# Early Detection of Extreme Storm Tide Events Using Multimodal Data Processing

**Marcel Barros[1], Andressa Pinto[1], Andres Monroy[2], Felipe Moreno[1], Jefferson Coelho[1], Aldomar Pietro Silva[1], Caio Fabricio Deberaldini Netto[1], José Roberto Leite[1], Marlon Mathias[1], Eduardo Tannuri[1], Artur Jordão[1], Edson Gomi[1], Fábio Cozman[1], Marcelo Dottori[1], Anna Helena Reali Costa[1]**

[1]Universidade de São Paulo, São Paulo, SP, Brasil
[2]Massachusetts Institute of Technology, Cambridge, MA, USA
marcel.barros@usp.br

## Abstract

Sea-level rise is a well-known consequence of climate change. Several studies have estimated the social and economic impact of the increase in extreme flooding. An efficient way to mitigate its consequences is the development of a flood alert and prediction system, based on high-resolution numerical models and robust sensing networks. However, current models use various simplifying assumptions that compromise accuracy to ensure solvability within a reasonable timeframe, hindering more regular and cost-effective forecasts for various locations along the shoreline. To address these issues, this work proposes a hybrid model for multimodal data processing that combines physics-based numerical simulations, data obtained from a network of sensors, and satellite images to provide refined wave and sea-surface height forecasts, with real results obtained in a critical location within the Port of Santos (the largest port in Latin America). Our approach exhibits faster convergence than data-driven models while achieving more accurate predictions. Moreover, the model handles irregularly sampled time series and missing data without the need for complex preprocessing mechanisms or data imputation while keeping low computational costs through a combination of time encoding, recurrent and graph neural networks. Enabling raw sensor data to be easily combined with existing physics-based models opens up new possibilities for accurate extreme storm tide events forecast systems that enhance community safety and aid policymakers in their decision-making processes.

## Introduction

Sea-level rise is a well-known and documented trend directly associated with climate change. There is also mounting evidence in favor of related phenomena such as changes in the intensity and frequency of extreme meteorological events. The confluence of tide fluctuations and storm surges can cause severe flooding in coastal areas. In this context, storm tides are key phenomena that refer to the composition of high astronomical and meteorological tides.

This work focuses on improving early alerts for extreme storm tide events in the city of Santos, Brazil. The city is home to the largest port in Latin America and also hosts rich ecosystems like the Atlantic Rainforest and a vast mangrove system. Under an optimistic sea-level rise scenario, the cumulative damage to the Santos coastal area between 2010 and 2100 could reach 242 million USD (Marengo et al. 2019).

Researchers successfully implemented physical models for predicting extreme events in Santos (Costa et al. 2020). These models, considered state-of-the-art for the oceanography community, serve as baselines for forecasting. Given that we have access to those models' predictions and data from sensors and satellites, we exploit these modalities in a lightweight manner. Such a model aims to match the performance of physical models, while integrating observed and simulated time series with satellite imagery. Our novel approach ensures lower computational costs than numerical models and addresses real-world challenges, such as noise and missing data. Designed for multimodal data and with low complexity, it enables a significant societal and environmental impact, adaptable to regions with similar characteristics. Importantly, our objective is not to benchmark this task; instead, we emphasize the practical application and potential for widespread implementation of our model in problems with highly irregular multimodal time series.

Therefore, we propose a modularized model for multimodal data, including observational sensory data, numerical models, and satellite images. The model is divided into three modules: (1) a timestamp encoding for capturing temporal information of periodic signals, with the ability to smoothly handle missing data; (2) a Recurrent Neural Network (RNN)-based temporal encoding for time-series signals, providing flexibility to the model for differences in signals' sampling and missingness rates; (3) a signal information-sharing module using Graph Neural Networks (GNNs) (Scarselli et al. 2009) so as to extract connections among variables and use those connections profitably.

We state our main contributions as:

- Defining an architecture that can learn from multimodal data sources to generate accurate early detection of extreme storm tide events;
- Providing evidence that recurrent and graph neural networks (RNN+GNN) mixtures can be applied to reduce the complexity of handling irregular data sources;

Copyright © 2024, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

- Providing evidence that time encoding can increase a model robustness to missing data and provide flexibility to autoregressive inference;
- Providing evidence that data-driven approaches can be combined with physics-based models to increase community safety in coastal areas;
- Demonstrating a real application of the method showing results for historical data from Santos region, Brazil, a site of significant social importance that is regularly affected by storm tides.

Section *Background* contextualizes the storm tide phenomena and its interactions with local communities, lists existing efforts to improve numerical modeling for the region. We formalize the target task for this work, and highlight related work recently proposed for similar tasks in Section *Problem Statement and Related Work*. Section *Proposed Model* details the architecture, justifying technical choices based on the challenges inherent to this task. Section *Experiments* presents and comments on experiments results, providing evidence for the model's claimed capabilities. Finally, the *Conclusion* emphasizes how this work is both a technical innovation and a significant step towards increasing Santos' communities safety.

## Background

Among the phenomena that significantly impact coastal populations, storm tides cause some of the largest impacts on economy and society, since they can cause flooding (Costa et al. 2020), resulting in infrastructure damage, service disruption, and loss of human life. Coastal regions in developing countries are especially susceptible to these significant damages in the long term (Dasgupta et al. 2009; Edmonds et al. 2020). An example of a susceptible area is the metropolitan region of Santos, in Brazil, which suffers periodically from storm tides.

A storm tide is the total observed water elevation during a storm event, resulting from the combination of storm surge and the astronomical tide. Tides are usually caused by the gravitational forces of the sun and the moon and have their greatest effects on the sea level during new and full moons, known as spring tides. A storm surge is the abnormal sea level rise caused by a storm's winds pushing the sea water towards the coast (NOAA 2023). When a storm surge coincides with a high spring tide, coastal flooding is likely to occur.

The metropolitan region of Santos, located on the central part of the South Brazil Bight (SBB), is home to more than 1.7 million people (Ribeiro et al. 2019) and hosts the largest port complex in South America. In the SBB region (from 23ºS to 28.5ºS, approximately), storm surges are induced by the passage of cold fronts which reach the region every 6-10 days throughout the year (Castro and Lee 1995). However, due to climate changes, the frequency and magnitude of extreme sea level events is expected to increase by the end of the century. In addition, as the global mean sea level has been rising, coastal zones can be more vulnerable to weaker storms (Oppenheimer et al. 2019).

