



EURODINAME III - An International Symposium on Dynamic
Problems of Mechanics
June 8-11, 2026 - Giens Peninsula, FR

EURODINAME-2026-14495

Learning Model Discrepancy using Sparse Bayesian Polynomial NARX: Case study for Bouc-Wen Hysteresis model

Eyoel Atnafe

Université Marie et Louis Pasteur, CNRS, Institut FEMTO-ST
eyoel.atnafe@femto-st.fr

Rafael Teloli

Université Marie et Louis Pasteur, SUPMICROTECH, CNRS, Institut FEMTO-ST
rafael.teloli@femto-st.fr

Emeline Sadoulet-Reboul

Université Marie et Louis Pasteur, CNRS, Institut FEMTO-ST
emeline.sadoulet-reboul@unlp.fr

Scott Cogan

Université Marie et Louis Pasteur, CNRS, Institut FEMTO-ST
scott.cogan@univ-fcomte.fr

Abstract. Comparing experimental data with numerical simulation results may reveal discrepancies that cannot be directly explained by the underlying physical model. This issue is particularly pronounced in nonlinear regimes, where the system behavior is not fully understood or where the sources of nonlinearity remain unclear. The aim of this study is to model these discrepancy errors through a data-driven framework. The proposed approach combines the NARX (Nonlinear AutoRegressive with eXogenous inputs) model with the SINDy (Sparse Identification of Nonlinear Dynamical Systems) methodology. In this framework, NARX is employed to represent the discrepancy dynamics, while the nonlinear basis functions are identified using a Bayesian formulation of SINDy, which enables the quantification of uncertainty. The methodology is validated through a numerical application involving one of the nonlinear benchmarks, the Bouc-Wen hysteresis model. The approach learns an interpretable model that corresponds to the hidden hysteresis force from a noisy data and demonstrates the ability to derive an interpretable time-domain model that remains robust in the presence of measurement noise.

Keywords: Model discrepancy, SINDy, Bayesian regression, NARX

1. INTRODUCTION

Physics-based models of dynamical systems are crucial for the design, analysis, and optimization of complex engineering systems. These models provide interpretable, first-principles-based predictions of system behavior. However, even advanced models can diverge from experimental observations due to unaccounted phenomena, such as complex interactions and unmodeled nonlinearities caused by insufficient observation and physical knowledge. Mechanical joint designs, for example, may introduce frictional hysteresis, contact nonlinearities, or wear effects, which, if not explicitly modeled, can lead to discrepancies between the model and experimental observations. In such cases, parameter calibration alone cannot eliminate the disparity since the model is structurally inadequate. This disparity is known as model discrepancy or model error (Brynjarsdóttir and O'Hagan (2014)). To account for the missing information reflecting epistemic gaps, aspects of the system that physics-based models are unable to represent, it is crucial to model the discrepancy from data. This need has led to increased interest in hybrid modeling frameworks that mix physics-based models with data-driven corrections to increase fidelity.

In recent years, Machine-learning discrepancy models are increasingly used to complement physics-based simulators in dynamical systems. One of such examples is Neural Network (NN). Recurrent NN can capture memory-dependent



model-form errors within physics-based frameworks (Levine and Stuart (2022); Brajard *et al.* (2020)), while equation-discovery methods such as Sparse Identification of Nonlinear Dynamics (SINDy) (Brunton *et al.* (2016)) offer sparse, interpretable alternatives (Kaheman *et al.* (2019); Ebers *et al.* (2024)). Hybrid strategies can further combine NNs with SINDy structure (Lai *et al.* (2021); Najera-Flores *et al.* (2024)). However, these deterministic parametric approaches are noise-sensitive and provide no uncertainty quantification (UQ).

To address the challenges of UQ and model robustness, researchers have explored probabilistic frameworks such as Gaussian Processes (GPs) to model discrepancy. Notably, Kennedy and O’Hagan (2001) introduced a statistical discrepancy modeling approach, giving rise to the well-known KoH framework. Others have developed upon these approach with different kinds extensions such as using multi-response GPs (Arendt *et al.* (2012)), physics-informed priors (Spitieris and Steinsland (2022); Plumlee (2017)), and staged calibration (Maupin and Swiler (2020)). For interpretable UQ, Bayesian sparse regression has also been explored: Model Error Discovery with Interpretability and Data Assimilation (MEDIDA) recovers closed-form partial differential equation (PDE) discrepancy models via sparse Bayesian learning (Mojgani *et al.* (2022)), and Bayesian SINDy quantifies uncertainty over structures and coefficients (Hirsh *et al.* (2022)). Despite these advances, uncertainty-aware, interpretable discrepancy modeling for discrete-time, memory-dependent experimental systems remains an open gap.

This paper introduces Bayesian Sparse Identification of Nonlinear AutoRegressive with eXogenous input (B-SINARX) discrepancy modeling framework, a Bayesian SINDy framework built on discrete-time Nonlinear AutoRegressive with eXogenous input (NARX) models, to fill this gap. NARX naturally captures memory through autoregressive terms with exogenous inputs and admits linear-in-parameters regression over polynomial libraries, which align with physical intuition and have proven effective for identifying nonlinear hysteresis and oscillators (Billings (2013); Worden and Barthorpe (2012); Martins and Aguirre (2014)). While Bayesian NARX variants and sparse structure-selection methods exist (Cheng *et al.* (2011); Bhattacharyya (2020); Baldacchino *et al.* (2013)), none have targeted discrepancy modeling explicitly. B-SINARX addresses this by delivering parsimonious, interpretable discrete-time corrections with quantified uncertainty.

