

Data-driven stochastic model for cross-interacting processes with different time scales

Andrey Gavrilov, Aleksei Seleznev, Dmitry Mukhin, Alexander Feigin

Institute of Applied Physics of the Russian Academy of Sciences
Nizhny Novgorod, Russian Federation



Abstract

The problem of modeling interaction between processes with different time scales is very important in geoscience. In this report, we propose a new form of empirical evolution operator model based on the analysis of multiple time series representing processes with different time scales. We assume that the time series are given on the same time interval.

To construct the model, we extend the previously developed general form of nonlinear stochastic model based on artificial neural networks and designed for the case of time series with constant sampling interval [1]. This sampling interval is related to the main time scale of the process under consideration, which is described by the deterministic component of the model, while the faster time scales are modeled by its stochastic component, possibly depending on the system's state. This model also includes slower processes in the form of weak time-dependence, as well as external forcing. The structure of the model is optimized using Bayesian approach [1]. The model has proven its efficiency in a number of applications [2, 3, 4].

The idea of modeling time series with different time scales is to formulate the above-described model individually for each time scale, and then to include the parameterized influence of the other time scales in it. Particularly, the influence of “slower” time series is included in the form of parameter trends, and the influence of “faster” time series is included by time-averaging their statistics. The algorithm and first results of comparison between the new model and the model without cross-interactions will be discussed.

The work was supported by the Russian Science Foundation (Grant No. 20-62-46056).

Overview

- 1 Introduction
- 2 Multi-scale stochastic model
 - Two-scale data case
 - Generalizations
- 3 Example: slow Van der Pol and fast linear stochastic oscillators
 - Data generation
 - System with bi-directional connection
 - System with uni-directional connection: fast-to-slow
 - System with uni-directional connection: slow-to-fast
 - System with no connection
- 4 Conclusions

Overview

- 1 Introduction
- 2 Multi-scale stochastic model
 - Two-scale data case
 - Generalizations
- 3 Example: slow Van der Pol and fast linear stochastic oscillators
 - Data generation
 - System with bi-directional connection
 - System with uni-directional connection: fast-to-slow
 - System with uni-directional connection: slow-to-fast
 - System with no connection
- 4 Conclusions

Bayesian ANN-based stochastic model

- Consider the data – N observations of the variable x with the dimension D , with the constant sampling interval Δt : $x_n = x(t_0 + n\Delta t)$, $n = 1, \dots, N$.
- For the data $x_1, \dots, x_N$ the following stochastic model can be constructed:

$$x_n = F(x_{n-1}, \dots, x_{n-L}) + \hat{G}(x_n, \dots, x_{n-L}) \cdot \xi_n \quad (1)$$

- Nonlinear vector and matrix functions $F : \mathbb{R}^{D \cdot L} \rightarrow \mathbb{R}^D$ and $\hat{G} : \mathbb{R}^{D \cdot L} \rightarrow \mathbb{R}^{D \times D}$ are parameterized using simple artificial neural networks (ANNs)
- Gaussian white noise is assumed in the residuals $\xi_n \in \mathbb{R}^D$
- Bayesian approach is used to find optimal hyperparameters (memory depth L and numbers of neurons in ANNs) and train ANN coefficients in F and $\hat{G}$
- See more details and examples in references [1, 5]

Bayesian ANN-based stochastic model: modifications

- The model can be modified to take into account explicit time t_n (weak non-stationarity) or external forcing signal Q_n :

$$x_n = F(x_{n-1}, \dots, x_{n-L}; t_n, Q_n) + \hat{G}(x_n, \dots, x_{n-L}) \cdot \xi_n \quad (2)$$

- The Bayesian algorithm for model training and optimization remains the same.
- These modifications of the model were successfully:
 - tested on synthetic climate examples [2, 6],
 - used to make a real-time El Nino Southern Oscillation forecast [3], marked “IAP-NN” in the IRI/CPC ENSO Predictions Plume¹
 - investigate mechanisms of the Middle Pleistocene Transition [4]
- There are modifications using more complex ANN structure in F and $\hat{G}$, see [7, 8]

¹<https://iri.columbia.edu/our-expertise/climate/forecasts/enso/current/>

Bayesian ANN-based stochastic model: properties

$$x_n = F(x_{n-1}, \dots, x_{n-L}; t_n, Q_n) + \hat{G}(x_n, \dots, x_{n-L}) \cdot \xi_n$$

