Loss-based directed graph constraints for recurrent flow network dynamics

UNIVERSITY OF BATH

Konstantina Theodosiadou[1], Thomas Rodding Kjeldsen[2], Georgios Exarchakis[1]
{[1]Department of Computer Science, Faculty of Science, [2] Department of Architecture & Civil Engineering, Faculty of Engineering}, University of Bath, UK

1. Introduction

Task

Accurate **flow network forecasting** requires capturing both temporal dynamics and spatial connectivity.

Problems

i. Recurrent Neural Networks model non-linear temporal dependencies, but **not network topology.**

ii. Spatio-temporal models embed **graph representation** to temporal learning, leading to **high computational complexity.**

iii. Standard neural networks minimize data variance at the expense of physics, resulting to solutions that violate the network's mass conservation.

Solution

We propose **a loss regularization** that can be described as the discrete form of mass conservation encoding the elements of a flow network: spatial proximity & directionality of graph.

- We show that it improves predictive performance while remaining computationally strong in large-scale networks.

2. Methodology

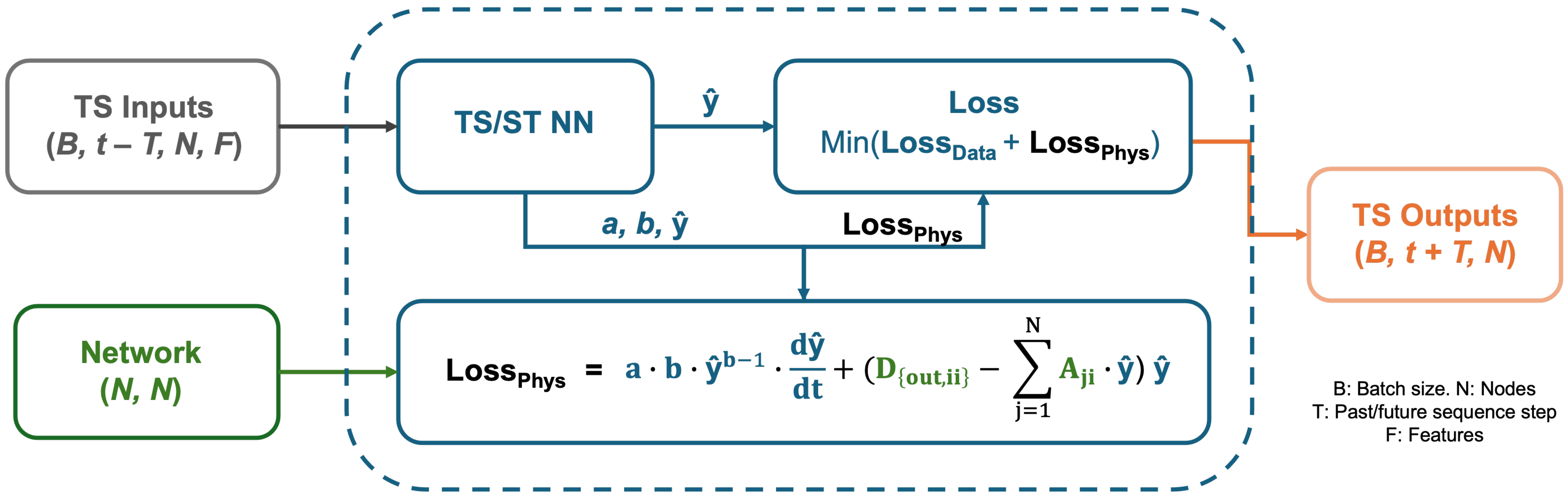

Fig. 1: Model architecture.

Proposed physics prior

Based on Mass-Conservation Error:

$\mathrm{MCE} = \frac{dS_i}{dt} - (I_i(t) - Q_i(t))$ (1) , now using:

i. Power-law storage: $S_i = a_i Q_i^{b_i} \Rightarrow \frac{dS_i}{dt} = a_i b_i Q_i^{b_i - 1} \frac{dQ_i}{dt}$ (2)

- a_i, b_i: Learnable storage constants per node

ii. Directed Laplacian: $I_i - Q_i = -(L_{dir}Q)_i = \sum_{j=1}^{N} A_{ji} Q_j - (D_{out})_{ii} Q_i$ (3)

