A Probabilistic Machine Learning Pipeline Using Topological Descriptors for Real-Time State Estimation of High-Rate Dynamic Systems

Yang Kang Chua[1], Daniel Coble[2], Arman Razmarashooli[3], Steve Paul[1], Daniel A. Salazar Martinez[3], Chao Hu[1], Austin R.J. Downey[2,4], Simon Laflamme[3,5]

[1] School of Mechanical, Aerospace, and Manufacturing Engineering, University of Connecticut, Storrs, CT 06269, USA
[2] Department of Mechanical Engineering, University of South Carolina, Columbia, USA
[3] Department of Civil, Construction, and Environmental Engineering, Iowa State University, Ames, IA, 50010, USA
[4] Department of Civil and Environmental Engineering, University of South Carolina, Columbia, USA
[5] Department of Electrical and Computer Engineering, Iowa State University, Ames, IA,50010, USA

Introduction to high-rate systems

Systems experiencing high-rate dynamics involve accelerations > 100 g_n ($g_n = 9.81 \frac{m}{s^2}$) within 1 ms.

Examples

Introduction to high-rate systems

Systems experiencing high-rate dynamics involve accelerations > 100 g_n ($g_n = 9.81\ \frac{m}{s^2}$) within 1 ms.

Examples



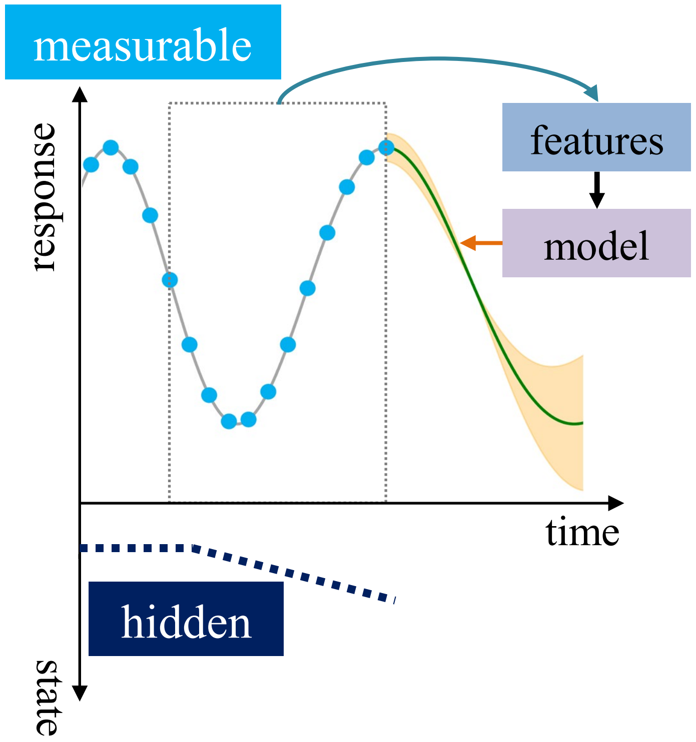
Characteristics

- Large uncertainties in external loads
- High levels of non-stationarities in the structure and heavy disturbances
- Unmodeled dynamics from changes in system configuration

Problem definition

Current State Estimation Models

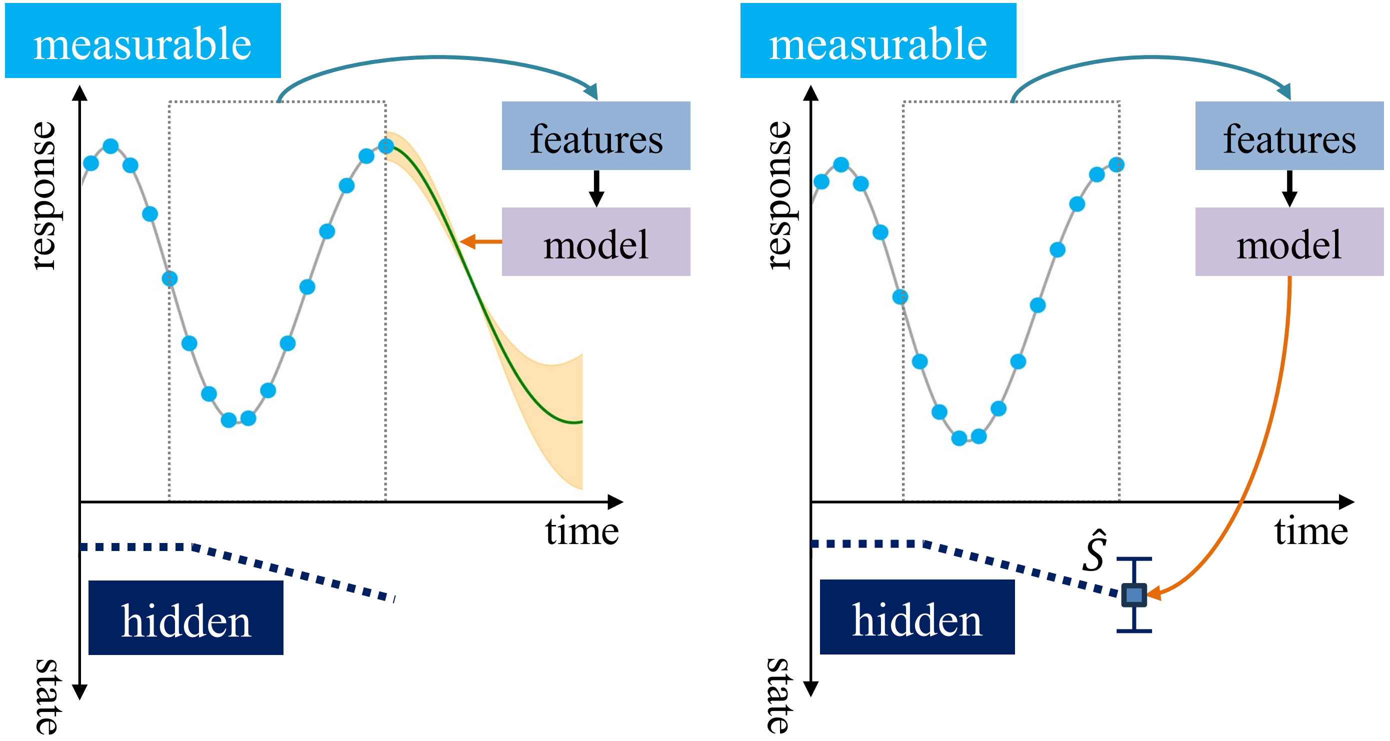
- **Forecasting Models:** Predict future states based on past data patterns.



Problem definition

Current State Estimation Models

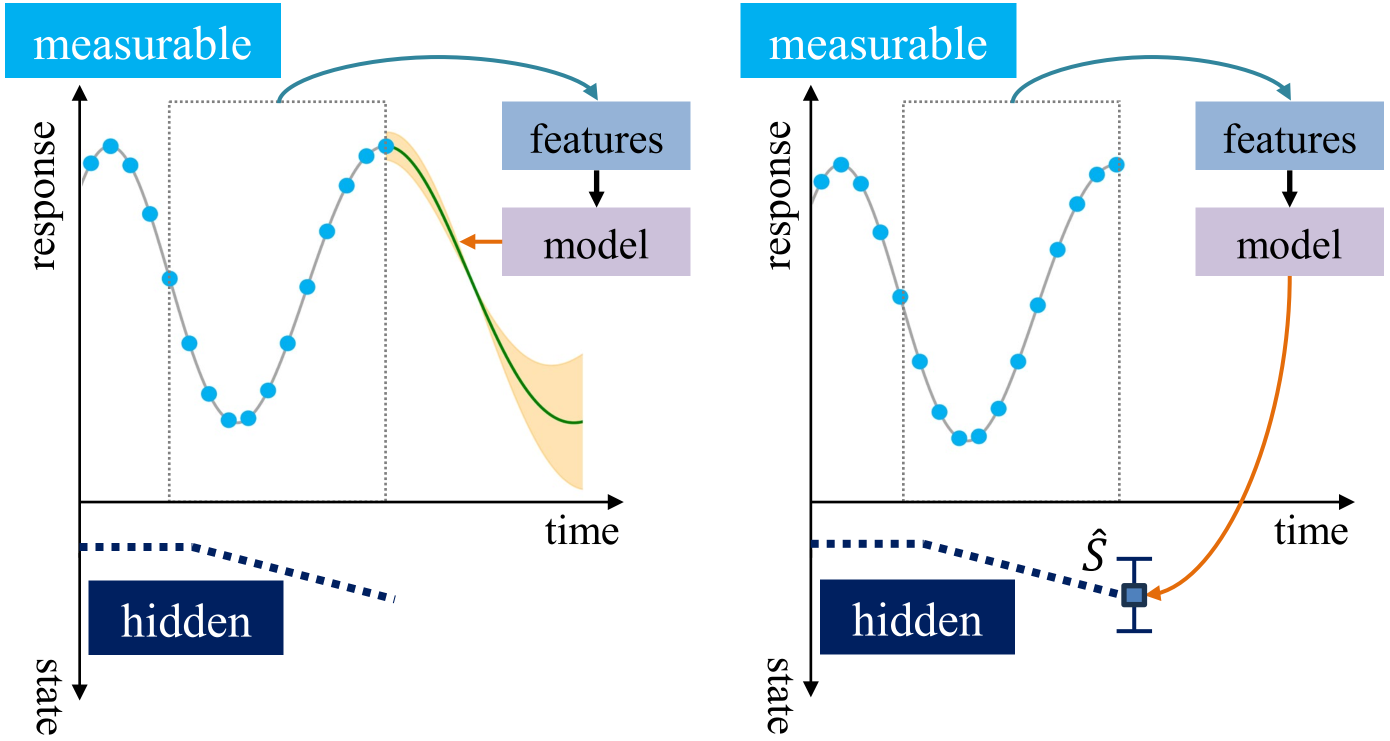
- **Forecasting Models:** Predict future states based on past data patterns.
- **Hidden State Estimation Models:** Estimate unobserved (hidden) internal states of the system that drive its behavior.



Problem definition

Current State Estimation Models

- **Forecasting Models:** Predict future states based on past data patterns.
- **Hidden State Estimation Models:** Estimate unobserved (hidden) internal states of the system that drive its behavior.



Limitation: lack effective uncertainty quantification.

Proposed solution: develop a probabilistic machine learning pipeline that integrates Topological Data Analysis (TDA) to enhance feature extraction, state estimation, and uncertainty quantification (UQ).

