JOURNAL OF
HYDRO-METEOROLOGY

Research Article

Exploring the training results of machine learning models using different batch sizes and epochs: A case study with GNSS time series data

Le Duc Tinh[1], Huynh Nguyen Dinh Quoc[2], Nguyen Gia Trong[1,3*]

[1] Faculty of Geomatics and Land Administration, Hanoi University of Mining and Geology; leductinh@humg.edu.vn , nguyengiatrong@humg.edu.vn
[2] Ho Chi Minh City University of Natural Resources and Environment; hndquoc@hcmunre.edu.vn
[3] Geodesy and Environment research group, Hanoi University of Mining and Geology;

*Corresponding author: nguyengiatrong@humg.edu.vn ; Tel.: +84–963124980

Received: 25 March 2024; Accepted: 03 May 2024; Published: 25 June 2024

Abstract: This study applies the GRU (Gated Recurrent Unit) model when selecting different values of batch-size, namely 16, 32, and 64, with varying epochs of 20, 50, 100, 150, and 200. The input data comprises observations collected by two GNSS CORS stations from the VNGEONET network, namely HYEN and CTHO, spanning from August 10, 2019, to March 18, 2022. Initially, GNSS CORS data is processed using Gamit/Globk software to obtain the Up-component, which serves as the input data for the GRU model. The research results indicate that the statistical performance metrics of the model, such as RMSE and MAE, decrease while the F-Score increases when the batch-size decreases and the epoch value increases. In cases where the Up-component exhibits irregular variations (seasonal fluctuations), the performance of the GRU model is subpar, with an F-Score of 0 observed when batch-size values are 32 and 64 and epoch value is 20. For data following the pattern of CTHO CORS station, the GRU model performs exceptionally well when batch-size is 16 and epoch is 200. However, the forecasting performance is low for data from HYEN CORS station, indicating the need for further investigation in the future.

Keywords: Artificial Intelligence; Batch size; Epoch; GRU; GNSS time-series.

1. Introduction

Due to technological advancements, data collection has been automated, continuous, or temporally dense, resulting in various types of time series data. Time series data in geospatial applications include GNSS, satellite altimetry, remote sensing data, etc. GNSS data has been applied in atmospheric layer research, oceanic observations, soil moisture monitoring, ice sheet observations [1], and tectonic plate movements [2], etc. Altimetry time series data is utilized in various marine activities, monitoring marine life, weather and climate forecasting, coastal inundation monitoring due to sea level rise or subsidence, natural disaster mitigation, etc. [3]. Remote sensing time series data applications include land cover classification [4], forest monitoring [5], erosion studies [6], etc. Research [7] has highlighted the extensive applications of artificial intelligence in large geospatial datasets, quality assessment, data modeling and structuring, data visualization and visual analytics, data mining, and knowledge discovery, etc. With the establishment continuosly operation reference station (CORS), users are provided with time-series data. The data provided by CORS station networks can be applied in various fields such as tectonic plate movement monitoring, sea

level monitoring, atmospheric research [8] etc. Data collected by CORS stations firstly need to be processed using high-precision GNSS data processing software such as Gamit/Globk [9], Bernese [10] etc. The result of this processing is the daily coordinate components of CORS stations. To analyze the daily time-series data obtained as mentioned above, various traditional solutions have been published, such as least-squares estimation, moving ordinary least-squares wavelet decomposition (WD), singular spectrum analysis (SSA), Kalman Filter (KF), adaptive wiener filter (AWF), or combinations thereof [11]. Additionally, artificial intelligence models have also been applied to analyze GNSS time-series data.

Five artificial intelligence models, namely attention mechanism with long short time memory neural network (AMLSTM), long short time memory neural network (LSTM), recurrent neural networks (RNN), support vector machine (SVM), and random forest (RF), have been utilized for landslide detection [12]. Additionally, the authors proposed combining the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) technique with the LSTM model. Experimental results demonstrated that the CEEMDAN-LSTM model could be recommended for other landslide prediction studies and has significant potential in landslide risk assessment. The study [13] applied Gradient Boosting Decision Tree (GBDT), LSTM, and SVM models to analyze GNSS time-series data. The results showed that the proposed models had RMSE values ranging from 3mm to 5mm, smaller than those of corresponding traditional methods. Moreover, artificial intelligence models enable the integration of various factors causing in land surface movement from GNSS time series data [13].

