Mini-Project: ML for Time Series

Graph-Based Time Series Forecasting

Soël Megdoud
Julien Delavande

MVA 2024/2025

Based on: *Learning Graphs from Data: A Signal Representation Perspective*
Xiaowen Dong*, Dorina Thanou*, Michael Rabbat, and Pascal Frossard

Introduction

- Goal: Learn meaningful graph representations and forecast time series.
- Data: COVID-19 weekly case counts from multiple countries.
- Models:
  - AR+Diffusion Model.
  - Structural Vector Autoregression (SVAR).

AR+Diffusion Model

$$X_i(t) = \sum_{k=1}^{p} \alpha_k X_i(t-k) + \beta \left[\exp(-\tau \mathbf{L}) \mathbf{X}(t-1) \right]_i,$$
$$\mathbf{X}(t) \in \mathbb{R}^N, \quad i = 1, \ldots, N.$$

- $\exp(-\tau \mathbf{L})$: Diffusion kernel over the graph.
- Loss function:

$$\text{Loss} = \sqrt{\sum_{t=p}^{K} \|\mathbf{X}(t) - \hat{\mathbf{X}}(t)\|^2 + \lambda \|\mathbf{L}\|}.$$

SVAR Model

$$\mathbf{x}_t \approx \sum_{m=1}^{M} \left(\sum_{j=0}^{m} c_{m,j} \, \mathbf{A}^j \right) \mathbf{x}_{t-m}.$$

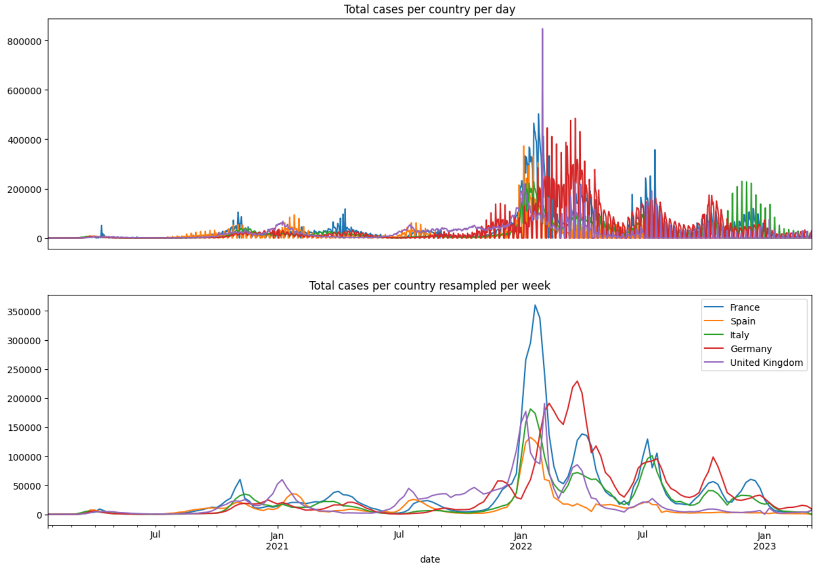
- Polynomial expansion of adjacency matrix $\mathbf{A}$.
- Loss function:

$$\text{Loss} = \frac{1}{2} \sum_{t=M+1}^{K} \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|^2 + \text{Reg}_{\mathbf{A}}(\lambda_1) + \text{Reg}_{\mathbf{c}}(\lambda_2).$$

Data Preparation

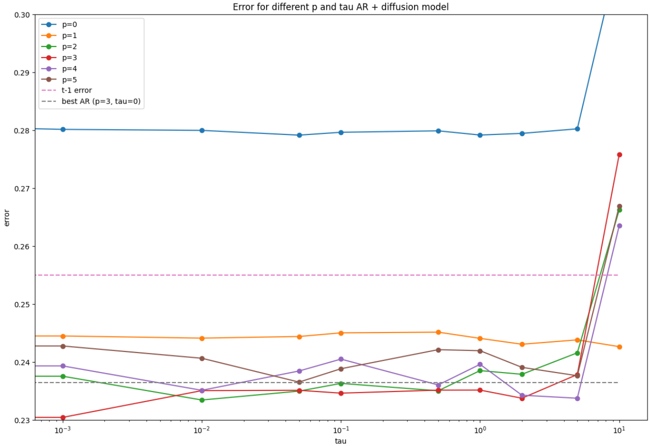
- Dataset: COVID-19 case counts from Our World in Data.
- Processing:
  - Aggregated daily data to weekly.
  - Imputed missing values with zeros.
- Scaling: Used `MinMaxScaler` for training and testing splits.
- Evaluation: L2 loss between true vector of case of one week and predicted one.

