Machine-learning prediction of fluid variables
-Reservoir computation-

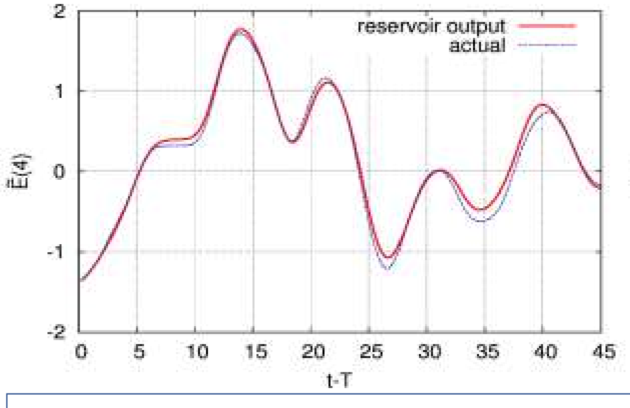
Yoshitaka Saiki
Hitotsubashi University, Japan
JST (Japan Science and Technology Agency) , Japan
University of Maryland, USA

in collaboration with Kengo Nakai (University of Tokyo, Japan)

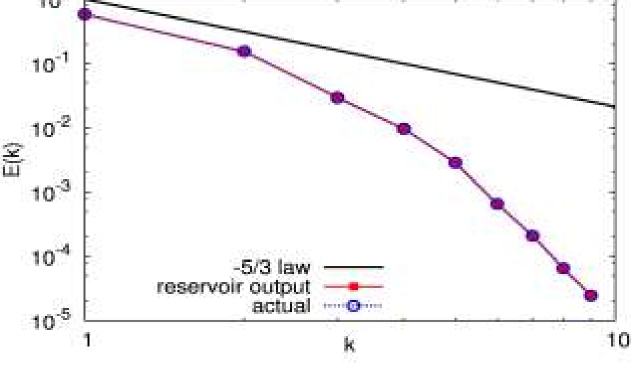
[arXiv:1805.09917]

Goal : Construction of a data-driven model (dynamical system) of a fluid flow

Time-series of a macroscopic variable (Energy function)



Statistical quantity (Energy spectrum)



We have succeeded in constructing a closed form equation describing a macroscopic behavior of a fluid flow only from macroscopic data without prior knowledge of physical process.

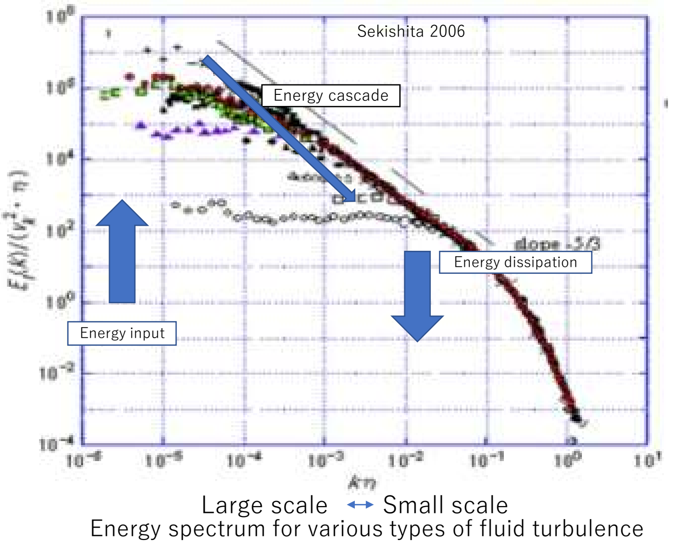
Outline of this talk

1. Introduction to a fluid flow
2. Motivation
3. Reservoir computation
4. Partial-prediction of a microscopic variable
5. Full-prediction of a macroscopic variable (main topic)
6. Summary

1-1. Introduction to a fluid flow

- Fluid flow appears everywhere around us: from coffee cup to space.