The GBDT model has been used as a benchmark against the XGBoost and RF models for interpolating coordinate values in GNSS time-series data. The computational results indicate that the Up-component is interpolated with up to 45% greater accuracy compared to traditional methods, with the XGBoost model yielding the poorest interpolation results [14]. For each different setting of batch size and epoch, different artificial intelligence models will yield different prediction results. Research [15] has compiled errors for both training and testing datasets using batch sizes ranging from 64 to 2048. The results indicate that a batch size of 128 yields the smallest error for both datasets. Additionally, research [16] has shown that the noise level increases as the loss decreases during training and largely depends on the model size, with model performance being improved.

Research [17] has demonstrated that training with small batches has been proven to improve generalization performance and allows for significantly smaller memory usage, which can also be leveraged to enhance machine throughput. Nesterov and Adam optimizers have been found to train more efficiently than baseline models when using large batch sizes. There have been several publications on the application of artificial intelligence in analyzing time-series data. Author [18] employed a recurrent neural network to forecast meaningful wave heights for disaster prevention efforts in Vietnam. Research [19] applied an ANN model to analyze GNSS time-series data, yielding an RMSE determination of 0.006m. RNN models were chosen by authors [20] to predict surface water quality with an accuracy ranging from 75% to 85%. Although some studies have been published, there has been no research in the earth science field specifically addressing experimentation with different batch sizes and epochs to analyze time-series data.

This study evaluates the performance of the GRU model by configuring different values of batch size and epoch during model execution, applied in a typical case of analyzing GNSS time-series data.

2. Data and Research methodology

2.1. Data

The data used in this paper was collected by continuosly operation reference station (CORS) receivers, belonging to the VNGEONET network, namely CTHO and HYEN,

provided by the Department of Survey and Mapping, Vietnam. Information regarding the GNSS data used in this study is provided as shown in Table 1.

Table 1. Information about measurement data at GNSS CORS stations.

Station name	Time		Receiver type	Antenna type	Interval (second)
	First epoch	Last epoch			
CTHO HYEN	2019/08/10	2022/03/18	LEICA GR50	LEIAR25.R4 LEIT	30

The positions of the HYEN CORS station and the CTHO CORS station are depicted as shown in Figure 1.

Figure 1. The positions of the HYEN CORS station and the CTHO CORS station.

2.2. Methodology

The data as described in Table 1 was processed using Gamit/Globk software to obtain daily coordinate components (including the North, East, and Up components) of the GNSS CORS stations. The up-component value series was then utilized as the input data for the GRU model. The research methodology of the paper is provided in Figure 2.

The GNSS data, once collected, will be converted into RINEX data format and analyzed using the Gamit/Globk software [21] to obtain daily coordinate components of the stations. The GNSS data processing procedure in this case has been presented in the study [22].

The Gated Recurrent Unit (GRU) is a type of artificial neural network model, particularly suited for sequential data processing tasks such as natural language processing and time series analysis. Developed as an enhancement of the traditional recurrent neural network (RNN), the GRU addresses some of the shortcomings of the standard RNN architecture, particularly in handling long-range dependencies and the vanishing gradient problem.

One of the key features of the GRU is its gating mechanism, which allows it to selectively update and forget information over time. This mechanism consists of update and reset gates, which regulate the flow of information within the network. By adaptively controlling the flow of information, the GRU is able to capture relevant patterns and dependencies in sequential data