- The model is fitted to reproduce the evolution with the time step Δt dictated by the **single dataset** $x_1, \dots, x_N$
- The main time scale is described by the deterministic part $F(x_{n-1}, \dots, x_{n-L}; t_n, Q_n)$ in the low-dimensional state space. To be well-captured, this scale must be comparable with the sampling interval Δt
- The contribution of subsample (fast) time scales less than Δt is modeled via the stochastic component $\hat{G}(x_n, \dots, x_{n-L}) \cdot \xi_n$
- The influence of slow trends can be modeled via explicit weak dependence on time t_n

We will call this model **single-scale stochastic model**

In this work, we consider the next step – **multi-scale stochastic model**

Overview

- 1 Introduction
- 2 Multi-scale stochastic model
 - Two-scale data case
 - Generalizations
- 3 Example: slow Van der Pol and fast linear stochastic oscillators
 - Data generation
 - System with bi-directional connection
 - System with uni-directional connection: fast-to-slow
 - System with uni-directional connection: slow-to-fast
 - System with no connection
- 4 Conclusions

Two-scale data

- Consider the two-scale data – time series on the same time interval $(0, N \cdot \Delta t]$, but with two different sampling rates
- “Fast” process – N observations **sampled with the interval Δt** :

$$u_1, u_2, \dots, u_N$$

- “Slow” process – $M = N/s$ observations **sampled with the interval $s \cdot \Delta t$** :

$$x_s, x_{2s}, \dots, x_{M \cdot s}$$

- Useful notation (upsampling rule):

$$\tilde{x}_k := x_{s \cdot \lfloor k/s \rfloor}$$

Here $\lfloor \cdot \rfloor$ means “integer part” operation.

Two-scale data: individual models

$u_1, u_2, \dots, u_N$ – “fast” process

$x_s, x_{2s}, \dots, x_{M \cdot s}$ – “slow” process

- First option – single-scale stochastic model (single-scale SM) for each process
- “Fast single-scale SM”:

$$u_n = F(u_{n-1}, \dots, u_{n-L_u}) + \hat{G}(u_{n-1}, \dots, u_{n-L_u}) \cdot \xi_n^u \quad (3)$$

- “Slow single-scale SM”:

$$x_n = f(x_{n-s}, \dots, x_{n-l_x \cdot s}) + \hat{g}(x_{n-s}, \dots, x_{n-l_x \cdot s}) \cdot \xi_n^x \quad (4)$$

- Parameterization, optimization and training procedure are the same as in Eq. (1)

Set of single-scale SMs

Two-scale data: coupled model

$u_1, u_2, \dots, u_N$ – “fast” process

$x_s, x_{2s}, \dots, x_{M \cdot s}$ – “slow” process

- Another option – including interactions between processes in the model
- Adding the influence of slow process to the fast single-scale SM – parametric slow forcing:

$$\begin{aligned} u_n &= F(u_{n-1}, \dots, u_{n-L_u}; Q_n) + \hat{G}(u_{n-1}, \dots, u_{n-L_u}; Q_n) \cdot \xi_n^u \\ Q_n &= (\tilde{x}_{n-1-s \cdot 0}, \dots, \tilde{x}_{n-1-s \cdot (L_x-1)}) \end{aligned} \quad (5)$$

- Adding the influence of fast process to the slow single-scale SM – time-averaged nonlinear mapping $q(\dots)$ of fast variables:

$$\begin{aligned} x_n &= f(x_{n-s}, \dots, x_{n-l_x \cdot s}; q_n) + \hat{g}(x_{n-s}, \dots, x_{n-l_x \cdot s}; q_n) \cdot \xi_n^x \\ q_n &= \langle q(u_k, \dots, u_{k-(l_u-1)}) \rangle_{k=n-l_q, \dots, n-1} \end{aligned} \quad (6)$$