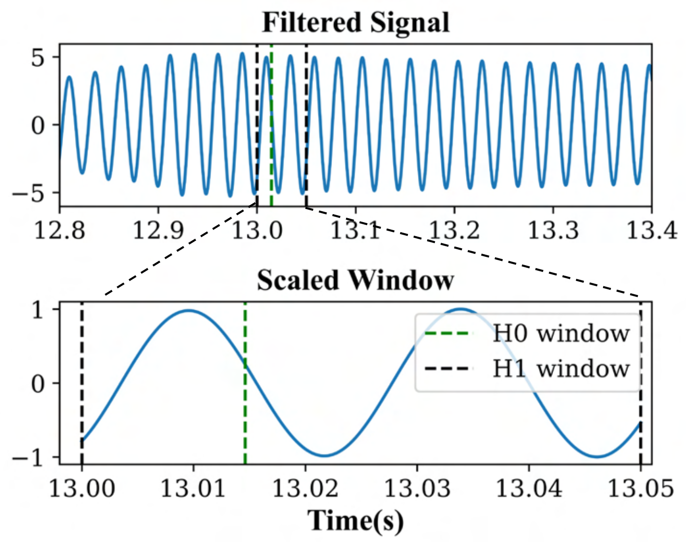
Methodology Overview

Offline Data

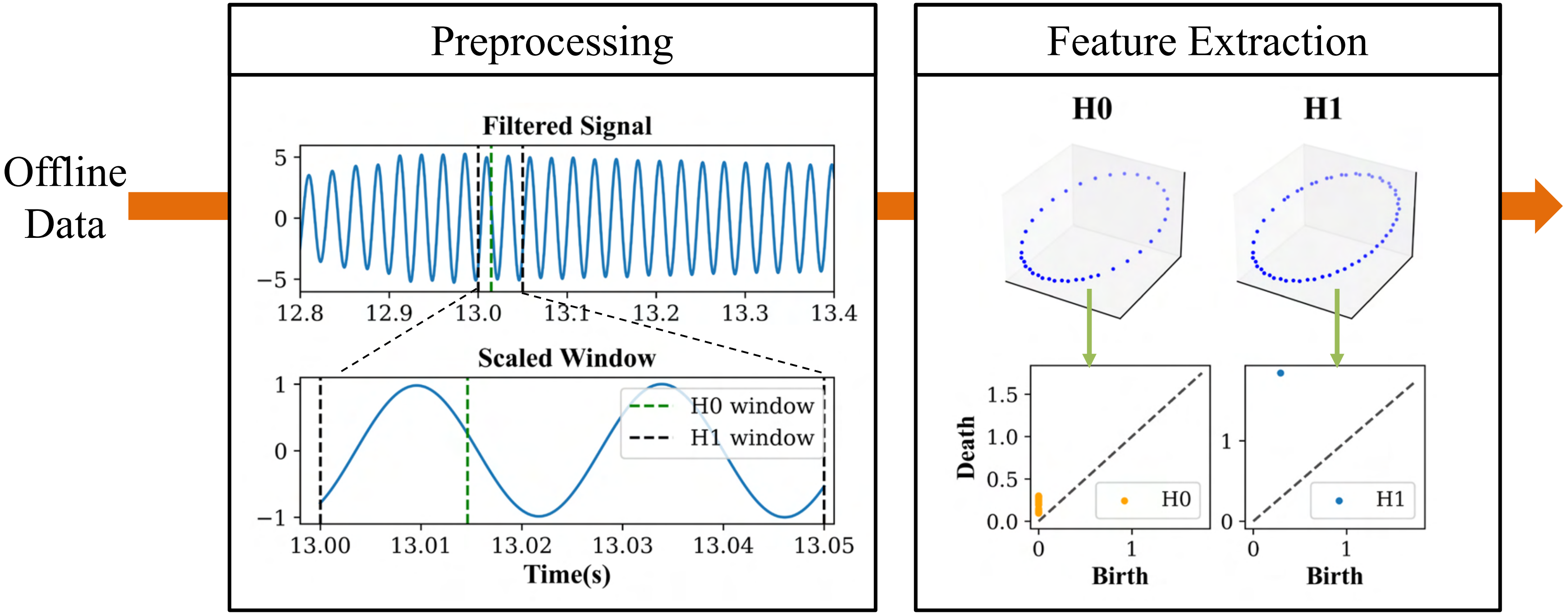
Methodology Overview

Offline Data

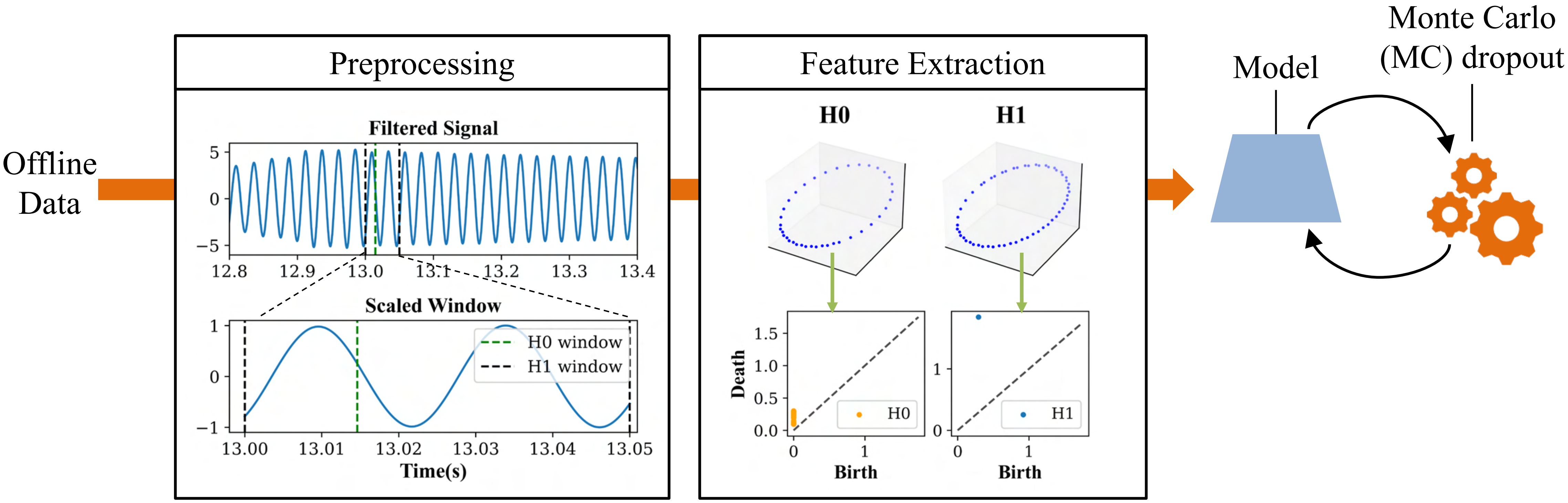
Preprocessing



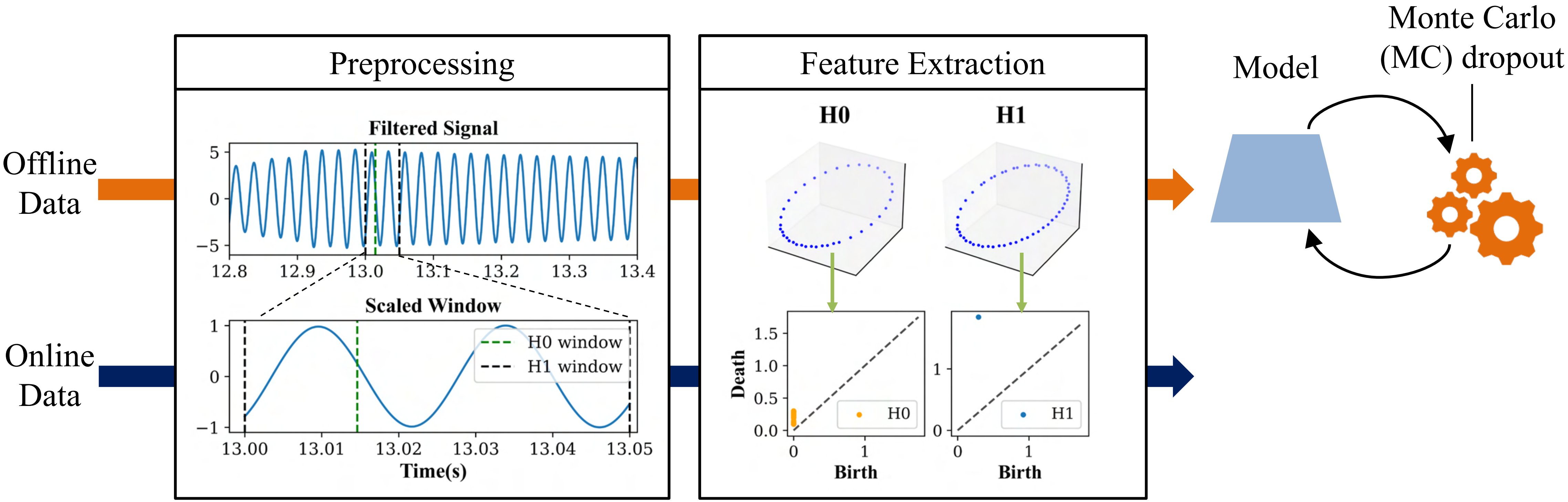
Methodology Overview



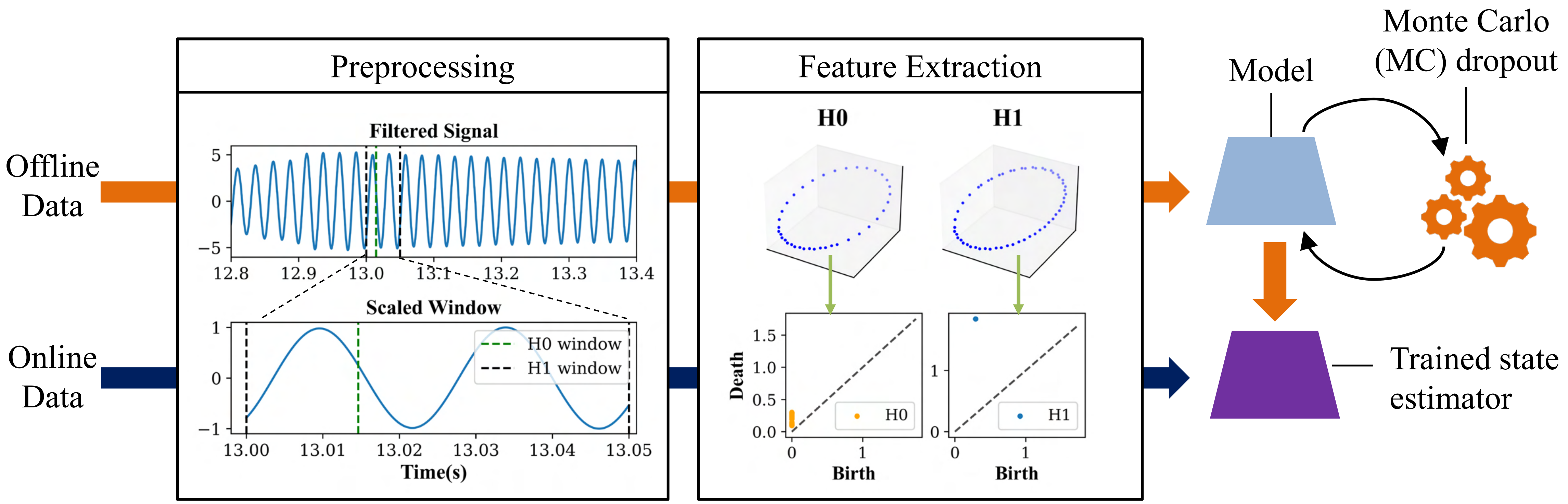
Methodology Overview



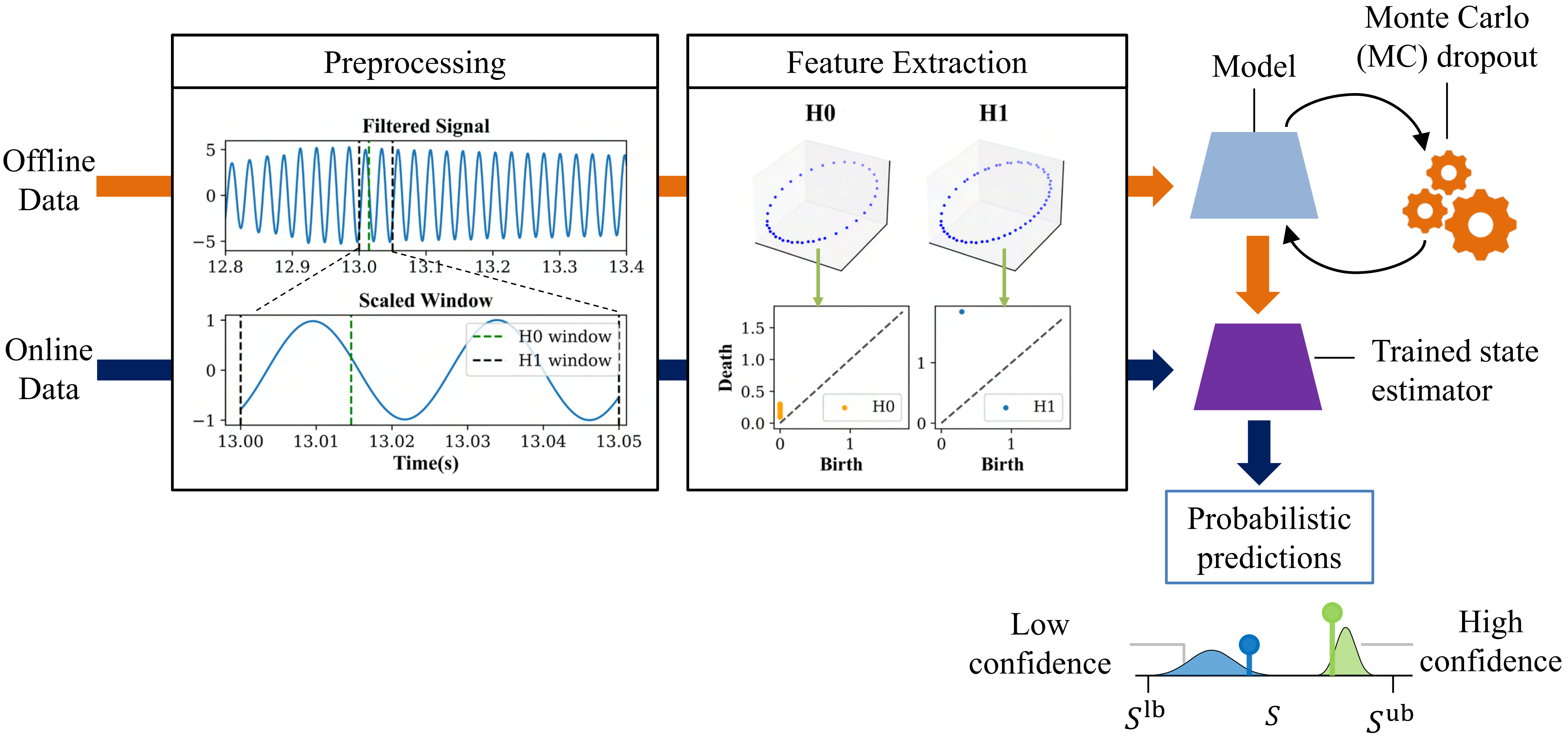
Methodology Overview



Methodology Overview

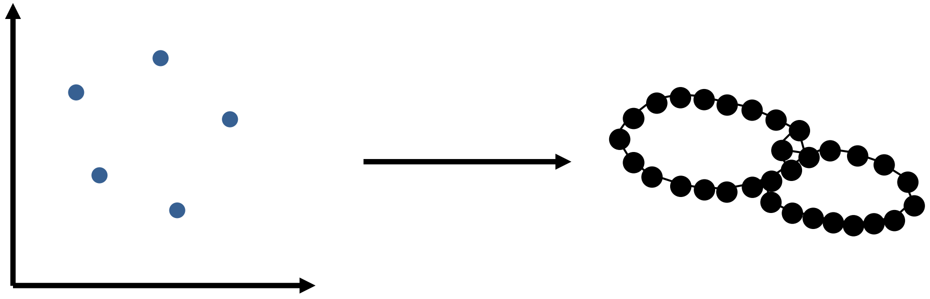


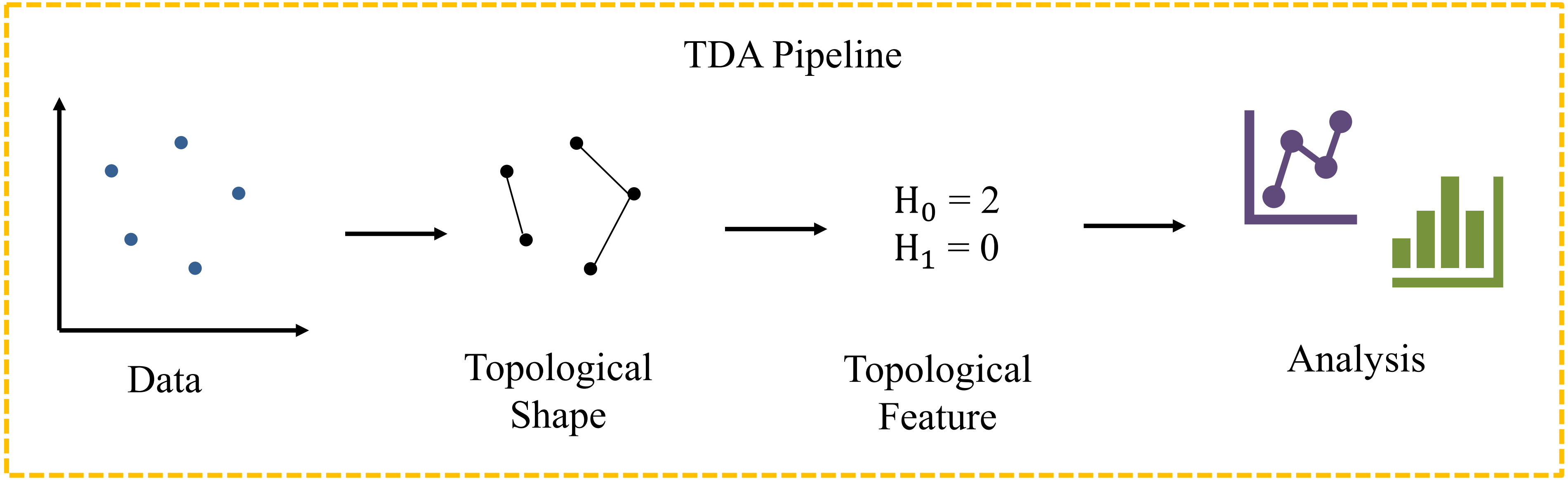
Methodology Overview



Generating Datasets: Topological Data Analysis (TDA)

TDA is a method that studies the shape and structure of data by identifying patterns and features that persist across multiple scales.

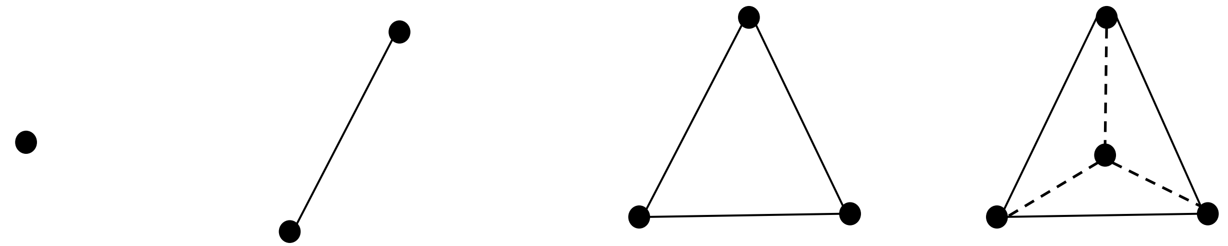




