SWOT-based Simulation of River Discharge with Temporal Graph Neural Networks

Kevin Osanlou[1] **Augusto Getirana**[2,3] **Thomas Holmes**[3] **Tristan Cazenave**[4]
[1] Talan Research & Innovation Center [2] Science Applications International Corporation
[3] NASA GSFC [4] LAMSADE, Université Paris Dauphine - PSL
kevin.osanlou@talan.com
{thomas.holmes,augusto.getirana}@nasa.gov
tristan.cazenave@lamsade.dauphine.fr

Abstract

The Surface Water and Ocean Topography (SWOT) satellite launched in December 2022 and has been providing high-accuracy monitoring of rivers' water surface elevation. Remote measurements do not, however, include discharge, an essential element for improved modelling of river dynamics. Furthermore, orbital characteristics of SWOT satellite overpasses result in spatial and temporal discontinuity of observations. The process of inferring hydrological information from one river portion to another is called regionalization and may help alleviate this issue. While most regionalization methods based on machine learning do not leverage geometrical river properties, recent works point to graph machine learning as a promising solution for simulation of river flow. This work applies state-of-the-art temporal graph machine learning techniques for regionalization and discharge estimation. More specifically, we leverage available SWOT data, in-situ gauge discharge and a temporal graph neural network to offer basin-scale simulation of river discharge in US rivers where the SWOT mission operates. We run experiments in different basins and against several in-situ gauges on which the proposed method offers convincing performance. The approach is compared to the drainage-area ratio method, a non-temporal graph neural network and a long short-term memory model, which it outperforms.

1 Introduction and Related Works

The Surface Water and Ocean Topography (SWOT) satellite monitors rivers wider than 100 m. It is especially useful in poorly gauged regions where it may enable better management of flood risks [Jongman et al., 2012], improvement of hydrological models [Sagarika et al., 2016] and better allocation of water resources [Kansakar and Hossain, 2016]. Each river reach monitored by SWOT is 10 km-long on average. When SWOT measurements are available for a reach on a given day, the reach is assigned a SWOT product comprising daily average Water Surface Elevation (WSE), width and slope [Fu et al., 2012]. While discharge is a critical part of the SWOT mission, it is not remotely measured. Moreover, SWOT follows an orbit with a 21-day repeat within ±78 degrees latitude and shorter repeat cycles near poles [Biancamaria et al., 2016]. Some river reaches are observed more often than others: observations are subject to spatial and temporal discontinuity, which is a challenge for the estimation of weekly or monthly hydrodynamic attributes [Biancamaria et al., 2009].

Inferring discharge from SWOT observations can be related to finding reach-scale flow law parameters in Gauckler–Manning's equation relying on observations over a time period. Significant reach-scale efforts have been devoted to this purpose [Durand et al., 2016, Gleason and Smith, 2014, Hagemann et al., 2017]. However, directly deploying these reach-scale algorithms at global scale is

