Getting the most from low data cardinality in time series

StiefelGen: A Novel Perspective on Time Series Data Augmentation

Prasad Cheema, Mahito Sugiyama
(National Institute of Informatics)

どんな研究？

- A method to perform time series data augmentation in low cardinality settings
- Important for many engineering fields!

🌉🚀🗼✈

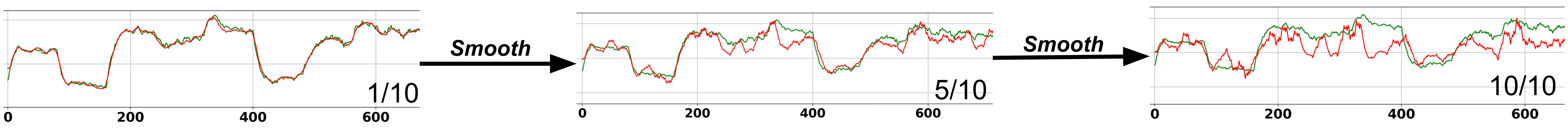
何がわかる？

- Generalization of deep learning models
- Linear forecasting with uncertainty
- Enhancing structural health monitoring

背景・目的

In many settings there is limited time series data. It can be hard to get more because: (i) Too expensive to create more (🚀), (ii) Not enough observed history (🌡) (iii) Inaccurate scientific / econometric models (💹).

We can solve (i) – (iii) using (Stiefel) differential geometry!



研究内容（方法・結果・結論）

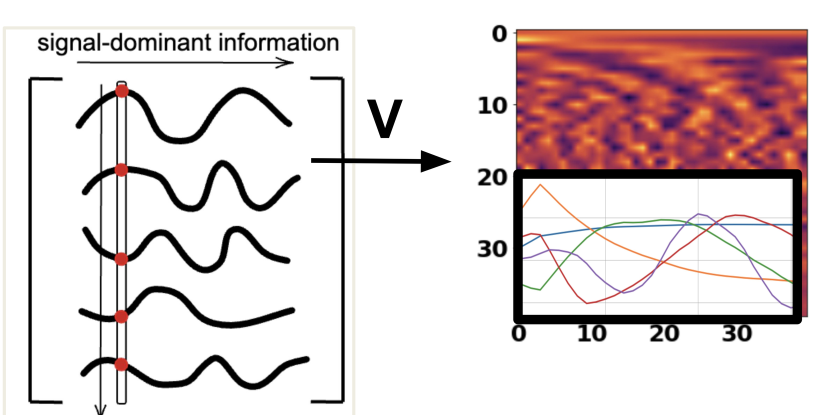
方法

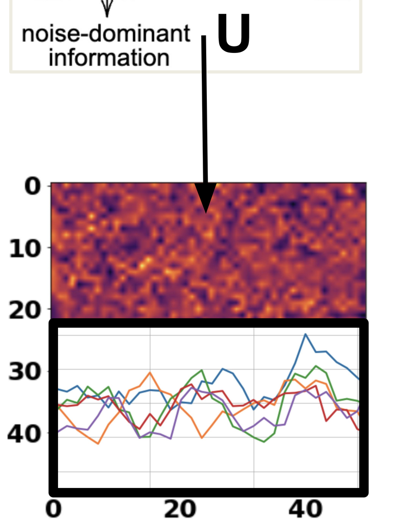
Basic Idea:

0. SVD decomposition (matrix representation, $X=USV^T$)
1. Perturbation (U and V matrices from SVD)
2. Linear Projection (onto tangent space)
3. "Pull Back" (to within radius of injectivity)
4. Exponential Map (onto Stiefel Manifold)
5. SVD reconstruction (obtain augmented time series)

