# Application of LSTM-Based GNSS Data Imputation for Structural Health Monitoring: A Case Study of a World's Top 10 Longest Long-Span Floating Bridge in Norway

YAN LI

**ABSTRACT**

This paper investigates an LSTM-based method to impute missing GNSS data from the Bergsøysund Bridge in Norway. Simulating 10% randomly data loss, the approach demonstrated high quality in restoring time-series displacement measurements, supporting improved structural health monitoring. Future work will address longer time gaps for the data loss, and integration with additional sensors.

## INTRODUCTION

A key gap in structural health monitoring (SHM) is that many existing structures requiring real-time monitoring were built without embedded sensors, while newly constructed structures that include such sensors typically do not yet require active maintenance or repair. When selecting sensors for these older structures, remote non-contact techniques such as optical and laser scanning, as well as contact but minimally invasive methods like GNSS, are often recommended.

Global Navigation Satellite System (GNSS) provides unparalleled capability for long-term, real-time, remote, and absolute displacement monitoring of bridges [1]. However, a common issue with GNSS is data loss, which can result from various factors such as signal blockage, satellite availability, and receiver hardware faults. GNSS data outage rate can vary from about 14% to 47%, with the longest outage duration varying between about 8 seconds and 60 seconds, depending on the environmental conditions [2]. Incomplete data impacts movement path monitoring, time-series trend analysis, and real-time monitoring. The main solutions can be categorised into data imputation techniques, multi-constellation GNSS receivers, and improved antenna design and placement. The first category is algorithm-based, while the latter two are hardware-based. This paper investigates a data imputation technique, a deep learning method called Long Short-Term Memory (LSTM), using on-site data from a long-span floating bridge to mitigate the effects of incomplete GNSS data.

Yan Li, Imperial College London, South Kensington, London SW7 2AZ, UK

## SITE INFORMATION

This research focuses on the Bergsøysund Bridge in Norway as shown in Figure 1 and Figure 2, one of the world's ten longest long-span floating bridges. The water depth beneath the Bergsøysund Bridge from simplified model can reach approximately 320 metres [3]. The site also experiences strong winds, waves, and ocean currents. The seabed consists of soft clay and sediments. These complex geological conditions make it impractical to construct traditional piers or foundations for the bridge. In such conditions, one possible solution is the use of floating bridges, which are typically supported by buoyant pontoons and provide engineering solutions for spanning deep or geologically unstable waters [4]. However, the complex underwater environment continues to exacerbate structural safety challenges.