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

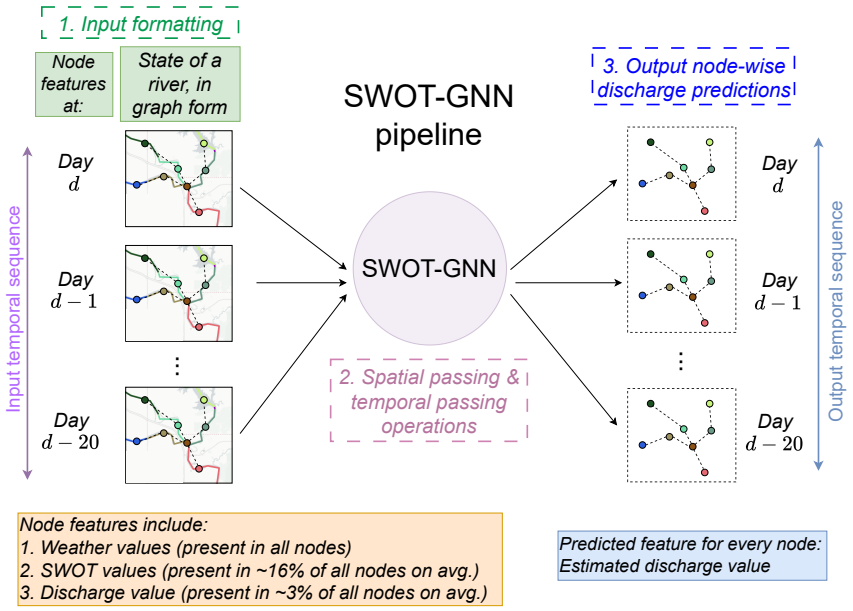

Figure 1: **Illustration of SWOT-GNN discharge simulation.**

computationally infeasible. Integrators, which enforce flow conservation at river confluences, are currently under development to enable these reach-scale algorithms at basin scale in the context of SWOT [Durand et al., 2023]. Other algorithms that infer flow law parameters on a higher-scale are also planned for use, such as geoBAM [Brinkerhoff et al., 2020] and MetroMan [Durand et al., 2014].

Concerning regionalization, a wide variety of methods exist and we refer the reader to [Guo et al., 2021] for a deep review. The drainage-area ratio method is one of the oldest methods [Archfield and Vogel, 2010]. It considers the drainage area of an ungauged location, the drainage area of a gauged location and uses a ratio of the two to obtain the ungauged discharge. Regionalization can also be carried out via Data Assimilation (DA). Observations are combined with a hydrological model which aims to reproduce river physics and dynamics [Reichle, 2008]. An estimated discharge is computed from the SWOT observation for the reach in question, which the model regionalizes to neighboring reaches based on flow rules [Li et al., 2020]. While DA may carry out effective regionalization, a hydrological model is complex to create and computationally expensive, especially on SWOT river scale. Fully data-driven methods have also gained in popularity owing to their cheap computational requirements and increasingly high accuracy. Various models have been used, such as artificial neural networks [Besaw et al., 2010], support vector machines [Hwang et al., 2012], and Long Short-Term Memories (LSTMs) [Kratzert et al., 2019]. Authors in [Kratzert et al., 2019] report their rainfall-discharge LSTM to be more accurate on multiple basins than hydrological models calibrated to each basin. Recent works now leverage temporal Graph Neural Networks (GNN) as a follow-up to LSTMs, due to their ability to incorporate geometrical basin properties on top of temporal dependencies. GNNs are generalizations of convolutional neural networks to non-Euclidean graphs [Bronstein et al., 2017] and follow the idea of passing messages between graph nodes. We refer the reader to [Zhou et al., 2020] for a deep review of GNNs. Relevant works include Zhao et al., who combine a RNN with a GNN to predict discharge and outperform learning-based baselines [Zhao et al., 2020]. Sun et al. compare several recurrent GNN architectures and report almost all architectures to outperform the previous state-of-the-art LSTM model [Sun et al., 2021]. They also introduce a recurrent GNN trained with data from a hydrological model which is reportedly highly accurate for reproducing the dynamics of a basin of the Colorado River [Sun et al., 2022].

In this work, we provide a computationally inexpensive alternative to heavy flow-law modelling methods to estimate discharge under SWOT data discontinuity. Our main contributions are: **(1)** We design a pipeline to gather daily SWOT river basins with hydrological properties which we store as temporal sequences of graphs. **(2)** We design a temporal GNN, SWOT-GNN, tailored for processing these temporal sequences and train it to simulate daily reach-wise discharge. **(3)** We validate the simulation performance of SWOT-GNN on United States Geological Survey (USGS) gauges.