Figure 1: In 2016 an extreme storm tide event caused severe inundation in Santos (Martiniano 2016).

Between 1961 and 2011, 89 storm surge events were reported by the press in the Santos region (Fundespa 2014). Although not all of them characterize a storm tide event, because they did not coincide with the spring tide, serious coastal flooding took place many times. A very intense and destructive storm tide happened in Santos in 2016. In that event, streets, garages of buildings, and commercial points were taken by the sea, causing huge damage to the local economy (no detailed study on the level of these losses seems to be available). Figure 1 shows how the Santos coastal line is vulnerable to sudden increase of sea-level.

Numerical modeling is commonly employed to simulate the circulation in oceans and in coastal regions. Such modeling is applied more frequently in case studies, toward specific objectives and simulating past conditions (e.g. Leitão et al. 2005; Dube et al. 2009; Bunya et al. 2010). However, operational models for the prediction of extreme events, such as storm tides, represent another important application of numerical modeling. A solid and reliable numerical forecasting system is an important tool for preventing the effects of storm tides, supporting contingency planning and reducing the risks associated with these events (Ribeiro et al. 2019).

One such model is the Regional Ocean Modelling System (ROMS). It has been implemented and validated for the SBB region, by the Coastal Hydrodynamics Laboratory from the Oceanographic Institute of the University of São Paulo. A sea-surface height (SSH) series for past periods was simulated and is used to detect and understand past storm tide events.

Another prominent model currently deployed for that region is the Santos Operational Forecasting System (SOFS) (Costa et al. 2020), which is based on the Princeton Ocean Model (POM). SOFS has two nested grids, where the coarser one encompasses the SBB region, while the finer one encompasses the Santos-São Vicente-Bertioga Estuarine System. The model outputs forecasts for SSH, currents, salinity and temperature for three days ahead.

Figure 2 shows that SOFS and ROMS have high corre-

lations with the measured SSH signal, although their predicted values show a difference in level and phase due to differences in mesh resolution, simulation and conventions adopted for each numerical model. Also, a particularity of SSH is that astronomical tide, which is a quantity that can be very accurately forecast, accounts for a major component of it.

A set of collected in-situ data is essential to produce robust and accurate forecasts, since modeling techniques with data assimilation can produce results that are closer to the observations. However, despite being an important tool for forecasting extreme events, numerical modeling requires significant computational effort. High-resolution numerical models demand great computer hardware for faster results, which implies high equipment acquisition and maintenance costs. Consequently, developing ways to better and more flexibly combine existing numerical models with ever-increasing sensor data is a pressing need.

Answering to that, the Santos Storm Tide Dataset was assembled from eleven different data sources. The dataset contains local SSH measurements for the Praticagem station, located at the main entrance of the Port of Santos channel at a distance of 3 kilometers from the Palmas station. The SOFS and ROMS numerical simulations generate SSH predictions specifically for the Palmas station location. Although changing the discretization setup for the simulations could reduce this distance, it was purposefully kept as it is to better assess data assimilation capacities of new techniques. Some characteristics of the dataset are summarised in Table 1. For our experiments, we define data relative to 2022 as our test dataset, leaving the rest for model training.

## Problem Statement and Related Work

Both measured and simulated data can be collectively formulated as a Multivariate Time Series (MTS). We can define a MTS as a (finite) set of time series $\mathcal{S} = \{\mathbf{Z_i}\}_{i=1}^{|S|}$. Each element in this set is a sequence of events $\mathbf{Z_i} = [Z_{i,t_1}, Z_{i,t_2}, \cdots, Z_{i,t_T}]$, such that $Z_{i,t} \in \mathbb{R}^{K_i}$ for tabular data and $Z_{i,t} \in \mathbb{R}^{H_i \times W_i \times K_i}$ for images, where $H_i$ and $W_i$ represents the spatial dimensions of the image and $K_i$ represents the number of features. Note that we choose to represent a time series as a boldface matrix $\mathbf{Z_i}$ with $T$ observations, $t_1, \cdots, t_T \in \mathbb{R}$.

The task objective is to use all information available at a reference time $t_\phi$ to forecast a sequence of future measurements for a subset of time series $\mathcal{S}_o \subseteq \mathcal{S}$. For this work $\mathcal{S}_o$ contains two time series, SSH and wave height (WH). Crucially, measured data only exists for past events, i.e., $t < t_\phi$, while the forecasts from numerical models typically extend up to $t < t_\phi + O_i$, where $O_i$ represents the forecast window length specific to that numerical model.

Handling irregularly sampled multivariate time series represent a fundamental challenge for classical models from machine learning and statistics (Shukla and Marlin 2021). Several data imputation approaches have been proposed to regularize MTS in order to achieve a grid-like structure (Emmanuel et al. 2021). A grid-like MTS would allow the usage of a RNN or Transformer models directly; however, for highly irregular signals it becomes unfeasible to regularize the grid, as shown in Figure 3. Additionally, the fact that some data points are missing can be informative, which means data imputation is not always beneficial (Little and Rubin 2014).

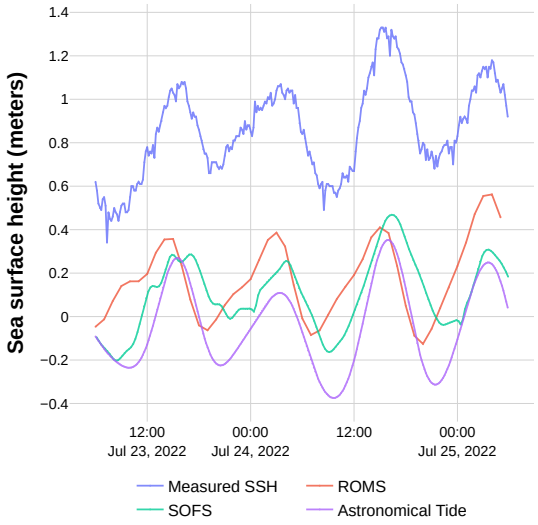


Figure 2: ROMS and SOFS, both numerical models, have high correlations to the measured SSH signal, even though the predicted values are out of phase and offset and have distinct sampling rates. That is due to differences in grid resolution, simulation setup, and conventions adopted for each numerical model. Also, it should be noted that the Astronomical tide is the main component of the SSH signal.

