Generalized Delay Embedding Theorem

Predicting and Tipping

Luonan Chen (陳洛南)

Chinese Academy of Sciences

statistics-based data science

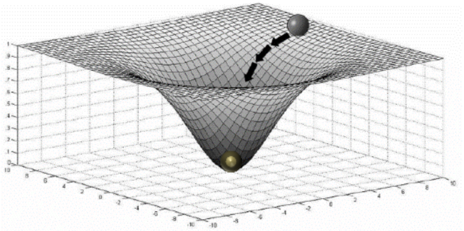
model-free approach

Dynamics-based Data Science

Luonan Chen (陳洛南)

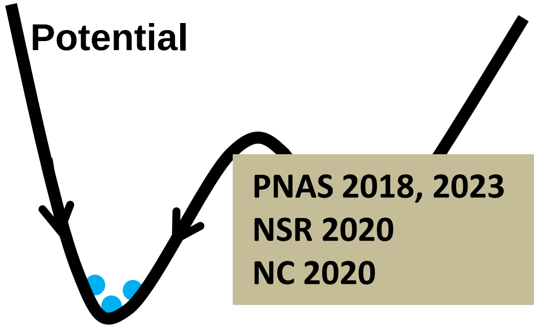
Chinese Academy of Sciences

1. Steady State → STI (预测)



Large basin or potential barrier

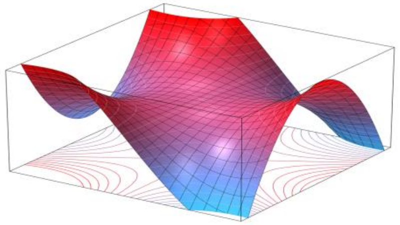
空间-时间信息
转换方程



PNAS 2018, 2023
NSR 2020
NC 2020

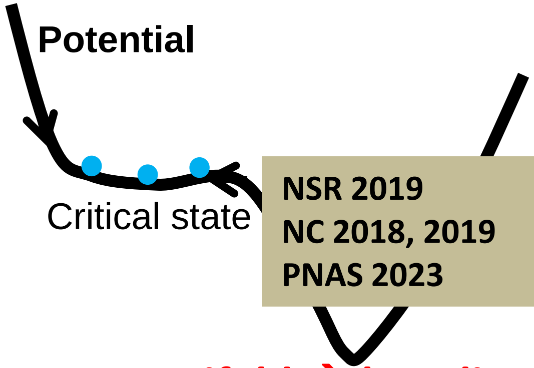
Attractor → low dim

2. Critical State → DNB (临界)



Small basin or potential barrier

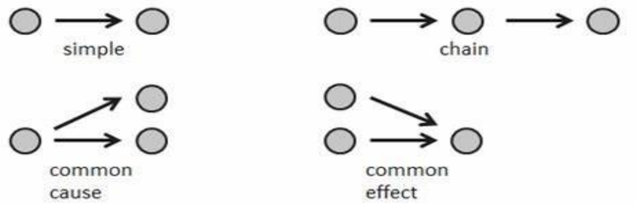
临界的数据性质
线性→非线性

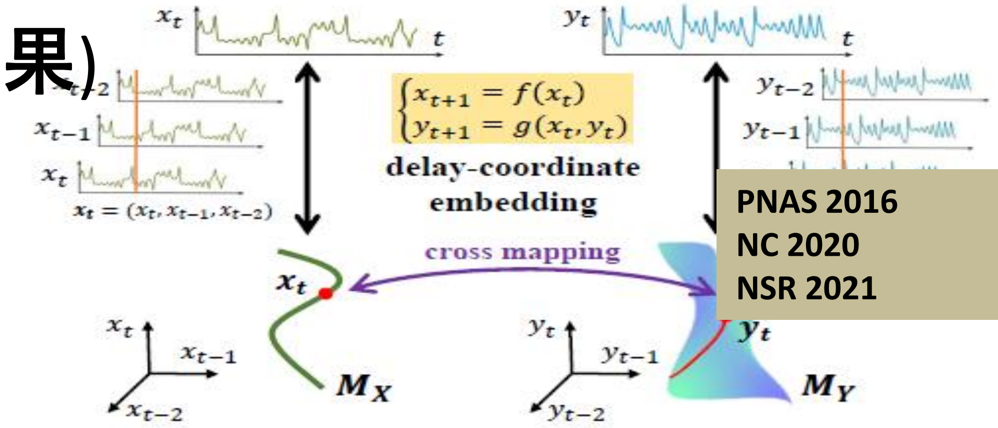


NSR 2019
NC 2018, 2019
PNAS 2023

Center manifold → low dim

3. Phase Space → PCM (因果)





PNAS 2016
NC 2020
NSR 2021

Embedding implies causality !

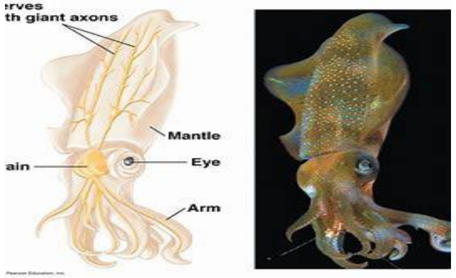
单变量相重构→拓扑→因果

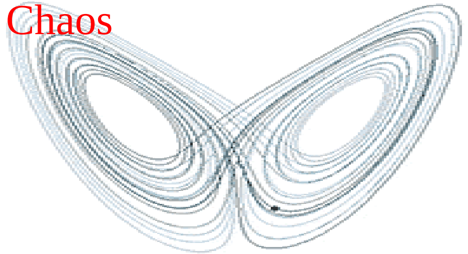
Statistic Data-Science → association and information

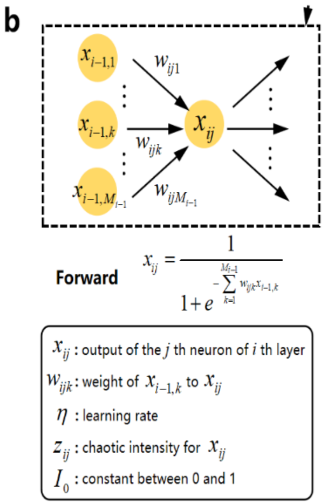
Dynamic Data-Science → causation and mechanism

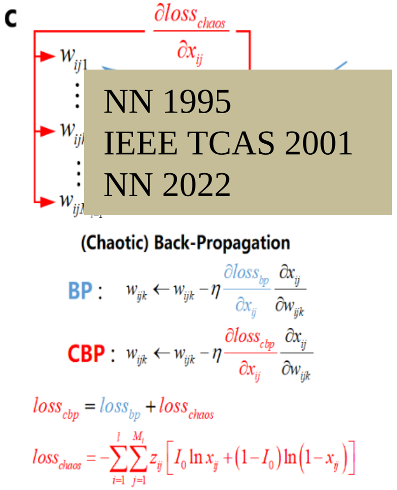
4. AI for Science → Science for AI　　梯度→混沌 类脑计算

Chaotic Back Propagation for AI









NN 1995
IEEE TCAS 2001
NN 2022

Science: long evolution makes brain explore chaotic dynamics

AI: local gradient dynamics → global chaotic dynamics

(类脑计算)

Short-term time-series prediction based on spatial-temporal information

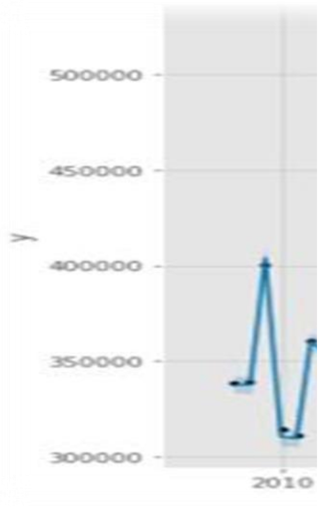
陳洛南(Luonan Chen)
Chinese Academy of Sciences

Time Series Data Prediction

Many methods

- Long-term data: statistic patterns → predictable
- **Short-term data: no patterns → unpredictable**

No methods



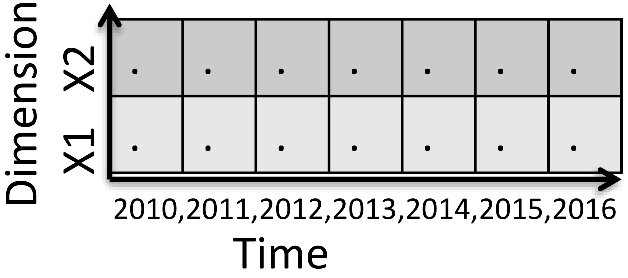
No math model

Time Series Data Prediction

Both measuring the same system !

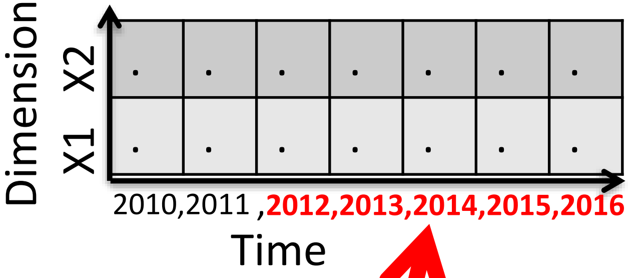
Spatial information → Temporal Information ?

数据间关联到时序列的信息转换



Equivalent ?

Method ?

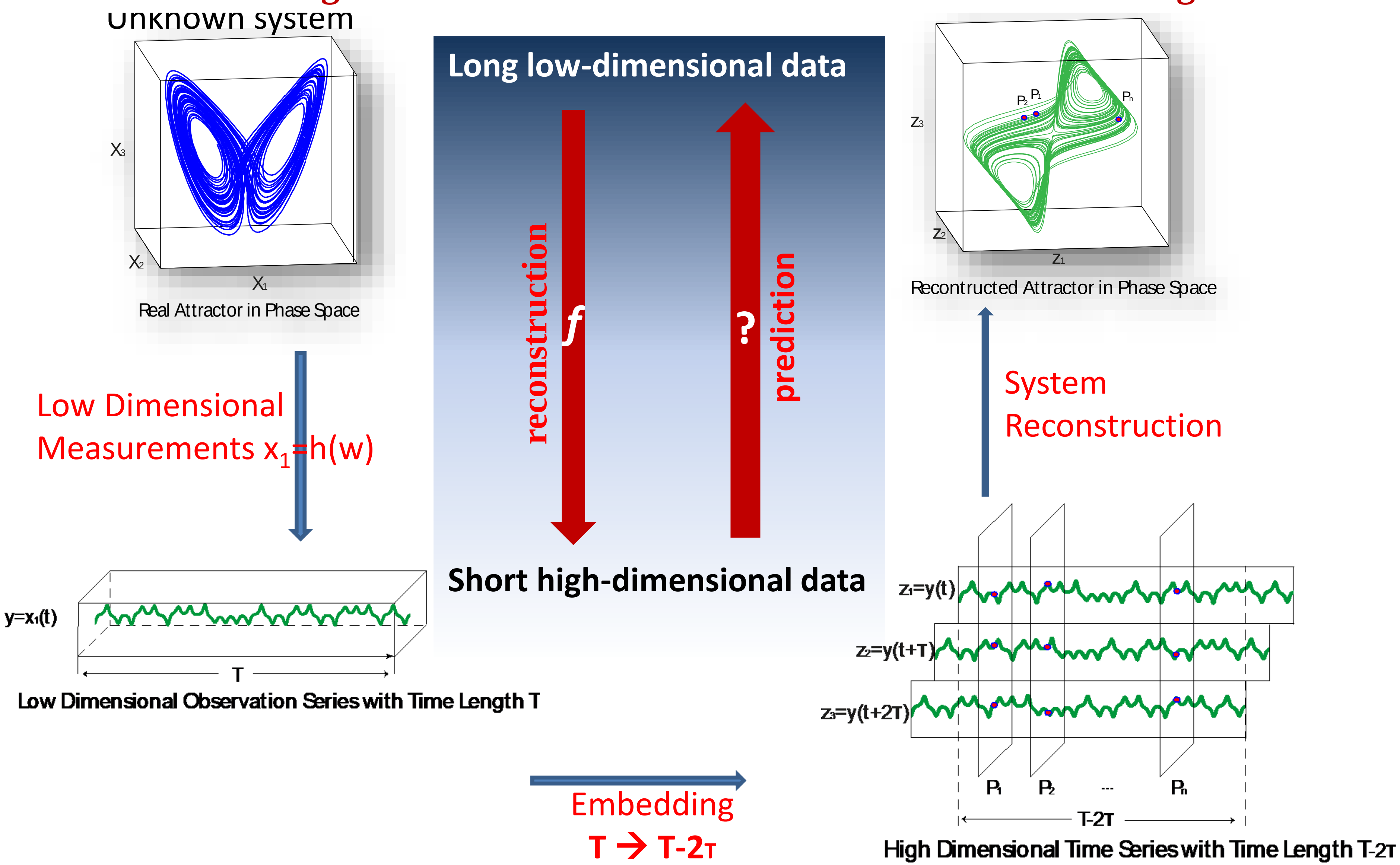


Predict future

Theoretical Foundation: milestone work in 1981

Long-term time-series of a single variable topologically reconstruct attractor of high-dim system

Can we construct long-term time-series of this variable from short-term high-dim data?