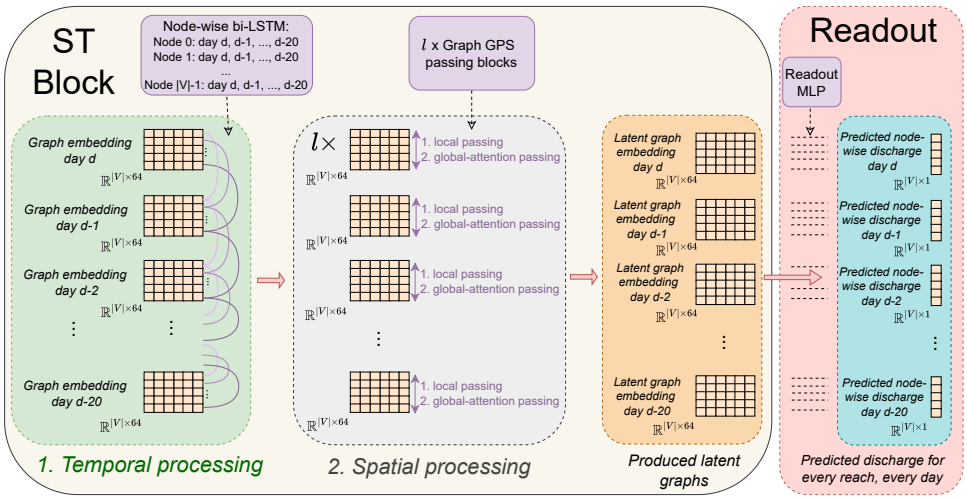

Figure 2: **Illustration of the SWOT-GNN architecture.**

2 Data Gathering and Temporal Graphs

The SWOT River Database (SWORD) serves as foundation for SWOT river products [Altenau et al., 2021]. Graphs built in this work follow SWORD structure. We format river basins into graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Graph nodes $v \in \mathcal{V}$ represent SWOT reaches, graph edges $e \in \mathcal{E}$ connections between adjacent reaches. Note we refer to reaches and nodes interchangeably.

Definitions. We refer to **the state of a river basin on a given day** d as a graph $\mathcal{G}$ that contains, for each reach v, the following node features. **(i)** A SWOT component: a binary variable to encode SWOT data availability on day d for reach v, followed by SWOT's measured average WSE, width and slope (zeroes if unavailable). **(ii)** A discharge component: a binary variable to encode USGS gauge discharge availability on day d for reach v, followed by the discharge value (zero if unavailable). **(iii)** A Climate component: average temperature, evapotranspiration and cumulative precipitation on day d for reach v. **(iv)** A static component independent of the day: a list of prior variables for v from the SWORD database. These include: prior WSE, width, distance to outlet, reach length, number of channels flowing into / out of v. Finally, we denote as a **temporal sequence** a sequence made of the states of the same river basin over 21 consecutive days, *i.e.* on day d, $d-1$, ..., $d-20$. A temporal sequence constitutes an input to the SWOT-GNN architecture. Note that while the river basin has to be the same in a temporal sequence to be temporally processable, river basins in different temporal sequences do not need to be the same and can all be processed by the same SWOT-GNN model.

Dataset creation. We consider 6 sub-areas from SWORD, each corresponding to a purely geometrical structure (*i.e.* a graph without the hydrological node variables present in state-of-river graphs). These areas are created from SWORD basins 78, 74 and 73 using bounding boxes and each totals $\sim 30,000 - 60,000$ sq. miles. We consider a study period from May 2023 to June 2024. We gather, for every reach on every day, SWOT variables (if available), USGS gauge discharge (if available) and climate variables. SWOT variables are obtained using the SWOT level 2 river single-pass products, discharge using USGS's WaterService daily value API and searching for USGS gauges on the corresponding SWORD reach, climate variables using Open-Meteo's historical weather API [Zippenfenig, 2023]. We create state-of-river graphs on every day of the study timespan for all 6 areas. For each state-of-river graph, $\sim 6\%$ of all nodes have a USGS discharge value available on average. For half of these, we mask their discharge value (*i.e.* remove their discharge value from node features and declare discharge unavailable). The masked values of these nodes become labels instead, which SWOT-GNN learns to predict using spatial and temporal connections.