The framework hybridizes physics-based and data-driven models, enforcing sparsity for parsimony and generalizability, and propagates parameter uncertainty to yield probabilistic free-run predictions. It is demonstrated on a Bouc-Wen hysteresis numerical case study, where unmodeled hysteretic damping constitutes the discrepancy. The paper is organized as follows: Section 2 covers mathematical background; Section 3 presents the B-SINARX framework and training algorithms; Section 4 details the case study and results; Section 5 concludes with discussion and future directions.

2. Background on Discrepancy Modeling

In this section, the framework for discrepancy modeling is introduced. It describes how discrepancy terms are introduced into the modeling scheme and demonstrates its implementation. This background paves the way for introducing the B-SINARX algorithm, later on Section 3.

2.1 Discrepancy modeling framework

Kennedy and O’Hagan (2001) proposed the discrepancy modeling framework to accommodate model-form errors. For a system modeled by an inadequate, low-fidelity model $f(x)$, the KoH discrepancy framework can be represented as:

$$y(x) = f(x) + \delta(x) + \varepsilon, \quad (1)$$

where ε is modeled as *i.i.d.* Gaussian noise with zero mean and variance σ^2 , i.e., $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. The discrepancy term, δ , represents the unmodeled structure of the system and can be approximated by a data-driven approach. Levine and Stuart (2022) present this type of discrepancy modeling as a *boosting* or additive hybrid model. Alternatively, the discrepancy can be modeled by *stacking* the models or combining the models one inside the other, yielding

$$y(x) = \delta(x, f(x)) + \varepsilon. \quad (2)$$

When model interpretability is significantly important and trust in the known model $f(\bullet)$ is high, additive models are preferred because the method hypothesizes some information about the nature of the discrepancy being additive. In contrast, when there is no prior assumption about the nature of the discrepancy, and a nonlinear transformation or a more complex modification of $f(\bullet)$ may be required, embedding $f(\bullet)$ in a discrepancy function is advantageous. This approach, however, could lead to cases where $\delta(\bullet)$ does not use $f(\bullet)$ at all, hence removing prior knowledge entirely. Therefore, in this paper, the KoH additive form is considered the best option.

A concise discrete-time hybrid discrepancy model can be modified to:

$$y_k = f(\mathbf{x}_k) + \delta(\mathbf{x}_k) + \varepsilon_k, \quad \varepsilon_k \sim \mathcal{N}(0, \sigma_y^2). \quad (3)$$

To generalize the formulation for partially observed systems, where the observable states are insufficient to determine the system's dynamics, Levine and Stuart (2022) suggest modeling the hidden state dynamics, denoted as z , in conjunction with the observable states. For this purpose, the state vector is augmented to consider the hidden state dynamics $\dot{z}$, yielding a discrete form:

$$y_k = f(\mathbf{x}_k, z_k) + \delta_1(\mathbf{x}_k, z_k) + \varepsilon_k, \quad (4)$$

$$z_k = g(\mathbf{x}_k, z_k) + \delta_2(\mathbf{x}_k, z_k) + \varepsilon_k, \quad (5)$$

where $f(\bullet)$ and $g(\bullet)$ are the known models for the observable states and hidden dynamics, respectively, whereas $\delta = [\delta_1(\bullet) \quad \delta_2(\bullet)]^\top$ is the discrepancy model.

3. Methodology: B-SINARX

After established the background of discrepancy modeling and the main motivations for the B-SINARX framework in mechanical systems, this section presents the key components of the methodology.

NARX is a nonlinear system modeling framework that uses time histories and exogenous inputs to predict future time steps through nonlinear regressors. This class of models is particularly well-suited for nonlinear time series problems, as it allows for a broad input space by incorporating time history information Billings (2013).

The discrete-time nonlinear dynamics of a system can be represented via a NARX model, which generalizes to a regression problem of the form:

$$y_{k+1} = f(y_k, y_{k-1}, \dots, y_{k-n_y}, u_k, u_{k-1}, \dots, u_{k-n_u}) + \varepsilon_k, \quad k = 1, 2, \dots, \text{ samples},$$

n_u and n_y denote the number of regressors in the input and output signals, respectively, where y_k is the vector of observed state variables and ε_k is *i.i.d.* noise at time step k . The function $f(\bullet)$ is a regressor that maps the input space, i.e., the time history of the state variables, to the output.

To efficiently identify the sparse structure of the NARX model within this framework, a Bayesian SINDy approach is employed, as described in the following subsection.