(1) $\xrightarrow{(2),(3)}$ $\mathrm{MCE} = a_i b_i Q_i^{b_i - 1} \frac{dQ_i}{dt} + (L_{dir}Q)_i$ NN loss: $L = L_{data}(Q, \hat{Q}) + \lambda \left| ab\hat{Q}^{b-1} \frac{d}{dt} + L_{dir}\hat{Q} \right|_2^2$

Computational cost

MCE term requires a sparse matrix-vector multiplication with time and memory complexity $\mathcal{O}(\mathcal{E})$ and $\mathcal{O}(\mathcal{E} + N)$ where N is the number of nodes. Gradient computation for $\hat{Q}$ leads to the same asymptotic cost.

Basic setup & evaluation metrics

The hyper-parameters of the baselines & λ (regularization constant) per model have been tuned. The eval. metrics access magnitude (MAE), high-peak magnitude and timing ($MAE_{0.95}$,PTB) & mass-conservation (MCE):

$\mathrm{MAE} = \frac{1}{N}\sum_{t=1}^{N} |y_t - \hat{y}_t|$, $\mathrm{MAE}_{0.95} = \frac{1}{N_H}\sum_{t: y_t > \zeta_{95}} |y_t - \hat{y}_t|$, $\mathrm{PTB} = |\frac{1}{P}\sum_{p=1}^{P}(\hat{t}_p - t_p)$, $\mathrm{MCE} = \frac{I_i(t) - Q_i(t) - \frac{dS_i}{dt}}{I(t)}$

3. Results

Real-world application: River flow forecasting

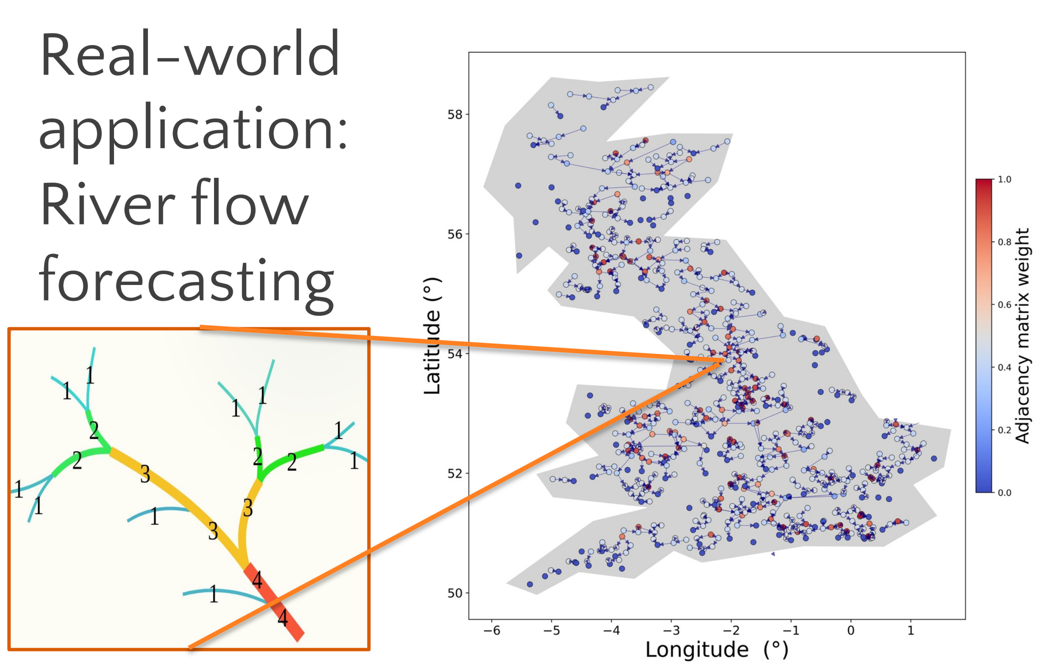

Fig. 2: River network in GB.

Table 1: Metrics per model.