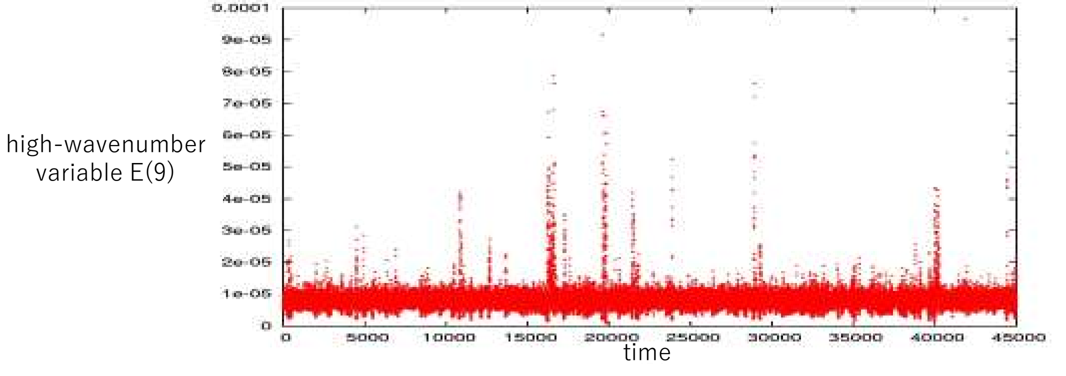
Universality in turbulence



Energy spectrum for various types of fluid turbulence

1-2. Introduction to a fluid flow: high-dimensional chaos

- Fluid flow shows chaos in most cases.
  - Deterministic
  - Sensitivity for initial conditions
- Fluid flow shows a high-dimensional chaos in most cases.
  - Coexistence of different number of unstable directions
  - Intermittent behavior especially in higher wavenumbers (smaller scale)



1-4. Introduction to a fluid flow: Navier-Stokes equation

- Most fluid flow is known to be modeled by three dimensional incompressible Navier-Stokes equation:

$$\begin{cases} \partial_t v - \nu \Delta v + (v \cdot \nabla) v + \nabla \pi = f, \ \nabla \cdot v = 0 \\ v\big|_{t=0} = v_0 \quad \text{with } \nabla \cdot v_0 = 0, \end{cases}$$

v : velocity

π : pressure

ν : viscosity (more turbulent for smaller ν)

- Lots of difficulties in this equation:
  - nonlinear
  - nonlocal interactions
  - high-dimensional (especially for small ν)

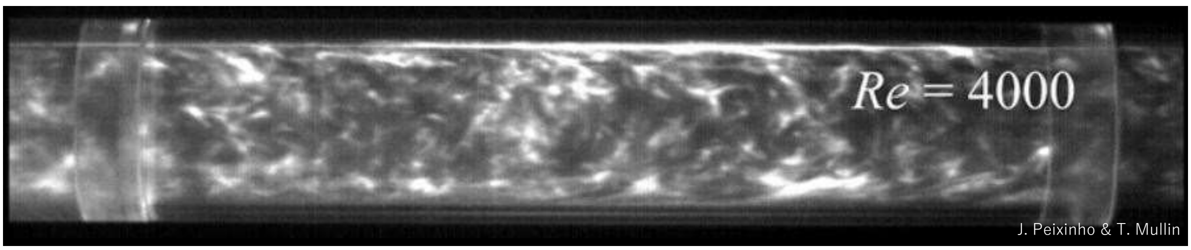
We do not even know the existence of a “global solution”.

(One of the Millennium prize problems by Clay Mathematics Institute $1,000,000)

1-5. Introduction to a fluid flow: Mean flow

- There are lots more open problems in the area of fluid mechanics such as the analytical derivation of mean flow velocity profile (=macroscopic quantity) of turbulence due to the "closure problem", that is, it is impossible to get a closed form equation of a mean flow (=1st order quantity) without knowing a higher order quantity.

[In order to write down the equation of the nth order quantity, we need the information of the n+1th order quantity for all n.]