Recent proposals have described techniques to inform classical machine learning architectures about missing data without data imputation. GRU-D (Che et al. 2018) proposes two additional structures to the classical GRU, namely, a masking tensor and a decay mechanism. RAINDROP (Zhang et al. 2022) proposes a mechanism that relies solely on GNN message passing by modeling each measurement as a node that creates ripples, therefore propagating information to nearby nodes.

## Proposed Model

To effectively merge numerical model forecasts with available measurement data, a model must represent events spanning both past and future timestamps. Moreover, given that sensors in the wild can be susceptible to malfunctions, it is essential for the model to remain functional and dependable even when one or more sensors are out of commission.

Furthermore, simplifying data preprocessing and reducing the model's size are paramount, since it allows smaller teams to efficiently deploy the model across various regions with minimal adjustments. This section describes the model in detail, but omit some minor information, such as the complete configuration for the CNN modules.

The proposed model is grounded on three primary principles:

Table 1: To develop more accurate forecast systems for Storm Tides in the Santos region, a multimodal dataset was assembled from different sources. The dataset includes data acquired from sensors installed in the region, numerical models (both global and regional), and satellite images.

| Source | Sampling interval (minutes) | Date Range | % Missing Data | Data shape | Number of Readings (train/test) |
|---|---|---|---|---|---|
| SSH (local measurement) | 60 / 10* | 2016-04, 2022-12 | 4.57/15.41 | [1] | 173607 / 52479 |
| Wave Height (local measurement) | 20 | 2016-01, 2022-12 | 6.84 | [1] | 148423 / 23070 |
| Astronomical Tide (calculated) | 10 | 2016-01, 2022-12 | 0 | [1] | 315648 / 52560 |
| ROMS (numerical simulation) | 60 | 2016-01, 2022-12 | 0 | [1] | 51876 / 8760 |
| SOFS (numerical simulation) | 15 | 2019-01, 2022-12 | 0 | [3] | 105216 / 35040 |
| SSH (numerical model - Copernicus) | 60 | 2020-11, 2022-12 | 0 | [1, 133, 133] | 10224 / 8760 |
| ADT (satellite - Copernicus) | 1440 | 2016-01, 2022-12 | 0 | [1, 45, 45] | 2192 / 365 |
| Waves (ERA5) | 60 | 2016-01, 2022-12 | 0 | [3, 23, 23] | 52608 / 8760 |
| Waves (satellite - Copernicus) | 1440 | 2016-01, 2022-12 | 0 | [2, 7, 6] | 2192 / 366 |
| Wind (local measurement) | 20 | 2016-01, 2022-12 | 4.53 | [2] | 149855 / 25905 |
| Wind (ERA5) | 60 | 2016-01, 2022-12 | 0 | [2, 45, 45] | 52608 / 8760 |

*SSH data has a sampling interval of 1 hour until 2019-03 and 10 minutes afterwards.

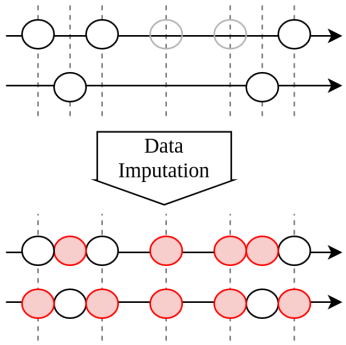


Figure 3: Shifted time series with missing data points are difficult to regularize. In this example, 65% of the data comes from interpolation after the data imputation process.

1. Timestamp Encoding using Periodic Functions: This approach ensures that time-related information is consistently encoded across time series, capturing patterns that occur periodically.
2. Independent Temporal Encoding of Time Series: This uses an independent Recurrent Neural Networks (RNN) to process each time series data stream, ensuring the model's capability to operate without one or more sensors.
3. Information Diffusion with Regularized Heterogeneous Graph Attention Network (RHGAT): This concept facilitates the dissemination of information across the model's nodes, allowing for each representation to be enriched based on neighbourhood information.

A visual representation of these concepts from a forward pass perspective is illustrated in Figure 4. The proposed architecture is trained following the supervised learning paradigm.

To build training batches, a set of $t_\phi$ are sampled uniformly from the time interval over which data is available. It is important to note that the sampling is performed over the entire time interval, rather than just over the set of timestamps that have associated measurements. Actually, there is no need for any $Z_{i,t_\phi}$ to exist.

Recall that for measured data, the context window includes data collected in $t < t_\phi$, while for estimated data (i.e. numerical models and astronomical tide), the context window will be in $t < t_\phi + O_i$. Thus, for each $t_\phi$, both a context window and a forecast window are determined for each time series,

$$t_\phi^{'} = t_\phi + O_i, \tag{1}$$

$$\mathbf{Z_i^c} = \left[ Z_{i,t} \in \mathbf{Z_i} \mid t_\phi^{'} - C_i \leq t < t_\phi^{'} \right], \tag{2}$$

$$\mathbf{Z_i^f} = \left[ Z_{i,t} \in \mathbf{Z_i} \mid t_\phi^{'} \leq t < t_\phi^{'} + F_i \right], \tag{3}$$

where: $O_i$ denotes the offset allowed for a specific time series. It is configured to be 48 hours for both numerical models and for the astronomical tide signal, and zero for others. This positive offset allows the context window for these types to contain future events, i.e., forecasts; $C_i$ represents the size of the context window and is set to 7 days for all time series; $F_i$ is the size of the forecast window and only applies to the target time series SSH and WH, both set to 2 days. These three parameters are problem specific and can be tuned without affecting our formulation.

