Persistent Homology for World Modeling

One-Page Summary

Key Questions Addressed

Theme	Question
Topological Seasonality Detection	Can persistent homology quantify and model seasonal structure in city-level weather time series?
Predictability of Loop Strength	Can loop strength—capturing Betti-1 features—be predicted from raw temperature profiles without computing persistent homology at inference time?
Cross-Domain Extension	Can this framework generalize to financial and economic time series, where latent cyclicity and regime shifts also play a central role?

Conceptual Ideas Proposed

- **Persistent homology** measures loop strength in city-year temperature profiles, capturing seasonal cycles as Betti-1 features from 3D window embeddings.
- Loop strength is defined as the **maximum Betti-1 persistence**, providing a scalar summary of cyclicity that is robust to noise and agnostic to calendar structure.
- A **Random Forest model** predicts loop strength from monthly temperature vectors, showing topological structure is learnable from raw profiles.

Key Results

- **Loop strength varies widely** across cities and years, with a heavy-tailed distribution capturing both cyclical and irregular patterns.
- The Random Forest model achieves $\mathbf{R^2 = 0.9653}$, **MAE = 0.1770**, and **median absolute error = 0.0272**, confirming loop strength is highly predictable.
- **December is the most important month** for predicting loop strength, reflecting the role of year-end closure in forming strong seasonal loops.
- **City clustering by loop strength** reveals distinct climate groups, offering a topological signature of seasonality.

Illustrative Figures and Tables

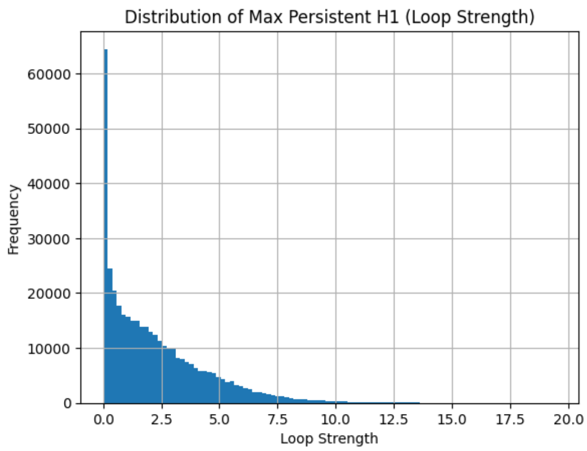

Figure 1: **Distribution of loop strength (maximum Betti-1 persistence) across city-year samples.** The heavy-tailed distribution captures both highly cyclical and irregular weather patterns.

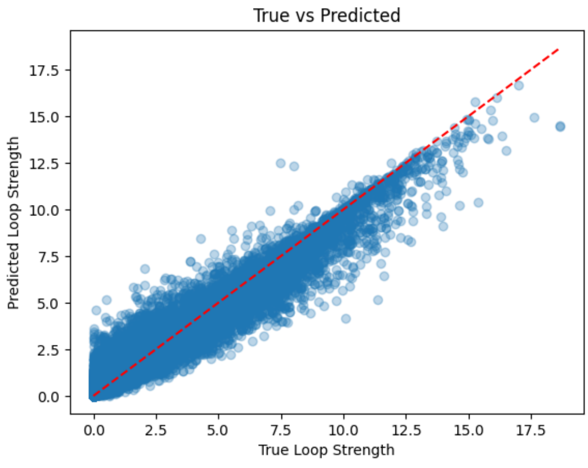

Figure 2: **Predicted vs. true loop strength using Random Forest regression.** High alignment confirms that topological structure is learnable from monthly temperature vectors without direct persistent homology computation.