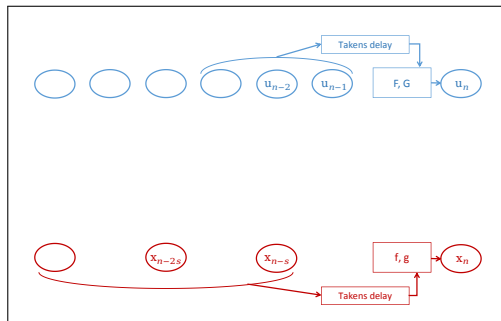
A multi-scale stochastic model

Two-scale data: graphical scheme of the models

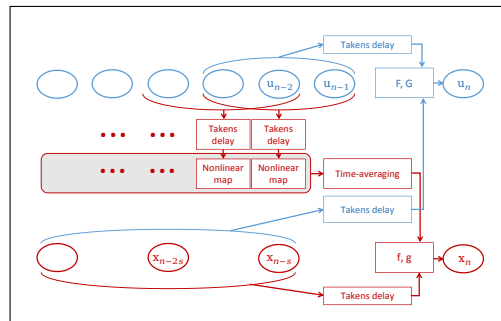
$u_1, u_2, \dots, u_N$ – “fast” process

$x_s, x_{2s}, \dots, x_{M \cdot s}$ – “slow” process

Set of single-scale SMs



Multi-scale SM



“Takens delay” refers to the method of embedding by multiple time-shifting of a variable to the past up to a certain lag [9].

Two-scale data: exact ANN parameterization, full Bayesian training and optimization algorithm

**The information will become available
in the upcoming paper!**

Generalizations

- The proposed two-scale model can be readily generalized to the case of any number of time scales by including all pairs of interactions into individual single-scale SMs
- Non-stationarity and external forcing can be included into multi-scale model in the same way as in Eq. 2.

Overview

- 1 Introduction
- 2 Multi-scale stochastic model
 - Two-scale data case
 - Generalizations
- 3 Example: slow Van der Pol and fast linear stochastic oscillators
 - Data generation
 - System with bi-directional connection
 - System with uni-directional connection: fast-to-slow
 - System with uni-directional connection: slow-to-fast
 - System with no connection
- 4 Conclusions

Fast process

- A discrete stochastic linear oscillator ($\Delta > 0, \Omega > 0$):

$$\begin{cases} u_{n+1} = e^{-\Delta} \cdot (u_n \cos \Omega + v_n \sin \Omega) + \sigma_u \xi_n^u, \\ v_{n+1} = e^{-\Delta} \cdot (v_n \cos \Omega - u_n \sin \Omega) + \sigma_v \xi_n^v. \end{cases} \quad (7)$$

- The attractor of the deterministic part is stable focus ($x = 0, y = 0$) with decrement Δ and frequency $\frac{\Omega}{2\pi}$

Slow process

- A discrete stochastic Van der Pol oscillator ($\delta > 0$, $\omega > 0$, $r > 0$):

$$\begin{cases} x_{n+1} = e^{\delta \cdot \left[1 - \frac{x_n^2 + y_n^2}{r^2}\right]} \cdot (x_n \cos \omega + y_n \sin \omega) + \sigma_x \xi_n^x, \\ y_{n+1} = e^{\delta \cdot \left[1 - \frac{x_n^2 + y_n^2}{r^2}\right]} \cdot (y_n \cos \omega - x_n \sin \omega) + \sigma_y \xi_n^y. \end{cases} \quad (8)$$

- The attractor of the deterministic part is a stable limit cycle with the center $(x = 0, y = 0)$, radius r , frequency $\frac{\omega}{2\pi}$ and radius decrement δ
- The point $(x = 0, y = 0)$ is an unstable focus of the deterministic part with the increment δ and frequency $\frac{\omega}{2\pi}$

Slow-to-fast connection

- Linear oscillator frequency $\frac{\Omega}{2\pi}$ depends on the slow variable x_n :

$$\begin{cases} u_{n+1} = e^{-\Delta} \cdot (u_n \cos \Omega_n + v_n \sin \Omega_n) + \sigma_u \xi_n^u, \\ v_{n+1} = e^{-\Delta} \cdot (v_n \cos \Omega_n - u_n \sin \Omega_n) + \sigma_v \xi_n^v. \end{cases} \quad (9)$$

- If slow-to-fast connection is **ON**:

$$\begin{aligned} \Omega_n &= \bar{\Omega} \cdot (1 + 0.7 \cdot \tanh(0.5x_n)), \\ \Delta &= 0.1, \quad \bar{\Omega} = \frac{2\pi}{10}, \quad \sigma_u = \sigma_v = 0.2. \end{aligned} \quad (10)$$

- If slow-to-fast connection is **OFF**:

$$\Omega_n = \bar{\Omega} = \frac{2\pi}{10}, \quad \Delta = 0.1, \quad \sigma_u = \sigma_v = 0.2. \quad (11)$$

- In both cases, the characteristic time scale of the process is ~ 10 (oscillation period)