3 SWOT-GNN Architecture and Experiments

SWOT-GNN takes as input a temporal sequence and outputs a temporal sequence for which every reach has a discharge prediction (Fig. 1). The model is trained on nodes for which discharge is available as label and masked from node features, by minimizing the square difference between the

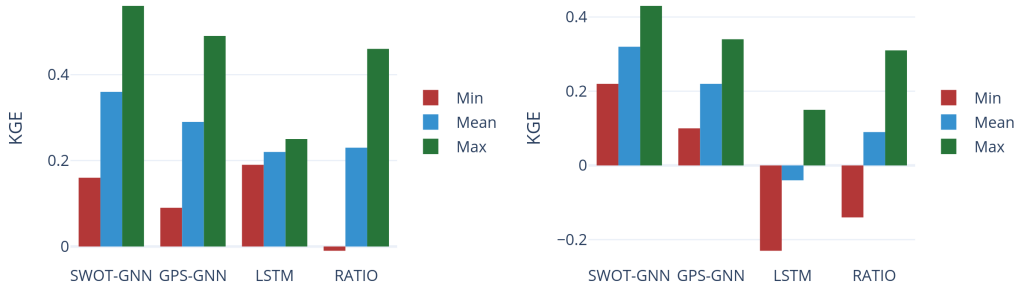

Figure 3: **Discharge KGE simulation scores on test basin 1 & 2**

label value and predicted discharge. When processing a temporal sequence, SWOT-GNN first applies a Multi-Layer Perceptron (MLP) to every node. Each state-of-river graph in the temporal sequence becomes a graph embedding (*i.e.* a graph in which nodes have 64-feature embeddings computed from the initial hydrological features). Next, multiple ST-blocks are applied. An ST-block is comprised of **(i)** a temporal processing phase, in which a two-layer bi-directional LSTM is applied node-wise (*i.e.* aggregates features only temporally for each node, through the entire temporal sequence), and **(ii)** a spatial processing phase in which graph transformers are used. This is done by applying l GraphGPS layers on node features [Rampášek et al., 2022]. Each GraphGPS layer is a graph transformer mainly consisting of a local message passing phase where node features are aggregated based on direct graph neighbors, followed by a global-attention message passing phase where every node receives information from any other node in the graph. Finally, the last ST-block is followed by a readout layer, which applies a readout MLP to every node of every graph, mapping their features to a single value, the predicted discharge. Figure 2 depicts the ST-block and readout layer. Architecture experiments lead us to retain a SWOT-GNN architecture with 2 ST-blocks, each with 3 GraphGPS layers. We conduct performance experiments of the trained SWOT-GNN model across different test basin graphs from SWORD on USGS gauges (Figure 3). The models we compare are: **(i)** SWOT-GNN. **(ii)** GPS-GNN: a GNN made of 6 GraphGPS layers, applied to state-of-river graphs. **(iii)** LSTM: a 4-layer bi-directional LSTM, applied temporally to reach features. **(iv)** Drainage-area ratio method using neighbors with non-masked discharge. GPS-GNN and LSTM are specifically designed as independent baselines for experiments and do not provide ablation insights into SWOT-GNN. Results are single-run. The metric used is the Kling–Gupta Efficiency (KGE) [Gupta et al., 2009]. KGE values are comprised in $[-\infty, 1]$ (higher is better). We provide hyperparameter details in the appendix. Regardless of the model, we note performance varies depending on the tested gauge, as shown with varying min/max KGE scores on all models. No model provides accurate discharge for a few gauges with the worst scores, which tend to be those with the least SWOT coverage. Overall, we note SWOT-GNN outperforms all baselines. GPS-GNN offers reasonable performance and outperforms the LSTM and drainage-area ratio method. The low performance of the LSTM may be attributed in part to its sole reliance on climate variables and the temporally-sparse SWOT variables which are not used spatially.