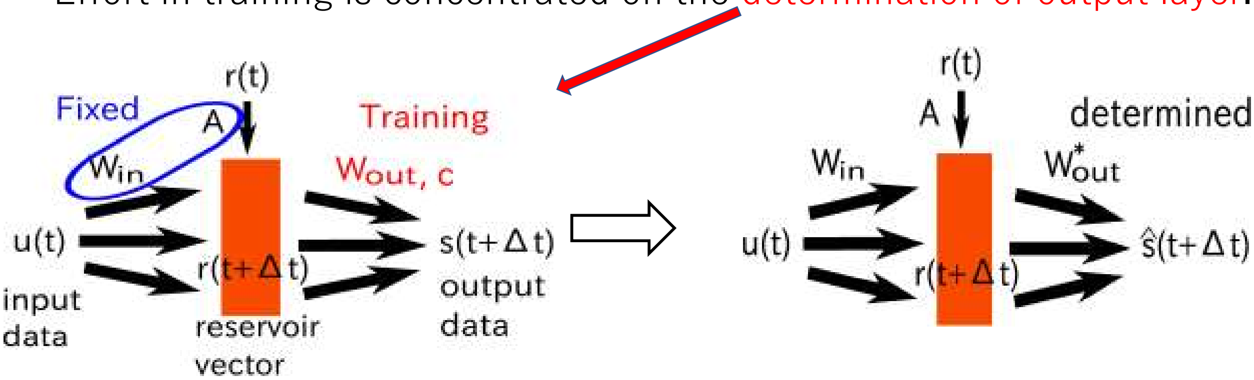
2. Motivation and outline of our study

- Data-driven modeling of fluid flow : Based on the limited time-series data of a fluid flow, we would like to construct a model. Especially, we are interested in constructing a model for macroscopic variables.

- We employ reservoir computation technique, which is proved to be very powerful for constructing models from data.
  Lu et al. (2017), Pathak et al. (2017, 2018): Lorenz & Kuramoto-Sivashinsky systems
- Training data as well as a reference data are generated from the direct numerical simulation of the three dimensional incompressible Navier-Stokes equation with periodic-boundary conditions.
  - Fourier variables of velocity: microscopic
  - Energy functions: macroscopic
- We would like to construct a data-driven model which can predict future behaviors of both microscopic and macroscopic fluid variables without a prior knowledge of physical process in relatively small computational costs.

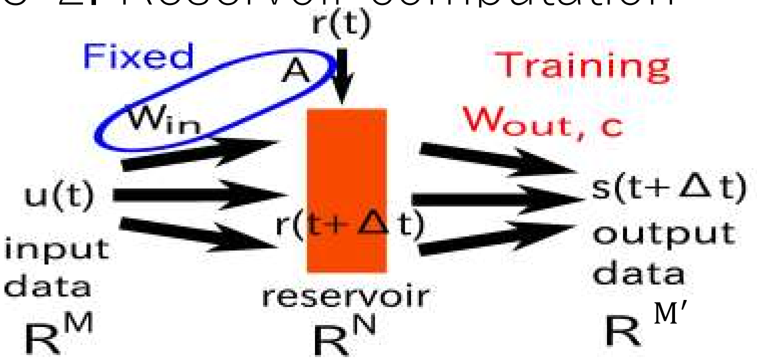
3-1. Reservoir computation

- Machine-learning technique that uses a neural-network composed of simple nonlinear dynamical systems.
- The framework was proposed as echo-state network (Jaeger 2001, 2004) and liquid-state machine (Maass et al. 2002).
- Effort in training is concentrated on the determination of output layer.



Training phase (t<T) **Prediction phase** (t>T)

3-2. Reservoir computation



$M, M' \ll N$

$10^0 \sim 10^3 \quad 10^3 \sim 10^5$

($M = M'$ for full-prediction)

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

$\mathbf{A}, \mathbf{W}_{\text{in}}$: sparse random matrix, whose maximal eigenvalue is controlled

α: nonlinearity parameter ($\alpha = 0.3$).

We determine W_{out} and c s.t.

$$t^{\forall} < T \qquad \mathbf{W}_{\text{out}}\mathbf{r}(t+\Delta t) + c \approx \mathbf{s}(t+\Delta t).$$

3-3. Reservoir computation: How to determine **Wout** and **c** ?

- Minimizing the quadratic form with respect to **Wout** and **c:**

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

regularization term to avoid overfitting

- Solution: $\hat{\mathbf{s}}(t)=\mathbf{W}^*_{\text{out}}\mathbf{r}(t)+c^*$ (Lukosevivsius and Jaeger, 2009)

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

where $\overline{r}=\sum_{l=1}^{L}\mathbf{r}(l\Delta t)/\text{L}$, $\overline{s}=\sum_{l=1}^{L}\mathbf{s}(l\Delta t)/L$, and $\mathbf{I}$ is the $N\times N$ identity matrix, $\delta\mathbf{R}$ (respectively, $\delta\mathbf{S}$) is the matrix whose l-th column is $\mathbf{r}(l\Delta t)-\overline{r}$ (respectively, $\mathbf{s}(l\Delta t)-\overline{s}$).