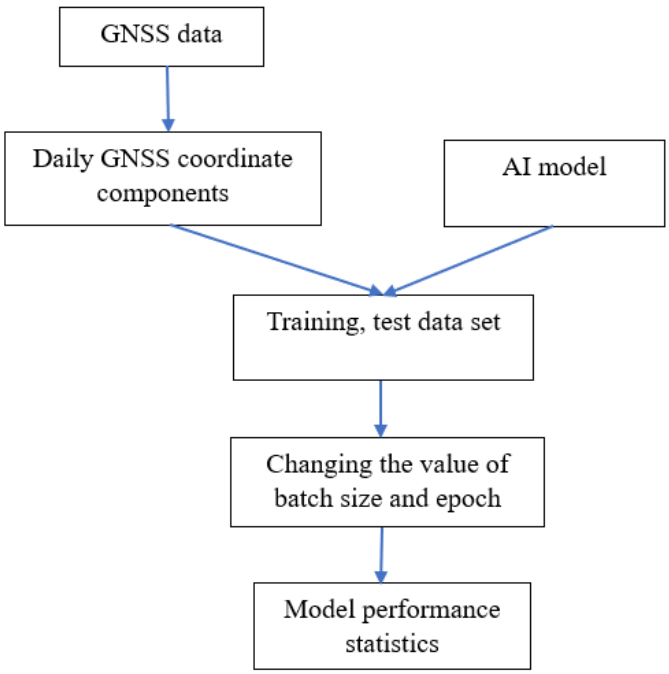

Figure 2. Experimental method of artificial intelligence model testing with different batch sizes and epochs.

more effectively. Compared to the long short-term memory (LSTM) model, another popular variant of the RNN architecture, the GRU offers similar performance with a simpler structure, requiring fewer parameters to train and often achieving faster convergence during training. Its computational efficiency and competitive performance make it a popular choice for various sequence modeling tasks in both research and practical applications.

Based on the selected model, Python language and library functions were utilized to construct the experimental computation program [23, 24], etc.

To achieve the desired results, the research team conducted experiments with batch sizes set to 16, 32, and 64, and for epochs, values were assigned as 10, 50, 100, 150, and 200, respectively. Model evaluation was performed by statistically analyzing performance metrics including RMSE, MAE, and F-score. To assess the performance of the model, evaluation methods similar to those used in studies [18, 19] were employed. The operation method of the GRU model in this case is depicted as shown in Figure 3.

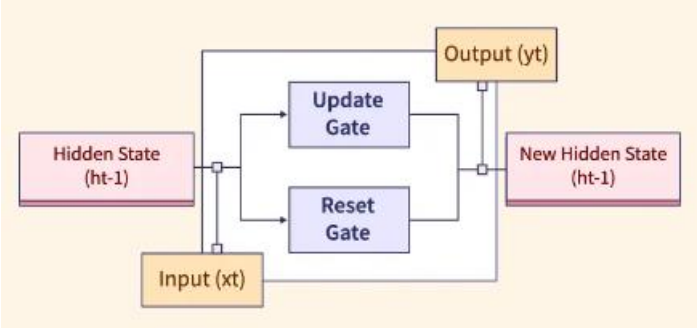

Figure 3. Prediction method with the GRU model.

3. Results and Discussion

3.1. Results obtained with the dataset from HYEN CORS station

From the data in Table 2, it can be observed that the RMSE value increases rapidly when the epoch value is small and the batch-size value increases from 16 to 64. For the same batch-size value, as the epoch value increases, the RMSE value decreases. In the case of the largest batch-size value (with a value of 64), the RMSE value decreases very rapidly. The minimum RMSE value for HYEN CORS station is achieved at 0.010 when batch-size = 16 and epoch = 200.

Table 2. RMSE determination results with data from HYEN CORS station.

Batch size	Epoch				
	20	**50**	**100**	**150**	**200**
16	0.543	0.487	0.248	0.017	0.010
32	5.626	0.557	0.541	0.337	0.103
64	13.394	0.611	0.651	0.606	0.512

Table 3. MAE determination results with data from HYEN CORS station.

Batch size	Epoch				
	20	**50**	**100**	**150**	**200**
16	0.469	0.379	0.192	0.014	0.006
32	5.618	0.429	0.420	0.261	0.079
64	13.393	0.468	0.506	0.471	0.398

The variation in the MAE values for the input data from HYEN CORS station is similar to RMSE. The minimum MAE value achieved is 0.006, corresponding to a batch-size of 16 and epoch of 200.