Fast-to-slow connection

- Limit cycle radius (amplitude) r depends on the fast variable u_n :

$$\begin{cases} x_{n+1} = e^{\delta \cdot \left[1 - \frac{x_n^2 + y_n^2}{r_n^2}\right]} \cdot (x_n \cos \omega + y_n \sin \omega) + \sigma_x \xi_n^x, \\ y_{n+1} = e^{\delta \cdot \left[1 - \frac{x_n^2 + y_n^2}{r_n^2}\right]} \cdot (y_n \cos \omega - x_n \sin \omega) + \sigma_y \xi_n^y. \end{cases} \quad (12)$$

- If fast-to-slow connection is **ON**:

$$\begin{aligned} r_n &= 1 + 1.2 \cdot \langle u_k^2 \rangle_{k=n-30+1}^n, \\ \delta &= 0.05, \quad \omega = \frac{2\pi}{100}, \quad \sigma_x = \sigma_y = 0.01. \end{aligned} \quad (13)$$

- If fast-to-slow connection is **OFF**:

$$r_n = 1.27, \quad \delta = 0.004, \quad \omega = \frac{2\pi}{100}, \quad \sigma_x = \sigma_y = 0.01. \quad (14)$$

- In both cases, the characteristic time scale of the process is ~ 100 (limit cycle period)

Data for the stochastic model construction

- We simulate the observed data by the above-described coupled system on the interval of $N = 10000$ steps, turning different connections **ON/OFF**
- “Fast” data – variable u sampled with the step 1 (10000 time samples):

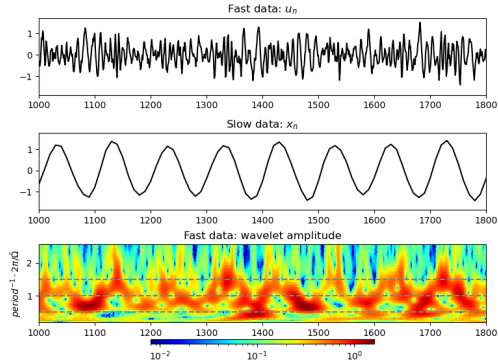
$$u_1, u_2, \dots, u_{10000} \quad (15)$$

- “Slow” data – variable x sampled with the step 10 (1000 time samples):

$$x_{10}, x_{20}, \dots, x_{10000} \quad (16)$$

Bi-directional connection:

Slow-to-fast connection is **ON**, Fast-to-slow connection is **ON**



Fragments of fast time series, slow time series,
and wavelet spectrum of the fast time series

Bi-directional connection:

Slow-to-fast connection is **ON**, Fast-to-slow connection is **ON**

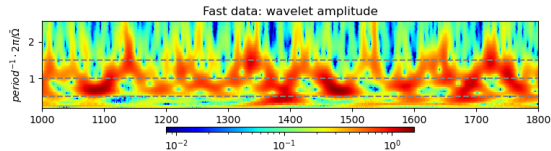


Figure: Fragment of wavelet spectrum of the fast time series

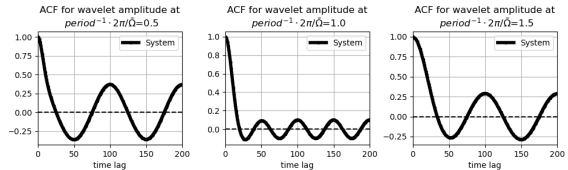


Figure: Autocorrelation function (ACF) of the wavelet amplitudes of the fast process at $0.5 \frac{\bar{\Omega}}{2\pi}$, $1.0 \frac{\bar{\Omega}}{2\pi}$ and $1.5 \frac{\bar{\Omega}}{2\pi}$ frequencies, computed using 10^7 fast time samples

The instant frequency of the fast process is oscillating with the slow limit cycle period.
This modulation is reflected by the periodic behaviour of the wavelet amplitude ACF.

Bi-directional connection:

Slow-to-fast connection is **ON**, Fast-to-slow connection is **ON**

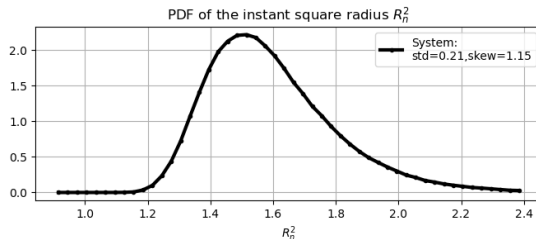


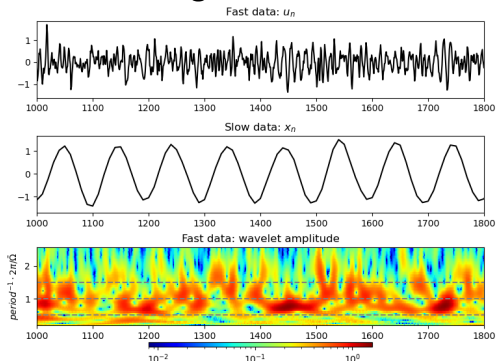
Figure: Probability density function (PDF) of period-average square radius of the slow limit cycle, defined as $R_n^2 = 2 \sum_{j=1}^{10} x_{s(n+j-1)}^2$, computed using 10^6 slow time samples.

