begin{cases} \partial_t v - \nu \Delta v + (v \cdot \nabla)v + \nabla \pi = f, & \nabla \cdot v = 0, & \mathbb{T}^3 \times (0, \infty), \\ v|_{t=0} = v_0 & \text{with } \nabla \cdot v_0 = 0, & \mathbb{T}^3, \end{cases}$$

ただし, $\mathbb{T} = [0, 2\pi)$, $\nu > 0$ は動粘性, $\pi(x, t)$ は圧力, $v(x, t) = (v_1(x, t), v_2(x, t), v_3(x, t))$ は速度とする. 切断波数を $(2 \times 9 + 1)^3$ としてフーリエスペクトル法 [5] を用いる. また, 4 次のルンゲクッタ法により時間積分をする. また, 低周波成分のエネルギーが一定になるように毎時刻外力を加える. さらに低周波成分にのみエネルギーを持つように初期値を定める.

レイノルズ数 R_λ . ここでは, 流れの複雑さを表すテーラーマイクロスケールレイノルズ数に注目する. まず, エネルギー $E(t)$ を以下で定義する,

$$E(t) = \sum_{\kappa \in D} \sum_{\zeta=1}^3 (\mathcal{F}_{[v_\zeta]}(\kappa, t))^2.$$

ただし,

$$\mathcal{F}_{[v_\zeta]}(\kappa, t) := \frac{1}{(2\pi)^3} \int_{\mathbb{T}^3} v_\zeta(x, t) e^{-i(\kappa \cdot x)} dx \quad (\zeta = 1, 2, 3),$$

$D = \{(\kappa_1, \kappa_2, \kappa_3) \in \mathbb{Z}^3 \mid \kappa_1, \kappa_2, \kappa_3 \in [-9, 9]\}$ とする. テーラーマイクロスケールレイノルズ数 $\tilde{R}_\lambda(t)$ [4] を次で定義する:

$$\tilde{R}_\lambda(t) := \frac{\sqrt{(2/3)E(t)\lambda}}{\nu} = \sqrt{\frac{20E(t)^2}{3\nu\epsilon(t)}},$$

ただし, ϵ は次で定義されるもので, 単位体積あたりの時間平均エネルギー損失量を表す.

$$\epsilon(t) = 2\nu \sum_{\kappa \in D} \sum_{\zeta=1}^3 |\kappa|^2 (\mathcal{F}_{[v_\zeta]}(\kappa, t))^2.$$

4 マクロ変数のモデルの構成

2 章で述べたリザーバーコンピューティングによりレイノルズ数 R_λ のモデルを構成する.

4.1 モデルの構成

時間遅れ座標系. M 次元の $\Delta\tau$ 時間遅れ座標系を入力変数 $\mathbf{u}(t) = (u_1(t), u_2(t), \dots, u_M(t))^T$, 出力変数 $\mathbf{s}(t) = (s_1(t), s_2(t), \dots, s_M(t))^T$ として導入する:

$$\mathbf{u}(t) = (\tilde{R}_\lambda(t), \tilde{R}_\lambda(t - \Delta\tau), \dots, \tilde{R}_\lambda(t - (M - 1)\Delta\tau))^T, \quad (5)$$

$$\mathbf{s}(t) = (\tilde{R}_\lambda(t), \tilde{R}_\lambda(t - \Delta\tau), \dots, \tilde{R}_\lambda(t - (M - 1)\Delta\tau))^T. \quad (6)$$

遅れ時間 $\Delta\tau$ や次元 M のとり方については次の章で述べる。

モデルの構成. 適当なパラメータ [8, Table 2] により定めたランダムな行列 $\mathbf{A}$, $\mathbf{W}_{\text{in}}$ に対して行列 $\mathbf{W}_{\text{out}}^*$, $\mathbf{c}^*$ を定める. これらの行列を用いて予想した時系列がある許容範囲 (5 章参照) に入っていた場合, それらの行列によって構成されたものをモデルとして用いる. ただし, 予想時間において t が $t - (M - 1)\Delta\tau < T$ を満たすとき, 入力変数として既知のデータを使うことができるが, ここでは使わないことにする. これ以降この章ではあるひとつの行列の組 $\mathbf{A}$, $\mathbf{W}_{\text{in}}$, $\mathbf{W}_{\text{out}}^*$, $\mathbf{c}^*$ からなるモデルを考える.

4.2 モデルの再現性

構成したモデルについて, いくつかの観点から評価する.