In AI applications, the F-score, also known as the F1-score, is a metric commonly used to evaluate the performance of a binary classification model. It is the harmonic mean of precision and recall, providing a single measure that balances between these two metrics. Precision measures the proportion of true positive predictions among all positive predictions, while recall measures the proportion of true positive predictions among all actual positives. The F-score ranges from 0 to 1, where a higher score indicates better performance. It's particularly useful when the class distribution is imbalanced, as it considers both false positives and false negatives. The determined F-Score results in this case are as follows:

Table 4. Results of F-Score determination with data from HYEN CORS station.

Batch size	Epoch				
	20	**50**	**100**	**150**	**200**
16	0.312	0.181	0.417	1.000	1.000
32	0.000	0.181	0.179	0.265	0.473
64	0.000	0.146	0.174	0.174	0.179

Table 4 demonstrates the very high performance of the model when selecting batch-size = 16 and epoch = 200; when batch-size is set to 32 or 64, with epoch = 20, the model's predictive performance equals 0. This aligns perfectly with the significantly large RMSE and MAE values. Figures 4, 5, 6 below represent the predicted values, actual values on the test dataset, the entire dataset, and the loss curve in the case of batch-size = 16 and epoch = 200 for the HYEN CORS station dataset.

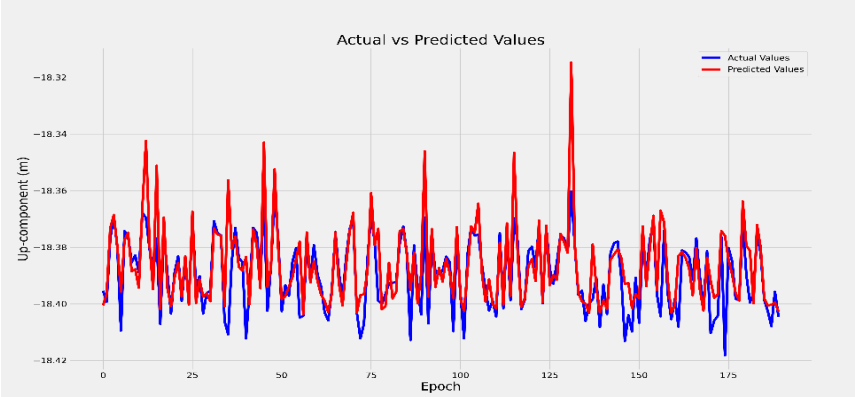

Figure 4. Predicted values on the test dataset of HYEN CORS station

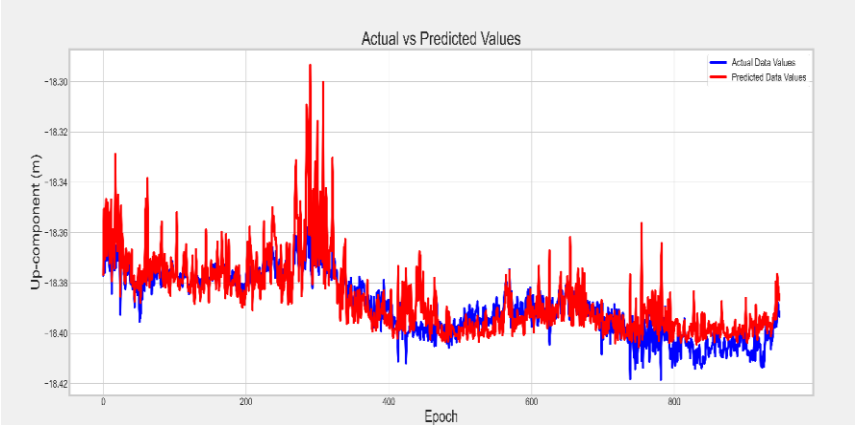

Figure 5. Actual and predicted values on the entire dataset of HYEN CORS station.

From Figure 4, it can be observed that the predicted values are significantly higher than the actual values. There are several factors that may contribute to this phenomenon, including unusual fluctuations in the daily Up-component. To accurately conclude on the

aforementioned phenomenon, it is necessary to gather additional closely related data concerning the variations of the Up-component, such as meteorological data, hydrological and geological data, etc.

Figure 6. Loss curve on the test dataset of HYEN CORS station.

From Figure 5, it can be observed that the variation in the Up-component displacement of HYEN CORS station does not exhibit periodicity as some published results have indicated [25]. Figure 7 shows the Up-component results of the HYEN CORS station determined using the Gamit/Globk software, serving as evidence for the arguments presented above.