3.1 Sparse Bayesian SINDy with Trimming and L-Curve Model Selection

The standard SINDy formulation casts system identification as a sparse regression problem. To enable uncertainty quantification (UQ) without expensive sampling, we adopt a Bayesian formulation. The measurement model is:

$$\mathcal{D} = \Theta^\top \xi + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \mathbf{I}\sigma^2), \quad (6)$$

where $\Theta \in \mathbb{R}^{N \times M}$ is the library matrix and ξ the sparse weight vector. Following Tipping and Faul (2003), sparsity is enforced via independent zero-mean Gaussian priors $\pi(\xi|\alpha) = \prod_m \mathcal{N}(\xi_m; 0, \alpha_m^{-1})$, where large α_m drives $\xi_m \rightarrow 0$. The posterior is Gaussian $\mathcal{N}(\mu, \Sigma)$ with:

$$\Sigma = (\mathbf{A} + \sigma^{-2} \Theta^\top \Theta)^{-1}, \quad \mu = \sigma^{-2} \Sigma \Theta \mathcal{D}. \quad (7)$$

The hyperparameters are obtained by maximizing the log marginal likelihood $\mathcal{L}(\alpha)$. The Tipping and Faul Tipping and Faul (2003) sparse Bayesian algorithm outputs a most probable weight matrix that optimizes the loss function, but does not always yield the sparsest solution. To address this, a trimming operation is applied after Bayesian optimization to enforce a predetermined sparsity level.

The algorithm exploits the “well-determinedness” factor γ , defined as:

$$\gamma_i = 1 - \alpha_i \Sigma_{ii}, \quad i = 1, 2, \dots, p, \quad (8)$$

which quantifies the contribution of each basis function. Weights with γ_i below a given threshold are removed, projecting the solution onto a k -sparse space. This enables cross-validation directly over sparsity levels, rather than searching over penalization parameters.

Sparsity level selection via L-Curve. The target sparsity k is chosen using an L-curve on a validation set, plotting the residual norm:

$$\mathcal{L}_{\text{val}} = \|D_{\text{val}} - \Theta^\top \xi(\mathcal{I})\| \quad \text{vs} \quad \|\xi\|_0. \quad (9)$$

The optimal corner is identified by the triangular method (Castellanos *et al.* (2002)), the point P^* maximizing the triangle area formed with the curve endpoints:

$$P^* = \arg \max_{P_i \in \{P_2, \dots, P_{n-1}\}} \mathcal{A}(P_1, P_n, P_i), \quad (10)$$

If the relative range

$$\mathcal{L}_{\text{rel}} = \frac{\mathcal{L}_{\text{max}} - \mathcal{L}_{\text{min}}}{\mathcal{L}_{\text{max}}} \times 100\% \leq \kappa \quad (11)$$

with $\kappa = 2\%$, the curve lacks an “L-shape” and the most sparse solution is selected directly.

3.2 Finding NARX Regressors via Bayesian SINDy (B-SINARX)

The system state $\mathbf{s}(t) = [x, v, z]^\top$ comprises observable states $\mathbf{x} = [x, v]^\top$ and a hidden state z , with discrete-time regressor vector $\phi_{k:d}$ stacking d time-delayed observations and inputs (see Table 1).

Table 1. Variable Definition

Variable	Contains	Discrete	Description
$\mathbf{x}(t)$	$[x, v]^\top$	$\mathbf{x}_k$	Observable states*
$z(t)$	z	z_k	Hidden state
$\mathbf{s}(t)$	$[\mathbf{x}, z]^\top$	$\mathbf{s}_k$	Full system state
$\phi(t)$	$[\mathbf{s}(t), u(t)]^\top, \quad t \in [t_d, t]$	ϕ_k	System input regressor vector containing stacked time delays

* For simplicity, we denote velocity by $v = \dot{x}$; henceforth, this notation will be used in both equations and figures.

From the hybrid formulation, the discrepancy model corrects the known physics in discrete time within the NARX framework:

$$\mathbf{s}_{k+1} = \mathbf{f}(\phi_{k:d}) + \mathbf{\Theta}(\phi_{k:d}) \boldsymbol{\xi} + \boldsymbol{\varepsilon}_k, \quad (12)$$

where $\mathbf{f}(\cdot)$ is the known physics model, $\mathbf{\Theta}(\bullet)\boldsymbol{\xi}$ is the sparse discrepancy term to be identified, and $\boldsymbol{\varepsilon}_k \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_3)$. In matrix form over N samples:

$$\mathcal{S} = \mathcal{F} + \mathbf{\Theta}_\Phi \boldsymbol{\xi} + \boldsymbol{\varepsilon}, \quad (13)$$

where $\mathcal{S}, \mathcal{F} \in \mathbb{R}^{(N-d+1) \times 3}$ collect the state outputs and known-model predictions, and $\mathbf{\Theta}_\Phi \in \mathbb{R}^{(N-d+1) \times M}$ is the library matrix. The sparse coefficients $\boldsymbol{\xi}$ are estimated by regressing on the residual $\mathcal{R} := \mathcal{S} - \mathcal{F}$, yielding the likelihood:

$$\pi(\mathcal{R} | \boldsymbol{\xi}, \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{\|\mathcal{R} - \mathbf{\Theta}_\Phi \boldsymbol{\xi}\|^2}{2\sigma^2}\right). \quad (14)$$

This formulation shifts the sparse regression toward identifying the best residual representation. In the absence of prior knowledge, setting $\mathbf{f}(\bullet) = \mathbf{0}$ recovers standard Bayesian SINDy; alternatively, a known (possibly inaccurate) linear model can be supplied, allowing B-SINARX to correct its parameters and augment it with nonlinear terms. The training framework is illustrated in Fig. 1.