To provide the model with information about irregular intervals between measurements we apply the same Positional Encoding (PE) formulation originally proposed by (Vaswani et al. 2017). Although this technique has been commonly used in Natural Language Processing models to add positional information to the fully connected attention mechanism in Transformers, our hypothesis is that this mechanism associated with a GNN would allow the model to adjust for missing data windows. Firstly, each timestamp $t$ is internally represented relative to $t_\phi$ and then PE is applied to create a vector representation $\tau_{t^i}$ of this temporal distance that has size $\mathcal{T} = 50$,

$$t^{'} = t - t_\phi, \tag{4}$$

$$\tau_{t^{'},2k} = \sin\left(\frac{t^{'}}{1000^{2k/\mathcal{T}}}\right), \tag{5}$$

$$\tau_{t^{'},2k+1} = \cos\left(\frac{t^{'}}{1000^{2k/\mathcal{T}}}\right). \tag{6}$$

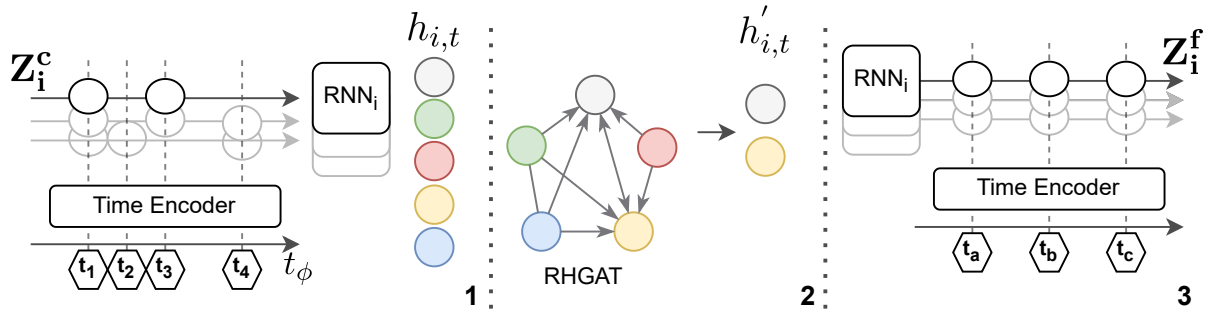


Figure 4: Architecture overview. 1) Measurements are enriched with encoded representation of subsequent timestamps. Each time series is encoded by a different RNN instance. 2) A Regularized Heterogeneous GAT updates $h_{i,t}$ the target nodes representations based on its neighbourhood. 3) $h'_{i,t}$ is used as the initial state for an autoregressive decoding process.

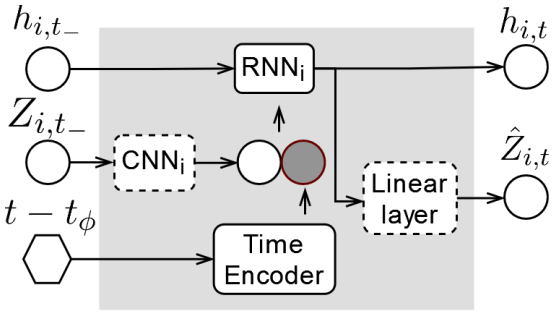


Figure 5: Time-aware RNN setup. Images are first encoded by a CNN into a vector of size $K_i$. The updated latent representations of the target time series are projected by a linear layer onto its original space.

This approach is convenient since it represents both past and future events relative to $t_\phi$ by leveraging periodic functions applied to $t^{'}$, which is negative for past events and positive otherwise. Finally, we denote $t^{'}_{-}$ the relative timestamp for the previous measurement of a time series.

Images are encoded into a vector with $K_i$ values using a two-layer convolutional neural network CNN. This step is necessary to allow temporal encoding RNNs to be defined uniformly for all time series. Associating CNNs and RNNs for image sequence encoding has been previously successfully applied in a multitude of use cases (Fan et al. 2016; Donahue et al. 2016; Sharma et al. 2021).

Information about the size of the gap ahead of each $Z_{i,t_-}$ is provided to the RNN by concatenating to it each encoded timestamp $\tau_{t'}$ of the next event $Z_{i,t}$,

$$h_{i,t} = \mathrm{RNN}_i\left(Z_{i,t_-} \parallel \tau_{t'}, h_{t_-}\right), \tag{7}$$

and $\parallel$ is the concatenation operator. This autogressive setup is depicted in Figure 5.

In our setup, all $h_{i,t}$ have size $\mathcal{H} = 200$; however, as the notation indicates, this hyperparameter could be different for each time series. Because these latent representations only contain information about their own time series, an information propagation mechanism is necessary to enrich the target nodes representations before the decoding step.

For that, we use a single-layered regularized variant of an Heterogeneous Graph Attention Network (HGAT)(Veličković et al. 2018; Wang et al. 2019). Each node type in the network has a single node, which means each graph has $S$ node types. Each target node has incoming edges from all other nodes, i.e., the total number of edges is $S \times S_o$, where $S_o$ is the number of target nodes.

In HGATs, each edge type $r \in \mathcal{R}$, also known as relation, is a unique pair of connected node types. To decouple the size of $\mathcal{R}$ from $S$ and $S_o$ we apply the regularization technique proposed by (Schlichtkrull et al. 2018), where a set of $B$ homogeneous GNNs and $\mathcal{R} \times B$ scalar coefficients are trained. The Regularized HGAT update can be defined as:

$$h'_{i,t} = \sum_{r \in \mathcal{R}} \sum_{b=1}^{B} \mathrm{GAT}(\{a_{r,b} \cdot \mathbf{h}_j : j \in \mathcal{N}^{(r)}(i)\}), \tag{8}$$

where GAT is the convolution proposed by Brody, Alon, and Yahav that builds upon the original GAT architecture (Veličković et al. 2018); $\mathcal{N}^{(r)}(i)$ represents the neighbourhood of a target node, which in our setup is the whole graph; $a_{r,b}$ is a trainable scalar coefficient; and $\mathbf{h}_j$ is the hidden representation of a neighbouring node. Applying techniques for latent graph inference, such as proposed by Kazi et al., is a relevant research direction to improve this formulation by addressing eventual over-squashing in the message passing mechanism (Alon and Yahav 2021).

Finally, to generate forecasts the same autoregressive setup described in Equation 7 is applied. However, for each step, a node specific linear layer with $W_i$ and $b_i$ is applied to $h'_{i,t}$ to generate a prediction from the latent representation, $\hat{Z}_{i,t}$,

$$\hat{Z}_{i,t} = W_i h'_{i,t} + b_i. \tag{9}$$

During the training regime, the timestamps are those of the labels available in the training dataset; however, in an inference setup it is possible to define for which timestamps predictions must be generated. Once again, the mechanism that represents both encoding and decoding is depicted in Figure 5.