Generating Datasets: Topological Data Analysis (TDA)

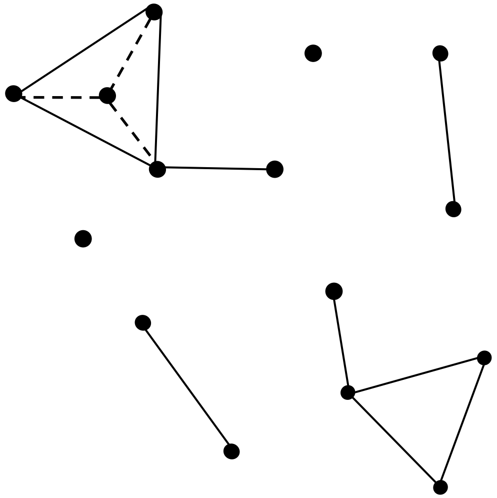
Topological shapes

Simplex: A **simplex** is the building block in topological data analysis. It's the simplest possible geometric object that represents a relationship between points



Point (0-simplex) | Line (1-simplex) | Triangle (2-simplex) | Tetrahedron (3-simplex)

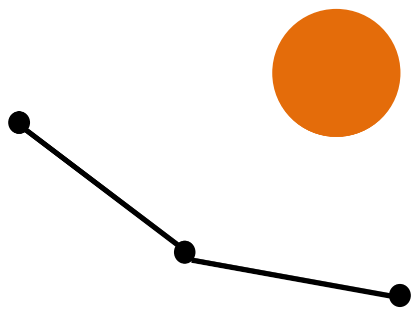
Collection of simplexes (simplicial complexes)



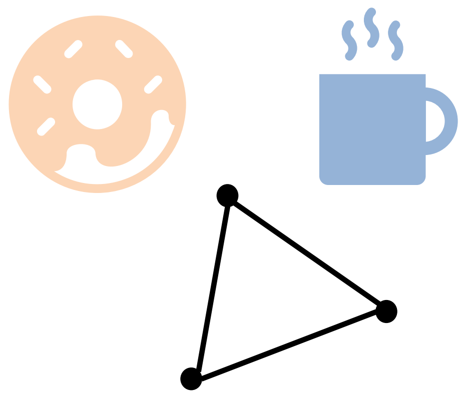
Generating Datasets: Topological Data Analysis (TDA)

Topological shapes

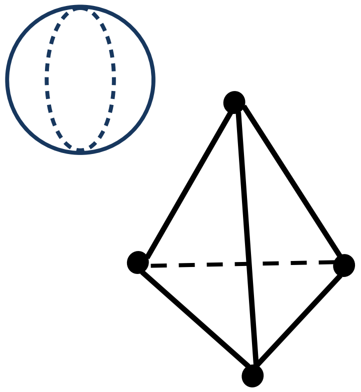
Homology: is a concept in topology that helps identify and categorize holes or voids within a topological space.