4. Numerical Validation: Bouc-Wen Identification via B-SINARX

4.1 Bouc-Wen Model

The single-degree-of-freedom Bouc-Wen oscillator is:

$$m\ddot{x} + c\dot{x} + kx + z = u, \quad (15)$$

$$\dot{z} = \alpha\dot{x} - \beta(\gamma|\dot{x}| |z|^{\nu-1} z - \delta\dot{x}|z|^\nu), \quad (16)$$

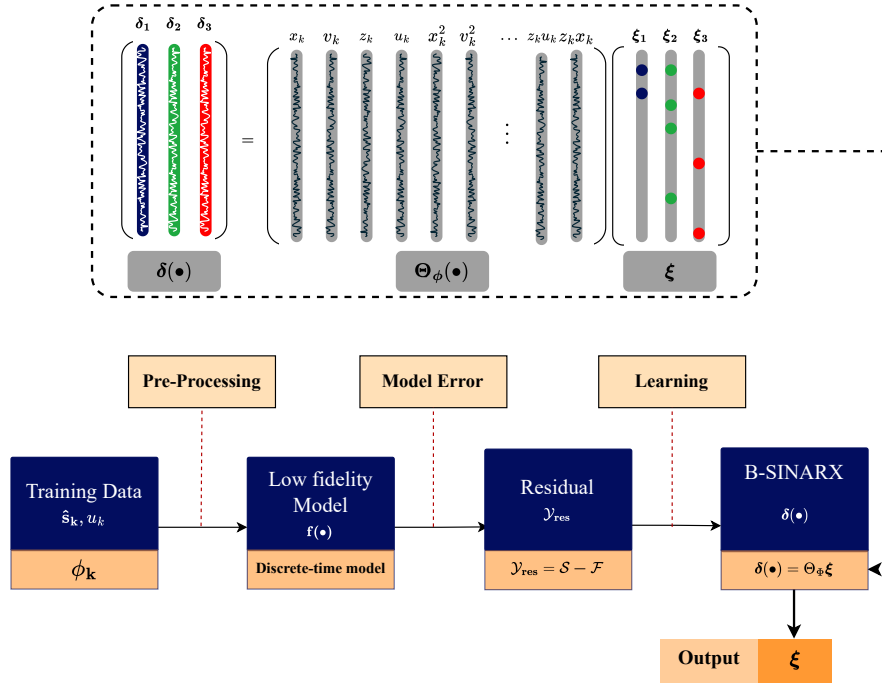


Figure 1. B-SINARX training framework: the sparse model is learned from the residual of the low-fidelity physics model.

where z is the hysteretic force and parameters are taken from the benchmark Schoukens and Noël (2017) (Table 2). This system is a standard benchmark due to its non-smooth, history-dependent nonlinearity Noël and Schoukens (2016); de O. Teloli and da Silva (2019). Equations (15)–(16) are integrated via a fourth-order Runge-Kutta scheme. The excitation is a random-phase multi-sine (excitation bandwidth 5–150 Hz, sampling frequency $f_s = 7500$ Hz, amplitude 50 N) with 40 dB SNR. Two 1-second simulations (15000 samples total) are used: 60% for training and 40% for validation. Testing uses a sweep-sine signal (5–150 Hz, 1 s).

Table 2. Bouc-Wen model parameters, Schoukens and Noël (2017).

m	c	k	α	β	γ	δ	ν
[kg]	[Ns/m]	[N/m]	[N/m]		[m ⁻¹]	[m ⁻¹]	
2	10	5×10^4	5×10^4	10^3	0.8	-1.1	1

4.2 Discrepancy Formulation

The state-space form of Eqs. (15)–(16) separates linear and nonlinear contributions:

$$\dot{\mathbf{s}} = \underbrace{\begin{bmatrix} \dot{x} \\ m^{-1}u - 2\zeta\omega_n\dot{x} - \omega_n^2x \\ 0 \end{bmatrix}}_{\text{Known Linear Physics}} + \underbrace{\begin{bmatrix} 0 \\ -m^{-1}z \\ \alpha\dot{x} - \beta(\gamma|\dot{x}|z - \delta\dot{x}|z|) \end{bmatrix}}_{\text{Nonlinear Contribution}}, \quad (17)$$

where $\mathbf{s} = [x, \dot{x}, z]^\top$. Assuming the hysteretic contribution is unknown, the linear part is discretized via forward finite differences, giving an ARX baseline:

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \Delta t \left(\begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\zeta\omega_n \end{bmatrix} \mathbf{y}_k + \begin{bmatrix} 0 \\ m^{-1} \end{bmatrix} u_k \right) + \varepsilon_k. \quad (18)$$

The hidden state z is augmented to capture the residual force, yielding the B-SINARX discrepancy model:

$$\mathbf{y}_{k+1} = \mathbf{f}_{ARX}(\mathbf{x}_k, z_k) + \delta_1(\mathbf{x}_k, z_k) + \varepsilon_k, \quad (19)$$

$$z_{k+1} = \delta_2(\mathbf{x}_k, z_k) + \varepsilon_k, \quad (20)$$

where δ_1 corrects the baseline parameters and δ_2 captures the nonlinear hysteretic dynamics. The library includes polynomial and absolute-value terms to handle the non-smooth nature of hysteresis.