Model	MAE	$MAE_{0.95}$	PTB	MCE
MC-LSTM	4.04 ± 0.06	31.08 ± 0.39	-1.27 ± 0.34	1.498 ± 0.004
GRU	4.12 ± 0.07	32.06 ± 0.68	-1.11 ± 0.21	1.494 ± 0.005
GRU + Reg.	**4.00 ± 0.04**	**31.48 ± 0.53**	**-1.04 ± 0.13**	**0.862 ± 0.000**
LSTM	4.11 ± 0.04	32.66 ± 0.53	-1.58 ± 0.22	1.494 ± 0.004
LSTM + Reg.	**3.99** ± 0.04	**31.64 ± 0.48**	**-1.12 ± 0.14**	**0.863 ± 0.000**
MPNN-LSTM	4.08 ± 0.02	31.95 ± 0.23	-1.53 ± 0.09	1.495 ± 0.003
MPNN-LSTM + Reg.	**3.97** ± 0.09	**30.97** ± 0.69	**-1.04** ± 0.22	**0.862 ± 0.000**

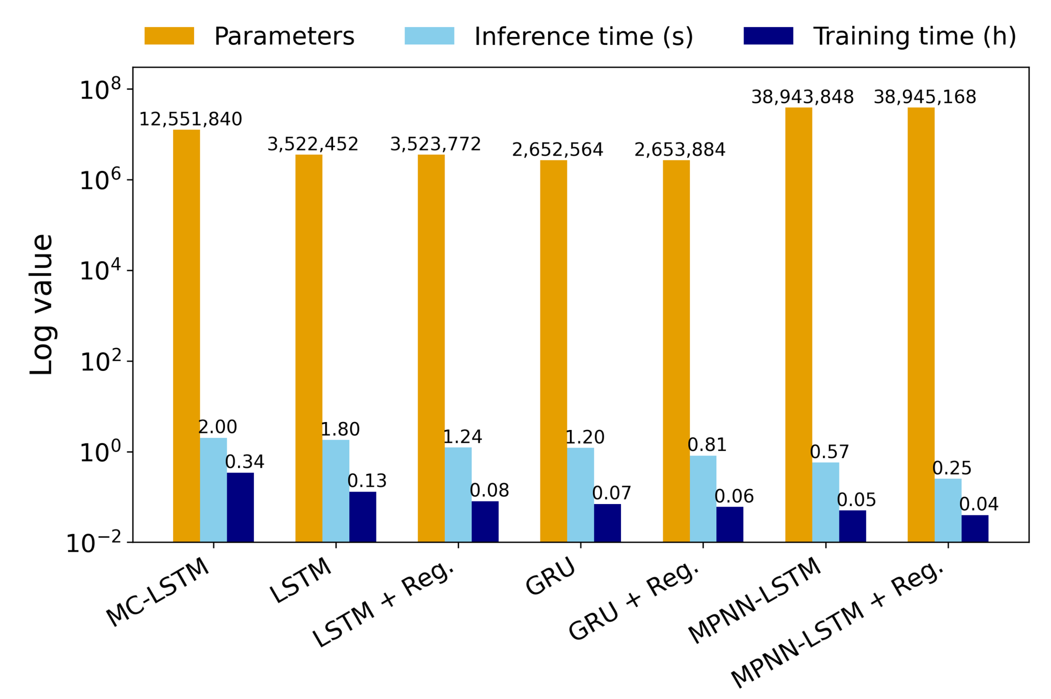

Fig. 3: Training/inference times & num. of params per model.

i. PTB in regularized MPNN-LSTM/GRU is ↓ by ~**13**% compared to MC-LSTM.

ii. The regularization leads to **2.84%** reduction in overall MAE across the respective plain architectures, with high-flow error ($MAE_{0.95}$) decreased up to **3.12%.**

iii. It drops the MCE down to **3.5**× reduction in residual mass imbalance.

iv. The regularized loss in GRU ↓ inference & training time by ~**32**% & ~**14**% .

v. The reg. models take less training time than MC-LSTM by **4.25**× more.

4. Conclusion

- The **results** indicate that the proposed regularization helps the NN models converge faster to a less violating mass-conserving solution with modest accuracy improvements, given the regularization constant is tuned.
- The **limitations** of the proposed regularization are **(i)** its sensitivity to hyper-parameter optimization **(ii)** mass-conservation may still be violated based on other exogenous drivers.
- **Future work** targets: **(i)** the extension of this work to other flow networks such as traffic flows **(ii)** using a mass-conserving architecture for this type of networks as an inductive bias.

Acknowledgements: This research is funded by a University Research Studentship Award through UKRI (UK Research and Innovation).