Connected
Components
H_0



Holes
H_1



Void
H_2

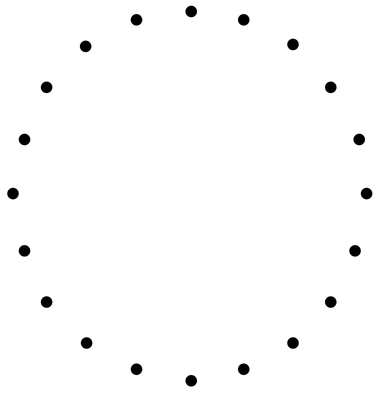
H_0, H_1, H_2 are the homology groups that capture different dimensional features of a topological space

Generating Datasets: Topological Data Analysis (TDA)

Persistent Homology is a technique in **Topological Data Analysis (TDA)** that captures and analyzes the topological features of data across multiple scales or resolutions. It helps in understanding how these features persist as the data is viewed at different levels of detail.

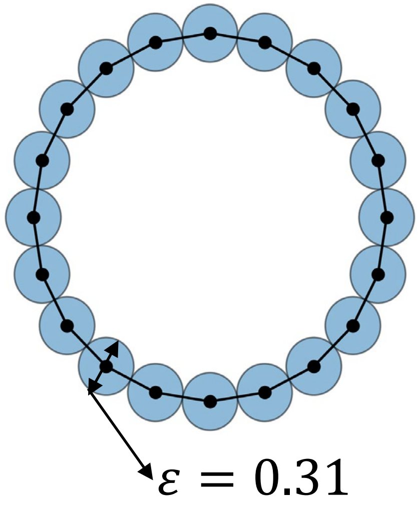
H_0 : Connected Components
H_1 : loops or cycle

