Automatic manifold identification for mNARX models

Author(s):
Schär, Styfen; Marelli, Stefano; Sudret, Bruno

Publication date:
2023-04-04

Permanent link:
https://doi.org/https://doi.org/10.3929/ethz-b-000621799

Funding acknowledgement:
- HIghly advanced Probabilistic design and Enhanced Reliability methods for high-value, cost-efficient offshore WIND ()

Automatic manifold identification for mNARX models

S. Schär, S. Marelli and B. Sudret
ETH Zürich, Chair of Risk, Safety and Uncertainty Quantification, styfen.schaer@ibk.baug.ethz.ch

1 Problem Statement

The challenge: build a surrogate $\widetilde{\mathcal{M}}$ that emulates the response of a complex time-dependent system $\mathcal{M}$ over long time periods:

$$y(t) = \mathcal{M}(\boldsymbol{x}(\mathcal{T} \le t)) \approx \widetilde{\mathcal{M}}(\boldsymbol{x}(\mathcal{T} \le t))$$

- Discretized time axis $\mathcal{T} = \{0, \delta t, 2\delta t, \ldots, (N-1)\delta t\}$
- System response $y : \mathcal{T} \to \mathbb{R}$
- High-dimensional exogenous excitation $\boldsymbol{x} : \mathcal{T} \to \mathbb{R}^M$

Our approach: Automated incremental construction of an exogenous input manifold suitable for autoregressive surrogates

2 Autoregressive Modelling

Nonlinear AutoRegressive with eXogenous input (NARX) models account for both the temporal coherence of the output and exogenous input:

$$y(t) = \widetilde{\mathcal{M}}(\boldsymbol{\varphi}(t), \mathbf{c})$$

Where:

- $\boldsymbol{c}$ is a finite set of model parameters/coefficients
- $\boldsymbol{\varphi}(t)$ collects current and past exogenous inputs and past outputs:

$$\begin{aligned}\boldsymbol{\varphi}(t) = \{&y(t-\ell_1^y), y(t-\ell_2^y), \ldots, y(t-\ell_{n_y}^y),\\ &x_1(t-\ell_1^{x_1}), x_1(t-\ell_2^{x_1}), \ldots, x_1(t-\ell_{n_{x_1}}^{x_1}),\\ &\ldots,\\ &x_{M_x}(t-\ell_1^{x_{M_x}}), x_{M_x}(t-\ell_2^{x_{M_x}}), \ldots, x_{M_x}(t-\ell_{n_{x_{M_x}}}^{x_{M_x}})\}\end{aligned}$$

- Autoregressive lags $\ell_i^y \in \{\delta t, 2\delta t, \ldots, (N-1)\delta t\}$
- Exogenous input lags $\ell_i^{x_j} \in \{0, \delta t, 2\delta t, \ldots, (N-1)\delta t\}$

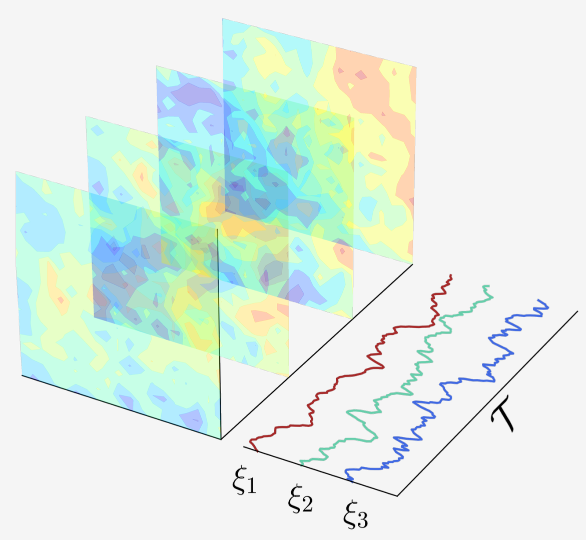
3 Exogenous input Manifold

Dimensionality reduction

Compression of high-dimensional exogenous input $\boldsymbol{x} : \mathcal{T} \to \mathbb{R}^M$ in its non-temporal coordinates:

$$\boldsymbol{\xi} = \mathcal{G}(\boldsymbol{x})$$

- $\boldsymbol{\xi} : \mathcal{T} \to \mathbb{R}^m$ such that $m \ll M$
- $\mathcal{G}$ preserves the original time scale