4-1. Partial-prediction of a microscopic variable:

Generation of learning data using direct numerical simulation of the Navier-Stokes equation

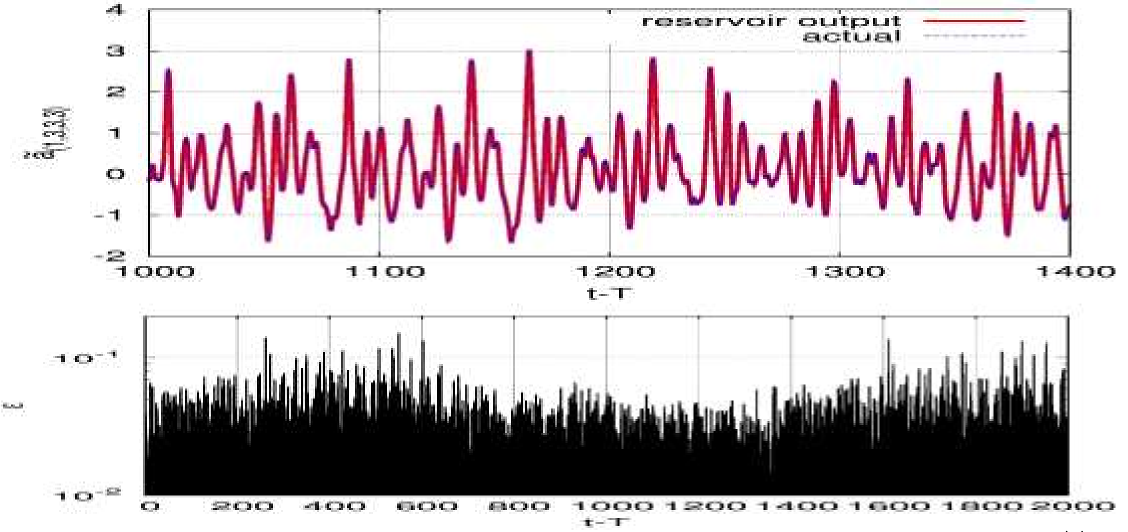
$$\begin{cases}\partial_t v-\nu\Delta v+(v\cdot\nabla)v+\nabla\pi=f,\ \nabla\cdot v=0, & \mathbb{T}^3\times(0,\infty)\\ v\big|_{t=0}=v_0\quad\text{with }\nabla\cdot v_0=0, & \mathbb{T}^3,\end{cases}$$

- We employ Fourier spectral method with N_0=9 modes, meaning that the system is approximated by $2(2N_0+1)^3$=13718 dimensional ODE.
- We focus on time-series data of 270 variables

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

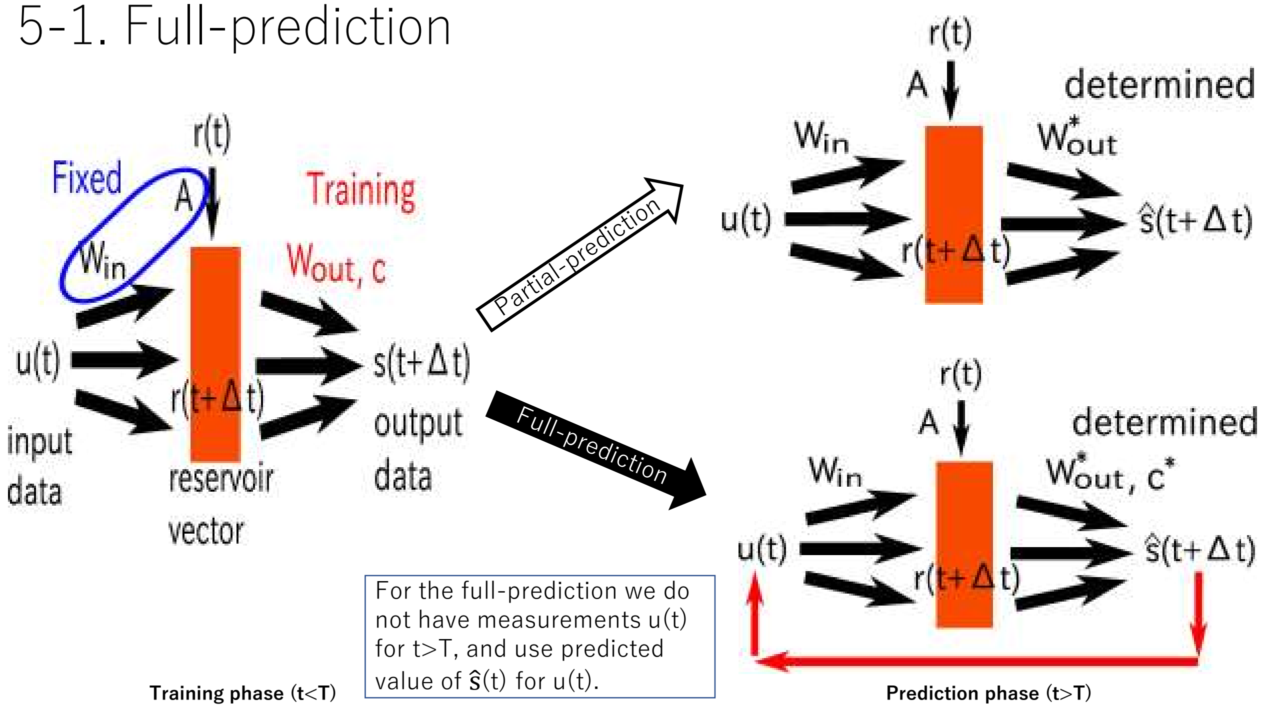
and predict another variable.