从原空间到嵌入空间的时间序列

Delay Embedding

From original space x to embedded space X

Original space in R : $x_1, x_2, \dots, x_t, \dots$

$\downarrow$

Embedded space in R^L: $X_1, X_2, \dots, X_t, \dots$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_L \end{bmatrix}, \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_{L+1} \end{bmatrix}, \dots, \begin{bmatrix} x_t \\ x_{t+1} \\ \vdots \\ x_{t+L-1} \end{bmatrix}, \dots$$

X_t represents the state of whole system, including information of other variable y_t

General Delay Embedding Theorem

The state space of the discrete time dynamical system is a *v-dim* manifold *M*. The dynamics is given by a smooth map

$\boldsymbol{f}$ **: M → M**

Assume that the dynamics *f* has a strange attractor A in M with box counting dimension d_A. *A* can be embedded in *k*-dim Euclidean space with $k > 2d_A$. That is, there is a diffeomorphism φ that maps *A* into M such that the derivative of φ has full rank (Whitney embedding theorem).

An observation function h : M → R must be twice-differentiable and associate a real number to any point of the attractor *A*. It must also be typical, so its derivative is of full rank and has no special symmetries in its components. Then the function

$$\mathbf{H(x) = (h(x), h(\mathit{f}(x)), \ldots, h(\mathit{f}^{2d}(x)))}$$

is an embedding of the strange attractor *A* in $\mathbf{R}^k$ (Takens embedding theorem).

General Delay Embedding Theorem

(1) h can be a vector, and then H is an embedding

(2) H = h, when h is a vector with its dim > $2d_A$

order-n derivative ←→ n-dim state equation

1-variable with order-n derivative

$$\frac{dx(t)}{dt} \approx \frac{(x(t)-x(t-1))}{\Delta t}, \quad \frac{d^2x(t)}{dt^2} \approx \frac{(x(t)-2x(t-1)+x(t-2))}{\Delta t^2}, \ldots$$

$\frac{d^n x(t)}{dt^n}$ can be represented by (n+1) delay-variables,
$x(t), x(t-1), \ldots, x(t-n)$

1. $x(t), x(t-1), \ldots, x(t-n)$ can represent a n-dim dynamical system;

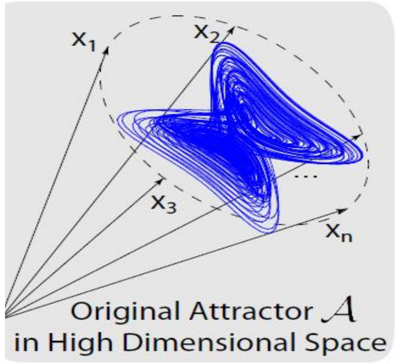
2. Taken Embedding theorem states that we need only $x(t), x(t-1), \ldots, x(t-2d)$ to represent n-dim system, with n>>2d.

with *n* variables $x(t), x_1(t), \ldots, x_{n-1}(t)$

n-dimensional state equation→ n-dimensional dynamical system

Constructing attractors by delay embedding

Ma et al. PNAS 2018



One by one projection ϕ between
“delay attractor” and “non-delay attractor”

Ma et al. PNAS 2018

Non-delay Attractor ←→ Delay Attractor

Nonlinear Mapping

observe n variables Φ predict one variable x_k

Hankel-matrix ↙

$$
\Phi:\begin{bmatrix}
x_1(t_1) & x_1(t_2) & \cdots & x_1(t_{m-1}) & x_1(t_m)\\
x_2(t_1) & x_2(t_2) & \cdots & x_2(t_{m-1}) & x_2(t_m)\\
x_3(t_1) & x_3(t_2) & \cdots & x_3(t_{m-1}) & x_3(t_m)\\
\vdots & \vdots & \ddots & \vdots & \vdots\\
x_n(t_1) & x_n(t_2) & \cdots & x_n(t_{m-1}) & x_n(t_m)
\end{bmatrix}
\rightarrow
\begin{bmatrix}
x_k(t_1) & x_k(t_2) & \cdots & x_k(t_{m-1}) & x_k(t_m)\\
x_k(t_2) & x_k(t_3) & \cdots & x_k(t_m) & x_k(t_{m+1})\\
x_k(t_3) & x_k(t_4) & \cdots & x_k(t_{m+1}) & x_k(t_{m+2})\\
\vdots & \vdots & \ddots & \vdots & \vdots\\
x_k(t_L) & x_k(t_{L+1}) & \cdots & x_k(t_{L+m-2}) & x_k(t_{L+m-1})
\end{bmatrix}
$$

Spatial → Temporal Information transformation

STI equation (“空间-时间”信息转换方程)

Linear Mapping

$$
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n}\\
a_{21} & a_{22} & \cdots & a_{2n}\\
\vdots & \vdots & \ddots & \vdots\\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix}
\begin{bmatrix}
x_1(t_1) & x_1(t_2) & \cdots & x_1(t_{m-1}) & x_1(t_m)\\
x_2(t_1) & x_2(t_2) & \cdots & x_2(t_{m-1}) & x_2(t_m)\\
x_3(t_1) & x_3(t_2) & \cdots & x_3(t_{m-1}) & x_3(t_m)\\
\vdots & \vdots & \ddots & \vdots & \vdots\\
x_n(t_1) & x_n(t_2) & \cdots & x_n(t_{m-1}) & x_n(t_m)
\end{bmatrix}
=
\begin{bmatrix}
x_k(t_1) & x_k(t_2) & \cdots & x_k(t_{m-1}) & x_k(t_m)\\
x_k(t_2) & x_k(t_3) & \cdots & x_k(t_m) & x_k(t_{m+1})\\
x_k(t_3) & x_k(t_4) & \cdots & x_k(t_{m+1}) & x_k(t_{m+2})\\
\vdots & \vdots & \ddots & \vdots & \vdots\\
x_k(t_L) & x_k(t_{L+1}) & \cdots & x_k(t_{L+m-2}) & x_k(t_{L+m-1})
\end{bmatrix}
$$

Past → **Future**

Non-Delay Attractor — Delay Attractor

Conjugated STI equations

$$\begin{cases} \Phi(\mathbf{X}^t) = \mathbf{Y}^t \\ \mathbf{X}^t = \Psi(\mathbf{Y}^t) \end{cases}$$

Primary STI equation

($\Psi\Phi = I$)

Conjugate STI equation

$$\mathbf{X}^t = \begin{pmatrix} \boldsymbol{x}_1^t \\ \boldsymbol{x}_2^t \\ \cdots \\ \boldsymbol{x}_n^t \end{pmatrix} \xrightarrow{\Phi} \mathbf{Y}^t = (\boldsymbol{y}^t, \boldsymbol{y}^{t+1}, \ldots, \boldsymbol{y}^{t+L-1})$$

Spatial information(空间信息) **Temporal information(时间信息)**

$y^t = x_k^t$

Neural network by STI equations

Conjugated STI equations
(Φ, Ψ) ?

$$\begin{cases} \Phi(\mathbf{X}^t) = \mathbf{Y}^t \\ \mathbf{X}^t = \Psi(\mathbf{Y}^t) \end{cases}$$

非线性版本!

$\Phi\Psi$=*I*

⇩

Linearized STI equations
(A,B) ?

$$\begin{cases} A\mathbf{X}^t = \mathbf{Y}^t \\ \mathbf{X}^t = B\mathbf{Y}^t \end{cases}$$

线性版本!

AB=***I***

⇩

Neural network STI equations
(A,B) ?

$$\begin{cases} AF(\mathbf{X}^t) = \mathbf{Y}^t \\ F(\mathbf{X}^t) = B\mathbf{Y}^t \end{cases}$$

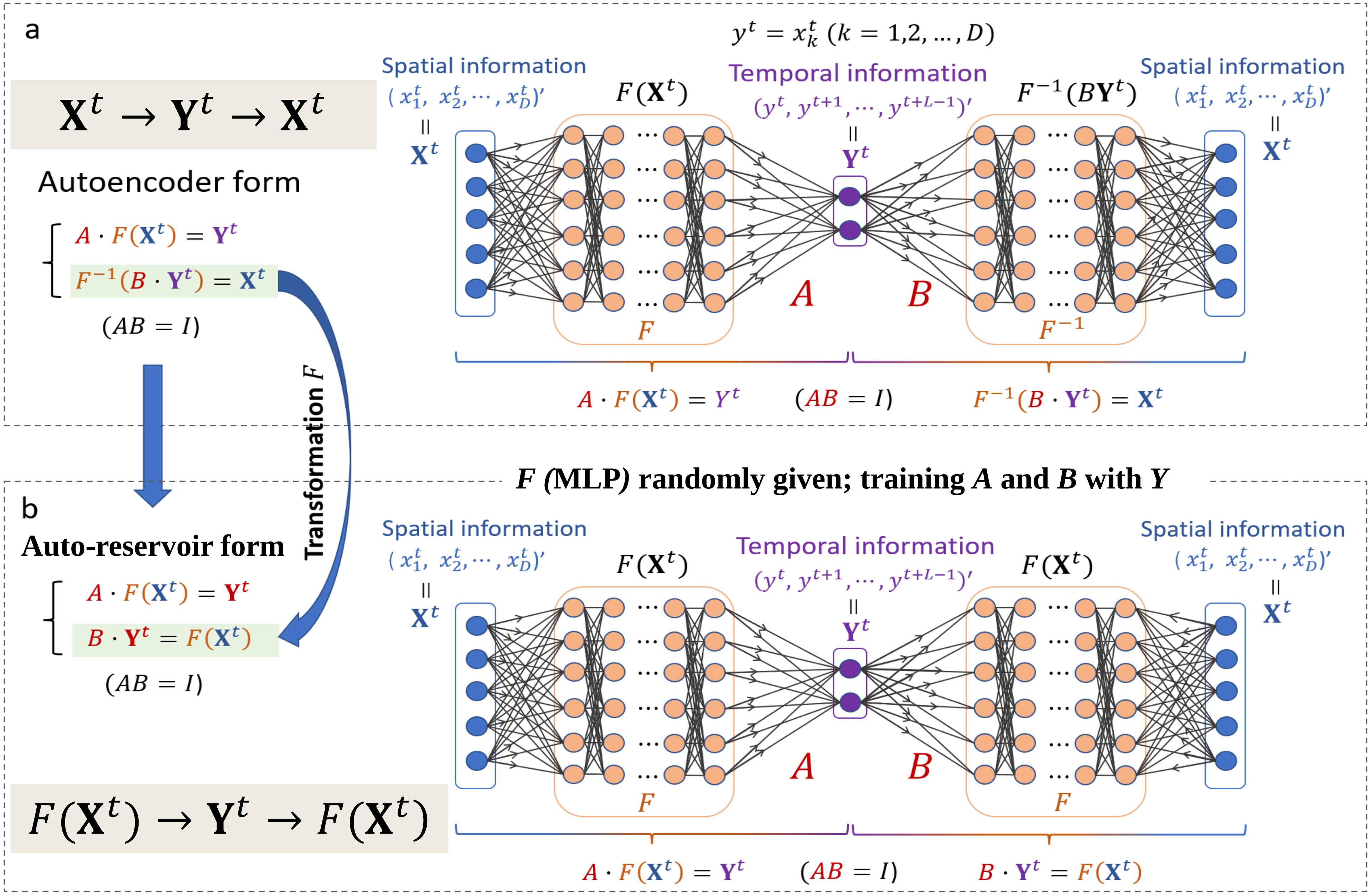
近线性版本!

