# My expectation: Why fluid-PDE researchers may need to learn machine learning?

Tsuyoshi Yoneda
Hitotsubashi University

# Closure problem

Let us briefly explain the closure problem.

**The Navier-Stokes equations**

$$\partial_t u + (u \cdot \nabla) u = -\nabla p + \nu \Delta u + f, \quad \nabla \cdot u = 0, \quad u|_{t=0} = u_0$$

Let $\hat{\cdot}$ be the Fourier transoform and $\mathcal{F}^{-1}$ be its inverse.

- Mean flow: $\langle u(x) \rangle = \mathcal{F}_\xi^{-1}[\hat{u}(\xi)\chi_{\{|\xi| \leq 1\}}](x)$
- Turbulent fluctuation: $u'(x) = 1 - \langle u(x) \rangle$

(for simplicity, the wavenumber scale of the mean flow is set to $1$)
The most important problem in the turbulence study (called closure problem) is the following:

**Can a turbulence model with high prediction accuracy be constructed for the mean flow $\langle u \rangle$?**

# Classical approach to the closure problem

Apply $\langle\cdot\rangle$ to both sides of the Navier-Stokes equations:

$$\partial_t\langle u\rangle + (\langle u\rangle\cdot\nabla)\langle u\rangle + \nabla\cdot\tau = -\nabla\langle p\rangle + \nu\Delta\langle u\rangle + \langle f\rangle, \quad \nabla\cdot\langle u\rangle = 0.$$

$\tau = \langle u\otimes u\rangle - \langle u\rangle\otimes\langle u\rangle$ is called the Reynolds stress, depends on the turbulent fluctuation $u'$. The classical closure problem is as follows:

**Can $\tau$ be approximated using only the mean flow $\langle u\rangle$?**

## Breakthrough in this study (at least I think)

The Google Science Team in 2023:

GraphCast: Learning skillful medium-range global weather forecasting

They constructed a machine learning model for weather forecasting (NOT physics-informed anymore) that surpasses conventional physical models, that is, without any use of the above fluid formula combining $\tau$.

So, I believe fluid-PDE researchers may need to learn why ML surpassed the fluid-PDE (physical model).

# The first thing for pure mathematicians to learn about ML

- input space $\mathcal{X}$, output space $\mathcal{D}$ and feature space $\mathcal{X}$
- $F : \mathcal{U} \to \mathcal{D}$ is the unknown function (we call "underlying system") allowed to be discontinuous, but really want to know
- $\dim(\mathcal{U}), \dim(\mathcal{D}) \ll \dim(\mathcal{X})$

Example:

- $u$: picture ($\mathcal{U}$ has $1000px \times 1000px \times 3colors$ dimension)
- $\underline{d = 1\text{: pic is cat}}$, $\underline{d = 0\text{: pic is NOT cat}}$ ($\mathcal{D}$ has only one dimension)

**Purpose of ML in pure math description**

Find continuous maps $\Phi : \mathcal{U} \to \mathcal{X}$ and $\Psi : \mathcal{X} \to \mathcal{D}$ such that

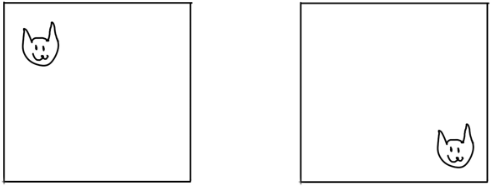
$$|\Psi \circ \Phi(u) - d| \quad \text{is small enough for } \forall (u, d) \in \mathcal{U} \times F(\mathcal{U}).$$

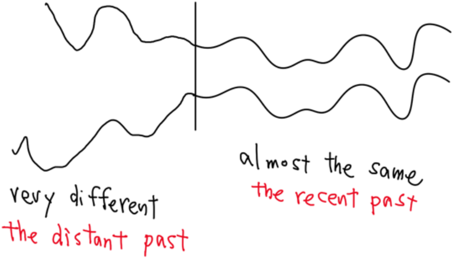
The crucial structure of ML (black box is NOT black box):

There may exist $(u_1, d_1), (u_2, d_2) \in \mathcal{U} \times F(\mathcal{U})$ satisfying $|u_1 - u_2| \gtrsim 1$ such that, the corresponding feature vectors are similar: $|\Phi(u_1) - \Phi(u_2)| \ll 1$.
How can we structure this into $\Phi$?