時系列データ.

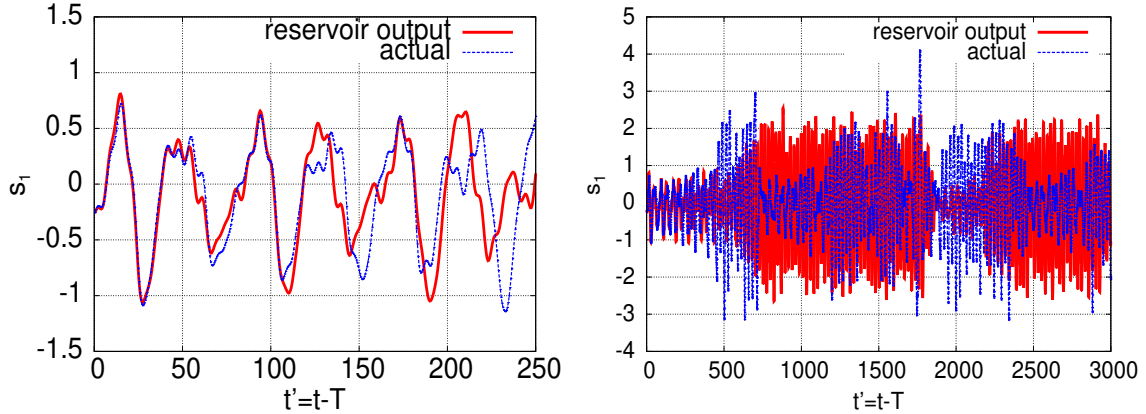


図1 Inference of a time-series of the Reynolds number of a fluid flow. Time-series of $s_1 = \tilde{R}_\lambda$ is inferred from the reservoir model in comparison with that of a reference data obtained by the direct numerical simulation of the Navier-Stokes equation. In the bottom figure switching between laminar state with a small amplitude fluctuation and bursting state with a large amplitude fluctuation appear in an inferred time-series of $s_1 = \tilde{R}_\lambda$, which are observed in the actual time-series. Remark that the model trajectory shows an intermittent behavior on a chaotic set, but after a long transient it will diverge eventually at $t' \approx 290000$.

学習時間直後におけるレイノルズ数 $s_1 = \tilde{R}_\lambda$ の時系列データの予想がうまく行っていることを確認した. 図 1 の左図に, 予測した変数 $\hat{s}_1 = \hat{\tilde{R}}_\lambda(t)$ ($t > T$) の時系列データを Navier-Stokes 方程式の直接数値計算から得られた正解の時系列 $s_1 = \tilde{R}_\lambda(t)$ とあわせて書きだした. 更に, 図 1 の右図に, $\hat{\tilde{R}}_\lambda(t)$ の振る舞いを長時間書きだした. 実際の直接数値計算結果に現れるような, 速度場が穏やかな流れ (ラミナー) から複雑な流れ (バースト) に時間間欠的に移り変わる様子も確認できた.

時間遅れ成分. 時間遅れ座標系の各成分について学習時間では任意の m ($m = 2, \dots, M$) に対して $s_1(t) = s_m(t + (m - 1)\Delta\tau)$ という関係が成立している. 予想時間でも対応する関係が成り立つべきである. そこで, 図 2 に $\hat{s}_1(t)$ と $\hat{s}_{14}(t + 13\Delta\tau)$ を書きだした. この図より $\hat{s}_1(t) \approx \hat{s}_{14}(t + 13\Delta\tau)$ となっていることが見て取れる. また, 別の時刻 t やその他の m についても同様に $\hat{s}_1(t) \approx \hat{s}_m(t + (m - 1)\Delta\tau)$ の関係を満たしていることを確認した. これらの結果から, 我々

