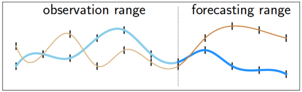
Physiome-ODE: A Benchmark for Irregularly Sampled Multivariate Time Series Forecasting Based on Biological ODEs

Christian Klötergens[1], Vijaya Krishna Yalavarthi[1], Randolf Scholz[1], Maximilian Stubbemann[1], Stefan Born[2], Lars Schmidt-Thieme[1]

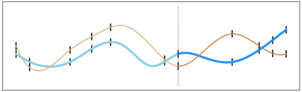
[1]University of Hildesheim, Germany; [2]TU Berlin, Germany

Introduction

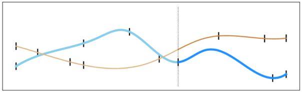
- Most ML-based time series forecasting literature focuses on regular MTS
- However, IMTS forecasting is an important problem and emerging area of research
- Nevertheless, there is no benchmark for IMTS forecasting



(a) Forecasting regular multivariate time series (MTS)



(b) Forecasting irregular multivariate time series



(c) Forecasting irregularly sampled time series with missing values (IMTS)

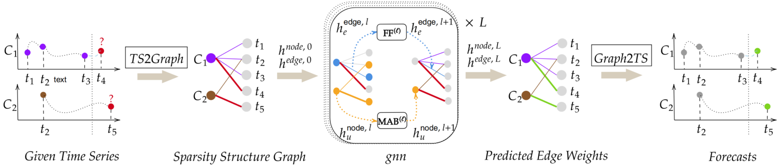
[Yalavarthi et al., 2024]

Neural ODEs:

- Neural ODEs model continuous dynamics using neural networks:
  $$\frac{d\mathbf{h}(t)}{dt} = f_\theta(\mathbf{h}(t), t),$$
  where $\mathbf{h}(t)$ is the hidden state and f_θ is a neural network parameterized by θ.
- Members of the model family include:
  - GRU-ODE-Bayes [De Brouwer et al., 2019]
  - Neural Flows [Biloš et al., 2021]
  - Continuous Recurrent Units [Schirmer et al., 2022]
  - LinODEnet [Scholz et al., 2022]

GraFITi: Graphs for Forecasting Irregularly Sampled Time Series

GraFITi:



[Yalavarthi et al., 2024]

- GraFITi [Yalavarthi et al., 2024] is a graph-based approach for forecasting irregularly sampled time series
- outperforms Neural ODEs on established evaluation datasets

Time-Constant Approaches Outperform ODE-based Approaches in the Current Evaluation Scenario

Test MSE for forecasting next 50% after 50% observation time. OOM refers to out of memory. We highlight the best model in **bold** and <u>underline</u> the second best. † indicates that we show the results from [Yalavarthi et al., 2024]

Model	USHCN	PhysioNet-2012	MIMIC-III	MIMIC-IV
GRU-ODE	1.017±0.325	0.653±0.023†	0.653±0.023†	0.439±0.003†
LinODEnet	<u>0.662±0.126</u>	0.411±0.001†	<u>0.531 ± 0.022</u>†	0.336±0.002†
CRU	0.730±0.264	0.467±0.002†	0.619±0.028†	OOM†
Neural Flow	1.014±0.336	0.506±0.002†	0.651±0.017†	0.465±0.003†
GraFITi	**0.636±0.161**	**0.401±0.001**†	**0.491 ± 0.014**†	**0.285±0.002**†
GraFITi-C	0.875±0.204	<u>0.407±0.001</u>	0.543±0.024	<u>0.324±0.002</u>

- GraFITi-C is a modification of GraFITi that is restricted to forecast a constant value across the complete forecasting horizon
- Despite its severe limitation it outperforms Neural ODE-based models

Physiome-ODE

- Are the currently established evaluation datasets well-suited?
- Physiome Model Repository (**PMR**) contains a collection **200+** of Biological ODE systems
- We create: **Physiome-ODE** - a benchmark for IMTS forecasting created with ODE systems contained PMR

How can we IMTS forecasting datasets from ODE models?