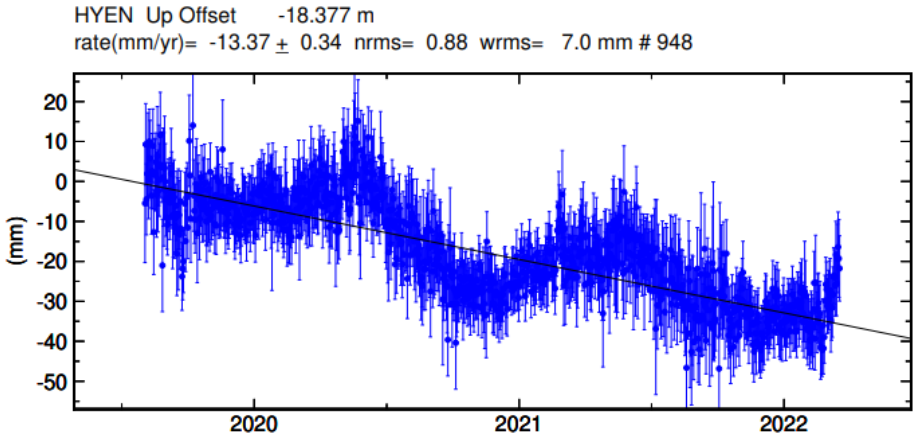

Figure 7. Up-component results of the HYEN CORS station were determined by the Gamit/Globk software.

In Figure 7, the vertical axis represents the daily changes in the up-component (measured in mm), while the horizontal axis represents time (measured in years).

This may lead to the predictive performance of the artificial intelligence model in this case not being as high as the results achieved even with the use of a simple artificial intelligence model [19].

3.2. Results obtained with the dataset from CTHO CORS station

Table 5. RMSE determination results with data from CTHO CORS station.

Batch size	Epoch				
	20	**50**	**100**	**150**	**200**
16	0.1523	0.0870	0.00154	0.00098	0.00074
32	0.1632	0.1513	0.0469	0.00136	0.00129
64	0.2668	0.1846	0.1410	0.07414	0.00213

From Table 5, it can be observed that the RMSE values decrease only slightly as the batch-size varies from 16 to 64. This indicates the suitability of the artificial intelligence model being employed.

Table 6. MAE determination results with data from CTHO CORS station.

Batch size	Epoch				
	20	**50**	**100**	**150**	**200**
16	0.1175	0.0686	0.0010	0.00062	0.00053
32	0.1255	0.1179	0.0370	0.00084	0.00058
64	0.2303	0.1424	0.1110	0.05866	0.00091

Table 7. F-Score determination results with data from CTHO CORS station.

Batch size	Epoch				
	20	**50**	**100**	**150**	**200**
16	0.5573	0.5653	0.95216	0.9621	0.9589
32	0.5490	0.5447	0.55288	0.9522	0.9504
64	0.2312	0.5367	0.54468	0.5778	0.9256

From the data in Table 5 to Table 7, it can be observed that the fluctuation trends of RMSE, MAE, and F-Score values for the data from CTHO CORS stattion are similar to those for HYEN CORS station. However, the variation of the Up-component for CTHO CORS station is similar to previous publications (Figure 8).