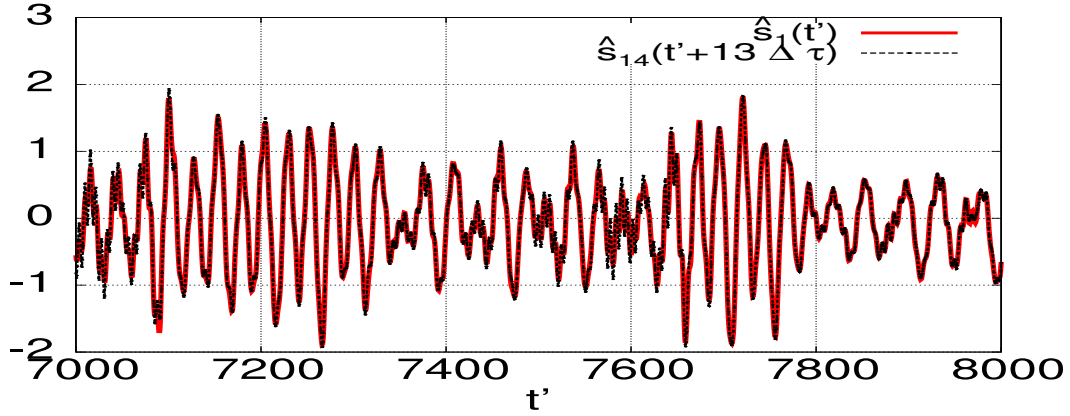


図2 Reproducing the delay property which is to be satisfied for the successfully inferred time-series $\hat{s}$. We observe that for all values of $m = 2, \dots, 14$ and for most t' , $\hat{s}_1(t') \approx \hat{s}_m(t' + (m - 1)\Delta\tau)$, although the time-series of only $\hat{s}_1(t')$ and $\hat{s}_{14}(t' + 13\Delta\tau)$ ($7000 \leq t' \leq 8000$) are shown. The delay property is reproduced only from the training data.

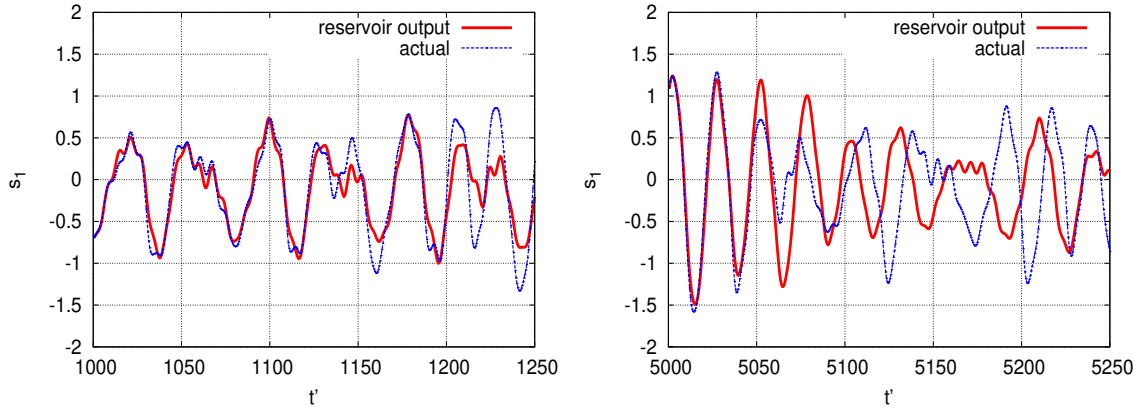


図3 Inference of time-series of the Reynolds number in many time-intervals $T_{\text{out}} < t' < T_{\text{out}} + 250$ ($T_{\text{out}} = 1000, 5000$) using the same reservoir model constructed by using the training data for $t' \leq 0$ (see Fig. 1.) We only change the initial condition for each case, while the model is fixed after the appropriate choice of $\mathbf{W}_{\text{in}}, \mathbf{A}, \mathbf{W}_{\text{out}}^*$ and $\mathbf{c}^*$ is determined by using the training data for $t' < 0$. Each panel shows that the time-series inference is successful by using the same model, although the time-interval for the successful inference is limited to a relatively short time especially for the bursting phase with large fluctuations maybe due to the high instability. These panels suggest that the model can describe the dynamics of the Reynolds number.

の構成したリザーバーモデルは時系列データのみから時間遅れ関係の学習に成功したことがわかる。

異なる時間に対する予測。ある学習時間から得られたリザーバーモデルを使って一定時間の時系列データの振る舞いを予想できることを確認した。ここではそのモデルが異なる時間の時系列データも予想できることを確認する。図3では、いくつかの時間から予測をスタートさせた時の予想した時系列と比較のため正解の時系列とともに書きだした。ある程度は予測が成功していることが見て取れる。

ただし、この予想には図 1 で用いたものと同じリザーバーモデルを用いている。これらの結果から構成したモデルはレイノルズ数の挙動をうまく表現できていることがわかる。

5 時間遅れ座標と遅れ時間の決め方.