Model accuracy is assessed via the EEARTH metric Zhan *et al.* (2011), a weighted combination of phase (E_p), magnitude (E_m), and slope (E_s) errors:

$$E = 0.2 E_p + 0.4 E_m + 0.4 E_s, \quad E \in [0, 1], \quad (21)$$

where higher values indicate better agreement.

4.3 Results for Numerical simulation

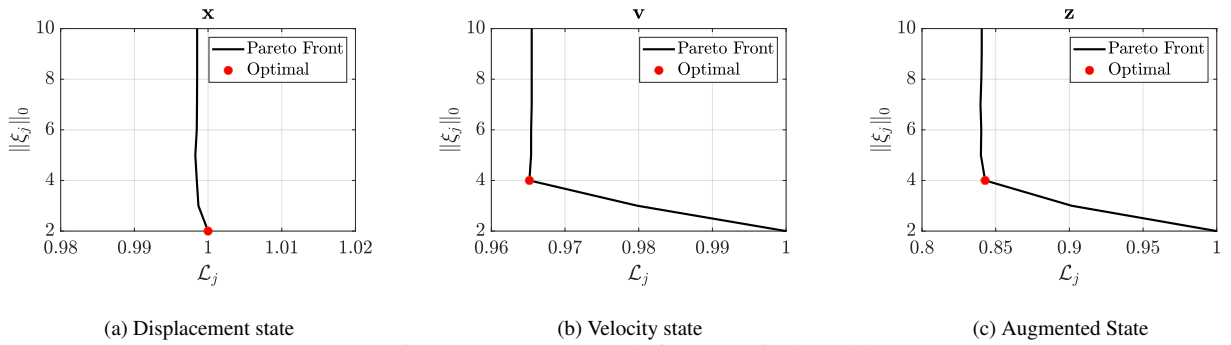


Figure 2. L-Curve result for numerical model

Figure 2 shows the L-Curve used to identify sparse models. Figure 2a shows that changing the number of functions does not affect the prediction accuracy since the relative loss is 2%. This indicates that the first function, x_{k+1} , as it can intuitively be deduced from inherent mathematical relationship between displacement and velocity, does not require an update. However, as can be seen on table 3, there is a third term, z_k , that is added in the identification. This results from the fact that the minimum number of functions considered in the L-curve is set to two. Since the method operates within a Bayesian framework, it provides access to posterior distributions of the identified parameters and their correlations. The augmented state term z_k is found to be linearly dependent on x_k , indicating that the correction term is unnecessary and can be replaced by x_k . The influence of these spurious terms is negligible, with a coefficient smaller than 10^{-7} , as indicated in Table 3. The four terms identified for the augmented state correspond to those of the true Bouc–Wen model. For v_{k+1} , the stiffness term ξ_{x_k} is estimated with a mean error of 7.6%, while the hysteresis force term ξ_{z_k} presents a mean error of 10.4%, with the two terms being negatively correlated. The residual term z_{k+1} is identified with a relative error of $< 1\%$. Overall, the B-SINARX framework produces a sparse and accurate representation of the discrepancy, demonstrating its capability to both correct existing parameters and identify new model structures directly from data.

The comparison between the identified model and the true dynamics is presented in Table 3. It indicates that the identified discrepancy model closely matches the true dynamics, with most parameters estimated with an error below 2%. The table also specifies the equations from which the correct model terms are obtained. The reported model parameters correspond to the mean values of Gaussian distributions.

Figure 3a and 3b, present the dynamic response prediction compared to the true response for the displacement and velocity, respectively, under the sweep-sine excitation. With 99% confidence interval, x and v are predicted accurately, correcting the linear model in discrete time. Figure 6 illustrates the hysteresis loop for one cycle in the nonlinear regime, showing that the identified model captures the hysteresis behavior within the confidence bounds. Figure 5a, 5b, and 5c show a longer time horizon of prediction for the full state vector. It shows that the discovered NARX model is stable and does not accumulate significant error for longer duration of prediction.

In the Table 4, the EEARTH performance for the predicted average of both displacement and velocity responses under different input signals is presented. The identified B-SINARX discrepancy model achieves a performance higher than 97% when considering sweep sine excitation (figure 3) and Multi-sine (figure 4 for 1 sec multi-sine excitation and extended prediction horizon figure 5 a 3 sec multi-sine excitation). Although prediction errors in the later case are more pronounced, they remain below 3%, confirming that the identified model maintains an acceptable level of generalization. It has been

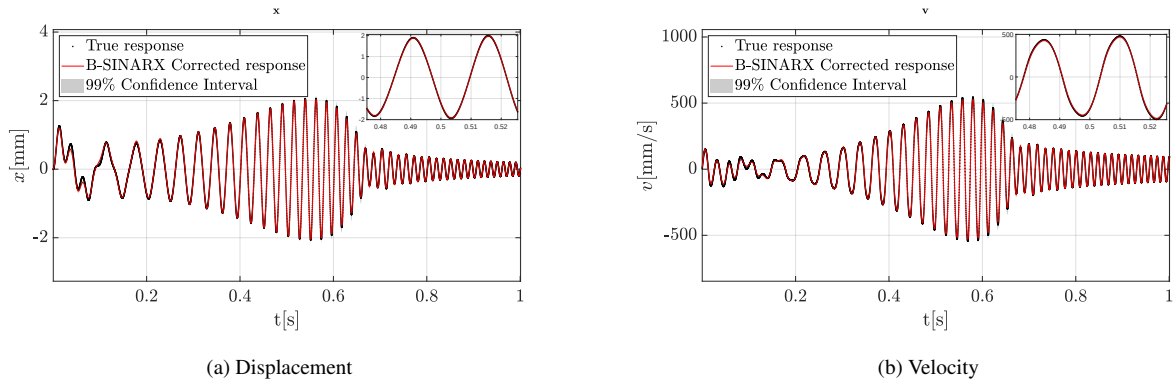


Figure 3. Comparison between the actual and predicted time response from B-SINARX for a chirp excitation