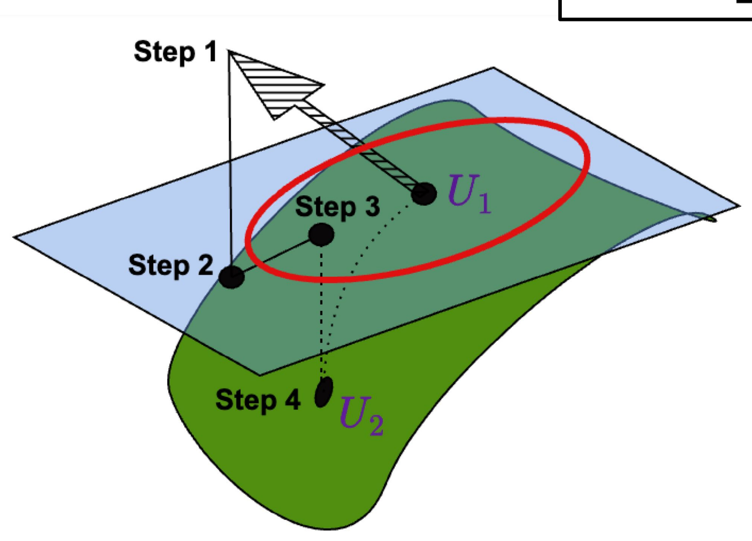
Univariate Case: Must reshape to a *page matrix*