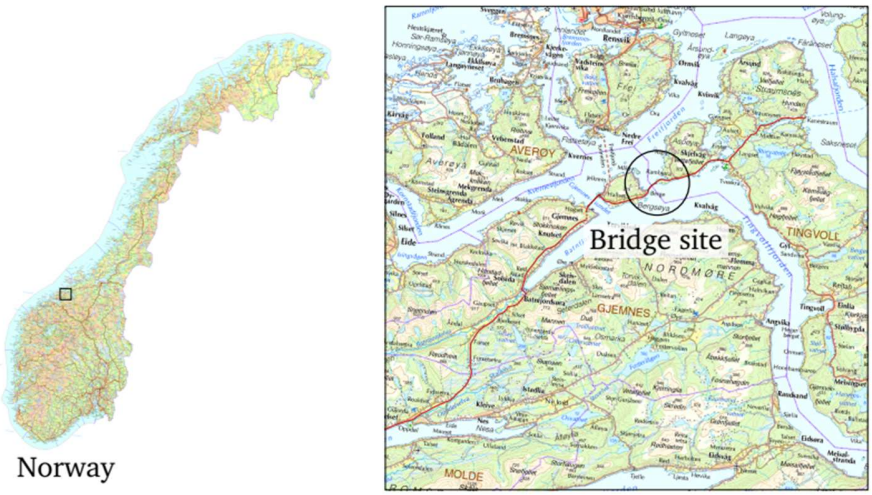
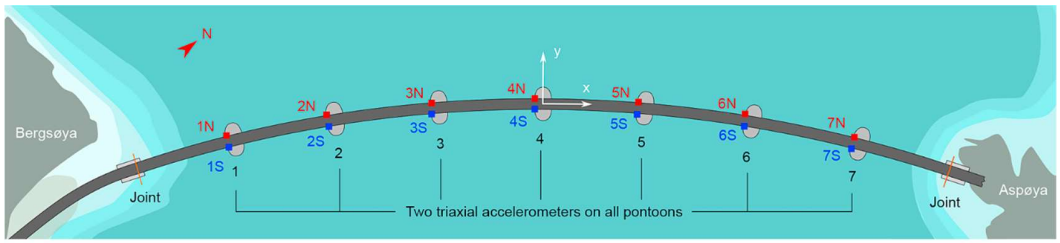
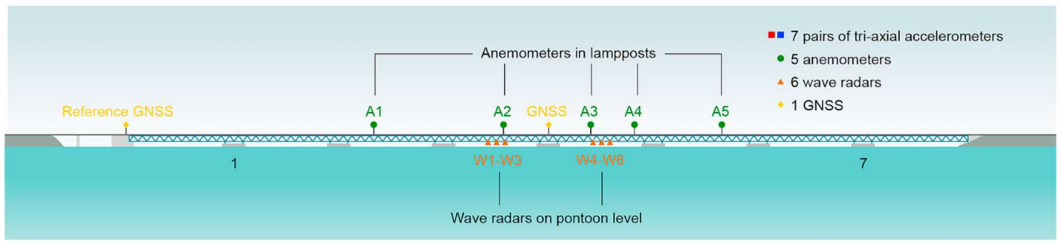



Figure 1. Location of the Bridge [5]




(a) View from the top, including annotations for the north direction and the global coordinate system. $z = 0$ on the bridge deck.




(b) View from the side.

Figure 2. All Installed Sensors on the Bridge [6]

It is also worth noting that among the world's ten longest floating bridges, the United States contributes four: the Evergreen Point Floating Bridge, the Hood Canal Bridge, the Lacey V. Murrow Memorial Bridge, and the Homer M. Hadley Memorial Bridge, all located in the Seattle area. Three of these were constructed in the 20th century, and all remain in active use. The main reasons for adopting the floating design in the Seattle region, similar to the conditions at the Bergsøysund Bridge, are the deep water depth, soft sediment, and additionally, the seismic risk. Long-term use often necessitates increased monitoring to ensure proper maintenance. The scenario presented in this research has strong potential for application to other floating bridges requiring real-time monitoring.

## DATA DESCRIPTION

The GNSS data were collected using a Trimble RTK GNSS receiver at a sampling rate of 20 Hz [5]. The location of GNSS receiver with the simplified bridge model is shown in Figure 3. In the GNSS dataset, the origins of the X and Z axes are set at the left corner of the bridge. The specific GNSS receiver model is not indicated in the dataset, although Trimble RTK GNSS products are among the most precise positioning technologies available. The vertical and horizontal accuracy of these products is generally 2 centimetres and 1 centimetre, respectively, under proper operating conditions. Precision can reach sub-centimetre levels. For long-term continuous monitoring, accuracy affects the ability to detect the magnitude of measurements, while precision influences the detection of trends over time. The accurate and precise RTK GNSS hardware used in this dataset ensures that the data can reliably represent both the changes and trends at the point where the receiver is mounted.