Manifold construction

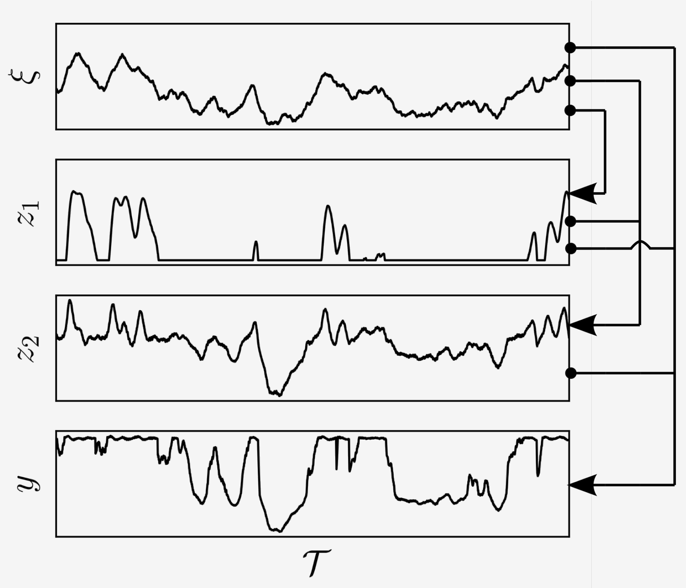
1. Collection of existing and derived time-dependent features $\{\boldsymbol{z_1} \ldots \boldsymbol{z_n}, \boldsymbol{z_i} : \mathcal{T} \to \mathbb{R}\}$, based on prior knowledge about the system
2. Feature selection based on a measure of association $\rho > \theta$, e.g. $\theta = 0.05$:

$$\rho_{\boldsymbol{z}_i} = \mathcal{Z}(\boldsymbol{z_i}, \boldsymbol{y}), \quad \rho_{\boldsymbol{z}_i} \in \mathbb{R}$$

3. Construction of NARX model onto exogenous input manifold $\boldsymbol{\zeta}$:

$$\hat{y}(t) = \widetilde{\mathcal{M}}(\boldsymbol{\zeta}(\mathcal{T} \le t), \hat{y}(\mathcal{T} < t)), \quad \boldsymbol{\zeta} = \{\boldsymbol{z_i} \mid \rho_{\boldsymbol{z}_i} > \theta, \boldsymbol{\xi_i} \mid \rho_{\boldsymbol{\xi}_i} > \theta\}$$

4. Incremental construction of auxiliary quantities during prediction phase:

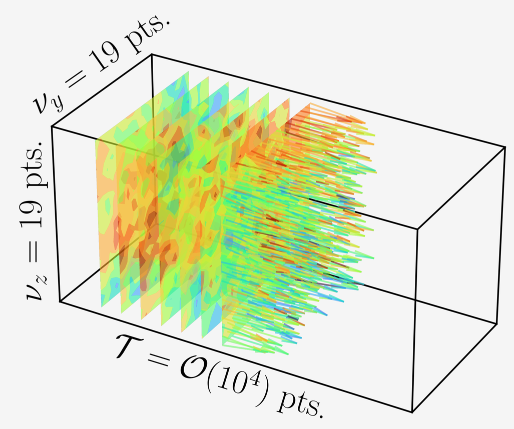


$$\begin{aligned}z_1(t) &= \mathcal{F}_1(\boldsymbol{\xi}(\mathcal{T} \le t), z_1(\mathcal{T} < t))\\ z_2(t) &= \mathcal{F}_2(z_1(\mathcal{T} \le t), \boldsymbol{\xi}(\mathcal{T} \le t), z_2(\mathcal{T} < t))\\ &\vdots\\ z_i(t) &= \mathcal{F}_i(z_1(\mathcal{T} \le t), \ldots, z_{i-1}(\mathcal{T} \le t),\\ &\qquad \boldsymbol{\xi}(\mathcal{T} \le t), z_i(\mathcal{T} < t))\end{aligned}$$

4 Case study

Complex **onshore wind turbine simulator** with control systems

- High-dimensional turbulent wind input: $\boldsymbol{v} : \mathcal{T} \to \mathbb{R}^{3 \times \nu_y \times \nu_z}$
- Quantity of interest: Power output $P : \mathcal{T} \to \mathbb{R}$



Computational model

Turbine	NREL 5-MW
Type	Onshore
Controller	ROSCO
Simulator	OpenFAST

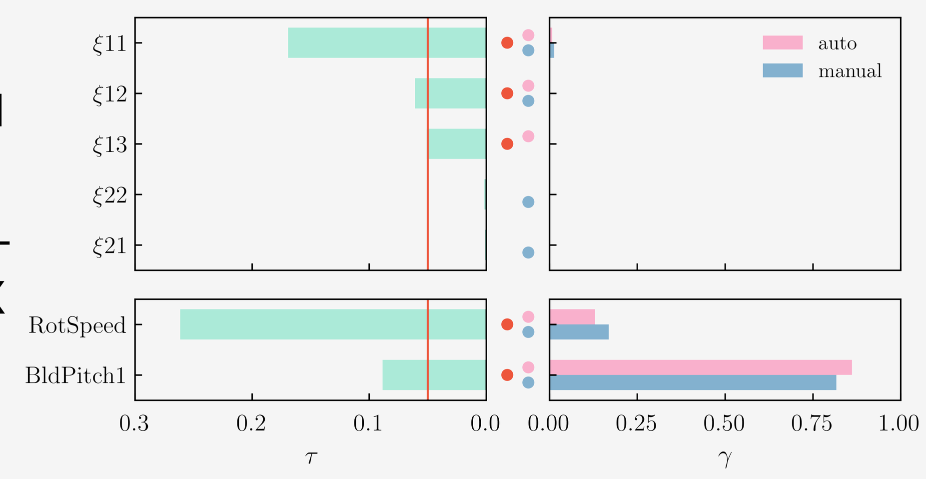
- Fast-to-construct and evaluate **polynomial NARX model**