The instant slow limit cycle square radius is perturbed by the noise and the fast process.
The uncertainty is reflected by the skewed PDF

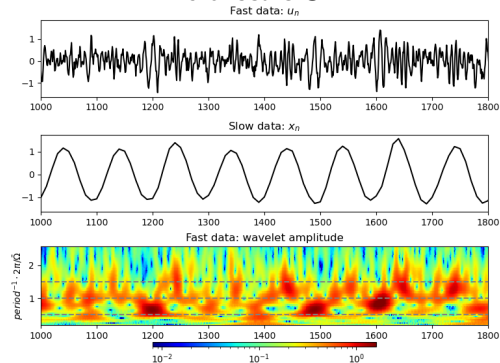
Bi-directional connection:

Slow-to-fast connection is **ON**, Fast-to-slow connection is **ON**

Single-scale SM



Multi-scale SM



Fragments of fast time series, slow time series, and wavelet spectrum of the fast time series

Both models reproduce the main time scales

Bi-directional connection:

Slow-to-fast connection is **ON**, Fast-to-slow connection is **ON**

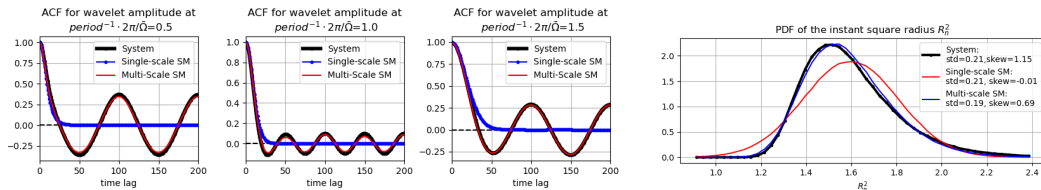


Figure: Comparison of the wavelet amplitude ACFs and square radius PDF for the optimal single-scale and multi-scale SMs, computed in the same way as for the system.

The features related to the interaction (ACF periodicity and PDF skewness) are only reflected by the multi-scale SM. This SM structure also has the best Bayesian optimality.

Uni-directional connection (summary):

Slow-to-fast connection is **OFF**, Fast-to-slow connection is **ON**

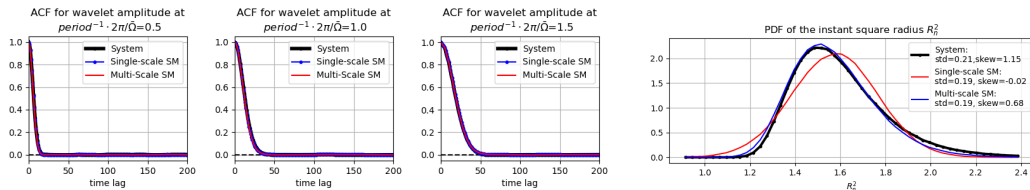


Figure: Comparison of the wavelet amplitude ACFs and square radius PDF for the optimal single-scale and multi-scale SMs, computed in the same way as for the system.

The features related to the fast-to-slow interaction (PDF skewness) are only reflected by the multi-scale SM. This SM structure also has the best Bayesian optimality.

Uni-directional connection (summary):

Slow-to-fast connection is **ON**, Fast-to-slow connection is **OFF**

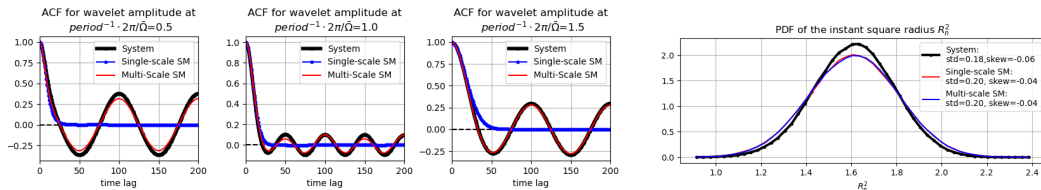


Figure: Comparison of the wavelet amplitude ACFs and square radius PDF for the optimal single-scale and multi-scale SMs, computed in the same way as for the system.