AB=***I***

Neural network *F* is randomly given

Nonlinear map

Auto-encoder and Auto-reservoir for STI

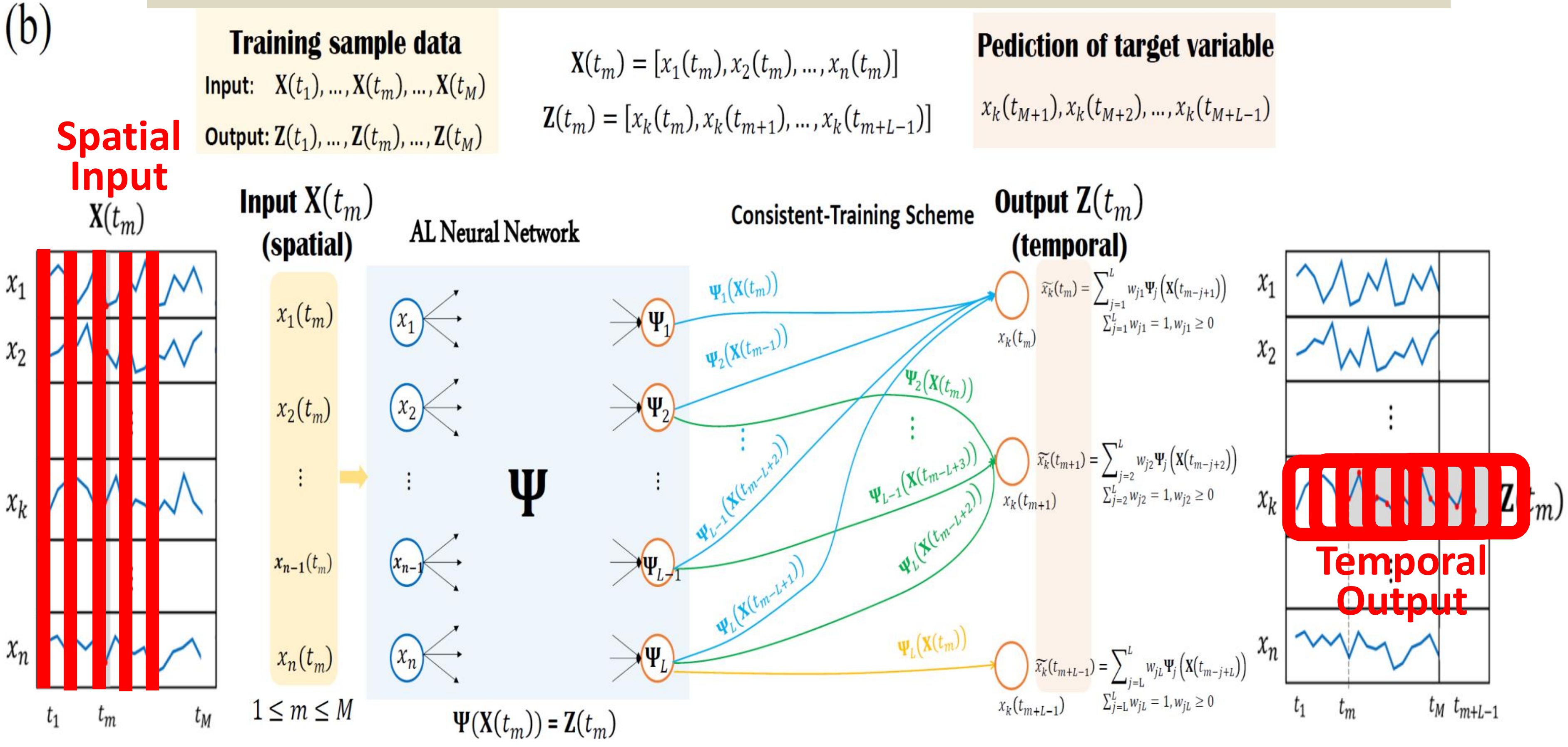


Chen et al. Nature Communications 2020; Chen et al. National Science Review 2020; Ma et al. PNAS 2018