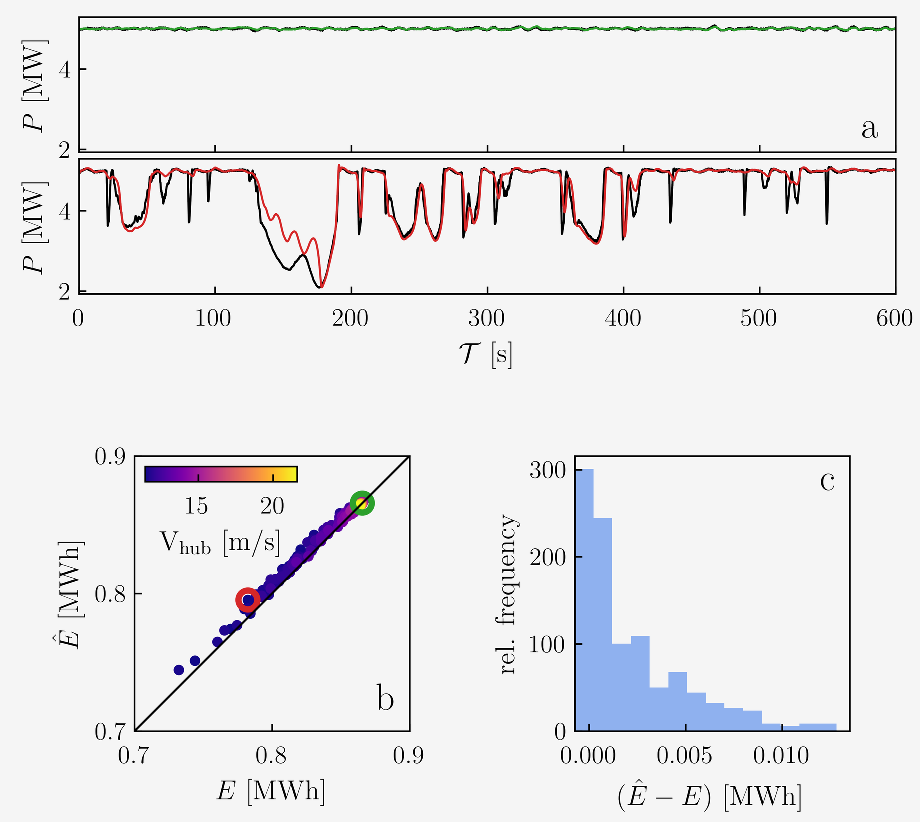
$$y(t) = \sum_{\alpha \in \mathcal{A}} c_{\boldsymbol{\alpha}} \mathcal{P}_{\boldsymbol{\alpha}}(\boldsymbol{\varphi}(t)), \qquad \mathcal{P}_{\boldsymbol{\alpha}}(\boldsymbol{\varphi}(t)) = \prod_{i=1}^{M_\varphi} \varphi_i(t)^{\alpha_i},$$

  where the output is represented as sum of monomials $\mathcal{P}$ weighted by real-valued coefficients $\boldsymbol{c_\alpha}$
- **Compression** of longitudinal wind speeds into spectral coefficients $\boldsymbol{\xi}$
- Identification of **important features** with Kendall's τ measure of association

5 Results

- Left: Kendall's τ of manually and automatically selected features
- Right: Corresponding relative coefficient magnitude of the NARX surrogate





a. Simulated power output (black) vs. the emulated one (colored)
a. Most accurate and least accurate prediction
b. Produced energy computed from simulator output (E) and emulated response ($\hat{E}$)
c. Discrepancy in the produced energy ($\hat{E} - E$)

6 Discussion and Outlook

Discussion

- Multistep approach allows accurate emulation of complex dynamical systems
- Relevant features can be automatically selected using a measure of association

Outlook

- Not only select important features but also determine ideal construction order
- Application to a broader range of problems

References

[1] Dimitrov, N., S. Marelli, and S. Schär (2022). Novel surrogate modelling approaches for wind turbine reliability assessment. H2020 Project HIPERWIND. Deliverable D4.1.

D BAUG

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 101006689