Data Preparation

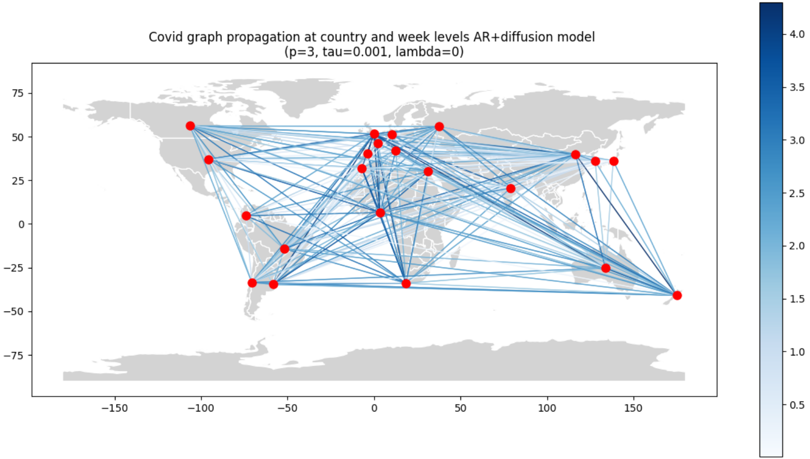


Results: AR+Diffusion Model

- Tested parameters:
  - $p \in \{1, \ldots, 5\}$.
  - $\tau \in \{0, 0.01, 0.1, 0.5, 1, 2, 5, 10\}$.
  - $\lambda_{\text{reg}} \in \{0, 0.01, 0.1, 1, 10\}$.
- Best configuration: $p = 3$, $\tau = 10^{-3}$.
- Observation: Inter-country interactions improve forecasting.



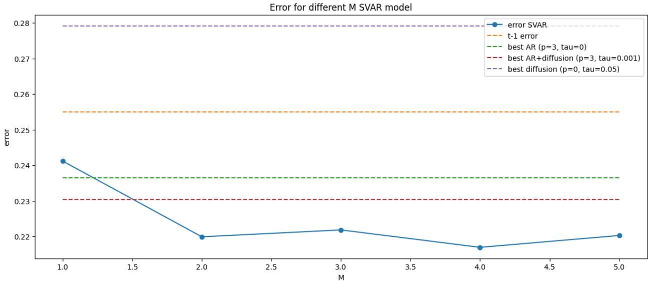
Graph Visualization: AR+Diffusion



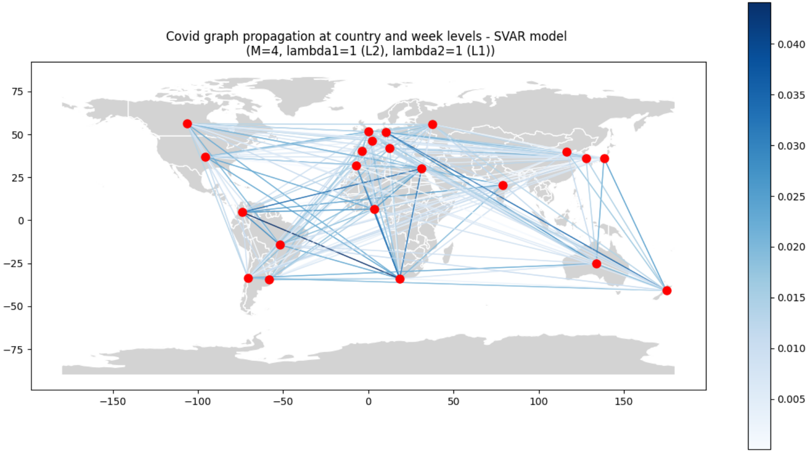
AR+Diffusion Graph

Results: SVAR Model

- Tested parameters:
  - $M \in \{1, \ldots, 5\}$.
  - $\lambda_1, \lambda_2 \in \{0, 0.01, 0.1, 1, 10\}$.
- Best configuration: $M = 4$.
- Sparse graph captures key relationships.



Graph Visualization: SVAR



SVAR Graph

Conclusion

- Best model: SVAR with $M = 4$.
- Graph-based models outperform baselines.
- Insights:
  - Learned graphs reveal meaningful relationships.
  - Models capture temporal dynamics effectively.
- Future Work:
  - Dynamic graphs for evolving relationships.
  - Incorporate additional data (e.g., mobility, vaccination).