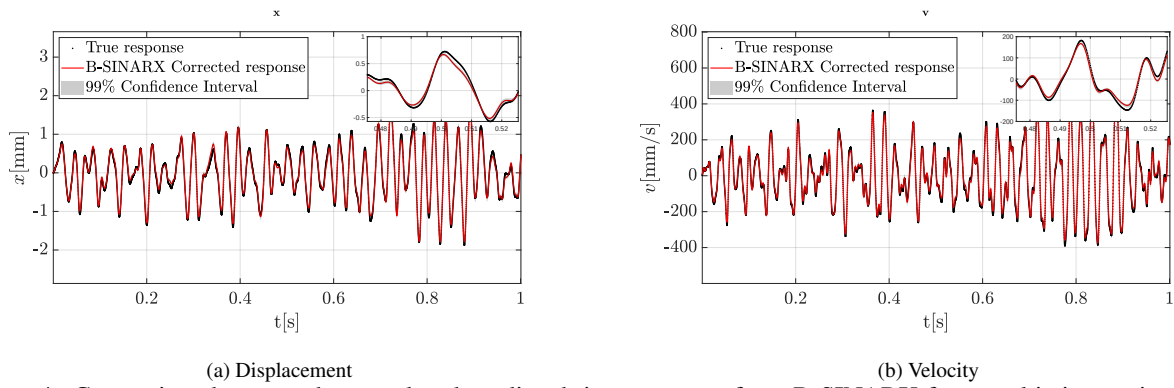


Figure 4. Comparison between the actual and predicted time response from B-SINARX for a multi-sine excitation for longer prediction horizon - 3 times longer the training datas

Table 3. Identified discrepancy model.

State	Correct Model	B-SINARX	Error
x_{k+1}	$x_k + 1.3 \times 10^{-4} v_k$	$0.997x_k + 1.3 \times 10^{-4}v_k + 2.16 \times 10^{-8}z_k$	0.3%
v_{k+1}	$0.999v_k - 3.3x_k + 6.67 \times 10^{-5}u_k - 6.67 \times 10^{-5}z_k$	$0.997v_k - 3.68x_k + 6.71 \times 10^{-5}u_k - 6.16 \times 10^{-5}z_k$	6.6%
z_{k+1}	$z_k + 6.67v_k - 0.107 v_k z_k + 0.147v_k z_k $	$0.998z_k + 6.63v_k - 0.105 v_k z_k + 0.148v_k z_k $	1.05%

— Incorrectly identified — Correctly identified (< 2% error) — Identified (> 2% error)

demonstrated that B-SINARX can extract an interpretable nonlinear discrete-time model with quantified uncertainty from noisy data. Informed by prior knowledge of the nonlinearity, absolute-valued functions were included in the library to represent the discontinuity observed in the hysteresis plot.

Table 4. EEARTH Performance of identified NARX model on different test data

Dataset	E_p	E_m	E_s	E_{Total}
Chirp	100	96.3	96.4	97.1
Multi-sine	100	97.9	98.3	98.5

5. Discussion and Conclusion

In this paper, a discrepancy modeling framework for discrete-time dynamical systems was presented, using a Bayesian formulation of the SINDy-NARX approach. The method aims to identify model form errors, *i.e.*, structural discrepancies