Forecasts are compared to the actual measured values by using a loss function derived from the Index of Agreement (Willmott 1981). This loss can detect additive and proportional deviations between time series and was successfully applied in ML modeling of oceanic phenomena by (Netto et al. 2022). In this formulation it is defined as:

$$L(\mathcal{S}_{\mathbf{o}}^{\mathbf{f}}, \hat{\mathcal{S}}_{\mathbf{o}}^{\mathbf{f}}) = \frac{1}{S_o} \sum_{i=1}^{S_o} \left( \frac{\sum_t (\mathbf{Z}_{\mathbf{i}}^{\mathbf{f}} - \hat{\mathbf{Z}}_{\mathbf{i}}^{\mathbf{f}})^2}{\sum_t (|\hat{\mathbf{Z}}_{\mathbf{i}}^{\mathbf{f}} - \bar{\mathbf{Z}}_{\mathbf{i}}^{\mathbf{f}}| + |\mathbf{Z}_{\mathbf{i}}^{\mathbf{f}} - \bar{\mathbf{Z}}_{\mathbf{i}}^{\mathbf{f}}|)^2} \right), \tag{10}$$

where $\mathcal{S}_{\mathbf{o}}^{\mathbf{f}}$ and $\hat{\mathcal{S}}_{\mathbf{o}}^{\mathbf{f}}$ are, respectively, the measured and forecast MTS in the forecast window, and $\bar{\mathbf{Z}}_{\mathbf{i}}^{\mathbf{f}}$ is the temporal mean of $\mathbf{Z}_{\mathbf{i}}^{\mathbf{f}}$.

## Experiments

Three variants of the model were trained: the complete model, as described in the previous section; a variant without the RHGAT; and a last variant without time encoding. This section analyses these variants by creating four different inference setups:

1. Input pairs: Evaluates the effect of each individual variable in the system's performance;
2. Input removal: Evaluates the effect of the eventual unavailability of each time series;
3. Missing data: Evaluates robustness to missing data by removing increasing amounts of data from all inputs;
4. Oversampling and undersampling: Evaluates the capacity of the model to produce forecasts with different sampling rates.

Each of these inference setups are described in the following, showing and discussing the results achieved.

### Inference setup 1: Input pairs

Table 2 presents results for the first setup, i.e., using pairs of time series as input instead of the full set of eleven input variables. These results show that the model, regardless of not being trained for this scenario, is capable of incorporating information from each time series independently.

As expected, the time series that increase SSH forecast performance the most are those most correlated to it, namely, ROMS, SOFS and Astronomical Tide. These results also show that when the inference is done with all the inputs, the performance rises above the level of any other pair, demonstrating that even a single layer of the information propagation mechanism is capable of taking advantage of synergies between nodes. Additionally, WH data from ERA5, which is a sequence of images, brings the most benefit for WH predictions, displaying evidence of multimodal data processing capabilities.

### Inference setup 2: Input removal

Table 3 shows the results for the second inference scenario where a single time series is removed at a time. Note that IoA remains high even when critical information, such as Astronomical Tide, is lost. This evidences the robustness of the proposed architecture for data failures.

Table 2: Results using PE as time encoding and a single additional time series during the inference. Same row and column are baselines with no extra time series. IoA stands for Index of Agreement and MAE for Mean Absolute Error. Most impactful additions in bold. (SC: Satellite - Copernicus; NMC: Numerical Model - Copernicus.)

| | **SSH (m)** | | **WH (m)** | |
|---|---|---|---|---|
| Additional Time Series | IoA ↑ | MAE ↓ | IoA ↑ | MAE ↓ |
| SSH* | $0.693 \pm 0.148$ | $0.22 \pm 0.07$ | $0.482 \pm 0.234$ | $0.30 \pm 0.29$ |
| WH* | $0.691 \pm 0.143$ | $0.22 \pm 0.07$ | $0.479 \pm 0.229$ | $0.31 \pm 0.29$ |
| Astron. Tide | $\mathbf{0.895 \pm 0.106}$ | $\mathbf{0.15 \pm 0.08}$ | $0.470 \pm 0.234$ | $0.30 \pm 0.29$ |
| ROMS | $0.801 \pm 0.124$ | $0.19 \pm 0.07$ | $0.471 \pm 0.236$ | $0.30 \pm 0.28$ |
| SOFS | $0.858 \pm 0.107$ | $0.15 \pm 0.04$ | $0.561 \pm 0.252$ | $0.25 \pm 0.20$ |
| ADT (SC) | $0.693 \pm 0.144$ | $0.22 \pm 0.07$ | $0.483 \pm 0.236$ | $0.30 \pm 0.29$ |
| SSH (NMC) | $0.688 \pm 0.140$ | $0.23 \pm 0.08$ | $0.492 \pm 0.243$ | $0.30 \pm 0.28$ |
| Waves (ERA5) | $0.709 \pm 0.142$ | $0.21 \pm 0.07$ | $\mathbf{0.621 \pm 0.239}$ | $\mathbf{0.20 \pm 0.19}$ |
| Waves (SC) | $0.663 \pm 0.150$ | $0.23 \pm 0.07$ | $0.452 \pm 0.233$ | $0.32 \pm 0.28$ |
| Wind* | $0.691 \pm 0.149$ | $0.22 \pm 0.07$ | $0.474 \pm 0.230$ | $0.30 \pm 0.29$ |
| Wind (ERA5) | $0.709 \pm 0.144$ | $0.21 \pm 0.07$ | $0.499 \pm 0.232$ | $0.29 \pm 0.28$ |
| All time series | $0.967 \pm 0.039$ | $0.08 \pm 0.03$ | $0.654 \pm 0.231$ | $0.18 \pm 0.14$ |

*local measurement

Table 3: Results using PE as time encoding and a single time series unavailable during the inference. Most impactful removals in bold. (SC: Satellite - Copernicus; NMC: Numerical Model - Copernicus.)