“空间-时间”信息转换

Spatial-temporal information transformation

Neural Network as a universal function



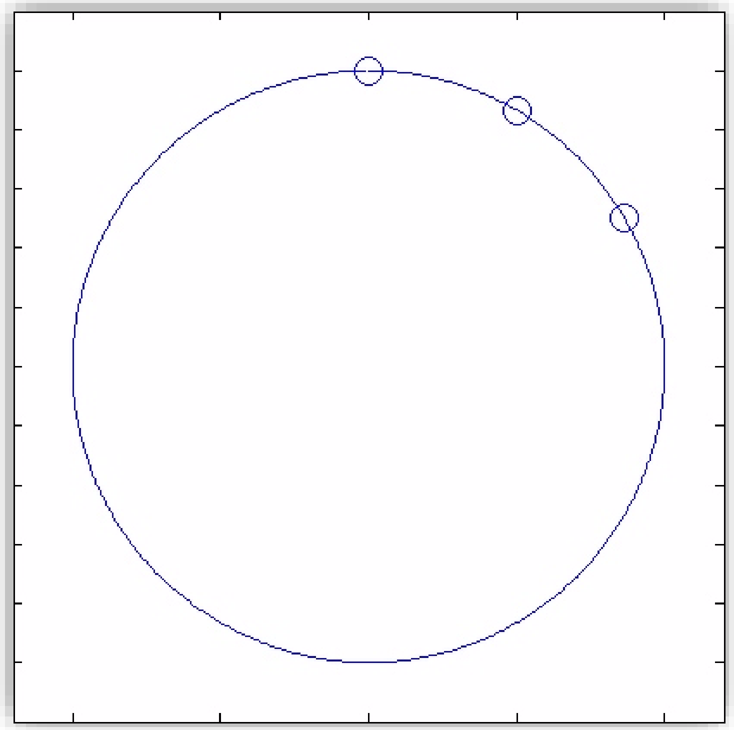
华为 合作

Multi-layer neural network

Two simulations

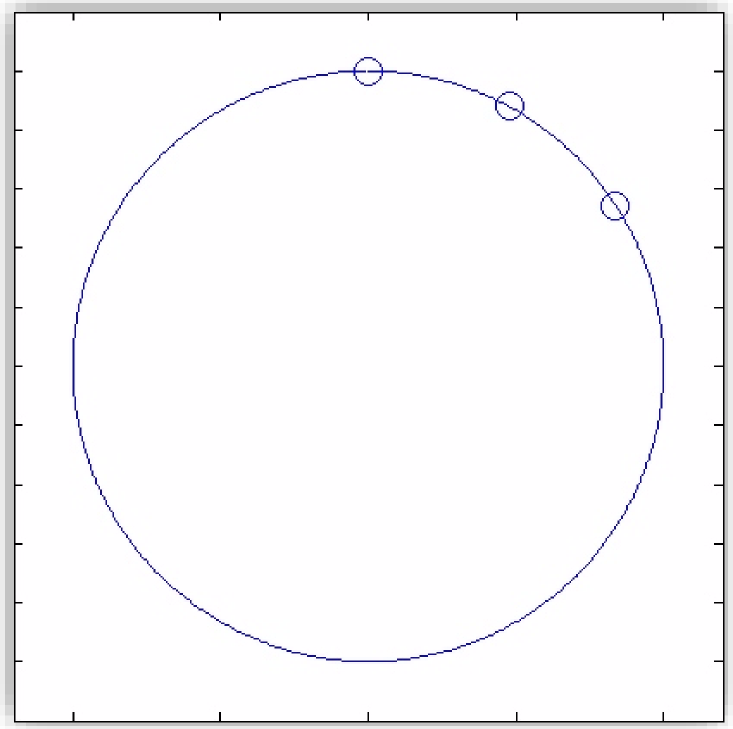
$$\tau = \pi/6$$

rational number for period



$$\tau = 0.5\pi/\sqrt{10}$$

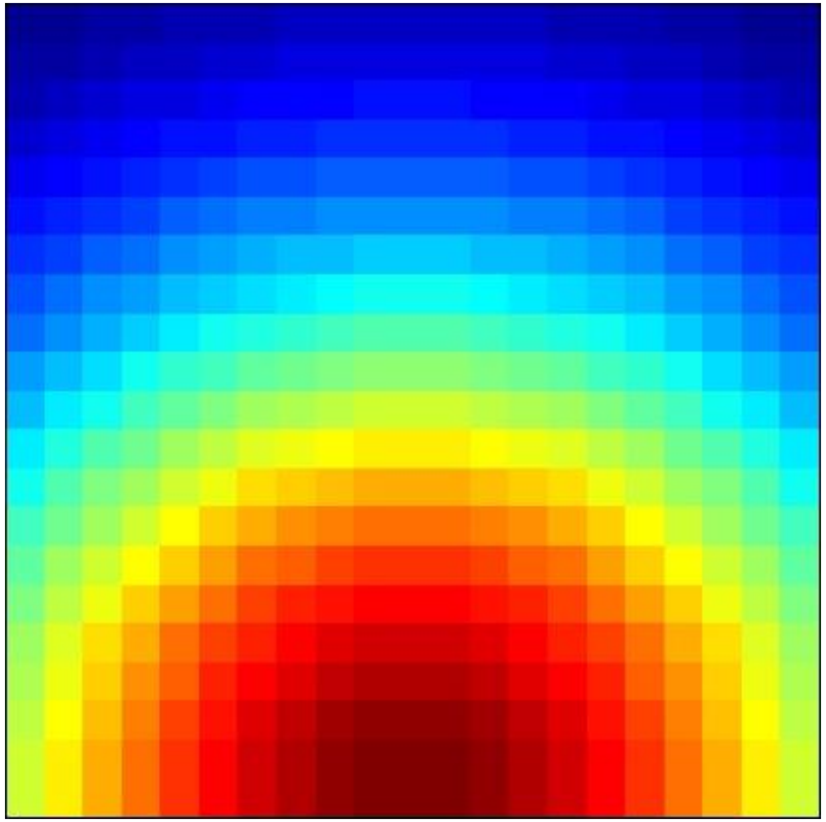
irrational number for period

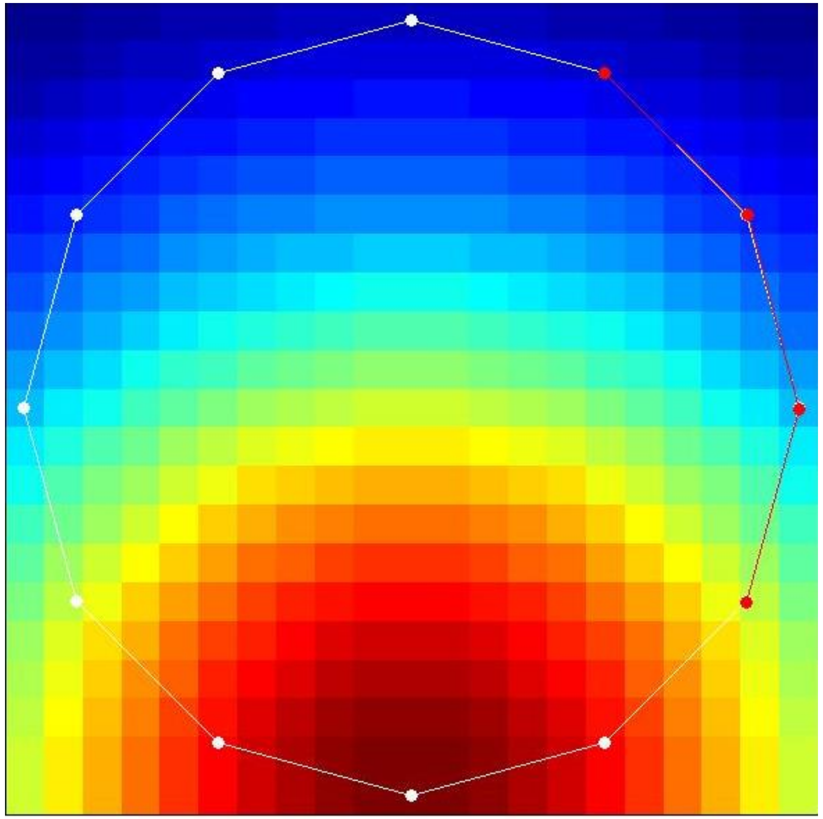


Application to Real Data 2

--RBF functional space

- Image prediction

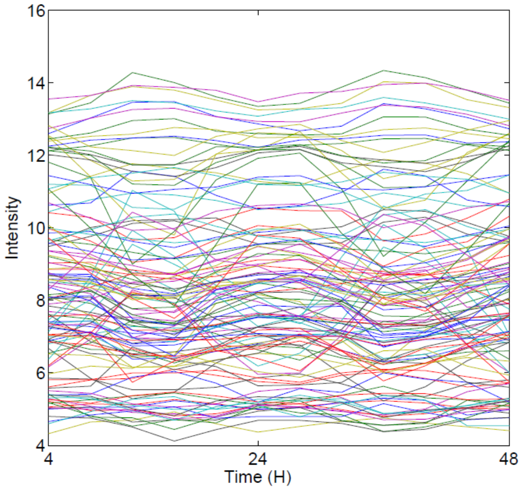


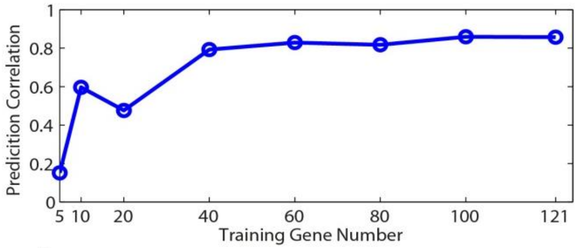


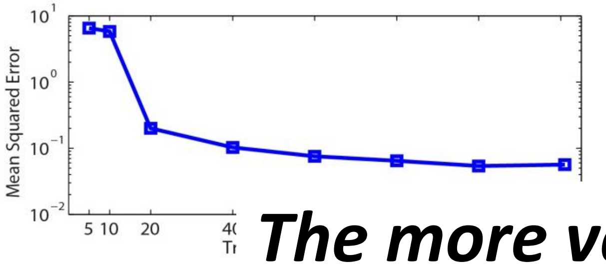
10 rotation pictures

Gene expression prediction

- Circadian microarray data

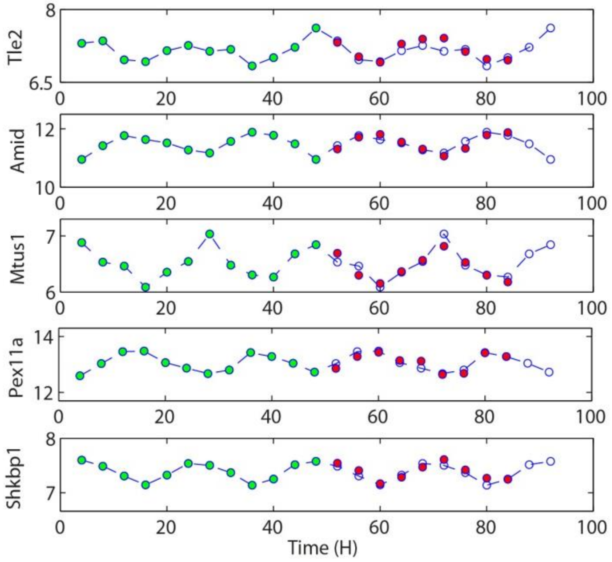






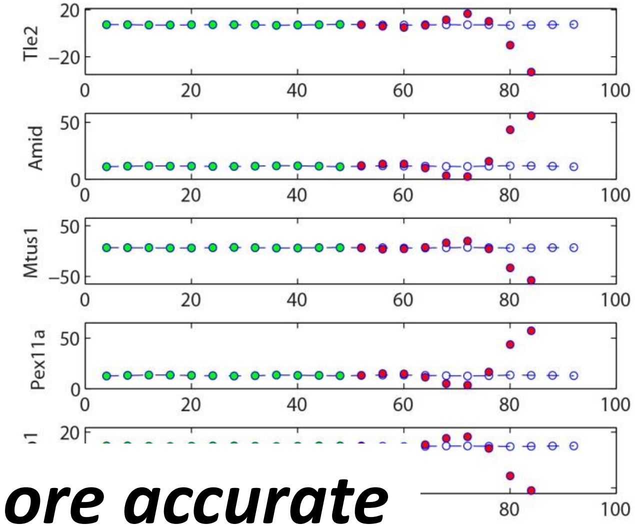
n=121,m=12,
L=10,k=2

High Dimension



n=10,m=12,
L=10,k=2

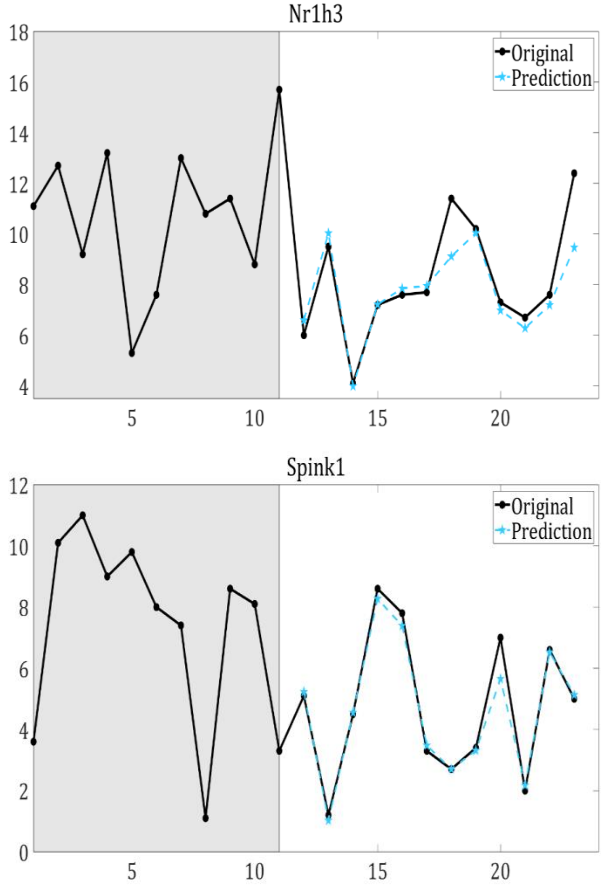
Low Dimension



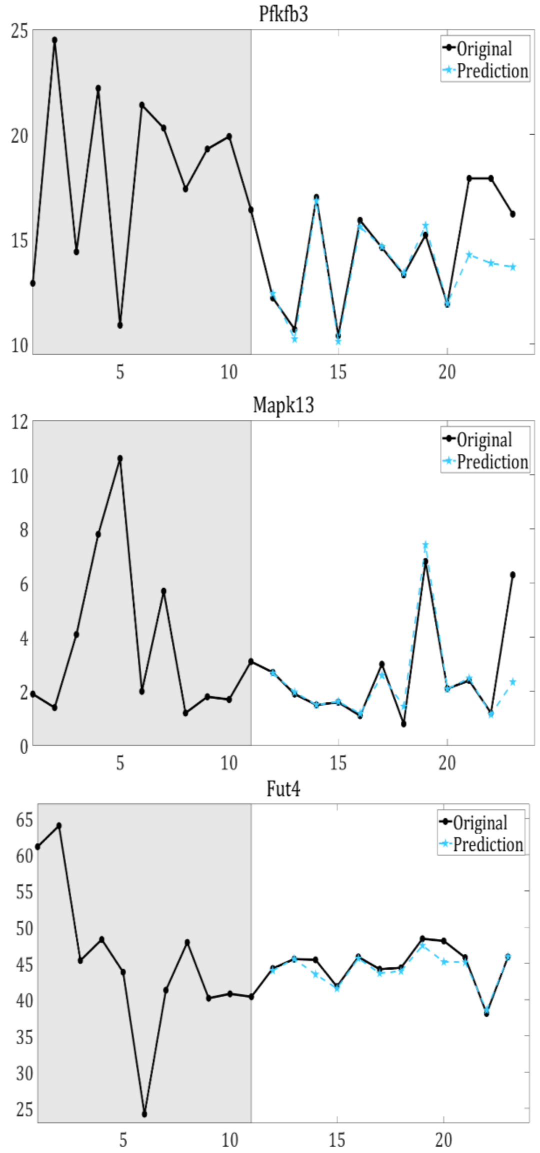
The more variables, the more accurate

- 31099D Gene Expression System (train 11 samples, predict 11 samples)

Non-periodic genes



Genes unrelated to circadian



Genes related to circadian

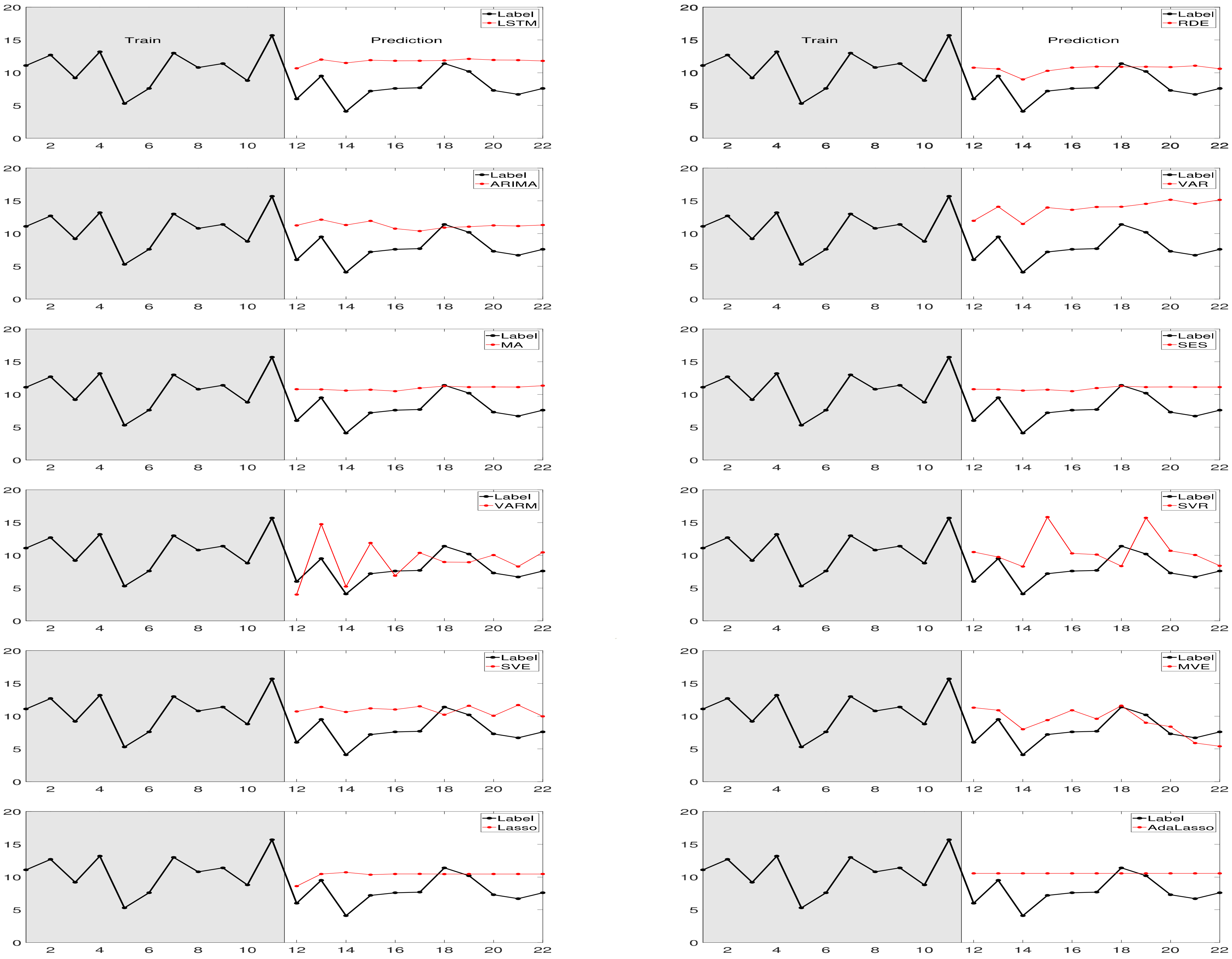
Goal:
Predict gene expression

Time interval:
2 hours

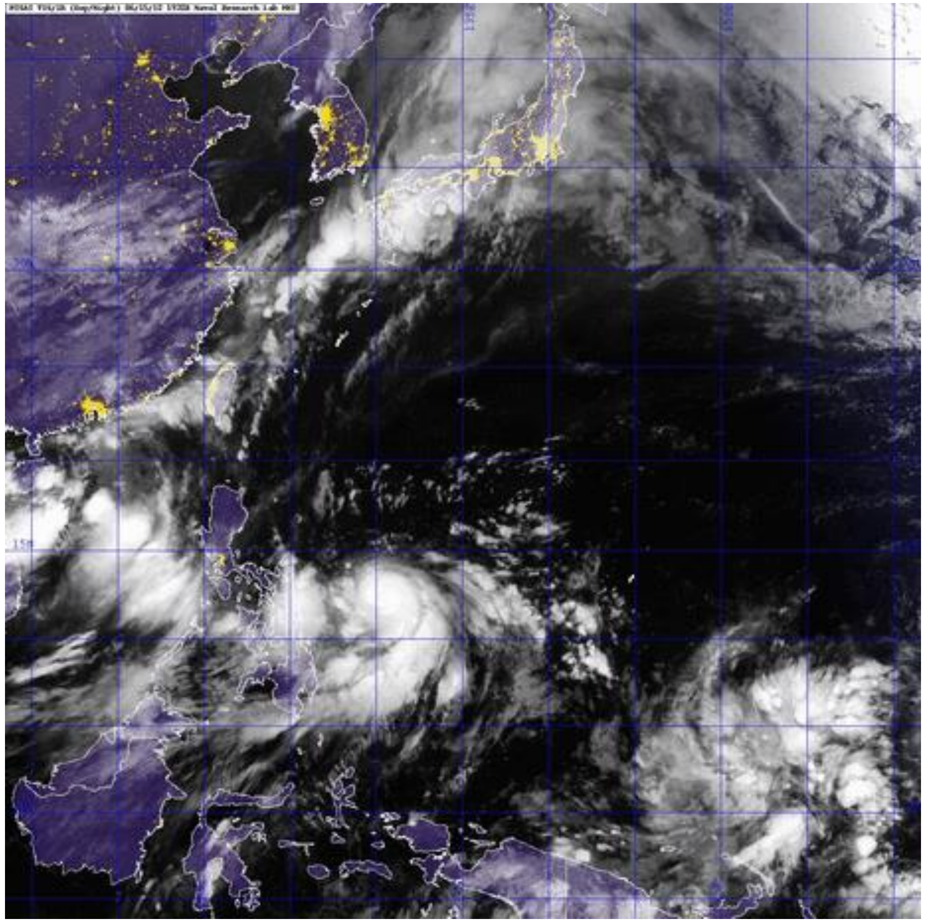
Exiting 12 Methods

- **ARIMA**: a well-known auto-regressive model for predicting future time series;
- **VAR**: a vector auto-regressive model;
- **MA**: moving average using the mean of previous data to make prediction;
- **SES**: Holt-Winters exponential smoothing which uses a moving average with exponentially decreasing weights of previous data to make predictions;
- **VARM**: Basic process of VARMAX (p,q,r) model which includes autoregressive process, moving average process, and independent exogenous terms;
- **SVR**: Support Vector Regression (SVR) which uses SVM to fit curves;
- **SVE**: The classic single-variable embedding;
- **MVE**: The recently proposed multi-view embedding;
- **RDE**: The method for short-term high-dimensional time series predictions;
- **LSTM**: A famous neural network which is widely used in the time series analysis;
- **Lasso**: Lasso procedure is used to estimate parameters of AR(p) for prediction;
- **Adalasso**: AdaLasso procedure is used to estimate parameters of AR(p) for prediction.