4 Conclusion

We introduced SWOT-GNN, a temporal graph neural network that estimates river discharge in SWOT reaches based on spatial and temporal data from the SWOT mission, weather variables and in-situ USGS gauges. SWOT-GNN is useful in that it is computationally inexpensive compared to hydrological models and tackles the problems of SWOT data discontinuity and absence of discharge measurements. We also provided a pipeline to gather data required for training SWOT-GNN. We carried out experiments on in-situ gauges in different basins to validate the performance of the model and demonstrated it outperforms the drainage-area ratio method, a non-temporal GNN and an LSTM. Future work may include incorporating additional hydrological parameters into SWOT-GNN to improve discharge modelling in areas with sparser SWOT coverage, as well as featuring upcoming SWOT reach-scale discharge estimations into the model for further fine-tuning.

References

Elizabeth H Altenau, Tamlin M Pavelsky, Michael T Durand, Xiao Yang, Renato Prata de Moraes Frasson, and Liam Bendezu. The surface water and ocean topography (swot) mission river database (sword): A global river network for satellite data products. *Water Resources Research*, 57(7), 2021.

Stacey A Archfield and Richard M Vogel. Map correlation method: Selection of a reference streamgage to estimate daily streamflow at ungaged catchments. *Water resources research*, 46(10), 2010.

Lance E Besaw, Donna M Rizzo, Paul R Bierman, and William R Hackett. Advances in ungauged streamflow prediction using artificial neural networks. *Journal of Hydrology*, 386(1-4):27–37, 2010.

Sylvain Biancamaria, Kostas M Andreadis, Michael Durand, Elizabeth A Clark, Ernesto Rodriguez, Nelly M Mognard, Doug E Alsdorf, Dennis P Lettenmaier, and Yannick Oudin. Preliminary characterization of swot hydrology error budget and global capabilities. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3(1):6–19, 2009.

Sylvain Biancamaria, Dennis P Lettenmaier, and Tamlin M Pavelsky. The swot mission and its capabilities for land hydrology. In *Remote sensing and water resources*, pages 117–147. Springer, 2016.

CB Brinkerhoff, CJ Gleason, D Feng, and P Lin. Constraining remote river discharge estimation using reach-scale geomorphology. *Water Resources Research*, 56(11):e2020WR027949, 2020.

Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

Krzysztof Marcin Choromanski, Valerii Likhosherstov, David Dohan, Xingyou Song, Andreea Gane, Tamas Sarlos, Peter Hawkins, Jared Quincy Davis, Afroz Mohiuddin, Lukasz Kaiser, et al. Rethinking attention with performers. In *International Conference on Learning Representations*, 2020.

Michael Durand, Jeffrey Neal, Ernesto Rodríguez, Konstantinos M Andreadis, Laurence C Smith, and Yeosang Yoon. Estimating reach-averaged discharge for the river severn from measurements of river water surface elevation and slope. *Journal of Hydrology*, 511:92–104, 2014.

Michael Durand, CJ Gleason, Pierre-André Garambois, David Bjerklie, LC Smith, Hélène Roux, Elizandro Rodriguez, Paul D Bates, Tamlin M Pavelsky, Jerome Monnier, et al. An intercomparison of remote sensing river discharge estimation algorithms from measurements of river height, width, and slope. *Water Resources Research*, 52(6):4527–4549, 2016.

Michael Durand, Colin J Gleason, Tamlin M Pavelsky, Renato Prata de Moraes Frasson, Michael Turmon, Cédric H David, Elizabeth H Altenau, Nikki Tebaldi, Kevin Larnier, Jerome Monnier, et al. A framework for estimating global river discharge from the surface water and ocean topography satellite mission. *Water Resources Research*, 59(4):e2021WR031614, 2023.

Lee-Lueng Fu, Douglas Alsdorf, Rosemary Morrow, Ernesto Rodriguez, and Nelly Mognard. Swot: the surface water and ocean topography mission: wide-swath altimetric elevation on earth. Technical report, Pasadena, CA: Jet Propulsion Laboratory, National Aeronautics and Space Administration, 2012.