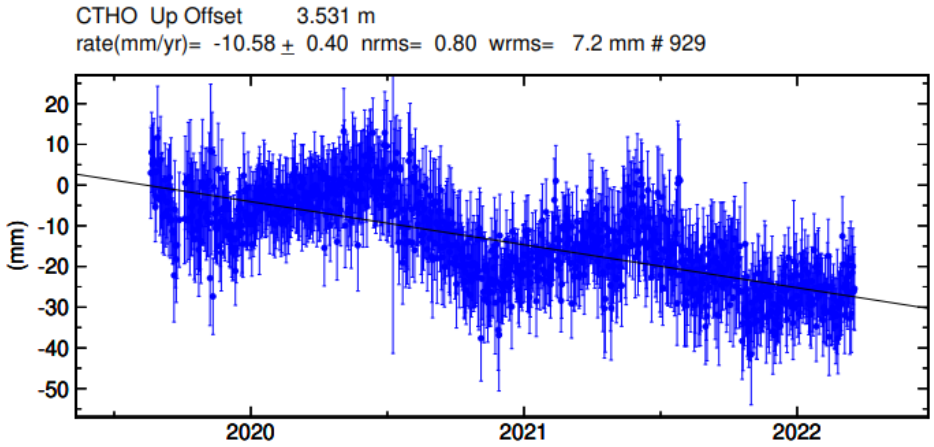

Figure 8. Graph of Up-component of CTHO CORS station determined by Gamit/Globk.

The up-component values in Figure 8 demonstrate a systematic variation of this component for the CTHO CORS station, explaining the F-score values of 0.5490 and 0.2312 respectively for batch-size = 32 and 64, epoch = 20. The RMSE = 0.00074 and MAE = 0.00053 for batch-size = 16 and epoch = 200. These indicate very high performance in predicting the determined quantity from the GNSS time-series data compared to existing publications [13,19,26]. These are promising preliminary results in the application of artificial intelligence models for analyzing GNSS time-series data.

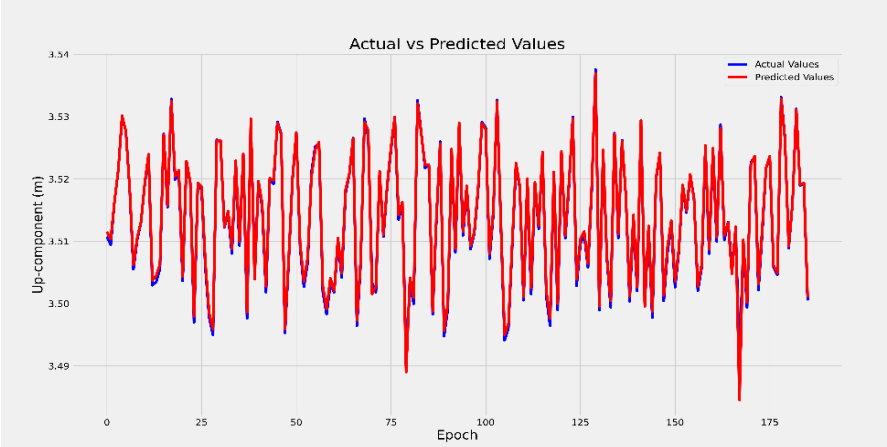

Figure 9. Predicted and actual values for the test dataset of the CTHO CORS station.

The graphs representing the actual values, predicted values on the test dataset, the entire dataset, and the loss curve when analyzing the CTHO CORS station data in the case of batch-size = 16 and epoch = 200 are shown in Figures 9 to 11.

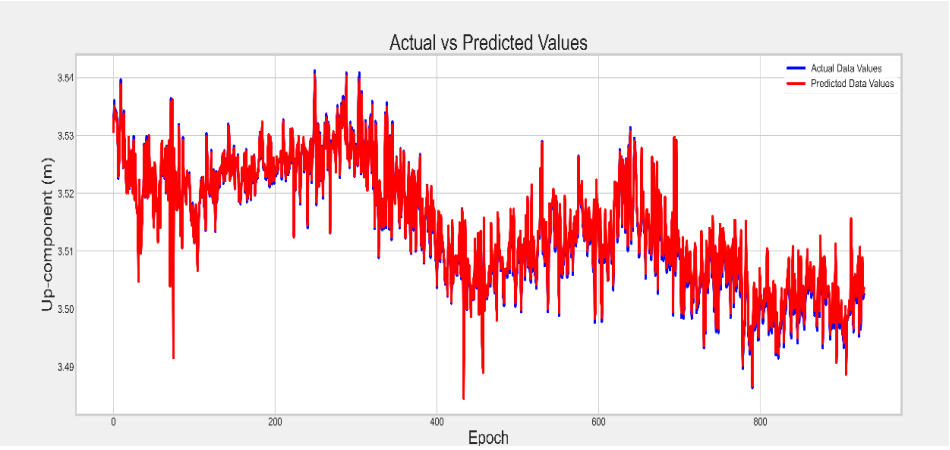

Figure 10. Predicted and actual values for the entire dataset of the CTHO CORS station.