The features related to the slow-to-fast interaction (ACF periodicity) are only reflected by the multi-scale SM. This SM structure also has the best Bayesian optimality.

No connection (summary):

Slow-to-fast connection is **OFF**, Fast-to-slow connection is **OFF**

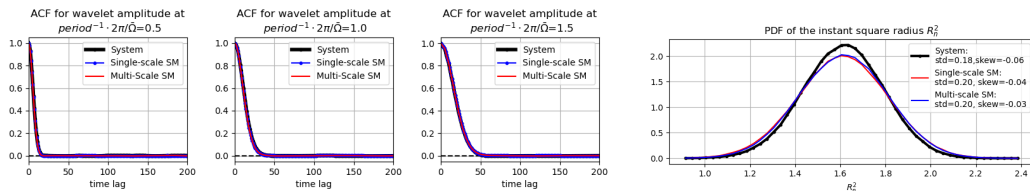


Figure: Comparison of the wavelet amplitude ACFs and square radius PDF for the optimal single-scale and multi-scale SMs, computed in the same way as for the system.

Set of single-scale SMs and multi-scale SM are equivalent.
 Set of single-scale SMs has the best Bayesian optimality.

Overview

- 1 Introduction
- 2 Multi-scale stochastic model
 - Two-scale data case
 - Generalizations
- 3 Example: slow Van der Pol and fast linear stochastic oscillators
 - Data generation
 - System with bi-directional connection
 - System with uni-directional connection: fast-to-slow
 - System with uni-directional connection: slow-to-fast
 - System with no connection
- 4 Conclusions

Conclusions

- ① A data-driven multi-scale nonlinear stochastic model is proposed. The model describes slow-to-fast and fast-to-slow interactions differently
- ② Bayesian optimization of all structural elements of the model is performed
- ③ The optimal model correctly addresses the interactions in all examples
- ④ Potential climate applications are ENSO and paleodata modeling:
 - Possible connection between monthly ocean data and daily atmospheric data (e.g. tropical SST and zonal wind)
 - Possible connection between 41/100 kyr deglaciation events and millennial processes in Pleistocene

References

- [1] A. Gavrilov, E. Loskutov, and D. Mukhin, "Bayesian optimization of empirical model with state-dependent stochastic forcing," *Chaos, Solitons and Fractals*, vol. 104, pp. 327–337, 2017.
- [2] D. Mukhin, D. Kondrashov, E. Loskutov, A. Gavrilov, A. Feigin, and M. Ghil, "Predicting Critical Transitions in ENSO models. Part II: Spatially Dependent Models," *Journal of Climate*, vol. 28, pp. 1962–1976, mar 2015.
- [3] A. Gavrilov, A. Seleznev, D. Mukhin, E. Loskutov, A. Feigin, and J. Kurths, "Linear dynamical modes as new variables for data-driven ENSO forecast," *Climate Dynamics*, vol. 52, pp. 2199–2216, feb 2019.
- [4] D. Mukhin, A. Gavrilov, E. Loskutov, J. Kurths, and A. Feigin, "Bayesian Data Analysis for Revealing Causes of the Middle Pleistocene Transition," *Scientific Reports*, vol. 9, p. 7328, may 2019.
- [5] Y. I. Molkov, E. M. Loskutov, D. N. Mukhin, and A. M. Feigin, "Random dynamical models from time series," *Physical Review E*, vol. 85, p. 036216, mar 2012.
- [6] D. Mukhin, E. Loskutov, A. Mukhina, A. Feigin, I. Zaliapin, and M. Ghil, "Predicting Critical Transitions in ENSO Models. Part I: Methodology and Simple Models with Memory," *Journal of Climate*, vol. 28, pp. 1940–1961, mar 2015.
- [7] A. Seleznev, D. Mukhin, A. Gavrilov, E. Loskutov, and A. Feigin, "Bayesian framework for simulation of dynamical systems from multidimensional data using recurrent neural network," *Chaos*, vol. 29, no. 12, 2019.
- [8] A. Seleznev, A. Gavrilov, D. Mukhin, E. Loskutov, and A. Feigin, "Simulation of Complex Systems Using the Observed Data Based on Recurrent Artificial Neural Networks," *Radiophysics and Quantum Electronics*, vol. 61, no. 12, 2019.
- [9] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence, Warwick 1980*, pp. 366–381, Springer Berlin Heidelberg.