Colin J Gleason and Laurence C Smith. Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geometry. *Proceedings of the National Academy of Sciences*, 111(13):4788–4791, 2014.

Yuhan Guo, Yongqiang Zhang, Lu Zhang, and Zhonggen Wang. Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A comprehensive review. *Wiley Interdisciplinary Reviews: Water*, 8(1):e1487, 2021.

Hoshin V Gupta, Harald Kling, Koray K Yilmaz, and Guillermo F Martinez. Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling. *Journal of hydrology*, 377(1-2):80–91, 2009.

MW Hagemann, CJ Gleason, and MT Durand. Bam: Bayesian amhg-manning inference of discharge using remotely sensed stream width, slope, and height. *Water Resources Research*, 53(11):9692–9707, 2017.

Seok Hwan Hwang, Dae Heon Ham, and Joong Hoon Kim. Forecasting performance of ls-svm for nonlinear hydrological time series. *KSCE Journal of Civil Engineering*, 16(5):870–882, 2012.

Brenden Jongman, Philip J Ward, and Jeroen CJH Aerts. Global exposure to river and coastal flooding: Long term trends and changes. *Global Environmental Change*, 22(4):823–835, 2012.

Pratistha Kansakar and Faisal Hossain. A review of applications of satellite earth observation data for global societal benefit and stewardship of planet earth. *Space Policy*, 36:46–54, 2016.

Frederik Kratzert, Daniel Klotz, Guy Shalev, Günter Klambauer, Sepp Hochreiter, and Grey Nearing. Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. *Hydrology and Earth System Sciences*, 23(12):5089–5110, 2019.

Dongyue Li, Konstantinos M Andreadis, Steven A Margulis, and Dennis P Lettenmaier. A data assimilation framework for generating space-time continuous daily swot river discharge data products. *Water Resources Research*, 56(6), 2020.

Edward Ott, Brian R Hunt, Istvan Szunyogh, Aleksey V Zimin, Eric J Kostelich, Matteo Corazza, Eugenia Kalnay, DJ Patil, and James A Yorke. A local ensemble kalman filter for atmospheric data assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, 56(5):415–428, 2004.

Ladislav Rampášek, Michael Galkin, Vijay Prakash Dwivedi, Anh Tuan Luu, Guy Wolf, and Dominique Beaini. Recipe for a general, powerful, scalable graph transformer. *Advances in Neural Information Processing Systems*, 35:14501–14515, 2022.

Rolf H Reichle. Data assimilation methods in the earth sciences. *Advances in water resources*, 31 (11):1411–1418, 2008.

Soumya Sagarika, Ajay Kalra, and Sajjad Ahmad. Pacific ocean sst and z500 climate variability and western us seasonal streamflow. *International Journal of Climatology*, 36(3):1515–1533, 2016.

Alexander Y Sun, Peishi Jiang, Maruti K Mudunuru, and Xingyuan Chen. Explore spatio-temporal learning of large sample hydrology using graph neural networks. *Water Resources Research*, 57 (12), 2021.

Alexander Y Sun, Peishi Jiang, Zong-Liang Yang, Yangxinyu Xie, and Xingyuan Chen. A graph neural network (gnn) approach to basin-scale river network learning: the role of physics-based connectivity and data fusion. *Hydrology and Earth System Sciences*, 26(19):5163–5184, 2022.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations*, 2018. URL `https://openreview.net/forum?id=rJXMpikCZ`.

Qun Zhao, Yuelong Zhu, Kai Shu, Dingsheng Wan, Yufeng Yu, Xudong Zhou, and Huan Liu. Joint spatial and temporal modeling for hydrological prediction. *IEEE Access*, 8:78492–78503, 2020.

Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. Graph neural networks: A review of methods and applications. *AI Open*, 1:57–81, 2020. ISSN 2666-6510.

Patrick Zippenfenig. Open-meteo.com weather api, 2023. URL `https://open-meteo.com/`.