リザーバーモデルを構成する際に入力変数 u , 出力変数 s として遅れ時間 $\Delta\tau$ で M 次元の遅れ時間座標系を用いている (eqs. (5),(6)). この章では、遅れ時間 $\Delta\tau$ と次元 M の決め方について述べる。相関係数、レイノルズ数 $\{R_\lambda(t)\}$ の自己相関関数 $C(x)$ を次のように定義する;

$$C(x) = \frac{\frac{1}{J} \sum_{j=0}^{J-1} (R_\lambda(A_j) - \bar{R}_\lambda)(R_\lambda(B_j) - \bar{R}_\lambda)}{\sqrt{\frac{1}{J} \sum_{j=0}^{J-1} (R_\lambda(A_j) - \bar{R}_\lambda)^2} \sqrt{\frac{1}{J} \sum_{j=0}^{J-1} (R_\lambda(B_j) - \bar{R}_\lambda)^2}}, \quad (7)$$

ただし、 $\bar{R}_\lambda$ は $R_\lambda(t)$ の時間平均値、 $(A_j, B_j) = (t_0 + j\Delta t^*, t_0 + j\Delta t^* + x)$ として、 Δt^* を軌道の時間ステップとする。また、時刻 t_0 を軌道の初期時刻とする。図 4 に、自己相関関数 $C(x)$ を書きだした。図 4 (左図) より、 $C(x) = 0.7$ から 0.3 の値をとるのは、 $x \approx 3.0$ から 5.0 の時である。

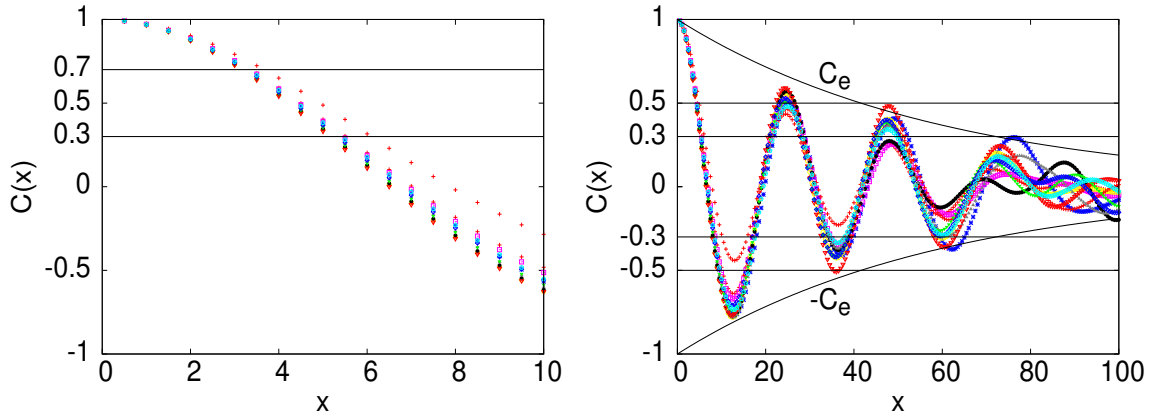


図 4 Auto-correlation function $C(x)$ for a trajectory $\{R_\lambda(t)\}$ with respect to the value of time-delay x (left), and its enlarged figure (right). Auto-correlation function $C(x)$ is shown together with the straight lines $\pm 0.3, \pm 0.5$ (left), and $0.3, 0.7$ (right). Each of the different colors represents $C(x)$ computed from a trajectory from a different initial condition with time-lengths 5000. The difference is mainly due to the intermittent property of the dynamics. In the left panel the envelope $C_e(x) (= \exp(-x/60))$ is shown to go below 0.5 when $x \approx 40$, and also go below 0.3 when $x \approx 75$. From the right panel $C(x)$ is shown to go below 0.7 at the first time, when $x \approx 3.0$, and go below 0.3 at the first time, when $x \approx 5.0$.