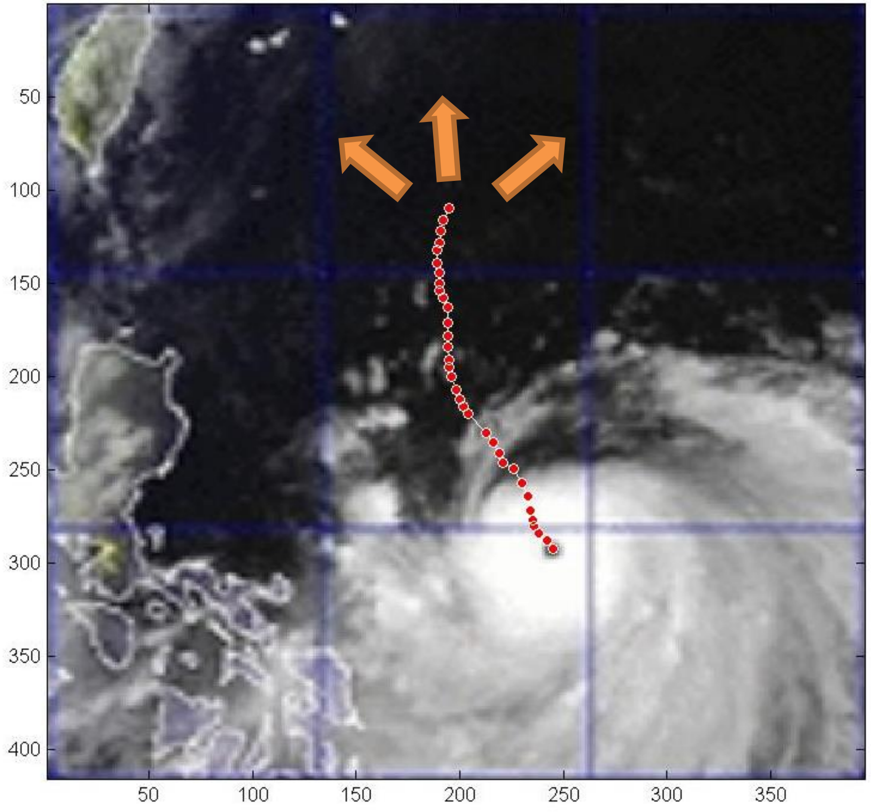
31099D Gene Expression System (all 12 existing methods)



Typhoon prediction

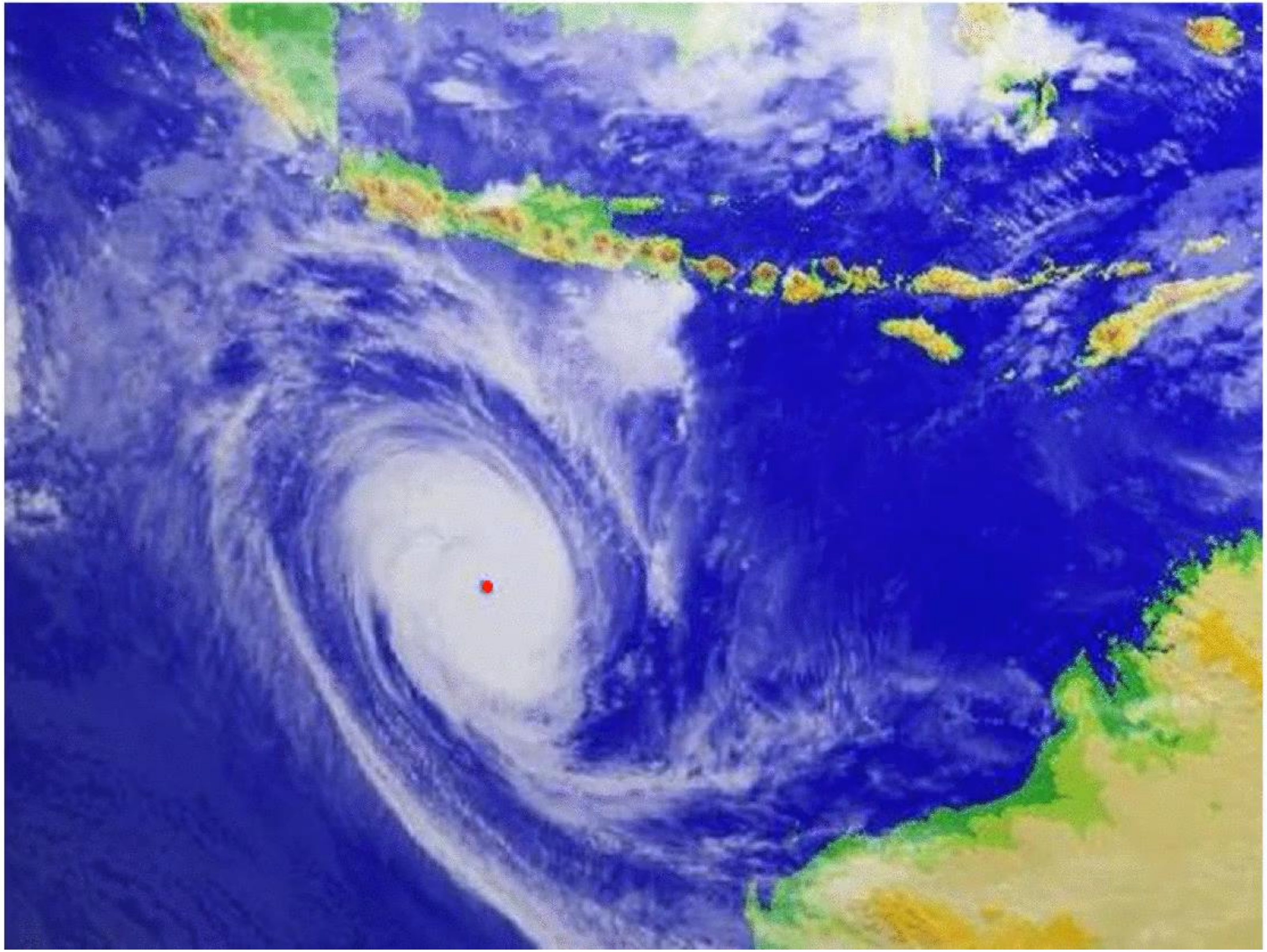


21*21 dimensions



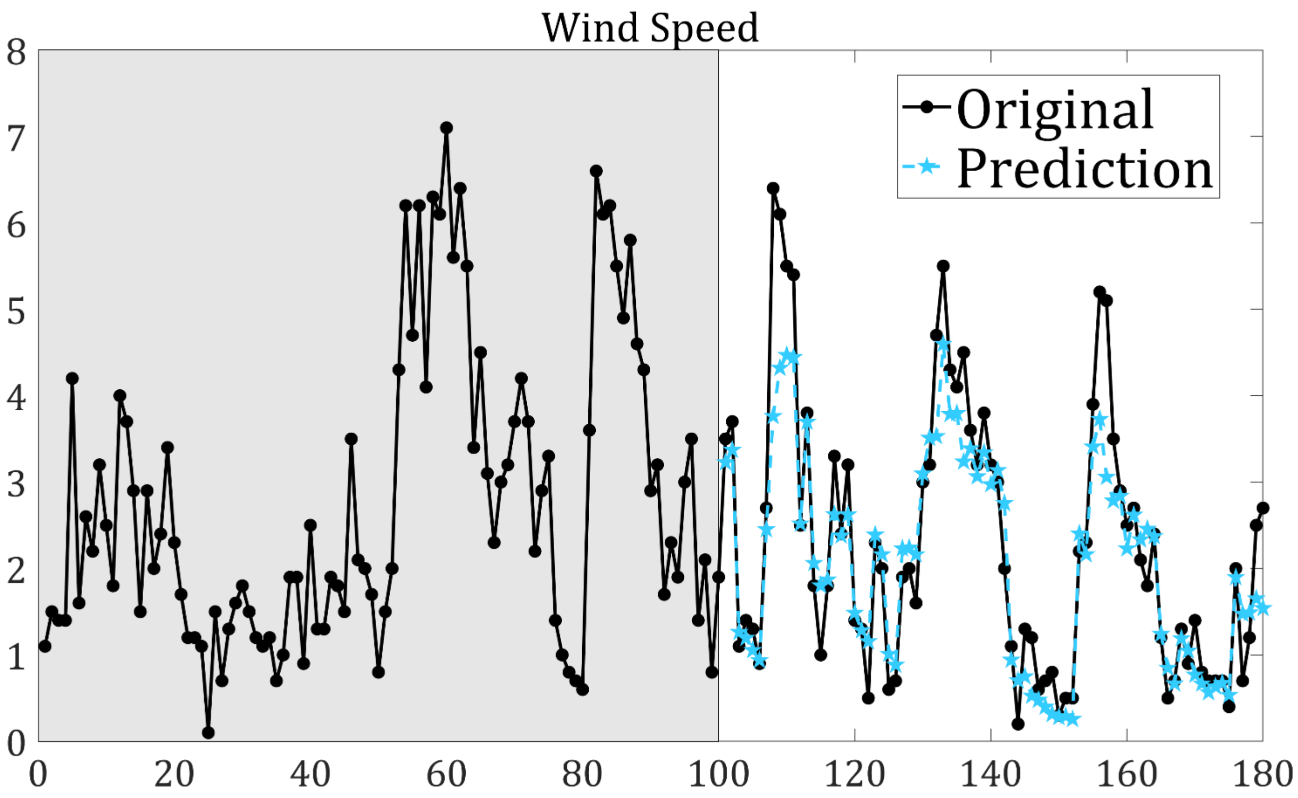
Typhoon in northern hemisphere

Marcus *Time: 18:00:00, March 21, 2018*



Wind Speed Prediction in Wakkanai, Japan

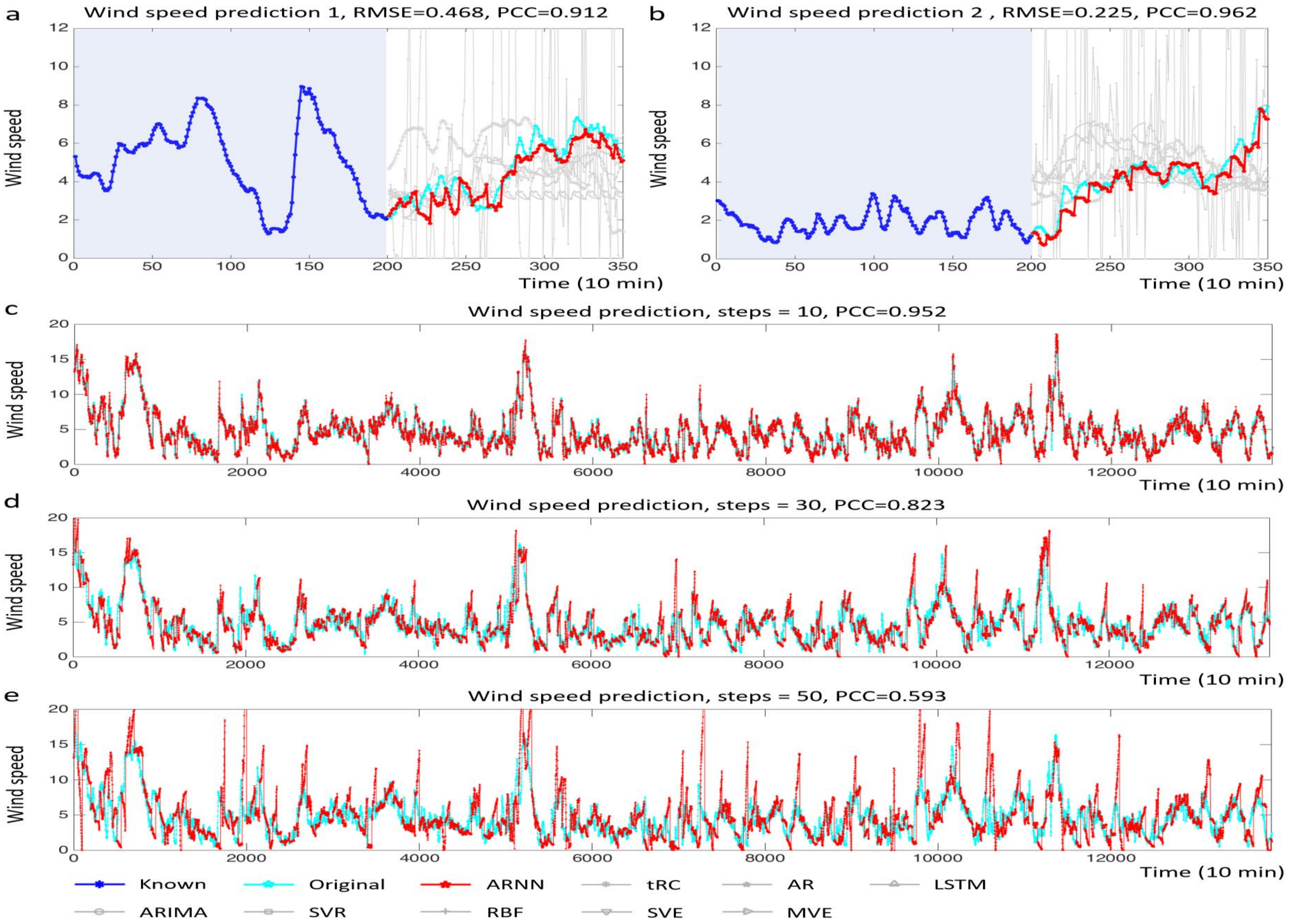
- 924Dim Wind Speed System (train 100 hours, predict 80 hours, with 155 locations)