Point Cloud Data



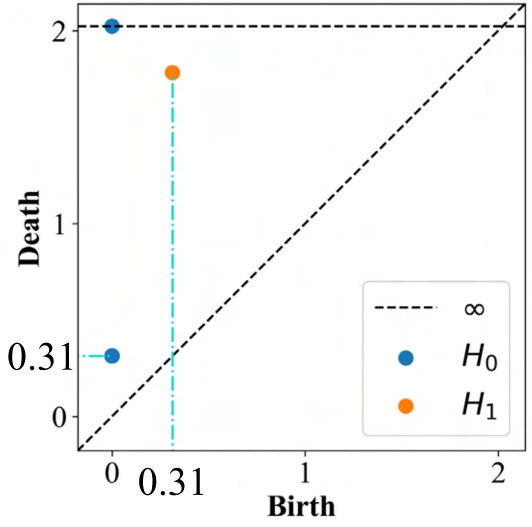
Filtration

Connected Component



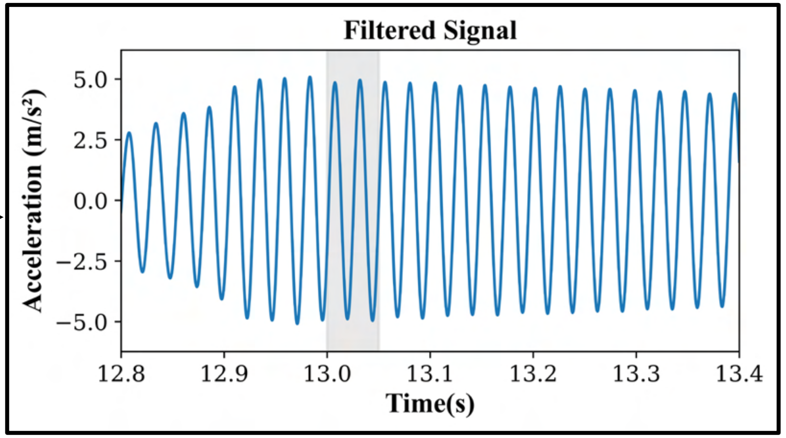
Record changes

Persistence Diagram

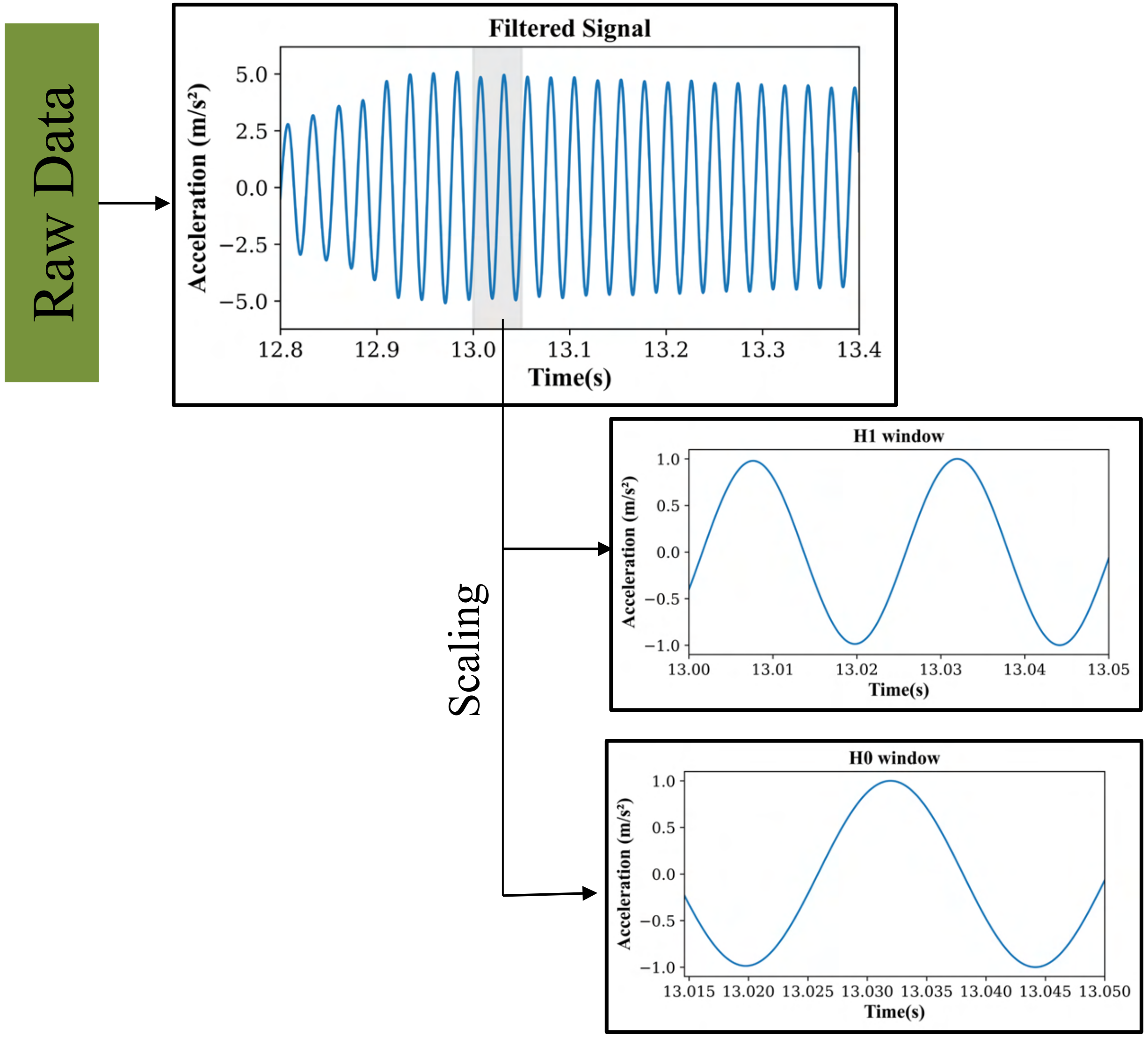


Topological Data Analysis (TDA) Feature Extraction

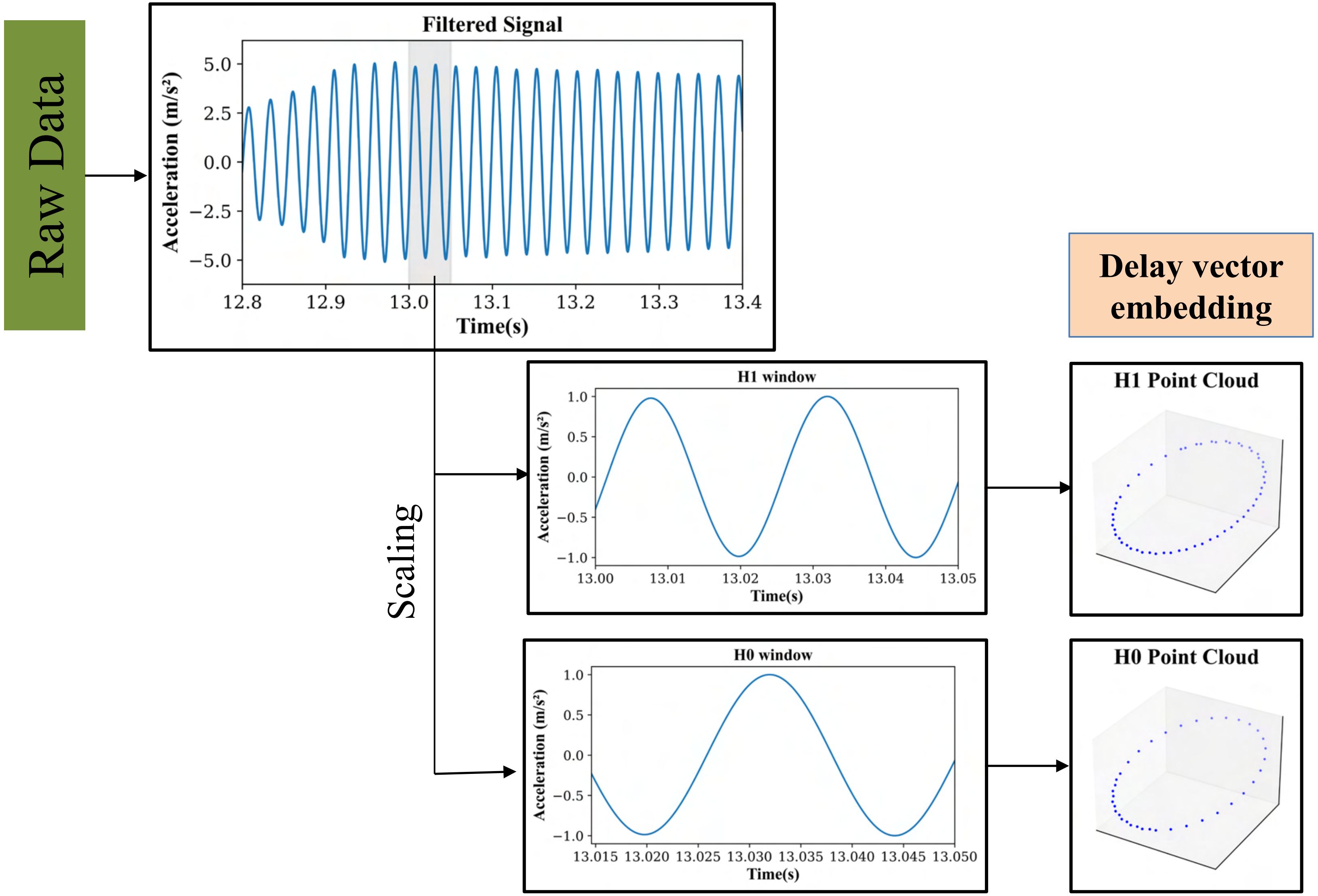
Raw Data



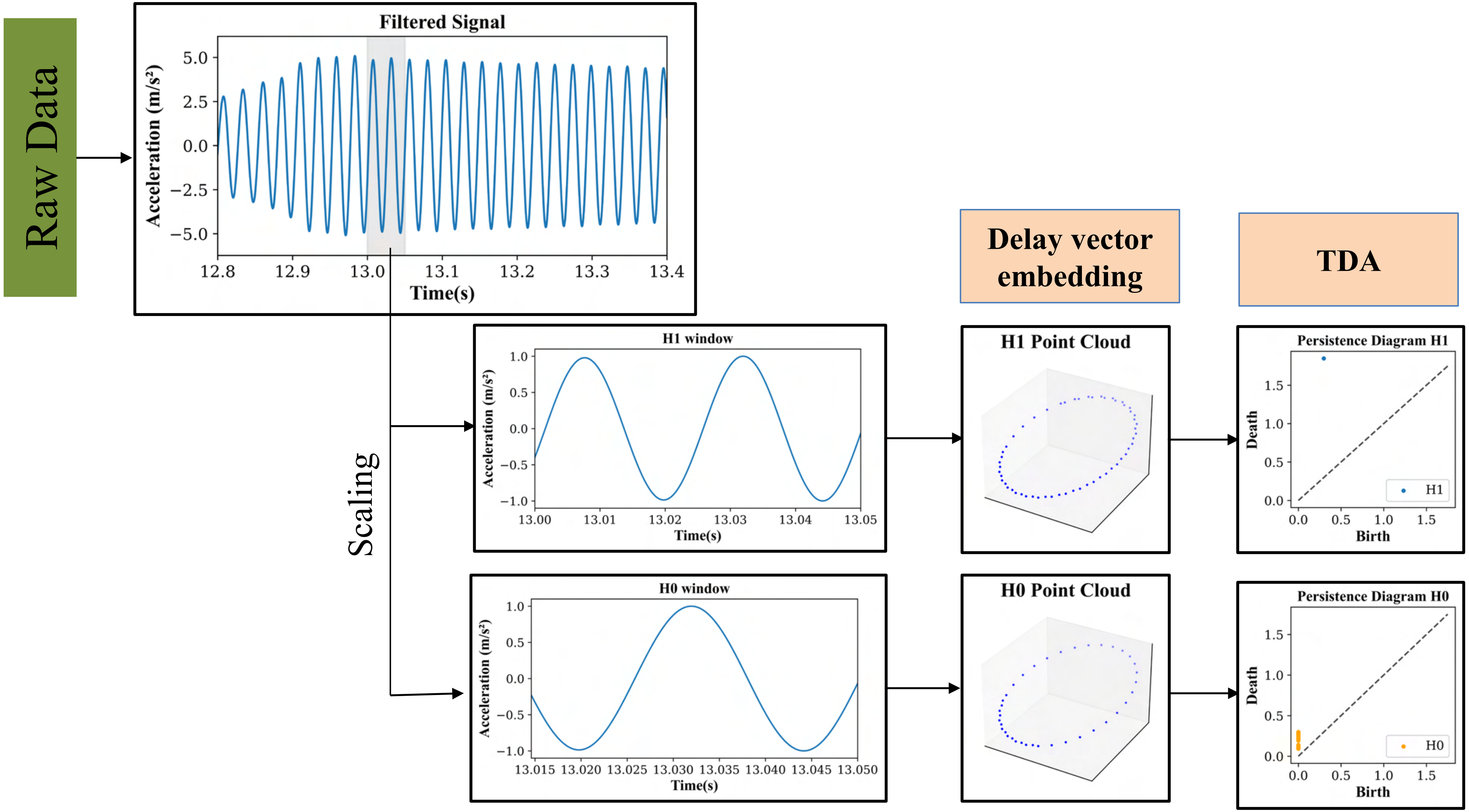
Topological Data Analysis (TDA) Feature Extraction



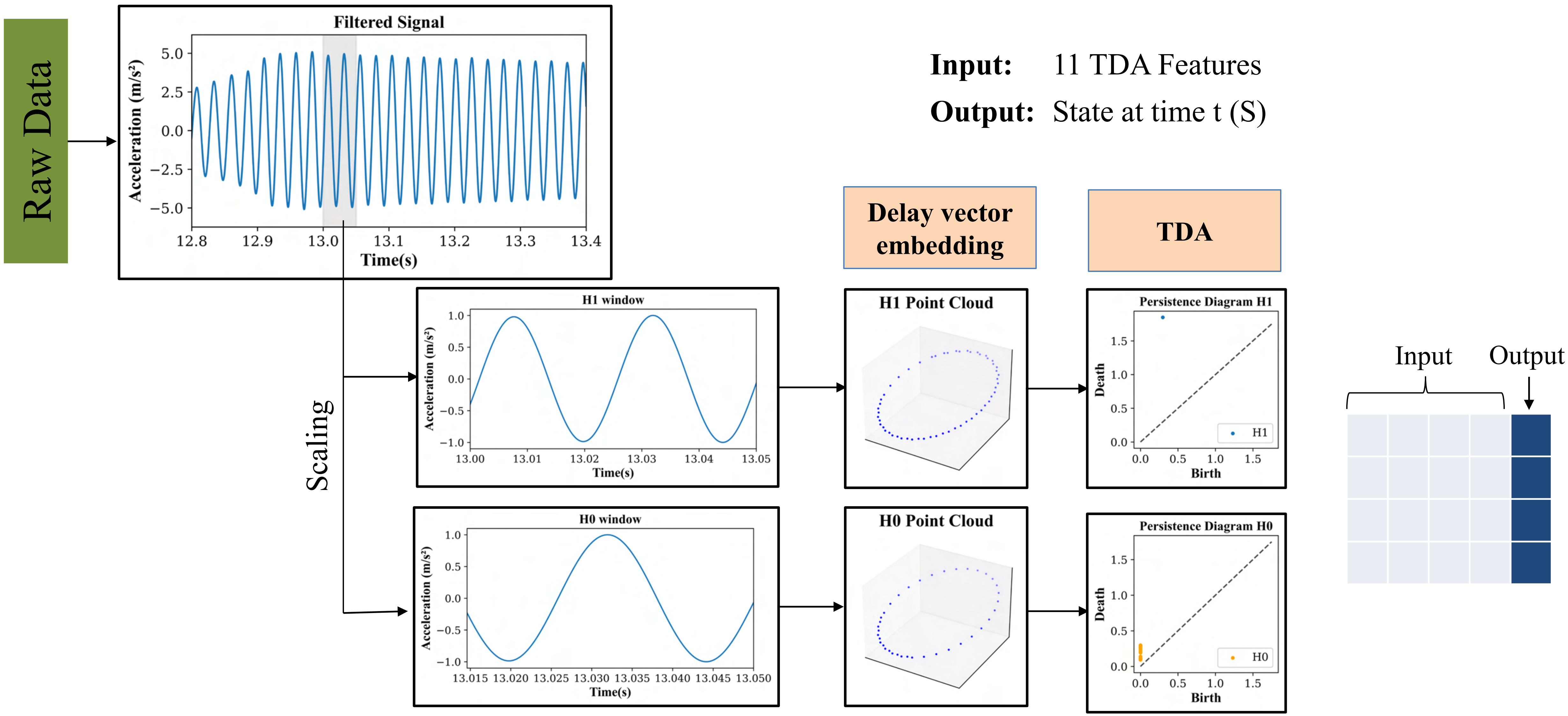
Topological Data Analysis (TDA) Feature Extraction



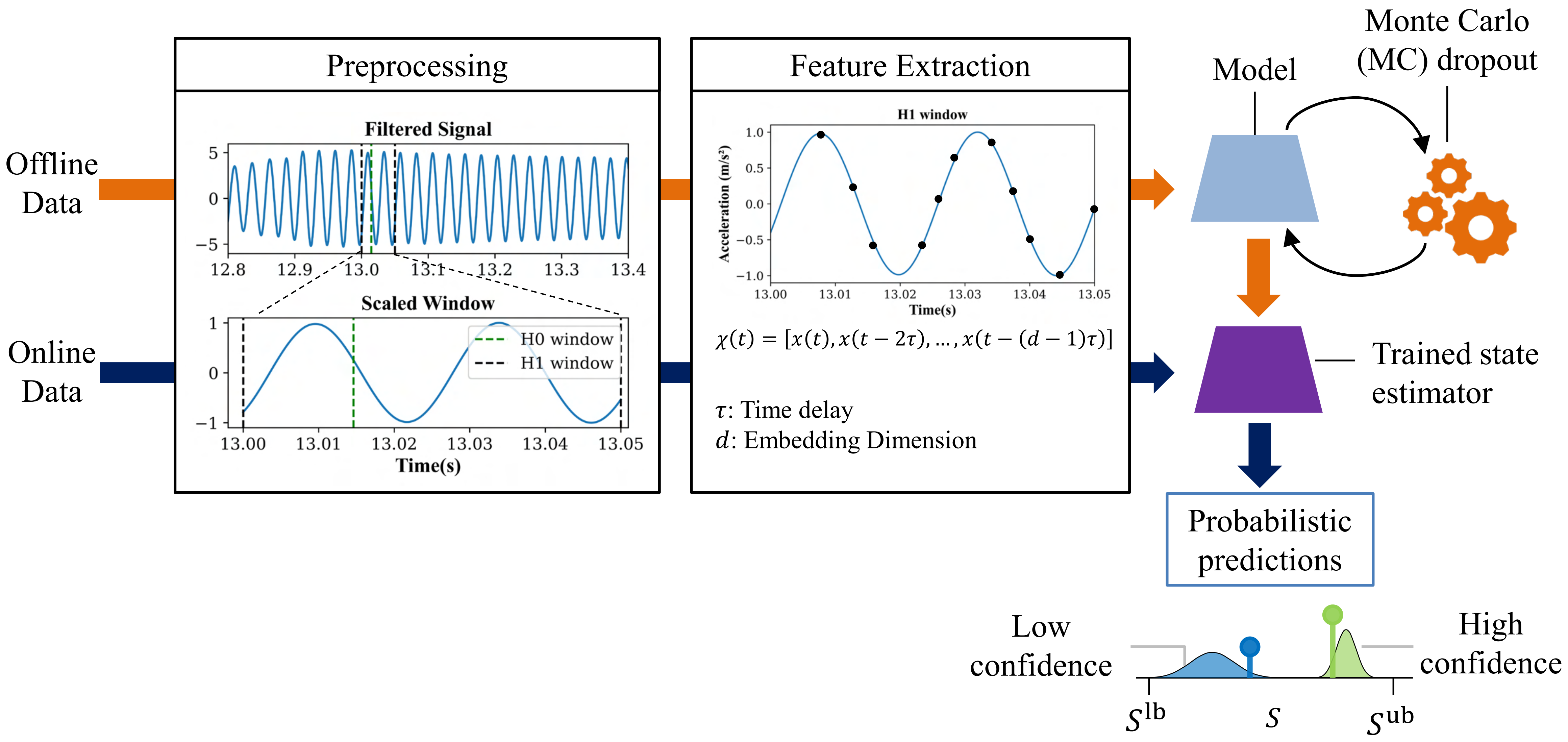
Topological Data Analysis (TDA) Feature Extraction