There may exist $(u_1, d_1), (u_2, d_2) \in \mathcal{U} \times F(\mathcal{U})$ satisfying $|u_1 - u_2| \gtrsim 1$ such that, the corresponding feature vectors are similar: $|\Phi(u_1) - \Phi(u_2)| \ll 1$. How can we structure this into $\Phi$?

We have three examples, convolutional NN, ESP, attention mechanism.



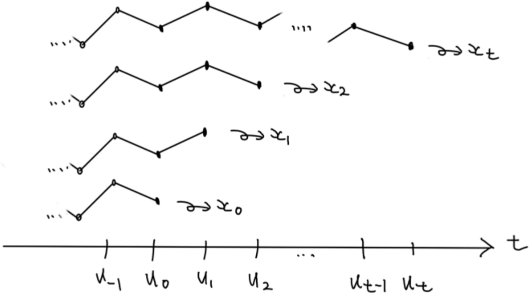


# ESP: Echo state property (for time series data)

The tricky point is the following definition: For $\Phi : \mathcal{U} \times \mathcal{X} \to \mathcal{X}$, let

$$x_t^1 = \Phi(u_t, \cdot) \circ \cdots \circ \Phi(u_2, \cdot) \circ \Phi(u_1, x_0^1),$$
$$x_t^2 = \Phi(u_t, \cdot) \circ \cdots \circ \Phi(u_2, \cdot) \circ \Phi(u_1, x_0^2).$$

$x_0^1$ and $x_0^2$ are expressing time series before $t = 0$.



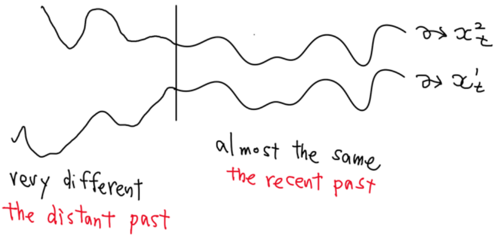
## Theorem

Assume there is $\lambda \in (0,1)$ such that

$$|\Phi(u, x^1) - \Phi(u, x^2)| \leq \lambda |x^1 - x^2|$$

for $(u, x^1), (u, x^2) \in \mathcal{U} \times \mathcal{X}$ with $x_1 \neq x_2$. Then for any $x_0^1, x_0^2 \in \mathcal{X}$, we have

$$|x_t^1 - x_t^2| \to 0 \quad (t \to \infty).$$




The proof is obvious. We have the following for $t \to \infty$:

$$|x_{t+1}^1 - x_{t+1}^2| \leq |\Phi(u_t, x_t^1) - \Phi(u_t, x_t^2)| \leq |x_t^1 - x_t^2| \leq \cdots \leq \lambda^t |x_0^1 - x_0^2| \to 0.$$

# Remark

## ML framework

Recall that, the underlying system $F$ could be discontinuous, on the other hand, $\Phi$ and $\Psi$ are continuous maps.

From this ML framework, We see that $\Psi$ and $\Phi$ should not be expressed as Fourier sums.

The evidence is the following: For the Fourier series of the indicator functions of such several dimensional balls, we observe

- Gibbs phenomena,
- Pinsky phenomena 1993 (diverging at the origin),
- Kuratsubo phenomena 2010 (preventing pointwise convergence).

To understand Kuratsubo phenomena intuitively, see numerical computations in Section 7 in Kuratsubo-Nakai-Ootsubo (2010).