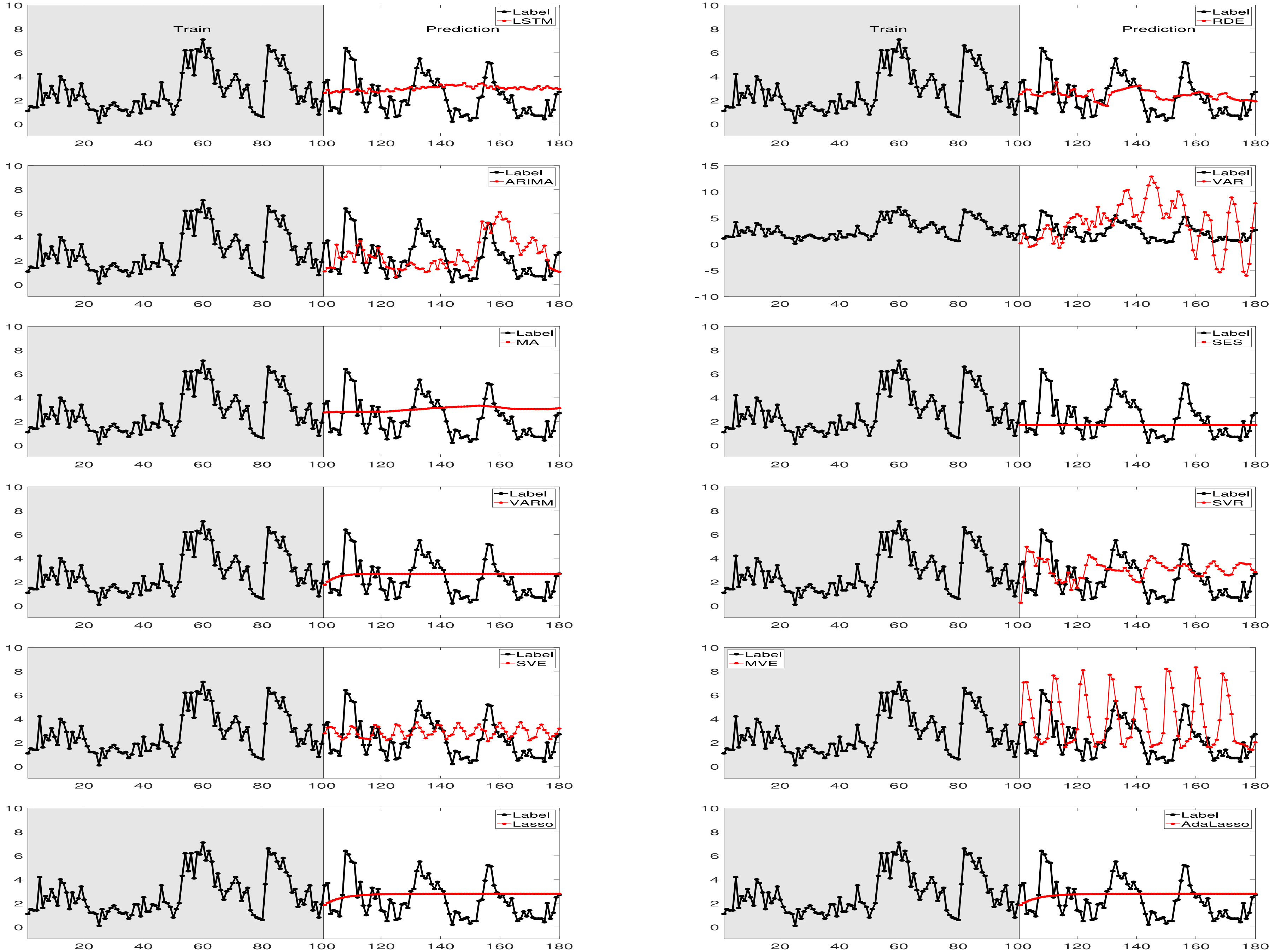
Goal:
Predict future wind speed of a wind station near Tokyo

Time interval:
1 hour

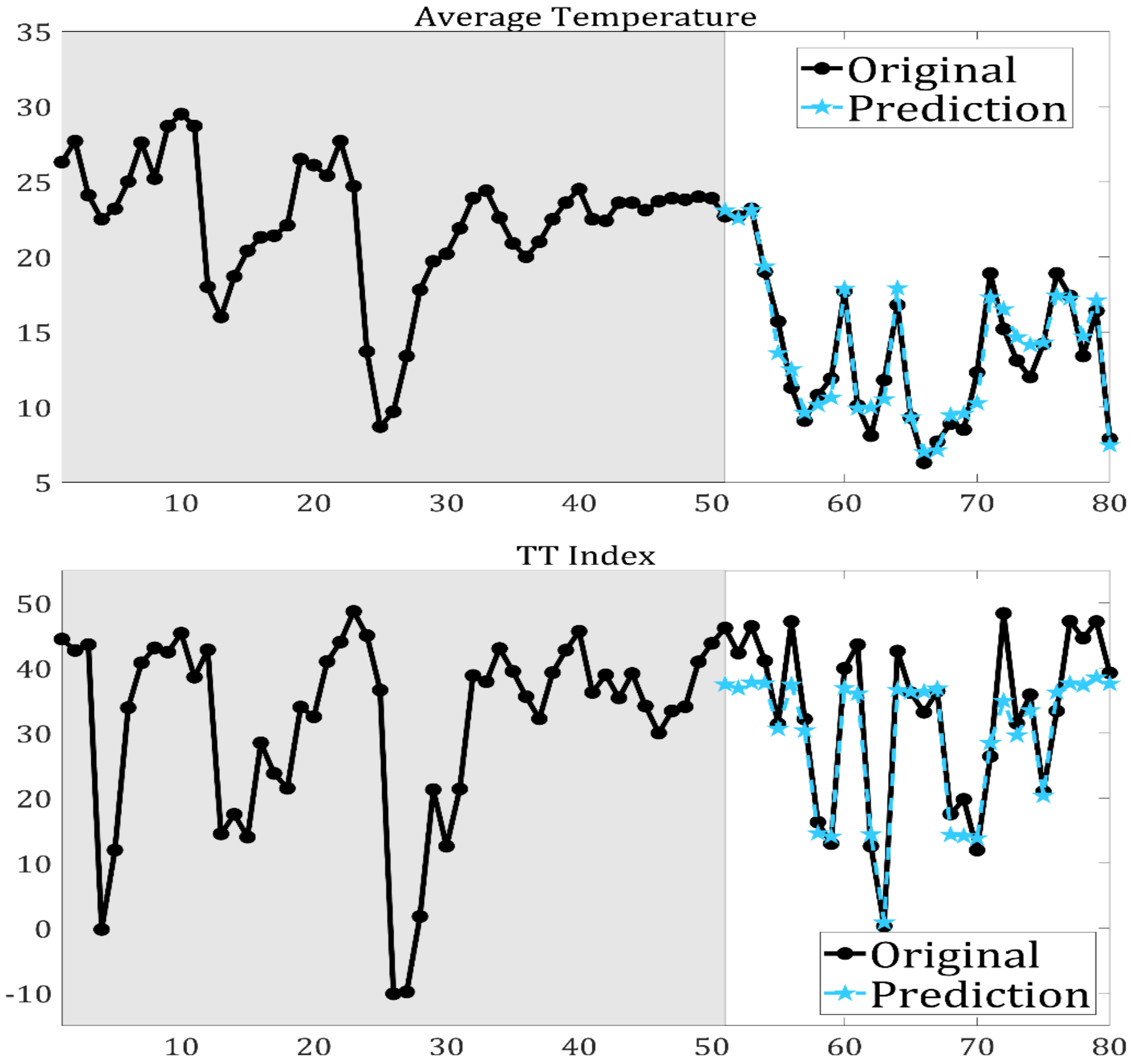
Wind speed prediction for whole 13,860 time points (96 days, interval 10 min) in Wakkanai, Japan



924Dim Wind Speed System (all existing methods)



- 72D Ocean System (train 50 days, predict 30 days)



Goal:
Predict pressure, average temperature and TT index of Ocean

Time interval:
1 day

Box-Counting Dimension

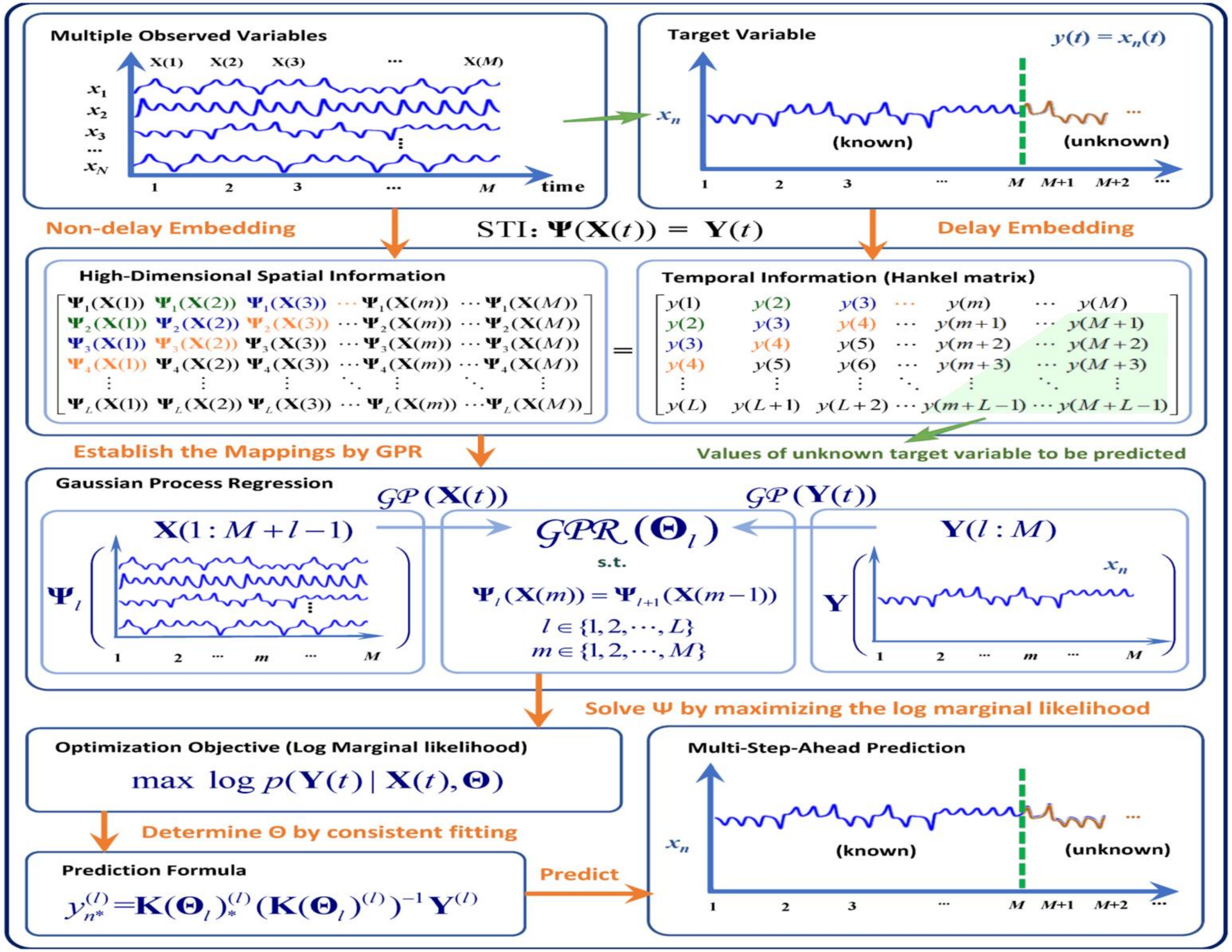
Number	Dataset	Estimated dimension
1	The 90D coupled Lorenz system	2.68 ± 0.335
2	Wind speed in Wakkanai, Japan	3.02 ± 0.224
3	Solar irradiance in Wakkanai, Japan	1.26 ± 0.138
4	Sea-level pressure & Average temperature in US	2.40 ± 0.206
5	Route of typhoon center in Indian Ocean	5.64 ± 0.381
6	Gene expressions related to circadian rhythm	1.86 ± 0.195
7	B-Share index in Shanghai Stock Exchange	1.92 ± 0.173
8	Daily number of cardiovascular inpatients	2.65 ± 0.168
9	Traffic speed in multiple locations in Los Angeles, CA	2.03 ± 0.206

进展

Tao et al. *Information Sciences* 2022

MT-GPRM: assume Gaussian process ?