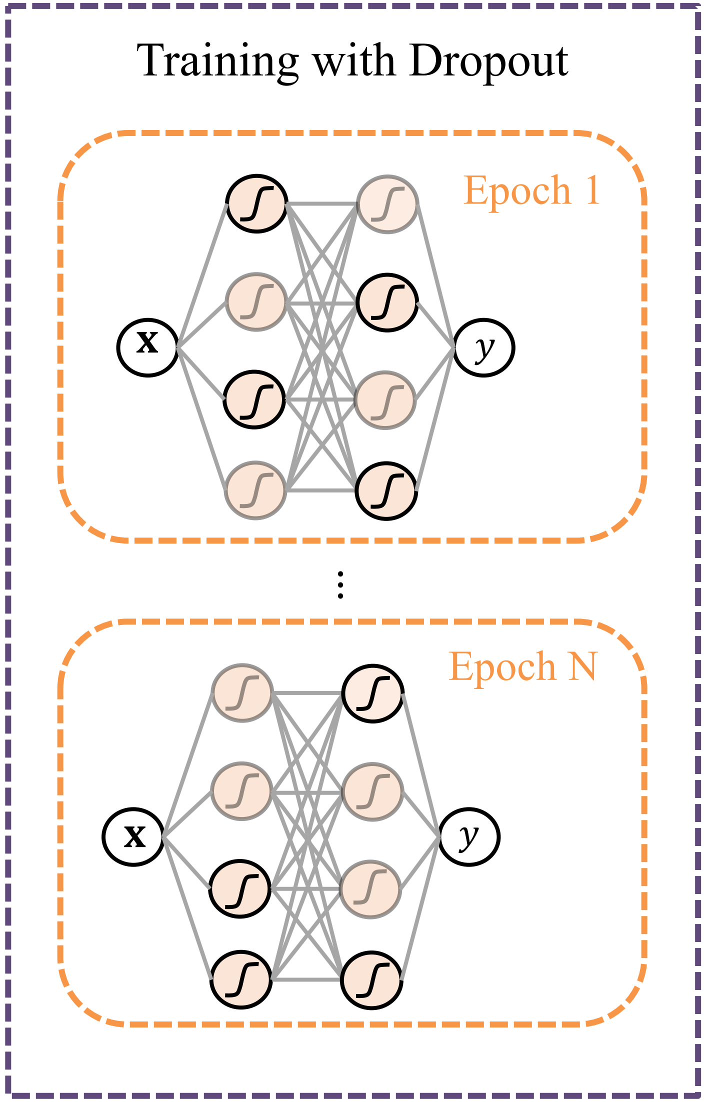
Topological Data Analysis (TDA) Feature Extraction



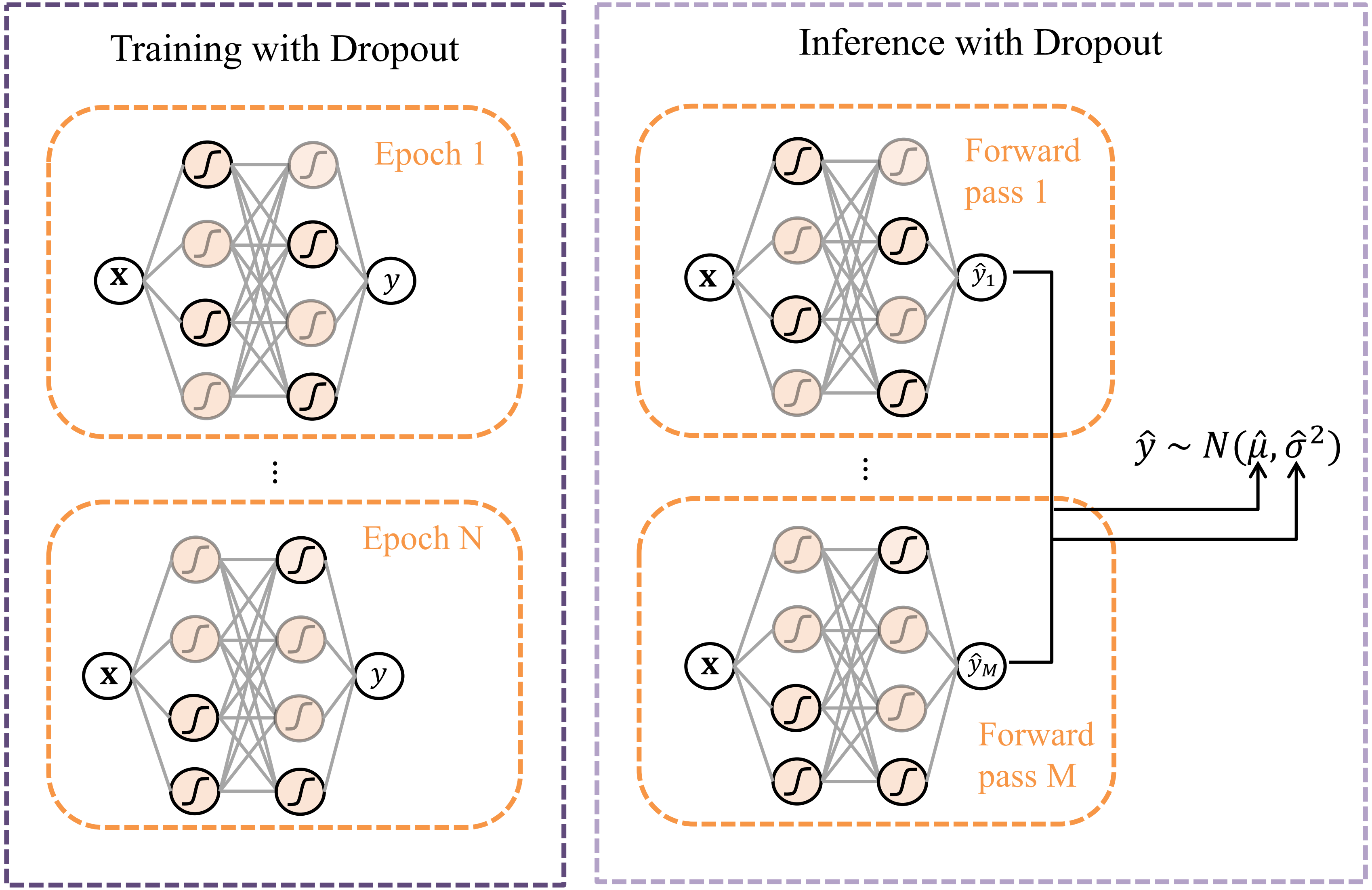
Methodology Overview: Baseline Comparison



Monte Carlo (MC) Dropout



Monte Carlo (MC) Dropout



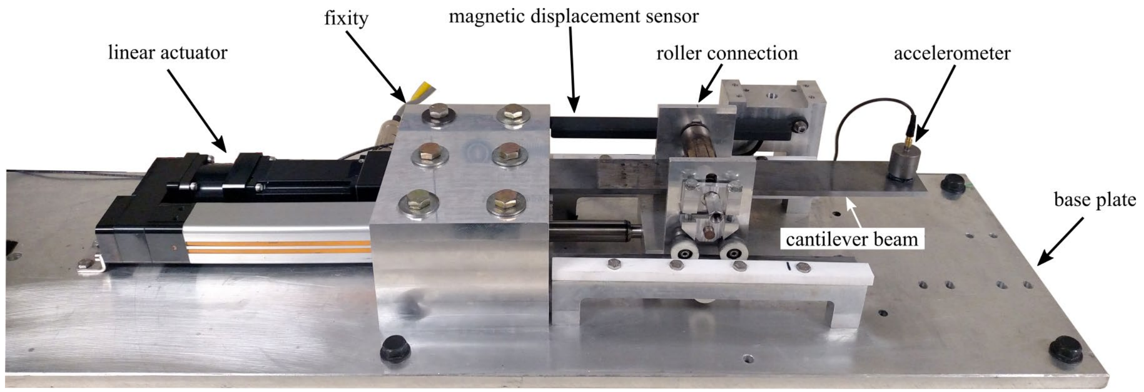
Hyperparameters

Number of forward passes (M)
Number of layer
Number of neurons
Activation function

Case Study : DROPBEAR Dataset

DROPBEAR experimental testbed:
The Dynamic Reproduction of Projectiles in Ballistic Environments for Advanced Research (DROPBEAR) was used to generate the experimental data.

Capabilities:
Reproduce fast boundary condition changes.
Mimic rapid mass changes.
Simulate the rapid dynamics typical of high-rate events.