4-2. Partial-prediction of a microscopic variable: Fourier variable of a velocity



Partial-prediction of a microscopic variable from a measurement u(t) of dimension 270 is quite successful.

5-1. Full-prediction



5-2. Full-prediction of a macroscopic variable

The energy function $E_0(k,t)$ for wavenumber $k \in \mathbb{N}$ is defined by

$$E_0(k,t) := \frac{1}{2}\int_{D_k} \sum_{\zeta=1}^{3} \left|\mathcal{F}_{[v_\zeta]}(\kappa,t)\right|^2 d\kappa,$$

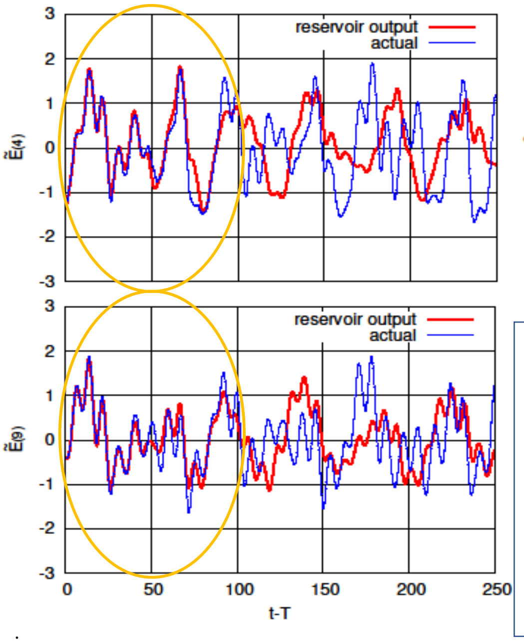
where $D_k := \{\kappa \in \mathbb{Z}^3 | k - 0.5 \le |\kappa| < k + 0.5\}$.

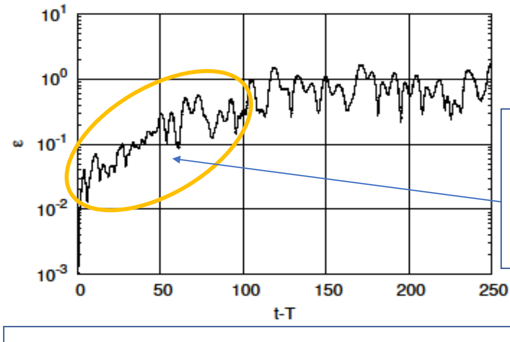
Short-time average of the energy function will be used later

$$E(k,t) = \sum_{s=t-49\Delta t}^{t+50\Delta t} E_0(k,s)/100$$

- By learning 9 time-series data of E(k,t) (k=1,⋯,9) for t≤T, we predict E(k,t) (k=1,⋯,9) for t>T. We do not learn any data during the prediction time for t>T