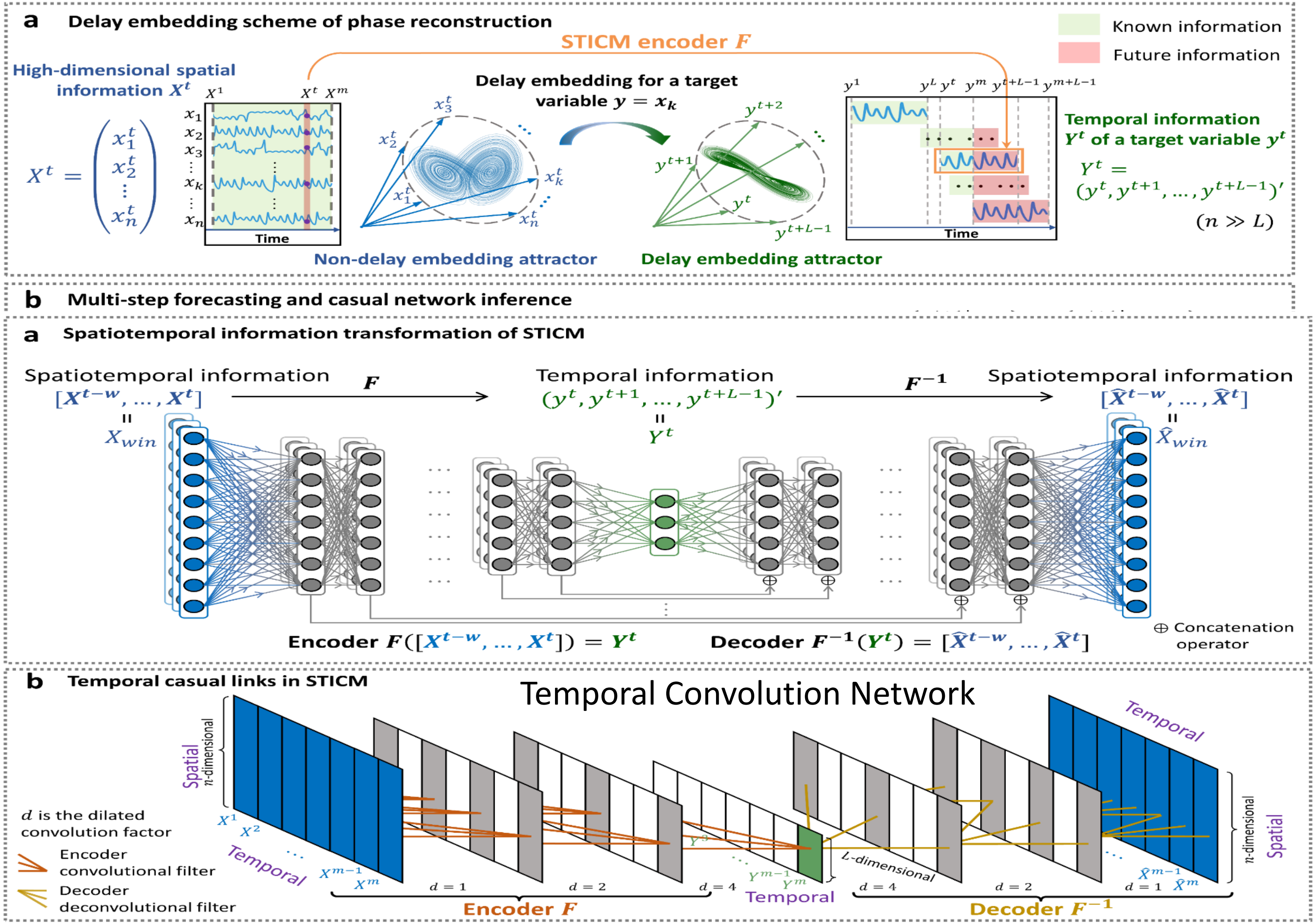
Gaussian Process Regression (GPR) with STI



进展

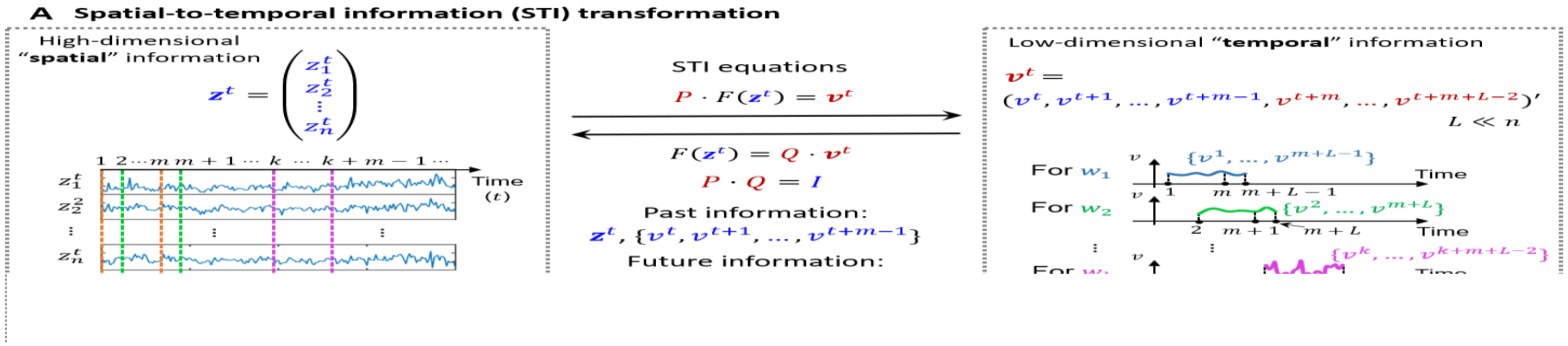
Peng et al. *Fundamental Research* 2023

Spatiotemporal information conversion machine (STICM)



Tong et al. PNAS 2023

地震预警　Earthquake alerting from spatial geodetic data by spatiotemporal information transformation: RSIT



STI

DNB

STI“预测” ？or DNB“预警”？

But

“Predicting” + “Tipping” → Earthquake
(STI)　　　　(DNB)

Tong et al. PNAS 2023 Web tool: https://earthquakepredictionrsit.com/eq_perdict

Performance statistics of RSIT in various regions within **14 days**

Indices / Region	Monitoring years	Number of sliding windows*	Earthquake magnitude (M)	Number of earthquakes	TPR	FPR	Specificity	Accuracy	Average days ahead
All regions									6.08
									6.23
Ibaraki	7								6.13
			$M \geq 5.0$	**133**	**0.82**	**0.0107**	**0.9893**	**0.9804**	**6.23**
West-central Hokkaido	6								6.16
									6.33
Miyagi	4	1465	$3.0 \leq M < 5.0$	200	0.60	0.0126	0.9874	0.9338	5.02
			$M \geq 5.0$	**157**	**0.85**	**0.0084**	**0.9916**	**0.9761**	**6.02**
Sichuan	5	1826	$3.0 \leq M < 5.0$	322	0.34	0.0033	0.9967	0.8817	6.56
			$M \geq 5.0$	**14**	**0.93**	**0.0039**	**0.9961**	**0.9956**	**7.31**
South-central Alaska	8	2922	$3.0 \leq M < 5.0$	371	0.56	0.0074	0.9926	0.9381	6.35
			$M \geq 5.0$	**61**	**0.80**	**0.0112**	**0.9888**	**0.9849**	**6.44**

Weak earthquakes : No

Strong earthquakes: Yes

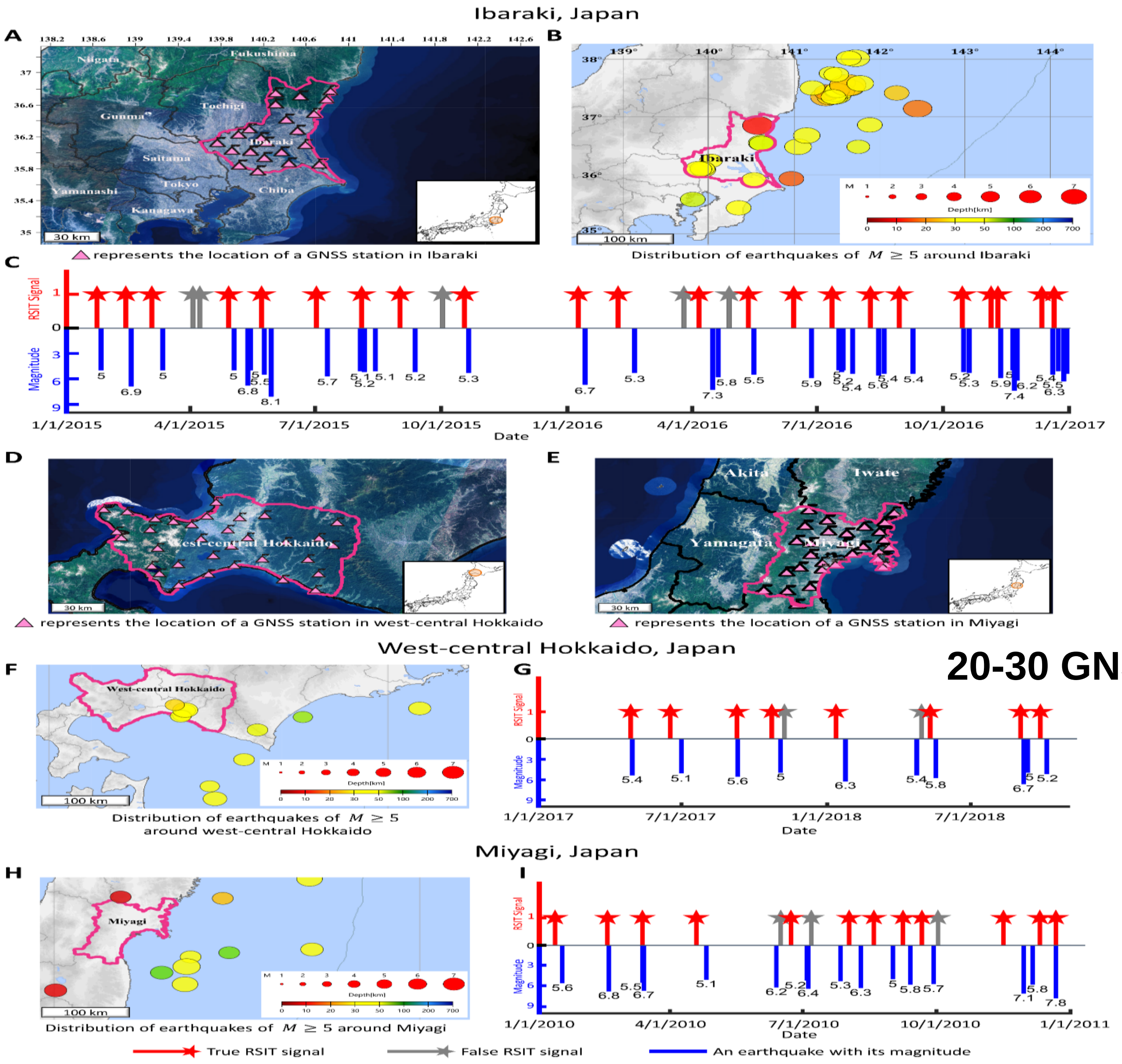
Accuracy is not so important !