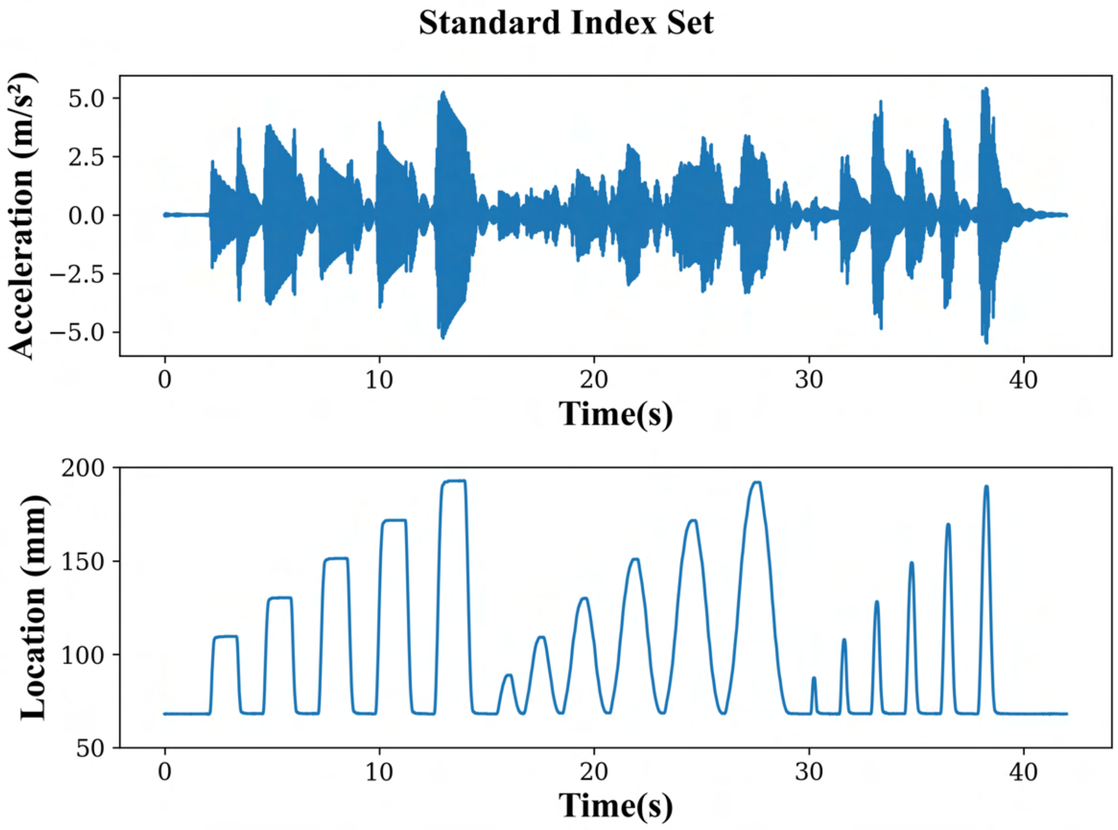


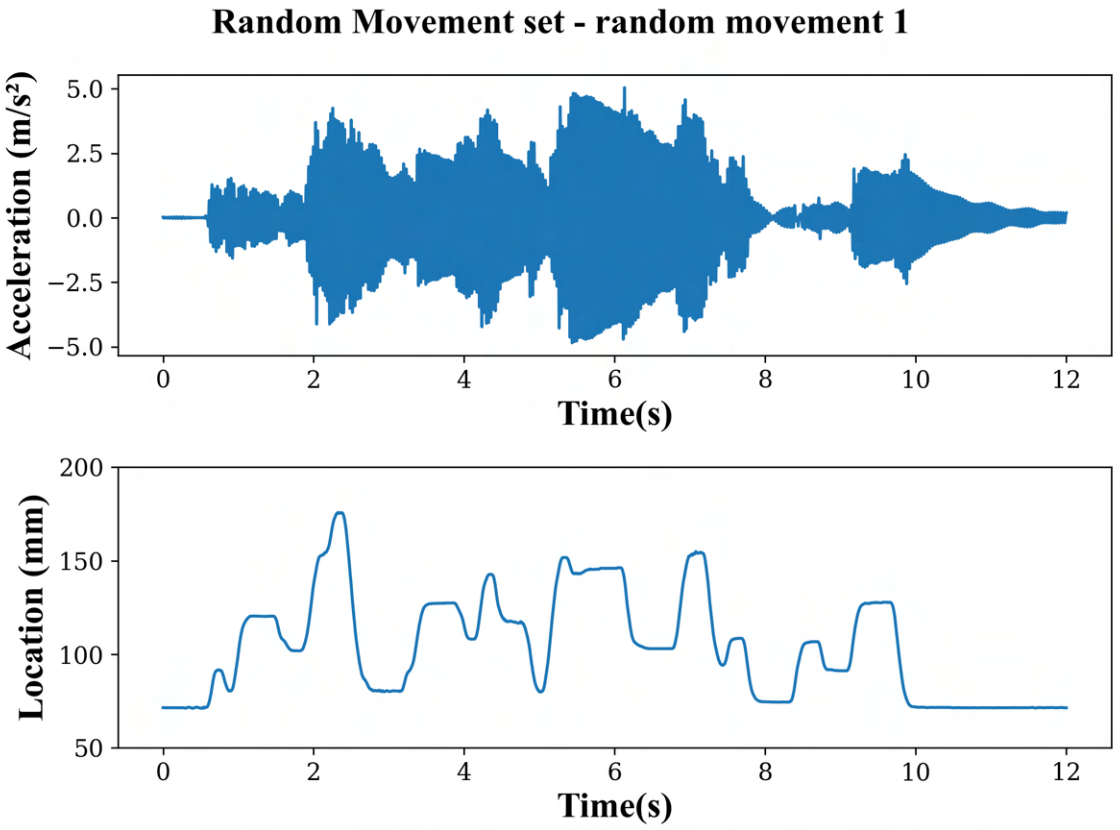
DROPBEAR Dataset 2 experimental video

Austin Downey, Jonathan Hong, Jacob Dodson, Michael Carroll, and James Scheppegrell, "Dataset-2-dropbearacceleration-vs-roller-displacement," Dec. 2021. [Online]. Available: https://github.com/High-Rate-SHM-Working-Group/Dataset-2-DROPBEAR-Acceleration-vs-Roller-Displacement