Multivariate Case: Simply stack the time series



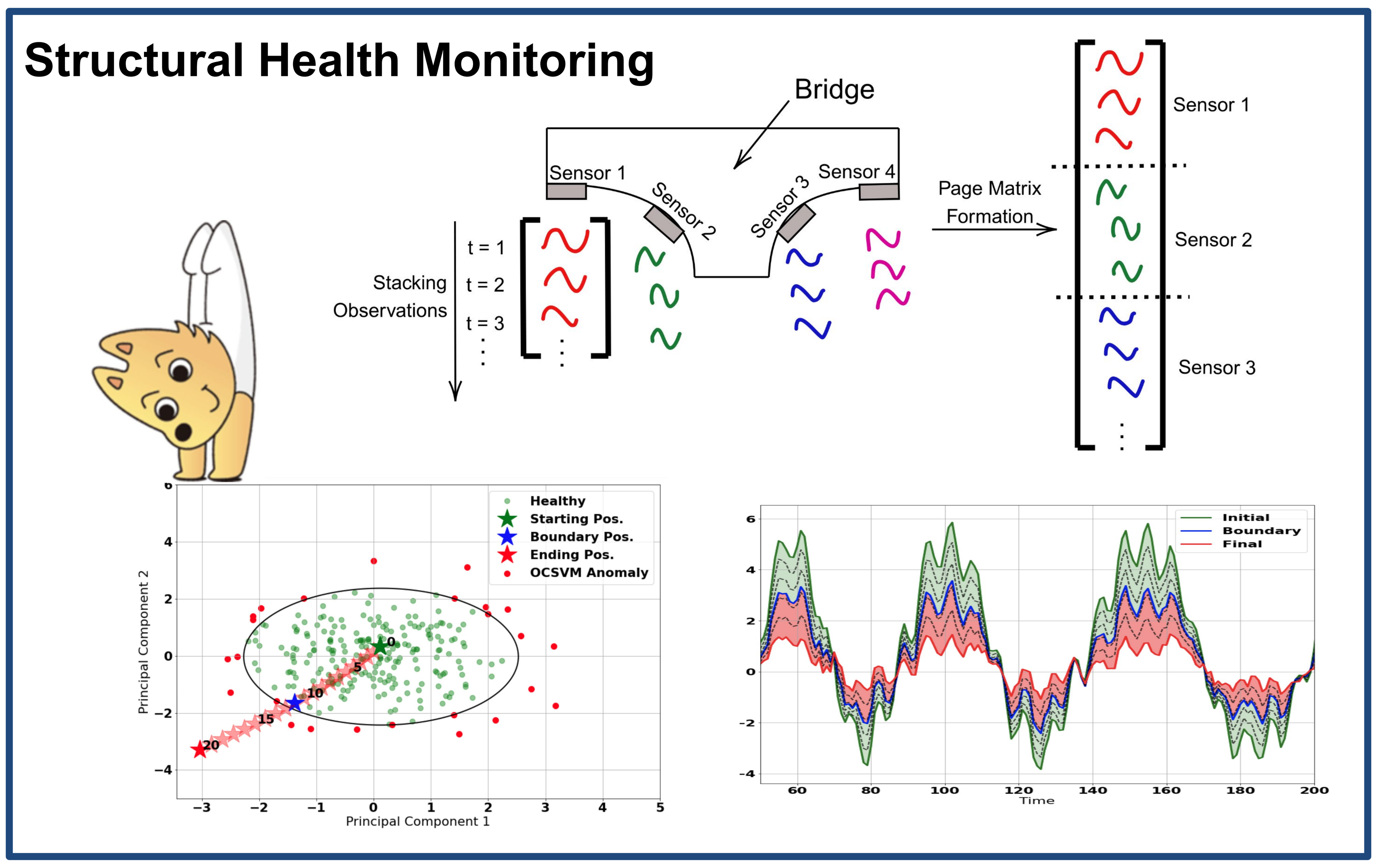


$$\mathcal{T}_{\text{mat}} = \begin{bmatrix} t_1 & \cdots & t_m \\ t_{m+1} & & t_{2m} \\ \vdots & \ddots & \vdots \\ t_{m(n-1)+1} & \cdots & t_N \end{bmatrix}$$



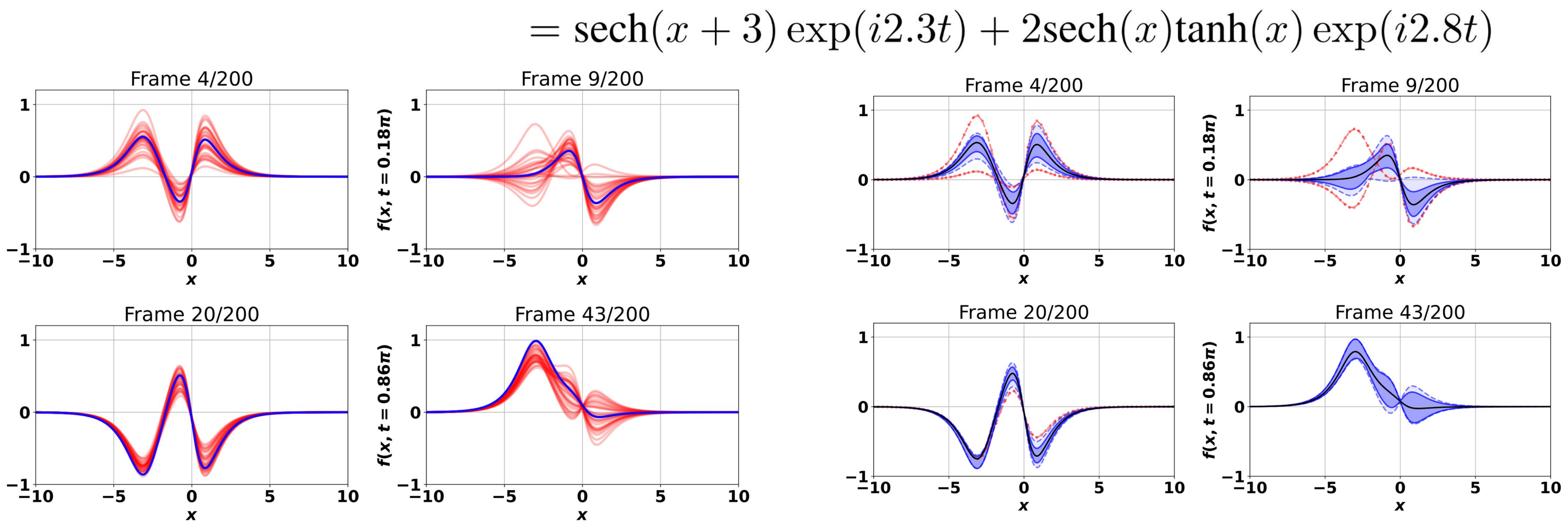
結果

Structural Health Monitoring



Spatiotemporal Uncertainty Quantification

$$f(x,t) = f_1(x,t) + f_2(x,t)$$
$$= \operatorname{sech}(x+3)\exp(i2.3t) + 2\operatorname{sech}(x)\tanh(x)\exp(i2.8t)$$



連絡先：国立情報学研究所 情報プリンシプル研究系 情報ビット
URL：https://www.nii.ac.jp/about/pr/jouhouken/ Email：{prasad, mahito}@nii.ac.jp