5-3. Full-prediction of a macroscopic variable: Energy function

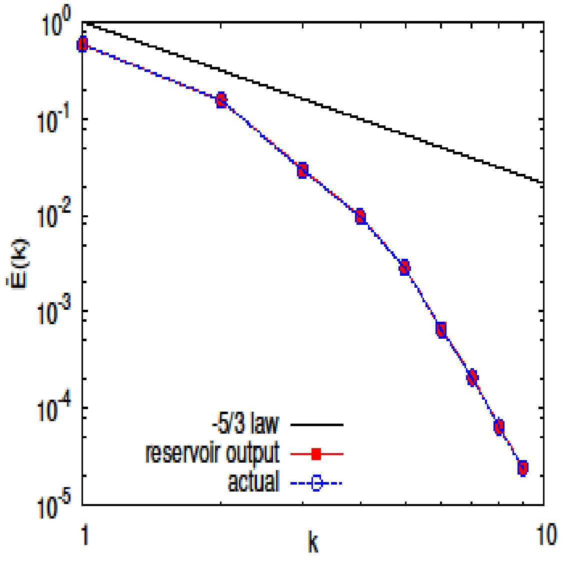




Exponential growth of error due to a chaotic property

- When t<T+100, predicted time-series data obtained from our reservoir system almost coincides with that of a reference data obtained from the DNS of the Navier-Stokes equation.
- The increase in the prediction error is due to the chaotic property of the fluid flow, which is inevitable.

5-4. Full-prediction of a macroscopic variable: Reproducing energy spectrum by the reservoir model



- The energy spectrum obtained from the full-prediction procedure of E(k,t) for t>T+100 (after the time-series prediction fails) coincides with that from a reference data obtained from the direct numerical simulation of the Navier-Stokes equation.
- This implies that the obtained reservoir system constructed without the knowledge of microscopic variables is equivalent to the dynamical system describing a macroscopic behavior of energy functions.

6-1. Summary

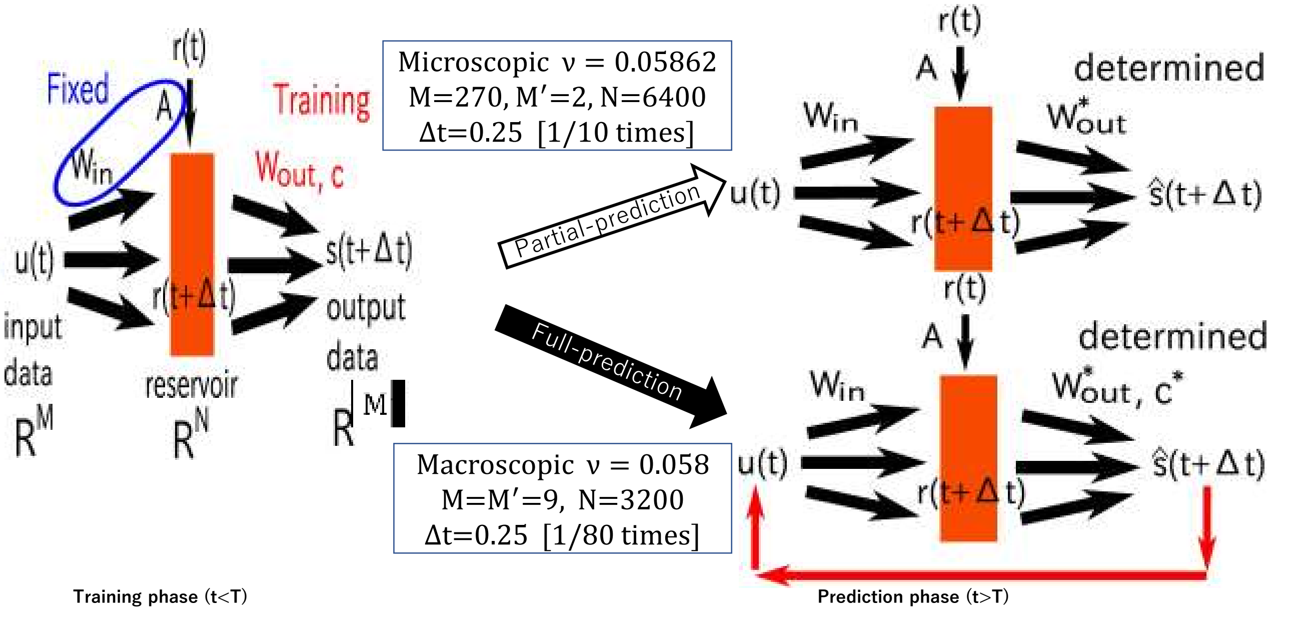
- We can predict time-series of both microscopic and macroscopic variables of fluid flow by machine-learning technique using reservoir computation without a prior knowledge of a physical process.
- In order to generate a time-series data of a macroscopic variable of a fluid flow, we do not need to go back to the microscopic dynamics.