Case Study : DROPBEAR Dataset

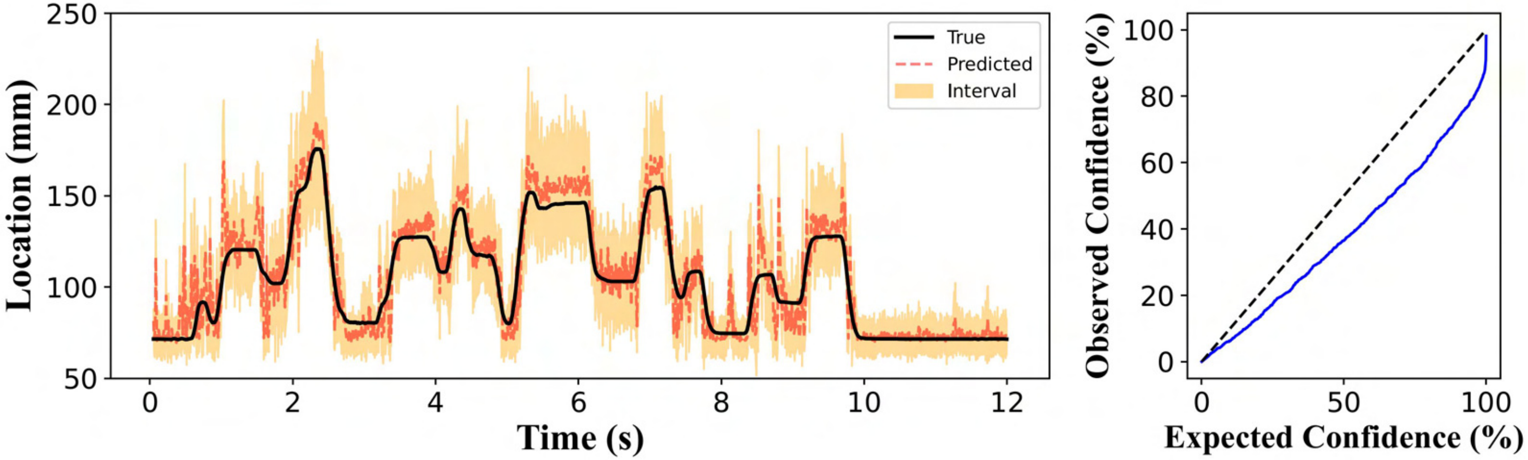
Training and validation	Testing



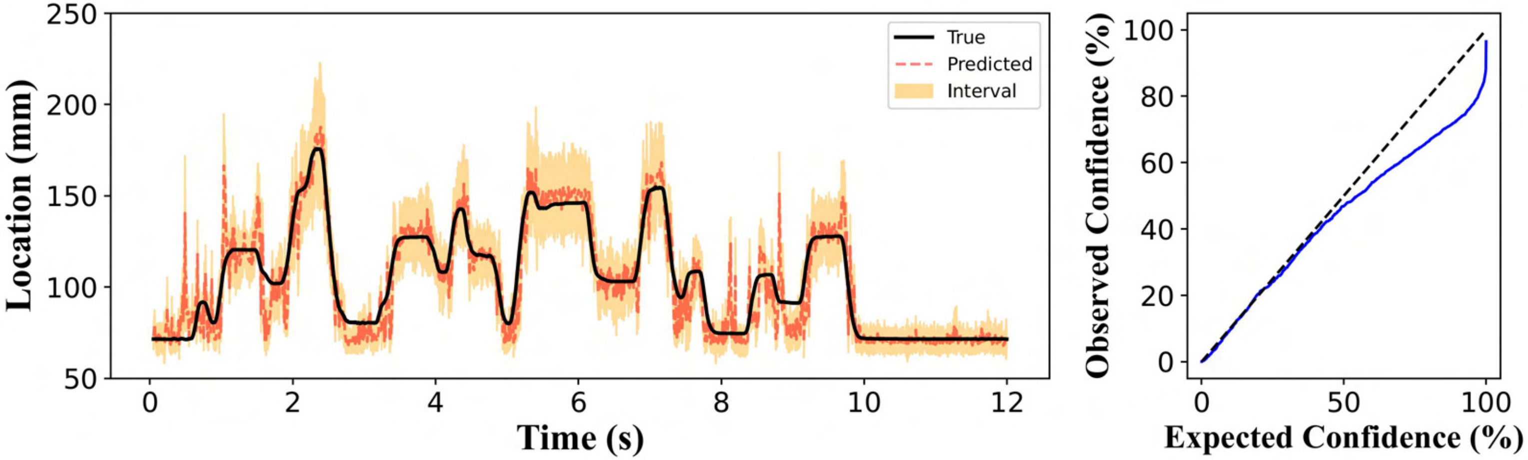


Case Study : Results

Result using Delay vector as input



Result using TDA features as input



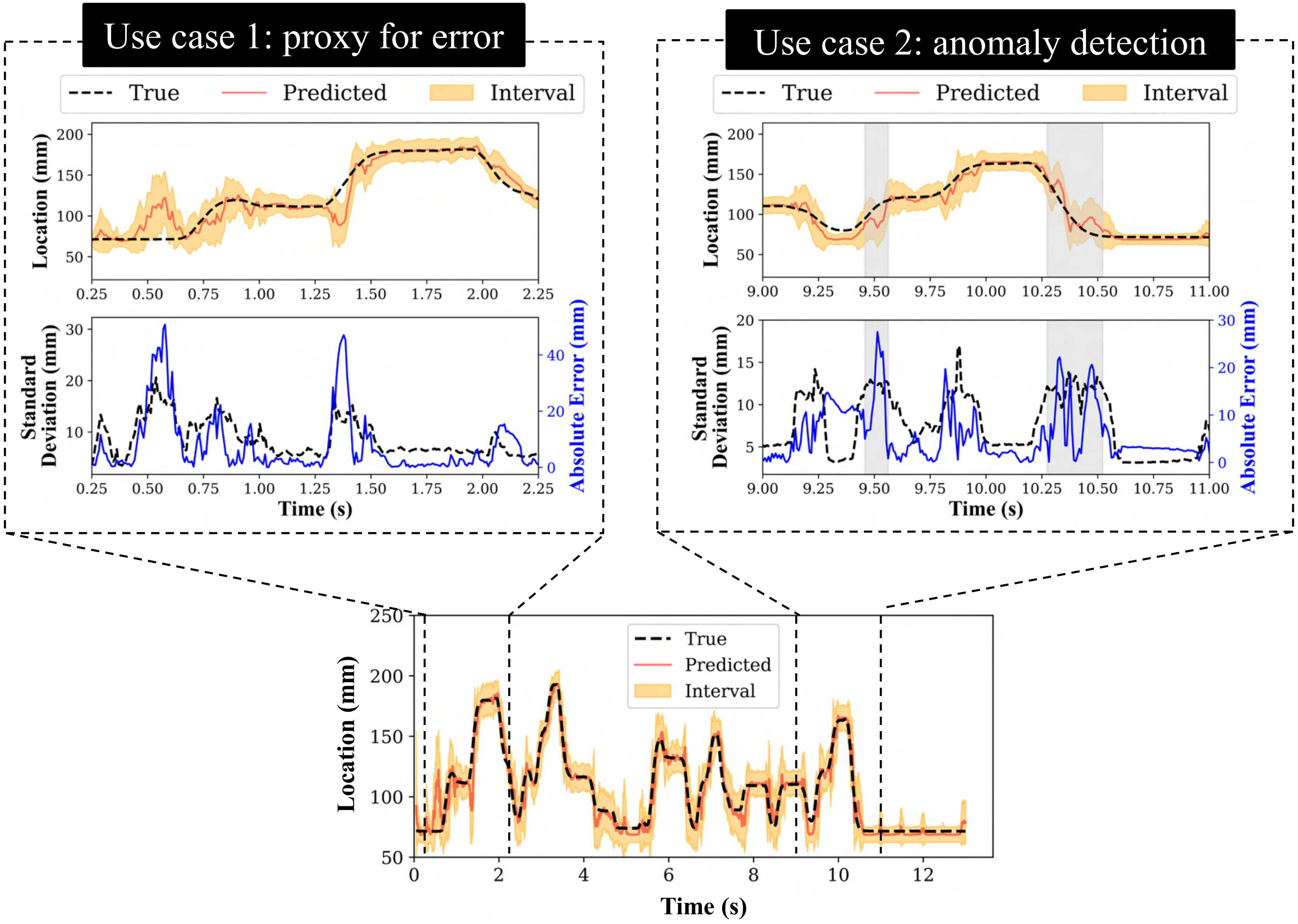
Case Study : Results

Metric Used

- **Mean Absolute Error (MAE)**: Measures average absolute error between predicted and actual values.
- **Time Response Assurance Criterion (TRAC)**: Assesses the correlation between predicted and actual time-series data.
- **Signal-to-Noise Ratio (SNR)**: Indicates the model's accuracy by comparing signal strength to noise level.
- **Expected Confidence Error (ECE)**: Evaluates the accuracy of the model's predictive uncertainty.

Data Type	MAE (%)	TRAC (%)	SNR (dB) (%)	ECE (%)
Standard	13.2	-0.1	3.2	60.3
Stepwise - 10	16.7	-0.1	2.4	13.1
Stepwise - 30	15.9	-0.2	0.2	72.6
Stepwise - 60	13	-0.2	-0.7	73.6
Random	12.1	0	3.3	47.6
Average	**14.2**	**-0.1**	**1.7**	**53.4**

Case Study: Practical Applications



Summary

Key Outcomes:
- Increased accuracy and reliability in high-rate state estimation.
- Enhanced feature extraction through TDA integration.
- Improved state estimation with robust uncertainty quantification (UQ).
- Better decision-making based on prediction confidence.

Future Work:
- Stabilize metrics for the probabilistic model by addressing run-to-run variation.
- Explore additional optimization and generalization methods for the proposed machine-learning pipeline.
- Incorporate forecasting capabilities into the probabilistic model.

Acknowledgements

- Air Force Office of Scientific Research (AFOSR).
- Defense Established Programs to Stimulate Competitive Research (DEPSCoR).
- Air Force Research Laboratory Munitions Directorate.
- National Science Foundation

Questions?

Backup Slides

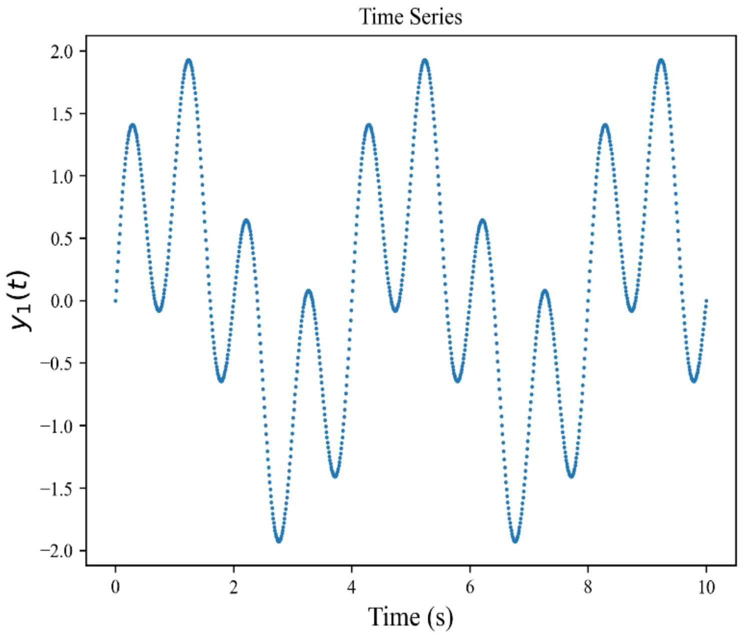
Generating Datasets: Conversion to Point Clouds

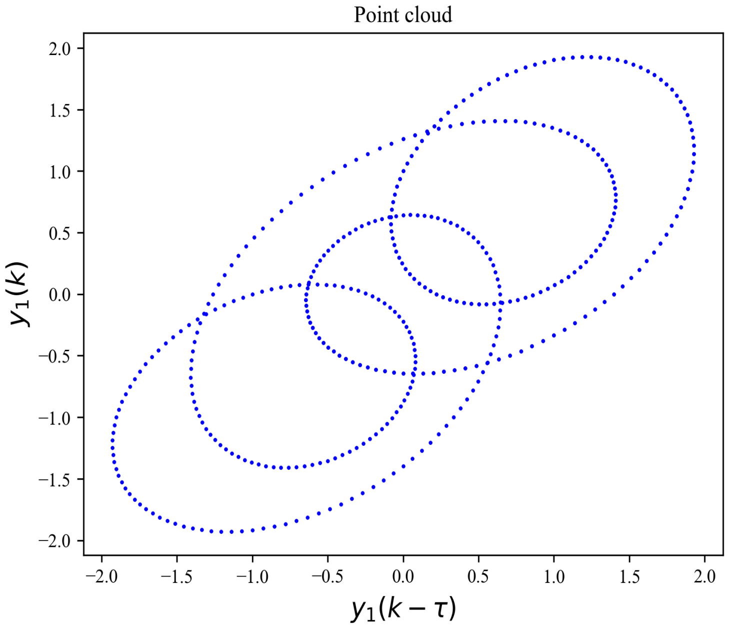
Time series to point cloud: Use delay vector embedding.

$$\chi(t) = [x(t), x(t - \tau), x(t - 2\tau), \dots, x(t - (d-1)\tau)]$$

Parameters:

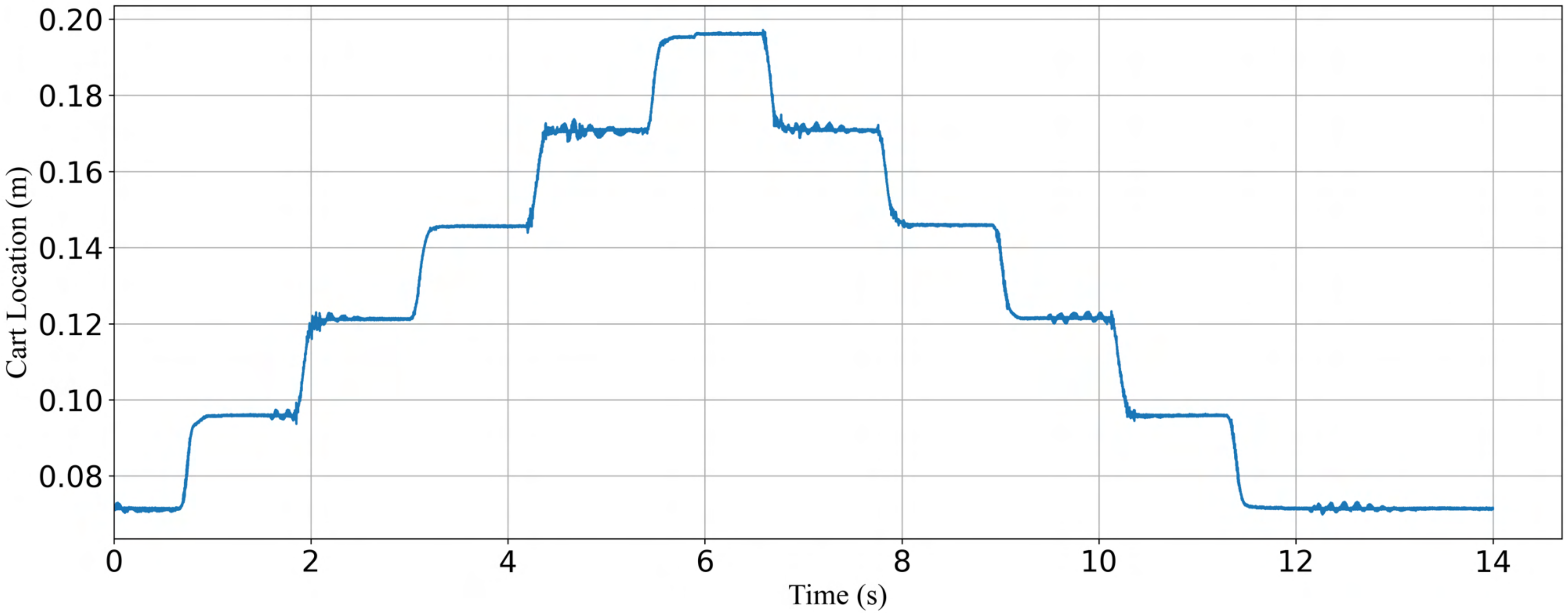
- τ = Time delay
- d = Embedding dimension





Case Study : DROPBEAR Dataset – Testing

- Moving the Cart in Incremental Steps:
  - Utilizing step sizes of 10, 30, and 60.
  - Each configuration comprises 10 trails.



Case Study : Results

Metric Used

- **Mean Absolute Error (MAE)**: Measures average absolute error between predicted and actual values.
- **Time Response Assurance Criterion (TRAC)**: Assesses the correlation between predicted and actual time-series data.
- **Signal-to-Noise Ratio (SNR)**: Indicates the model's accuracy by comparing signal strength to noise level.
- **Expected Confidence Error (ECE)**: Evaluates the accuracy of the model's predictive uncertainty.

	Delay Vector				TDA Features			
Data types	**MAE (mm)**	**TRAC**	**SNR (dB)**	**ECE (%)**	**MAE (mm)**	**TRAC**	**SNR (dB)**	**ECE (%)**
Standard	8.92	0.9875	18.15	22.03	7.74	0.987	18.73	8.75
Stepwise - 10	8.24	0.994	21.05	6.35	6.86	0.9933	21.55	5.52
Stepwise - 30	8.81	0.9938	20.75	10.8	7.41	0.9921	20.8	2.96
Stepwise - 60	9.53	0.9931	20.26	12.98	8.29	0.9909	20.11	3.43
Random	7.95	0.9905	19.65	11.76	6.99	0.9906	20.29	6.16