Appendix

Network Architecture, Hyperparameters and Input Formatting

We provide here more details about network architecture and hyperparameters. We justify the choice of a global graph transformer such as Graph-GPS by the fact that we have found local message passing graph neural networks to underperform at passing distant messages even with a sufficient number of layers, due to over-smoothing issues. Prior to processing inputs, a positional embedding based on the random walk algorithm and graph Laplacian is concatenated to each graph node, so as to provide positional information for the global attention mechanism present in each Graph-GPS layer. We select SWOT-GNN with 2 ST-Blocks, each containing 3 Graph-GPS layers and a LSTM with 2 layers. Deeper and wider architectures did not yield better results in experiments. Each hidden layer has a latent size of 32 neurons. The first LSTM layer is followed by a dropout layer with a *keep* rate of 0.5. Each graph-GPS layer first applies the GAT local message passing based on [Veličković et al., 2018], followed by the linear-cost FAVOR+ attention mechanism defined in [Choromanski et al., 2020], using 4 attention heads. A dropout layer with a *keep* rate of 0.5 follows. For training purposes, we use a batch size of 16 and a learning rate set to 0.005. Training was performed on a RTX A4000 with 16GB of RAM. We use a Scikit Learn's StandardScaler to normalize both input and output discharge values, while intermediate layers are normalized using LayerNorm. Figure 4 provides insight into how inputs are formatted, and fed into SWOT-GNN. Each data chunk, SWOT-based, weather-based, discharge and prior, is normalized with LayerNorm and fed into a different MLP. The embeddings obtained are concatenated before being normalized again, so as to create a graph embedding which is then fed into ST-blocks.

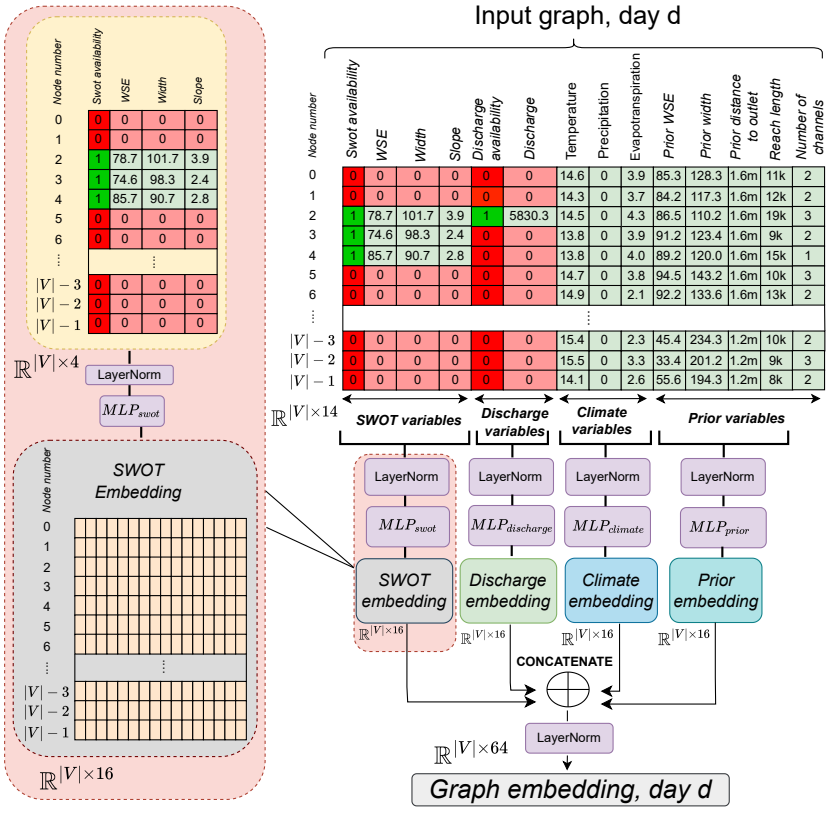

Figure 4: **Processing of inputs to produce a graph embedding.**