Comparison of earthquake-alerting performance of RSIT in five regions with ten existing methods within **14 days**

Methods / Region	Metrics	RSIT 真阳性	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Ibaraki	TPR	**0.82**	0.49	0.62	0.16	0.58	0.74	0.45	0.55	0.69	0.27	0.68
	Aggregative score	**0.87**	0.62	0.69	0.49	0.65	0.71	0.62	0.61	0.72	0.58	0.70
West-central Hokkaido	TPR	**0.81**	0.56	0.63	0.40	0.44	0.54	0.39	0.67	0.68	0.33	0.58
	Aggregative score	**0.80**	0.61	0.65	0.53	0.56	0.60	0.51	0.62	0.68	0.58	0.63
Miyagi	TPR	**0.85**	0.72	0.62	0.55	0.48	0.66	0.68	0.74	0.64	0.64	0.63
	Aggregative score	**0.92**	0.81	0.77	0.65	0.65	0.78	0.78	0.83	0.77	0.73	0.75
Sichuan	TPR	**0.93**	0.21	0.29	0.36	0.50	0.36	0.50	0.57	0.64	0.21	0.57
	Aggregative score	**0.86**	0.43	0.49	0.45	0.52	0.50	0.56	0.55	0.59	0.43	0.55
South-central Alaska	TPR	**0.80**	0.54	0.59	0.49	0.31	0.51	0.62	0.59	0.61	0.43	0.59
	Aggregative score	**0.80**	0.58	0.63	0.52	0.46	0.59	0.63	0.61	0.65	0.63	0.59

M1 is a real-time outlier detection method based on a window-based forecasting model
M2 is a supervised machine learning algorithm designed for GNSS positioning time-series prediction
M3 is based on the definition of the randomness of outliers in GPS time series
M4 is the GPS Interactive Time Series Analysis (GITSA) software program of GPS time series in geodetic and geodynamic studies
M5 represents an ANN for to alerting for earthquakes using real-time GNSS data
M6 is a deep learning algorithm called WANEH that combines wavelets and the Hilbert transform to detect anomalies in time-series data
M7 is a method based on martingale theory to extract anomalies from continuous GPS data as earthquake precursors
M8 represents a graphical method for detecting outliers in time series of continuous daily measurements
M9 represents support vector regression (SVR), a supervised machine learning algorithm
M10 is delayed long short-term memory (dLSTM), an anomaly detection method for time-series data

Early-warning signals of earthquakes with M ≥ 5.0 by RSIT from GNSS of three regions in Japan within 14 days

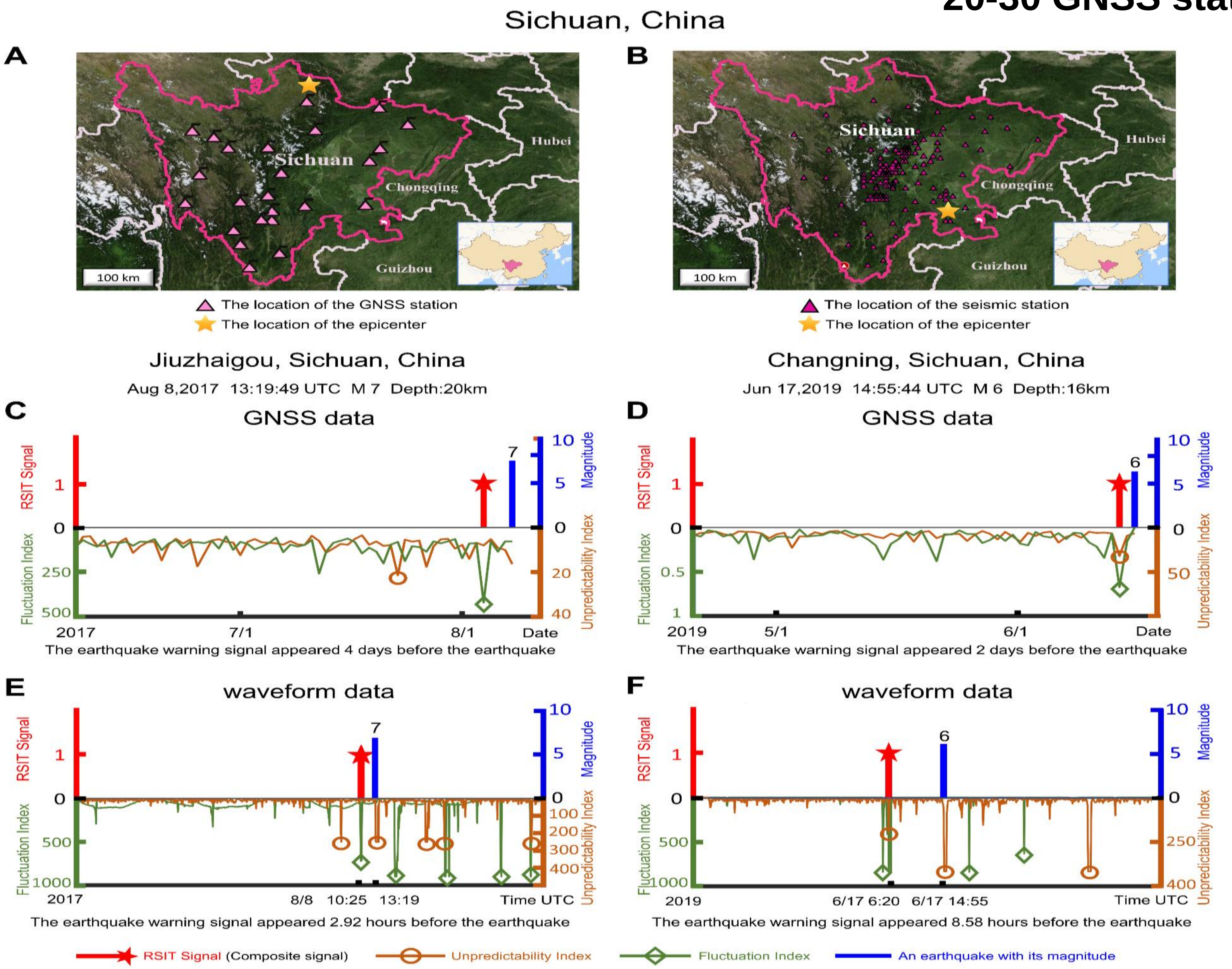


20-30 GNSS stations

Tong et al. PNAS 2023

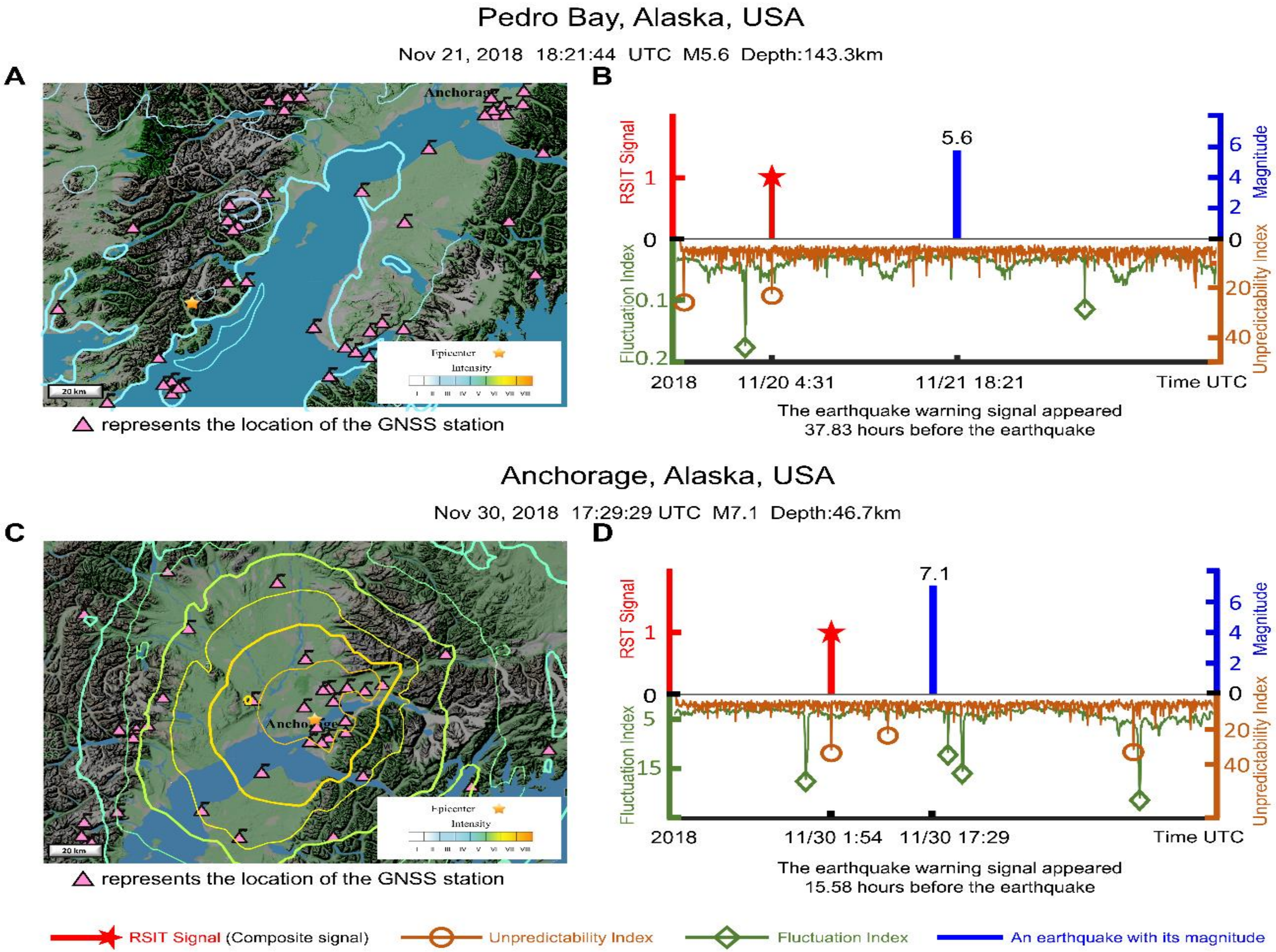
Alerting for earthquakes by RSIT in Sichuan, China from GNSS and seismic waveform data within 14 days

20-30 GNSS stations



Tong et al. PNAS 2023

RSIT alerts for earthquakes from GNSS in south-central Alaska of USA Within 14 days (with 20-30 GNSS stations)



PNAS

Randomly distributed embedding making short-term high-dimensional data predictable

Huanfei Ma[a], Siyang Leng[b,c,d], Kazuyuki Aihara[b,e,1], Wei Lin[c,d,f,g,h,1], and Luonan Chen[i,j,k,l,1]

PNAS 2018

Conclusions

- STI equation
- Prediction future with short high-dim data
- Back to future with short high-dim data
- Algorithms for linear and nonlinear forms
- Making small samples as equivalently large

Application to many areas, i.e. economics, ecological systems, bio-systems, AI

NSR National Science Review

RESEARCH ARTICLE

National Science Review
7: 1–13, 2020
doi: 10.1093/nsr/nwaa025
Advance access publication 19 February 2020

[1]School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510275, China; [2]Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai 200031, China; [3]School of

INFORMATION SCIENCE

Predicting future dynamics from short-term time series using an Anticipated Learning Machine

Chuan Chen[1], Rui Li[1], Lin Shu[1], Zhiyu He[1], Jining Wang[1], Chengming Zhang[2], Huanfei Ma[3], Kazuyuki Aihara[4,5] and Luonan Chen[2,6,7,8,*]

Delay Embedding → Reconstructing Attractor

Tong et al. PNAS 2023

Predicting + Tipping

PNAS RESEARCH ARTICLE | APPLIED MATHEMATICS

Earthquake alerting based on spatial geodetic data by spatiotemporal information transformation learning

Yuyan Tong[a], Renhao Hong[a], Ze Zhang[b], Kazuyuki Aihara[c], Pei Chen[a,1], Rui Liu[a,1], and Luonan Chen[b,d,e,f,1]

Edited by Peter Kelemen, Lamont-Doherty Earth Observatory, Palisades, NY; received February 18, 2023; accepted June 7, 2023 nature COMMUNICATIONS

Chen et al. NC 2020

ARTICLE

https://doi.org/10.1038/s41467-020-18381-0 OPEN

Autoreservoir computing for multistep ahead prediction based on the spatiotemporal information transformation

Pei Chen[1], Rui Liu[1✉], Kazuyuki Aihara[2,3] & Luonan Chen[4,5,6,7✉]

Conclusions

- Autoencoder X→Y→X

  Autoreservoir F(X) → Y → F(X)
- Reservoir computing → external dynamics X as reservoir;

  Autoreservoir → internal dynamics X as reservoir → save energy
- Statistically dimension reduction

  phase reconstruction of dynamics by a low dim system.
- spatiotemporal information transformation

Application to many areas, i.e. economics, ecological systems, bio-systems, AI

Acknowledgments

- Pei Chen, Rui Liu, Tong
  South China University of Science and Technology
- Peng Tao, Xiaohu Hao
  Hangzhou Institute for Advanced Study, UCAS
- Huanfei Ma
  Suzhou University
- Kazuyuki Aihara, Siyang Leng
  The University of Tokyo
- Wei Lin
  Fudan University
- Chuan Chen
  Sun Yat-sen University

Chen et al. Nature Communications 2020
Ma et al. PNAS 2018
Chen et al. National Science Review 2020
Leng et al. Nature Communications 2020
Peng et al. Fundamental Research 2023
Tao et al. Information Sciences 2022
Tong et al. PNAS 2023