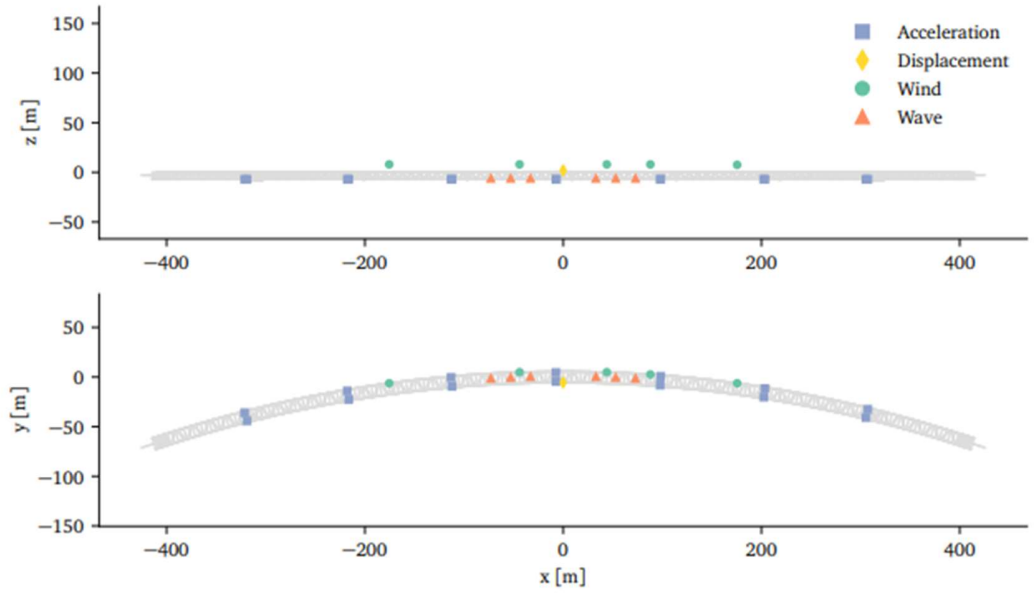



Figure 3. Sensor Positions with Simplified Bridge Model, where Displacment Refers to the GNSS Receiver [5]

## METHOD

GNSS data are continuously collected on-site, forming a time-series dataset with temporal dependencies. In practice, missing data usually account for only a small proportion of the entire dataset. Considering these temporal dependencies and the low proportion of missing data, LSTM-based methods could be effective for data imputation. The main principle of LSTM method as depicted in Figure 4, is to learn and remember long-term dependencies in the sequential data by using gates to decide what to keep, forget and output at each time step [7]. In this study, a standard unidirectional LSTM is adopted to process data in one direction, forward in time.

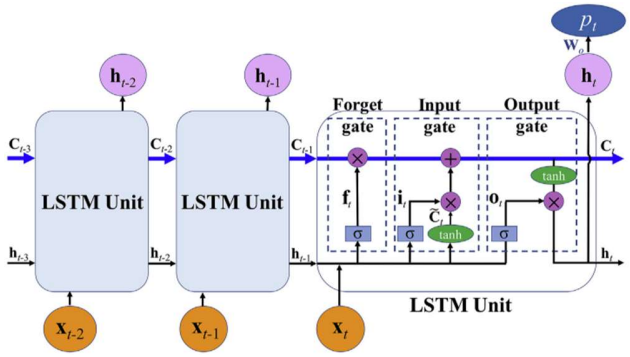



Figure 4 The Main Principle of LSTM Method *[8]*.

The complete GNSS dataset (2Hz version) from the bridge is then processed by randomly removing 10% of the data to simulate a severe random missing-data scenario. The data spans the period from 2015 to 2018. After filtering, the remaining 90% of the data are trained using the LSTM method. The imputed data are then compared with the original dataset prior to filtering to assess the performance of the LSTM method in imputing the bridge's GNSS data.