We have especially succeeded in constructing a closed form equation of a fluid flow describing macroscopic behavior only from data.

- The method is shown to be especially useful in generating a macroscopic time-series data with small computational costs.

6-2. Comparison about computational costs

DNS of Navier-Stokes eqns
13718-dim ODE, 4stage-RK4, Δt=0.05



6-3. Remarks on the choice of parameters

- As N ($\gg M$) increases, the error tends to decrease. Especially for the full prediction, N should be significantly larger than M to get an accurate prediction.
- As T increases, the error tends to decrease, but N should also be increased for getting a better result.
- If we choose measurements independently with each other, we can use smaller β (≥ 0). On the other hand, when we have a better set of measurements, smaller β is better.
- In order to obtain a "nonhomogeneous" behavior, the parameters D_1 and D_2 should be relatively small and avoid strong coupling among r. They should be proportional to N.

	parameter	(a)	(b)
τ	transient time	1000	2500
T	learning time	10000	20000
M	dimension of measurements	270	9
P	dimension of predicted variables	2	9
N	number of reservoir nodes	6400	3200
D_1	parameter of determining elements of $\mathbf{A}$	60	320
D_2	parameter of determining elements of $\mathbf{A}$	60	0
γ	scale of input weights in $\mathbf{A}$	0.1	0
ρ	maximal eigenvalue of $\mathbf{A}$	1.0	0.5
σ	scale of input weights in $\mathbf{W}_{in}$	0.4	0.3
α	nonlinearity degree of reservoir dynamics	0.7	0.3
Δt	time step for reservoir dynamics	0.1	0.25
β	regularization parameter	0	0.01

TABLE I. **Sets of parameters for the reservoir computations.** The set (a) is used for the partial-prediction of microscopic Fourier variables, whereas the set (b) is for the full-prediction of macroscopic variables of Energy function and Energy spectrum.

6-4.References

[1] Y. LeCun, Y. Bengio, and G. Hinton, Nature **521**, 436 (2015).
[2] M. Raissi and G. E. Karniadakis, J. Comp. Phys. **357**, 125 (2018).
[3] N. Nakano, M. Inatsu, S. Kusuoka, and Y. Saiki, Japan J. Indust. Appl. Math. , https://doi.org/10.1007/s13160 (2018).
[4] J. Ling, A. Kurzawski, and J. Templeton, J. Fluid Mech. **807**, 155 (1941).
[5] J. N. Kutz, J. Fluid Mech. **814**, 1 (2017).
[6] D. Verstraeten, B. Schrauwen, M. D'Haene, and D. A. Stroobandt, Neural Network **20**, 391 (2007).
[7] M. Inubushi and K. Yoshimura, Scientific Reports **7**, 10199 (2017).
[8] Z. Lu, J. Pathak, B. R. Hunt, M. Girvan, R. Brockett, and E. Ott, Chaos **27**, 041102 (2017).
[9] J. Pathak, Z. Lu, B. R. Hunt, M. Girvan, and E. Ott, Chaos **27**, 121102 (2017).
[10] D. Ibanez-Soria, J. Garcia-Ojalvo, A. Soria-Frisch, and G. Ruffini, preprint , https://arxiv.org/pdf/1710.08286.pdf (2017).
[11] J. Pathak, B. R. Hunt, M. Girvan, L. Zhixin, and E. Ott, Physical Review Letters **120**, 024102 (2018).
[12] P. Antonik, M. Gulina, J. Pauwels, and S. Massar, preprint , https://arxiv.org/pdf/1802.02844.pdf (2018).
[13] H. Jaeger, GMD Report **148**, 13 (2001).
[14] H. Jaeger and H. Haas, Scince **304**, 78 (2004).
[15] W. Maass, T. Natschläger, and H. Markram, Neural Computation **14**, 2531 (2002).
[16] M. Lukosevivcius and H. Jaeger, Computer Science Review **3**, 127 (2009).
[17] K. Ishioka, "ispack-0.4.1," `http://www.gfd-dennou.org/arch/ispack/`, (1999), GFD Dennou Club.