| | **SSH (m)** | | **WH (m)** | |
|---|---|---|---|---|
| Unavailable Time Series | IoA ↑ | MAE ↓ | IoA ↑ | MAE ↓ |
| SSH* | - | - | $0.652 \pm 0.231$ | $0.19 \pm 0.14$ |
| WH* | $0.963 \pm 0.037$ | $0.08 \pm 0.03$ | - | - |
| Astron. Tide | $\mathbf{0.857 \pm 0.092}$ | $\mathbf{0.15 \pm 0.04}$ | $0.653 \pm 0.232$ | $0.19 \pm 0.14$ |
| ROMS | $0.964 \pm 0.036$ | $0.08 \pm 0.03$ | $0.654 \pm 0.231$ | $0.18 \pm 0.14$ |
| SOFS | $0.922 \pm 0.079$ | $0.13 \pm 0.07$ | $0.600 \pm 0.236$ | $0.21 \pm 0.19$ |
| ADT (SC) | $0.964 \pm 0.036$ | $0.08 \pm 0.03$ | $0.654 \pm 0.231$ | $0.18 \pm 0.14$ |
| SSH (NMC) | $0.954 \pm 0.043$ | $0.09 \pm 0.03$ | $0.637 \pm 0.240$ | $0.20 \pm 0.15$ |
| Waves (ERA5) | $0.960 \pm 0.044$ | $0.09 \pm 0.03$ | $\mathbf{0.558 \pm 0.245}$ | $\mathbf{0.24 \pm 0.19}$ |
| Waves (SC) | $0.966 \pm 0.035$ | $0.08 \pm 0.03$ | $0.672 \pm 0.239$ | $0.17 \pm 0.13$ |
| Wind (ERA5) | $0.962 \pm 0.038$ | $0.08 \pm 0.03$ | $0.650 \pm 0.232$ | $0.19 \pm 0.14$ |
| Wind* | $0.964 \pm 0.036$ | $0.08 \pm 0.03$ | $0.652 \pm 0.230$ | $0.18 \pm 0.14$ |
| All time series | $0.967 \pm 0.039$ | $0.08 \pm 0.03$ | $0.654 \pm 0.231$ | $0.18 \pm 0.14$ |

*local measurement

### Inference setup 3: Missing data

Figure 6 shows qualitative results for the third inference setup where 2-hour long data windows are randomly removed until reaching a missing data ratio threshold. This setup illustrates how the variant without time encoding exhibits faster performance degradation as the missing ratio increases. Results shown in Table 4 also provide further evidence that the time encoding mechanism is crucial in order to achieve robustness against missing data.

### Inference setup 4: Oversampling and undersampling

The fourth inference setup aims to investigate if the RNN does indeed consider the encoded time gap $\tau_{t'}$ to estimate $Z_{i,t}$. Figure 7 clearly shows that this architecture is capable of producing time series forecasts in frequencies that it was not trained on. This is evidence that the introduction of time encoding is not only an efficient way to model varying time gaps in the input, but also represents a notable gain in

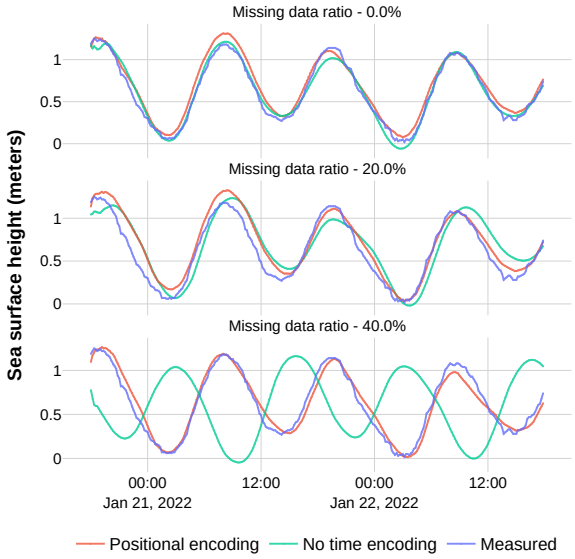


Figure 6: Time encoding is especially important as the missing data ratio increases as the information about the gaps between measurements allows the system to adapt.

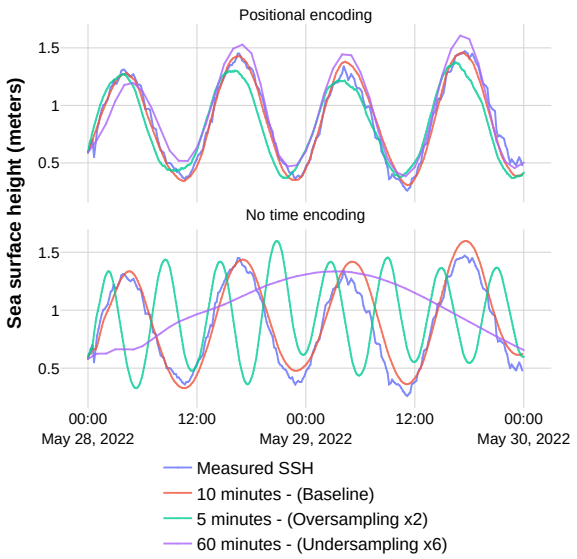


Figure 7: The model can of producing forecasts with different frequencies from those seen during training.

Table 4: Results for varying amounts of missing data. Models without time encoding exhibits faster performance degradation.

| Model variant / Missing ratio | SSH (m) | | WH (m) | |
|---|---|---|---|---|
| | IoA ↑ | MAE ↓ | IoA ↑ | MAE ↓ |
| Full model / 00% | $\mathbf{0.960 \pm 0.086}$ | $\mathbf{0.08 \pm 0.03}$ | $\mathbf{0.654 \pm 0.231}$ | $\mathbf{0.18 \pm 0.14}$ |
| No time encoding / 00% | $0.943 \pm 0.104$ | $0.10 \pm 0.03$ | $0.628 \pm 0.231$ | $0.20 \pm 0.17$ |
| No GNN / 00% | $0.917 \pm 0.115$ | $0.13 \pm 0.07$ | $0.521 \pm 0.250$ | $0.29 \pm 0.26$ |
| Full model / 10% | $\mathbf{0.954 \pm 0.087}$ | $\mathbf{0.09 \pm 0.03}$ | $\mathbf{0.645 \pm 0.235}$ | $\mathbf{0.19 \pm 0.14}$ |
| No time encoding / 10% | $0.891 \pm 0.138$ | $0.14 \pm 0.06$ | $0.605 \pm 0.224$ | $0.21 \pm 0.18$ |
| No GNN / 10% | $0.910 \pm 0.117$ | $0.14 \pm 0.07$ | $0.509 \pm 0.239$ | $0.30 \pm 0.27$ |
| Full model / 20% | $\mathbf{0.947 \pm 0.085}$ | $\mathbf{0.10 \pm 0.03}$ | $\mathbf{0.645 \pm 0.232}$ | $\mathbf{0.19 \pm 0.14}$ |
| No time encoding / 20% | $0.842 \pm 0.170$ | $0.16 \pm 0.08$ | $0.603 \pm 0.228$ | $0.21 \pm 0.19$ |
| No GNN / 20% | $0.906 \pm 0.119$ | $0.14 \pm 0.07$ | $0.508 \pm 0.239$ | $0.30 \pm 0.27$ |
| Full model / 40% | $\mathbf{0.917 \pm 0.097}$ | $\mathbf{0.12 \pm 0.04}$ | $\mathbf{0.631 \pm 0.241}$ | $\mathbf{0.20 \pm 0.15}$ |
| No time encoding / 40% | $0.720 \pm 0.238$ | $0.21 \pm 0.10$ | $0.587 \pm 0.230$ | $0.21 \pm 0.19$ |
| No GNN / 40% | $0.894 \pm 0.119$ | $0.15 \pm 0.07$ | $0.506 \pm 0.234$ | $0.30 \pm 0.27$ |

flexibility in how forecasts can be produced.