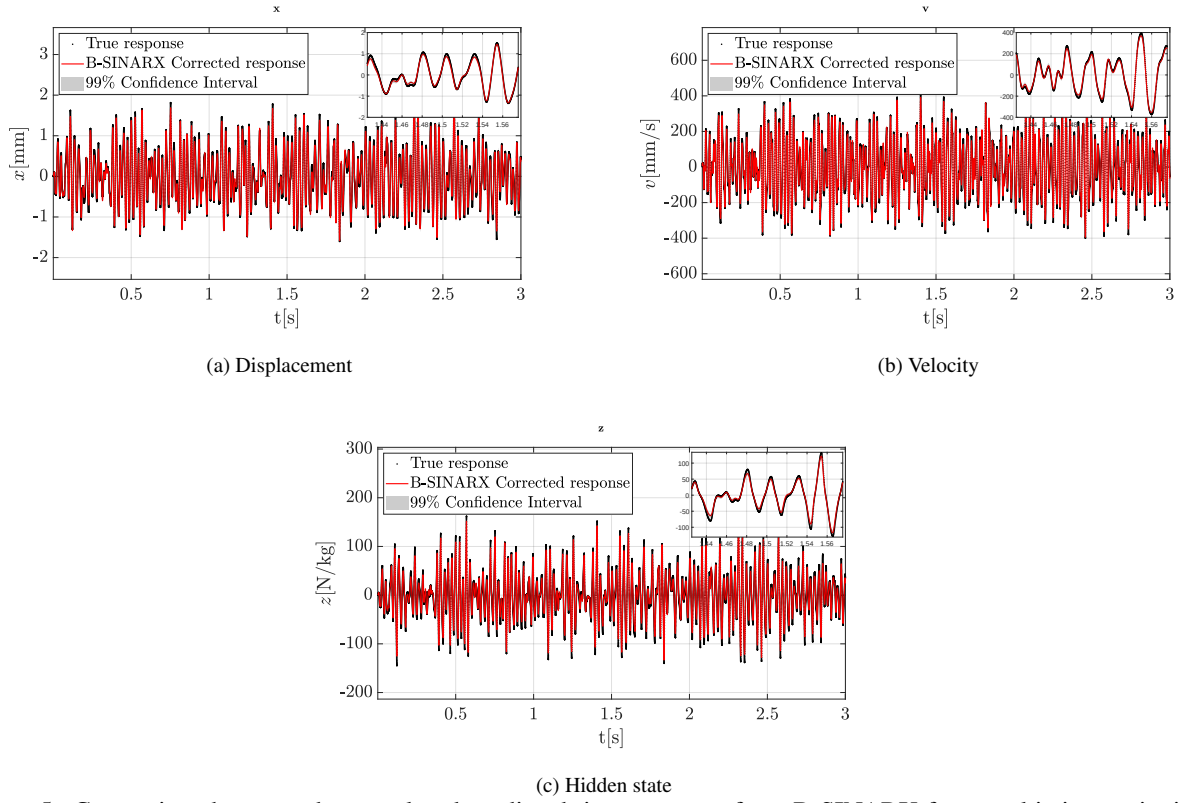


Figure 5. Comparison between the actual and predicted time response from B-SINARX for a multi-sine excitation for longer prediction horizon - 3 times longer the training data

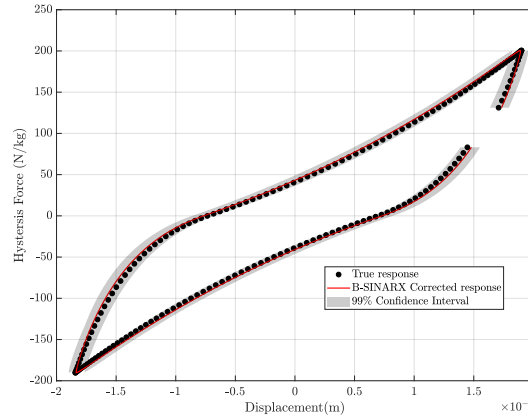


Figure 6. Hysteresis prediction for ≈ 1 period cycle.

learning sparse, interpretable correction terms directly from time-series observations. Through both numerical validation, the proposed framework has been shown to recover robust discrepancy models that effectively capture unmodeled nonlinear dynamics.

The discrepancy model successfully recovered the nonlinear Bouc-Wen hysteresis behavior from noisy observations. The discovered NARX model is able to correct linear model parameters and identify higher-order nonlinear terms that closely match the true dynamics. The hybrid model achieved a good predictive performance in closed-loop time response, and the results confirm that the B-SINARX approach can provide robust estimates even in low signal-to-noise conditions. The strength of the framework lies in its balance between interpretability and flexibility. Unlike black-box neural networks or Gaussian processes, the identified discrepancy functions are transparent and grounded in polynomial terms, making the results useful to physical interpretation, control design, and reduced-order modeling. Additionally, the Bayesian formulation provides an inherent mechanism for uncertainty quantification that offers confidence bounds on predictions

and facilitates robust decision-making.

However, there are limitations worth noting. First, the choice of library functions remains critical: while polynomials and absolute value functions worked well in the presented examples, other systems may require richer basis sets (e.g., trigonometric or exponential terms). Second, although the trimming algorithm used to enforce sparsity provides flexibility, it may lead to model selection inconsistencies in cases where multiple competing terms have similar contributions.

Further research can be done in including physics based equation, such as energy formulations, stability criterion that could assist the selection of better interpretable models.

6. ACKNOWLEDGEMENTS

The authors acknowledge the support of EUR EIPHI Graduate School. Part of the experimental investigations were conducted on the AMETISTE platform. The authors also express their gratitude to the Région Bourgogne–Franche-Comté for its financial support.