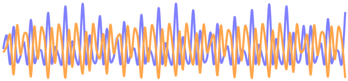
- For each ODE system we randomly vary:
  - initial conditions
  - parameters
- Solve the ODE for a pre-defined time interval T
- Select a time point from $[0, \frac{T}{2}]$ as start point t_0 and use only keep points from $[t_0, t_0 + \frac{T}{2}]$
- Randomly sample observation to create an IMTS
- Add Gaussian noise
- The forecasting objective is to forecast the second half based on the first half

How we created the Physiome ODE benchmark

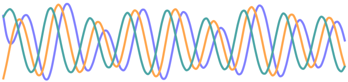
- Which of the 200+ ODE systems should we use ?
- In which range should we change constants and initial conditions?
- What would be a good time horizon T ?

$\rightarrow$ We develope a proxy-metric for the difficulty of the forecasting task and optimize that (**JGD**)

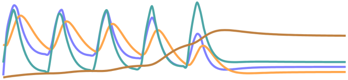
- Joint Gradient Deviation is the product of:
  - Mean point-wise gradient deviation: average variance of gradients in time series instances at each time point
  - Mean gradient deviation: average variance of gradients in each time series instance



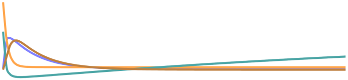
DUP01 — JGD: 2.697



LEN01 — JGD: 0.178



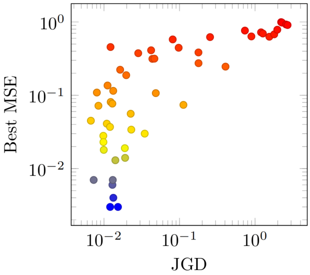
GUP01 — JGD: 0.014



NEL01 — JGD: 0.007

Results

Metric	# Wins	Rank
GRU-ODE	0	5.9
LinODEnet	**25**	**2.10**
CRU	8	2.82
Neural Flows	2	4.82
GraFITi	10	<u>2.20</u>
GraFITi-C	<u>15</u>	2.86



- Overall, LinODEnet is the best-performing model, winning the most datasets and having the highest average rank.
- The JGD appears to be a functioning proxy for MSE.

Thank You for Your Attention!

Check out the repository on GitHub!

https://github.com/kloetergensc/Physiome-ODE

References

M. Biloš, J. Sommer, S. S. Rangapuram, T. Januschowski, and S. Günnemann. Neural Flows: Efficient Alternative to Neural ODEs. In *Advances in Neural Information Processing Systems*, volume 34, pages 21325–21337. Curran Associates, Inc., 2021.

E. De Brouwer, J. Simm, A. Arany, and Y. Moreau. GRU-ODE-Bayes: Continuous Modeling of Sporadically-Observed Time Series. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

M. Schirmer, M. Eltayeb, S. Lessmann, and M. Rudolph. Modeling Irregular Time Series with Continuous Recurrent Units. In *Proceedings of the 39th International Conference on Machine Learning*, pages 19388–19405. PMLR, June 2022.

R. Scholz, S. Born, N. Duong-Trung, M. N. Cruz-Bournazou, and L. Schmidt-Thieme. Latent Linear ODEs with Neural Kalman Filtering for Irregular Time Series Forecasting. Sept. 2022.

V. K. Yalavarthi, K. Madhusudhanan, R. Scholz, N. Ahmed, J. Burchert, S. Jawed, S. Born, and L. Schmidt-Thieme. GraFITi: Graphs for Forecasting Irregularly Sampled Time Series. In M. J. Wooldridge, J. G. Dy, and S. Natarajan, editors, *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024, Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence, IAAI 2024, Fourteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2014, February 20-27, 2024, Vancouver, Canada*, pages 16255–16263. AAAI Press, 2024. doi: 10.1609/AAAI.V38I15.29560.