Here it is emphasized that the model was not trained with this capability in mind, however, based on these results, a dataset with more diverse time gaps within each time series could further improve this time generalization property.

Finally, Table 5 shows some statistics for the training and characteristics of the final model. Since all RNNs and the RHGAT have only one layer, the model is extremely lightweight. It was trained in a consumer level GPU and is capable of performing inference in CPUs.

## Conclusion

Results for the experiments demonstrate that the proposed model represents an innovative approach to handle irregular data sources. Despite being tested in adverse conditions (scenarios with high level of missing data, unavailable time series and noise, etc.), the model maintained itself operational with adequate performance. Evidence shows that this approach is capable of associating different types of time series, including images, numerical forecasts, and local measurements, in order to better represent the complex phenomena around storm tides.

Table 5: Training statistics relative to the full model variant.

| Training and model statistics | |
|---|---|
| Model size in disk | 48MB |
| Model size in GPU | 908MB |
| Number of parameters | 10,275,516 |
| GPU used | GeForce RTX 3060 |
| Average training time (hours) | 22.89 |
| Inference time in CPU (s) | 1.6 |

In particular, our experimental results highlight how schemes like Positional Encoding, which utilize periodic functions, can be applied in autoregressive RNN setups to encode timestamps. This allows models to account for varying time gaps, increasing robustness to missing data and allowing for flexible inference setups.

As the frequency of extreme meteorological events increases, we predict that efficient approaches for processing MTS, especially those capable of combining existing global and local numerical models to sensoring networks, will grow in importance.

As future research directions, our architecture can be used in a continuous operation setup, i.e., the autoregressive mechanism can be run without interruption with periodic information propagation events. There has been work in a similar direction, where a persistent memory vector is attached to GNN nodes, such as (Rossi et al. 2020).

## Acknowledgements

The authors would like to thank the Center for Artificial Intelligence (C4AI-USP), with support by the São Paulo Research Foundation (FAPESP) under grant number 2019/07665-4 and by the IBM Corporation. The authors would also like to thank the *Programa de Bolsas Itaú* (PBI) of the *Centro de Ciência de Dados* ($C^2D$), supported by *Itaú Unibanco S.A.*, and the São Paulo Pilots Association for providing the environmental and oceanic measured data used in this work. This work is also supported in part by the Brazilian National Council for Scientific and Technological Development (CNPq) under grant numbers 310085/2020-9, 310127/2020-3 and 305753/2022-3, and the Coordination for the Improvement of Higher Education Personnel (CAPES, Finance Code 001).

## References

Alon, U.; and Yahav, E. 2021. On the Bottleneck of Graph Neural Networks and its Practical Implications. ArXiv:2006.05205 [cs, stat].

Brody, S.; Alon, U.; and Yahav, E. 2022. How Attentive are Graph Attention Networks? In *International Conference on Learning Representations*.

Bunya, S.; Dietrich, J. C.; Westerink, J. J.; Ebersole, B. A.; Smith, J. M.; Atkinson, J. H.; Jensen, R.; Resio, D. T.; Luettich, R. A.; Dawson, C.; Cardone, V. J.; Cox, A. T.; Powell, M. D.; Westerink, H. J.; and Roberts, H. J. 2010. A High-Resolution Coupled Riverine Flow, Tide, Wind, Wind Wave, and Storm Surge Model for Southern Louisiana and Mississippi. Part I: Model Development and Validation. *Monthly Weather Review*, 138(2): 345 – 377.

Castro, B. M.; and Lee, T. N. 1995. Wind-forced sea level variability on the southeast Brazilian shelf. *Journal of Geophysical Research: Oceans*, 100(C8): 16045–16056.

Che, Z.; Purushotham, S.; Cho, K.; Sontag, D.; and Liu, Y. 2018. Recurrent Neural Networks for Multivariate Time Series with Missing Values. *Scientific Reports*, 8(1): 6085. Number: 1 Publisher: Nature Publishing Group.

Costa, C. G.; Leite, J. R. B.; Castro, B. M.; Blumberg, A. F.; Georgas, N.; Dottori, M.; and Jordi, A. 2020. An operational forecasting system for physical processes in the Santos-Sao Vicente-Bertioga Estuarine System, Southeast Brazil. *Ocean Dynamics*, 70(2): 257–271.

Dasgupta, S.; Laplante, B.; Murray, S.; and Wheeler, D. 2009. Sea-level rise and storm surges : a comparative analysis of impacts in developing countries. *The World Bank, Policy Research Working Paper Series*.

Donahue, J.; Hendricks, L. A.; Rohrbach, M.; Venugopalan, S.; Guadarrama, S.; Saenko, K.; and Darrell, T. 2016. Long-term Recurrent Convolutional Networks for Visual Recognition and Description. ArXiv:1411.4389 [cs].

Dube, S.; Jain, I.; Rao, A. D.; and Murty, T. 2009. Storm surge modelling for the Bay of Bengal and Arabian Sea. *Natural Hazards*, 51: 3–27.

Edmonds, D. A.; Caldwell, R. L.; Brondizio, E. S.; and Siani, S. M. O. 2020. Coastal flooding will disproportionately impact people on river deltas. *Nature Communications*, 11(1): 4741.

Emmanuel, T.; Maupong, T.; Mpoeleng, D.; Semong, T.; Mphago, B.; and Tabona, O. 2021. A survey on missing data in machine learning. *Journal of Big Data*, 8(1): 140.