7. REFERENCES

- Arendt, P.D., Apley, D.W., Chen, W., Lamb, D. and Gorsich, D., 2012. “Improving identifiability in model calibration using multiple responses”. *Journal of Mechanical Design*.
- Baldacchino, T., Anderson, S.R. and Kadiramanathan, V., 2013. “Computational system identification for bayesian narx modelling”. *Automatica*, Vol. 49, No. 9, pp. 2641–2651. ISSN 0005-1098. doi:<https://doi.org/10.1016/j.automatica.2013.05.023>.
- Bhattacharyya, B., 2020. “Global sensitivity analysis: A bayesian learning based polynomial chaos approach”. *Journal of Computational Physics*, Vol. 415, p. 109539. ISSN 0021-9991. doi:<https://doi.org/10.1016/j.jcp.2020.109539>.
- Billings, S.A., 2013. *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons.
- Brajard, J., Carrassi, A., Bocquet, M. and Bertino, L., 2020. “Combining data assimilation and machine learning to emulate a dynamical model from sparse and noisy observations: A case study with the lorenz 96 model”. *Journal of computational science*, Vol. 44, p. 101171.
- Brunton, S.L., Proctor, J.L. and Kutz, J.N., 2016. “Discovering governing equations from data by sparse identification of nonlinear dynamical systems”. *Proceedings of the national academy of sciences*, Vol. 113, No. 15, pp. 3932–3937.
- Brynjarsdóttir, J. and O’Hagan, A., 2014. “Learning about physical parameters: The importance of model discrepancy”. *Inverse problems*, Vol. 30, No. 11, p. 114007.
- Castellanos, J., Gómez, S. and Guerra, V., 2002. “The triangle method for finding the corner of the l-curve”. *Applied Numerical Mathematics*, Vol. 43, No. 4, pp. 359–373. ISSN 0168-9274. doi:[https://doi.org/10.1016/S0168-9274\(01\)00179-9](https://doi.org/10.1016/S0168-9274(01)00179-9).
- Cheng, Y., Wang, L., Yu, M. and Hu, J., 2011. “An efficient identification scheme for a nonlinear polynomial NARX model”. *Artificial Life and Robotics*, Vol. 16, No. 1, pp. 70–73. ISSN 1433-5298, 1614-7456.
- de O. Teloli, R. and da Silva, S., 2019. “A new way for harmonic probing of hysteretic systems through nonlinear smooth operators”. *Mechanical Systems and Signal Processing*, Vol. 121, pp. 856–875. ISSN 0888-3270. doi:<https://doi.org/10.1016/j.ymssp.2018.11.044>.
- Ebers, M.R., Steele, K.M. and Kutz, J.N., 2024. “Discrepancy modeling framework: Learning missing physics, modeling systematic residuals, and disambiguating between deterministic and random effects”. *SIAM Journal on Applied Dynamical Systems*, Vol. 23, No. 1, pp. 440–469.
- Hirsh, S.M., Barajas-Solano, D.A. and Kutz, J.N., 2022. “Sparsifying priors for bayesian uncertainty quantification in model discovery”. *Royal Society open science*, Vol. 9, No. 2, p. 211823.
- Kaheman, K., Kaiser, E., Strom, B., Kutz, J.N. and Brunton, S.L., 2019. “Learning discrepancy models from experimental data”. *arXiv preprint arXiv:1909.08574*.
- Kennedy, M.C. and O’Hagan, A., 2001. “Bayesian calibration of computer models”. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 63, No. 3, pp. 425–464.
- Lai, Z., Mylonas, C., Nagarajaiah, S. and Chatzi, E., 2021. “Structural identification with physics-informed neural ordinary differential equations”. *Journal of Sound and Vibration*, Vol. 508, p. 116196.
- Levine, M. and Stuart, A., 2022. “A framework for machine learning of model error in dynamical systems”. *Communications of the American Mathematical Society*, Vol. 2, No. 07, pp. 283–344.
- Martins, S.A.M. and Aguirre, L.A., 2014. “Narx modelling of the bouc-wen model”.
- Maupin, K.A. and Swiler, L.P., 2020. “Model discrepancy calibration across experimental settings”. *Reliability Engineering & System Safety*, Vol. 200, p. 106818.
- Mojgani, R., Chattopadhyay, A. and Hassanzadeh, P., 2022. “Discovery of interpretable structural model errors by combining bayesian sparse regression and data assimilation: A chaotic kuramoto–sivashinsky test case”. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, Vol. 32, No. 6.

- Najera-Flores, D.A., Quinn, D.D., Garland, A., Vlachas, K., Chatzi, E. and Todd, M.D., 2024. “A structure-preserving machine learning framework for accurate prediction of structural dynamics for systems with isolated nonlinearities”. *Mechanical Systems and Signal Processing*, Vol. 213, p. 111340.
- Noel, J.P. and Schoukens, M., 2016. “Hysteretic benchmark with a dynamic nonlinearity”. In *Workshop on nonlinear system identification benchmarks*. pp. 7–14.
- Plumlee, M., 2017. “Bayesian calibration of inexact computer models”. *Journal of the American Statistical Association*, Vol. 112, No. 519, pp. 1274–1285.
- Schoukens, M. and Noël, J.P., 2017. “Three benchmarks addressing open challenges in nonlinear system identification”. *IFAC-PapersOnLine*, Vol. 50, No. 1, pp. 446–451.
- Spitieris, M. and Steinsland, I., 2022. “Learning physics between digital twins with low-fidelity models and physics-informed gaussian processes”. *arXiv preprint arXiv:2206.08201*.
- Tipping, M.E. and Faul, A.C., 2003. “Fast marginal likelihood maximisation for sparse bayesian models”. In *International workshop on artificial intelligence and statistics*. PMLR, pp. 276–283.
- Worden, K. and Barthorpe, R., 2012. “Identification of hysteretic systems using narx models, part i: evolutionary identification”. In *Topics in Model Validation and Uncertainty Quantification, Volume 4: Proceedings of the 30th IMAC, A Conference on Structural Dynamics, 2012*. Springer, pp. 49–56.
- Zhan, Z., Fu, Y. and Yang, R.J., 2011. “Enhanced error assessment of response time histories (earth) metric and calibration process”. Technical report, SAE Technical Paper.

8. RESPONSIBILITY NOTICE

The author(s) is (are) solely responsible for the printed material included in this paper.