## RESULTS AND DISCUSSION

The key indicators for assessing the performance of the imputation model are summarised in Table 1. The model demonstrates high accuracy across all coordinate axes, although performance on the Z-axis is comparatively lower. The Mean Absolute Error (*MAE*) for the X-axis is as low as 0.0003 metres. The *MAE* values for the Y- and Z-axes are within the millimetre range, with the Z-axis exhibiting a larger average error of 6.4 millimetres. The Root Mean Squared Error (*RMSE*) follows a similar pattern to the *MAE*, ranking the axes from most to least accurate in the same order. Notably, the *RMSE* for the Z-axis occasionally exceeds 1 centimetre, reaching up to several centimetres in some cases. The coefficient of determination ($R^2$) for all axes exceeds 0.97, approaching the ideal value of 1, which signifies a strong fit and high imputation quality. The percentage of errors within 1 centimetre (*P1*) shows a decreasing trend

from the X-axis to the Z-axis, indicating that the Z-axis has a broader distribution of larger errors. Nevertheless, the *P1* values for all three axes exceed 95%, confirming that the vast majority of imputed values are highly accurate and fall within 1 centimetre of the ground truth. Figure 5 illustrates the plot of imputed values compared against the actual values. The observations in Figure 5 are consistent with the findings presented in Table 1. It is noteworthy that the trained model demonstrates significantly better imputation quality for data exhibiting periodic behaviour, while showing limited ability to impute data with large variance. Besides, a simple and fast baseline method, K-Nearest Neighbours (KNN) imputation, was applied to the same dataset. The LSTM model demonstrated significantly better performance than KNN imputation across all performance metrics.

Table 1 Performance Metrics for Each Axis

| | ***MAE*** | ***RMSE*** | $R^2$ | **Percentage of error within 1 cm (*P1*)** |
|---|---|---|---|---|
| **X** | 0.0003 | 0.0014 | 0.9733 | 99.60% |
| **Y** | 0.0014 | 0.0118 | 0.9924 | 98.16% |
| **Z** | 0.0064 | 0.0618 | 0.9895 | 95.42% |

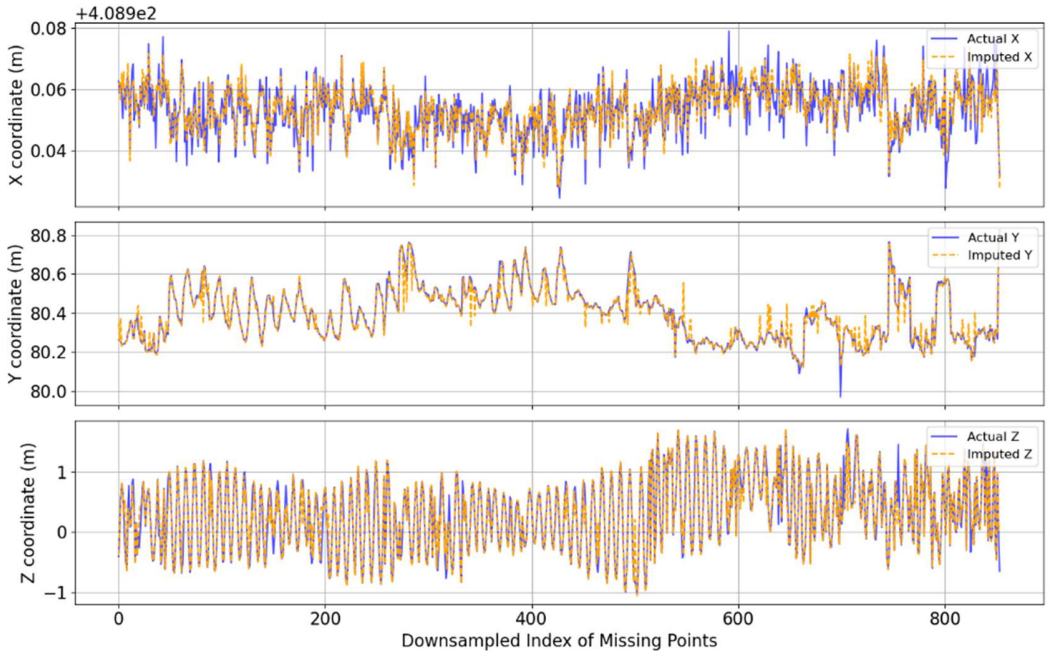


Figure 5 LSTM-Predicted versus Actual GNSS Coordinates

One main limitation of this research is that the method for simulating missing GNSS data involved randomly omitting data from the complete GNSS dataset. The scenario of long-duration data loss (e.g., lasting hours or days) was not simulated. Further investigation is needed to address cases involving extended continuous periods of missing data.