Fan, Y.; Lu, X.; Li, D.; and Liu, Y. 2016. Video-based emotion recognition using CNN-RNN and C3D hybrid networks. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, ICMI '16, 445–450. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-4556-9.

Fundespa. 2014. Plano básico ambiental da dragagem do Aprofundamento do Porto de Santos – Análise Histórica dos Registros de Ressacas ocorridos na região da Baía e Estuário de Santos. Technical report, Fundação para Estudos Aquáticos.

Kazi, A.; Cosmo, L.; Ahmadi, S.-A.; Navab, N.; and Bronstein, M. 2023. Differentiable Graph Module (DGM) for Graph Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2): 1606–1617. ArXiv:2002.04999 [cs, stat].

Leitão, P.; Coelho, H.; Santos, A.; and Neves, R. 2005. Modelling the main features of the Algarve coastal circulation during July 2004: A downscaling approach. *Journal of Atmospheric & Ocean Science*, 10(4): 421–462.

Little, R. J. A.; and Rubin, D. B. 2014. *Statistical Analysis with Missing Data*. John Wiley & Sons. ISBN 978-1-118-62588-0. Google-Books-ID: AyVeBAAAQBAJ.

Marengo, J. A.; Muller-Karger, F.; Pelling, M.; and Reynolds, C. J. 2019. The METROPOLE Project – An Integrated Framework to Analyse Local Decision Making and Adaptive Capacity to Large-Scale Environmental Change: Decision Making and Adaptation to Sea Level Rise in Santos, Brazil. In Nunes, L. H.; Greco, R.; and Marengo, J. A., eds., *Climate Change in Santos Brazil: Projections, Impacts and Adaptation Options*, 3–15. Cham: Springer International Publishing. ISBN 978-3-319-96535-2.

Martiniano, C. 2016. Extreme storm tide event in Santos. [Online; accessed August 12, 2023].

Netto, C. F. D.; de Barros, M. R.; Coelho, J. F.; de Freitas, L. P.; Moreno, F. M.; Mathias, M. S.; Dottori, M.; Cozman, F. G.; Costa, A. H. R.; Gomi, E. S.; and Tannuri, E. A. 2022. Modeling Oceanic Variables with Dynamic Graph Neural Networks. ArXiv:2206.12746 [cs].

NOAA. 2023. What is storm surge? National Ocean Service website, https://oceanservice.noaa.gov/facts/stormsurge-stormtide.html, 01/20/23.

Oppenheimer, M.; Glavovic, B. C.; Hinkel, J.; Wal, R. v. d.; Magnan, A. K.; Abd-Elgawad, A.; Cai, R.; Cifuentes-Jara, M.; DeConto, R. M.; Ghosh, T.; Hay, J.; Isla, F.; Marzeion, B.; Meyssignac, B.; and Sebesvari, Z. 2019. Sea Level Rise and Implications for Low Lying Islands, Coasts and Communities. In Pörtner, H.-O.; Roberts, D. C.; Masson-Delmotte, V.; Zhai, P.; Tignor, M.; Poloczanska, E.; Mintenbeck, K.; Alegría, A.; Nicolai, M.; Okem, A.; Petzold, J.; Rama, B.; and Weyer, N. M., eds., *The Ocean and*

*Cryosphere in a Changing Climate: Special Report of the Intergovernmental Panel on Climate Change*, 321–445. Cambridge, UK and New York, NY, USA: Cambridge University Press.

Ribeiro, R. B.; Sampaio, A. F. P.; Ruiz, M. S.; Leitão, J. C.; and Leitão, P. C. 2019. First Approach of a Storm Surge Early Warning System for Santos Region. In Nunes, L. H.; Greco, R.; and Marengo, J. A., eds., *Climate Change in Santos Brazil: Projections, Impacts and Adaptation Options*, 135–157. Cham, Switzerland: Springer International Publishing. ISBN 978-3-319-96535-2.

Rossi, E.; Chamberlain, B.; Frasca, F.; Eynard, D.; Monti, F.; and Bronstein, M. 2020. Temporal Graph Networks for Deep Learning on Dynamic Graphs. *arXiv:2006.10637 [cs, stat]*. ArXiv: 2006.10637.

Scarselli, F.; Gori, M.; Ah Chung Tsoi; Hagenbuchner, M.; and Monfardini, G. 2009. The Graph Neural Network Model. *IEEE Transactions on Neural Networks*, 20(1): 61–80.

Schlichtkrull, M.; Kipf, T. N.; Bloem, P.; van den Berg, R.; Titov, I.; and Welling, M. 2018. Modeling Relational Data with Graph Convolutional Networks. In Gangemi, A.; Navigli, R.; Vidal, M.-E.; Hitzler, P.; Troncy, R.; Hollink, L.; Tordai, A.; and Alam, M., eds., *The Semantic Web*, Lecture Notes in Computer Science, 593–607. Cham: Springer International Publishing. ISBN 978-3-319-93417-4.

Sharma, V.; Gupta, M.; Kumar, A.; and Mishra, D. 2021. Video Processing Using Deep Learning Techniques: A Systematic Literature Review. *IEEE Access*, 9: 139489–139507. Conference Name: IEEE Access.

Shukla, S. N.; and Marlin, B. M. 2021. A Survey on Principles, Models and Methods for Learning from Irregularly Sampled Time Series. ArXiv:2012.00168 [cs, stat].

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; and Polosukhin, I. 2017. Attention is All you Need. In Guyon, I.; Luxburg, U. V.; Bengio, S.; Wallach, H.; Fergus, R.; Vishwanathan, S.; and Garnett, R., eds., *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.

Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; and Bengio, Y. 2018. Graph Attention Networks. *arXiv:1710.10903 [cs, stat]*. ArXiv: 1710.10903.

Wang, X.; Ji, H.; Shi, C.; Wang, B.; Ye, Y.; Cui, P.; and Yu, P. S. 2019. Heterogeneous Graph Attention Network. In *The World Wide Web Conference*, WWW '19, 2022–2032. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-6674-8.

Willmott, C. J. 1981. On the Validation of Models. *Physical Geography*, 2(2): 184–194. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/02723646.1981.10642213.

Zhang, X.; Zeman, M.; Tsiligkaridis, T.; and Zitnik, M. 2022. Graph-Guided Network for Irregularly Sampled Multivariate Time Series. ArXiv:2110.05357 [cs].