$\Delta\tau < 3.0$, ならば、 $R_\lambda(t)$ と $R_\lambda(t - \Delta\tau)$ が似た動きをすることになり、一方で、 $\Delta\tau > 5.0$ ならば、両者の相関が小さくなる。このことから、遅れ時間 $\Delta\tau$ は 3.0 - 5.0 から選ぶと良いと推測される。

実際、 $\Delta\tau = 3.0$ - 5.0 について、最適な遅れ時間 $\Delta\tau$ と遅れ次元 M の組み合わせについて調査した (Table 1). 各 $(\Delta\tau, M)$ について、各々 8160 個の行列 A と W_{in} について学習を行い、予想した

変数 $\hat{s}_1(t)$ と $s_1(t)$ の時刻 $t \in [0, 60](t \in [0, 90])$ の差の最大値が $e_{60} = 0.14(e_{90} = 0.30)$ 以下になるものの個数を数えた。

(a) $(e_{60}, e_{90}) = (0.14, 0.30)$

$\Delta\tau \setminus M$	11	12	13	14	15	16	17	18	19
3.0	0	0	0	0	1	19	24	43	37
3.5	0	0	11	20	28	57	48	21	11
4.0	3	18	43	<u>107</u>	59	21	14	2	4
4.5	14	43	54	21	15	8	1	1	1
5.0	24	26	19	9	1	1	1	0	0

表 1 The number of successful trials for each choice of the delay-time $\Delta\tau$ and the dimension M of the delay-coordinate. The matrices $\mathbf{A}$ and $\mathbf{W}_{\text{in}}$ are chosen randomly, and the number of successful cases are counted. See [8, Table 2] for the parameter values. For each set of values $(\Delta\tau, M)$ we tried 8160 cases of $\mathbf{A}$ and $\mathbf{W}_{\text{in}}$. For each value of $\Delta\tau$, the best choice of M is identified by the bold number(s) (blue), and the best among each criterion is identified by the underlined bold number(s) (red).

$\Delta\tau \approx 4.0$ かつ $M\Delta\tau \approx 55-60$ となるよう $(\Delta\tau, M)$ を定めると、予想のエラーがしきい値以下になる個数が最も多くなることが見て取れる。これらの値はそれぞれ、 $C(\Delta\tau) \approx 0.45-0.55$ と包絡線 $C_e(M\Delta\tau) \approx 0.35-0.40$ に相当する (図 4 (右図))。

6 まとめ

流体のマクロ変数として一変数時系列データであるレイノルズ数の挙動を表現するリザーバーモデルを時間遅れ座標系を使うことで構成した。また、この時間遅れ座標系の遅れ時間 $\Delta\tau$ と遅れ次元 M について、自己相関係数 $C(\Delta\tau)$ が $0.45-0.55$ かつ、包絡線 $C_e(M\Delta\tau)$ が $0.35-0.40$ を満たすように $(\Delta\tau, M)$ を決めると最適であることがわかった。また、このようにして決めた遅れ時間座標系を元に学習したモデルは、時系列データの予想ができ、速度場が穏やかな流れから複雑な流れに時間間欠的に遷移する様子も再現することが確認できた。

参考文献

- [1] P. Antonik, M. Gulina, J. Pauwels, and S. Massar, preprint, (2018), <https://arxiv.org/pdf/1802.02844.pdf>.
- [2] D. Ibáñez-Soria, J. Garcia-Ojalvo, A. Soria-Frisch, and G. Ruffini, Chaos, 28, (2018), pp. 033118.
- [3] M. Inubushi, and K. Yoshimura, Scientific Reports, 7, (2017), pp. 10199.
- [4] T. Ishihara and Y. Kaneda, In Statistical Theories and Computational Approaches to Turbulence (2003), pp. 177–188.
- [5] K. Ishioka, GFD Dennou Club, (1999), <http://www.gfd-dennou.org/arch/ispack/>.

- [6] H. Jaeger, and H. Haas, *Science*, 304, (2004), pp. 78-80.
- [7] K. Nakai, and Y. Saiki, *Physical Review E* 98, 023111:1-6.
- [8] K. Nakai, and Y. Saiki, *Discrete and Continuous Dynamical Systems Series S* (accepted for publications).
- [9] Z. Lu, J. Pathak, B. Hunt, M. Girvan, R. Brockett, and E. Ott, *Chaos*, 27, (2017), pp. 041102.
- [10] M. Lukosevicius, and H. Jaeger, *Computer Science Review*, 3, (2009), pp. 127-149.
- [11] J. Pathak, Z. Lu, B. Hunt, M. Girvan, and E. Ott, *Chaos*, 27, (2017), pp. 121102.
- [12] J. Pathak, B. Hunt, M. Girvan, Z. Lu, and E. Ott, *Physical Review Letters*, 120, (2018), pp. 024102.
- [13] D. Verstraeten, B. Schrauwen, M. D'Haene, and D. A. Stroobandt, *Neural Network*, 20, (2007), pp. 391-403.