Another limitation is that the raw GNSS data, such as pseudorange, carrier phase, and elevation angle, are not available, which limits the potential for GNSS-specific

enhancements to imputation accuracy using LSTM. If such data were available, these GNSS-related features could be utilised to help the model understand satellite geometry and signal quality, thereby filtering noise, downweighting unreliable signals, and improving imputation accuracy.

The final limitation that needs to be emphasised is the necessity for further investigation to compare the overall SHM performance when using complete GNSS data versus imputed data. Although the imputation in this research has demonstrated satisfactory performance in terms of *MAE*, *RMSE*, $R^2$ and *P1*, additional validation is required by incorporating data from other sensors on the bridge to assess the overall SHM performance.

## CONCLUSION

This study assessed the LSTM method for GNSS data imputation using on-site data from a receiver mounted on the Bergsøysund Bridge in Norway. Missing data were simulated by randomly removing observations, with a missing rate set at 10%. The LSTM method demonstrated effective imputation performance on the dataset, with key performance indicators summarised in Table 1. Further work is recommended to investigate the longer data gaps and integrating additional sensor data to improve imputation quality.

## REFERENCES


[1] R. Xi, Q. He, and X. Meng, 'Bridge monitoring using multi-GNSS observations with high cutoff elevations: A case study', *MEASUREMENT*, vol. 168, Jan. 2021, doi: 10.1016/j.measurement.2020.108303.

[2] A. Angrisano, M. Petovello, and G. Pugliano, 'Benefits of Combined GPS/GLONASS with Low-Cost MEMS IMUs for Vehicular Urban Navigation', *SENSORS*, vol. 12, no. 4, pp. 5134–5158, Apr. 2012, doi: 10.3390/s120405134.

[3] C. Zhao, Z. Jiang, B. J. Leira, and Z. Shi Xuzhao and Zheng, 'Dynamic response characteristics of the 67-type railway pontoon bridge', *DISCOVER APPLIED SCIENCES*, vol. 4, no. 11, Nov. 2022, doi: 10.1007/s42452-022-05178-7.

[4] C. Wang, M. Cui, Z. Cheng, and T. Moan, 'A review on design and analysis of floating bridges: Numerical simulations, model tests and field measurements', *OCEAN ENGINEERING*, vol. 306, Aug. 2024, doi: 10.1016/j.oceaneng.2024.118065.

[5] K. A. Kvale, A. Fenerci, O. W. Petersen, A. Ronnquist, and O. Oiseth, 'Data Set from Long-Term Wave, Wind, and Response Monitoring of the Bergsoysund Bridge', *JOURNAL OF STRUCTURAL ENGINEERING*, vol. 149, no. 9, Sep. 2023, doi: 10.1061/JSENDH.STENG-12095.

[6] K. A. Kvale and O. Oiseth, 'Structural monitoring of an end-supported pontoon bridge', *MARINE STRUCTURES*, vol. 52, pp. 188–207, Mar. 2017, doi: 10.1016/j.marstruc.2016.12.004.

[7] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, 'LSTM: A Search Space Odyssey', *IEEE Trans Neural Netw Learn Syst*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017, doi: 10.1109/TNNLS.2016.2582924.

[8] X. Wei, L. Zhang, H.-Q. Yang, L. Zhang, and Y.-P. Yao, 'Machine learning for pore-water pressure time-series prediction: Application of recurrent neural networks', *GEOSCIENCE FRONTIERS*, vol. 12, no. 1, pp. 453–467, 2021, doi: https://doi.org/10.1016/j.gsf.2020.04.011.