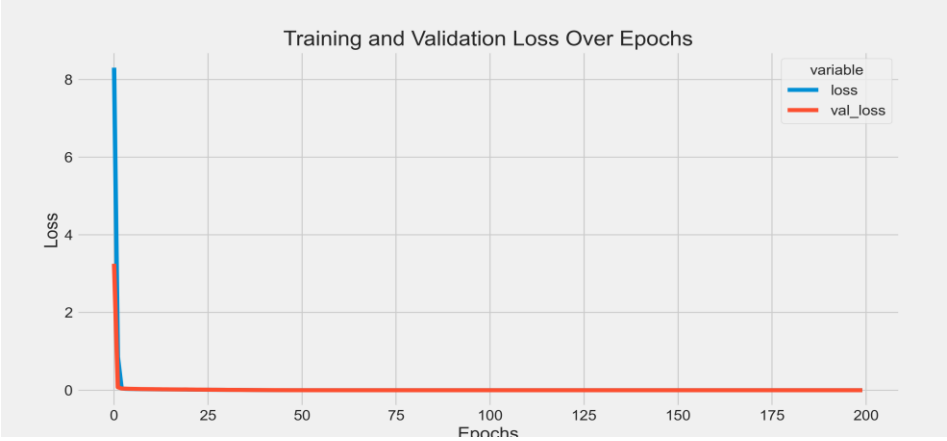

Figure 11. Loss curve of the test dataset of the CTHO CORS station.

Figures 9 and 10 demonstrate the very high suitability of the GRU model for the data of the CTHO CORS station, as evidenced by the predicted values closely matching the actual values across the datasets. Additionally, the loss value nearly approaches zero after a few epochs (Figure 11).

4. Conclusion

This study successfully experimented with different values of batch-size and epoch when analyzing GNSS time-series data using the artificial model. When applying artificial intelligence (specifically the GRU model) to analyze GNSS time series, it is necessary to select a small batch size value (specifically 16) to achieve the best forecasting performance with the model.

The experimental results demonstrate that when the up-component varies irregularly, unlike in the case of HYEN CORS station, predicting using artificial intelligence models is not very effective. Specifically, when the epoch value is 20 and the batch-size is 32 and 64 respectively, the GRU model cannot predict the Up-component values in this case, as indicated by an F-Score of 0.

Statistical metrics such as RMSE and MAE decrease, while F-Score increases (indicating improved prediction performance with artificial intelligence models) as the batch-size decreases and epoch increases. The prediction results show very high performance for the

CTHO CORS station dataset when batch-size = 16 and epoch = 200, demonstrated by metrics such as RMSE = 0.00074, MAE = 0.00053, and F-Score = 0.9589. This performance is excellent compared to existing publications.

One limitation of this study is that it does not propose a solution for forecasting or analyzing the Up-component in cases where the variation is irregular, as in the case of the HYEN CORS station. This is an issue that requires further investigation in the future.

Author contributions: Conceptualization: L.D.T., N.G.T., H.N.D.Q.; Methodology: L.D.T., N.G.T., H.N.D.Q.; Data processing: N.G.T., N.N.D.Q.; Writing - original draft: L.D.T., N.G.T., H.N.D.Q.; Writing - review and editing: N.G.T.

Declaration: The authors collectively declare that this article is the result of their research, not previously published elsewhere, and not copied from previous studies; there is no conflict of interest among the authors.

Acknowledgments: The authors of this paper sincerely thank the Department Of Survey, Mapping and Geographic Information Vietnam for providing the data; the Ministry of Education and Training under project code B2022-MDA-09 for providing the funding to conduct this research.

References

1. Yu, K.; Rizos C.; Burage, D.; Dempster, A.G.; Zhang, K.; Markgraf, M. An overview of GNSS remote sensing. *EURASIP J. Adv. Signal Process.* **2014**, *134*, 1–14.
2. Bastos, L.; Bos, M.; Fernandes, R.M. Deformation and tectonics: contribution of GPS measurements to plate tectonics–overview and recent developments. *Sci. Geodesy-I.* **2010**, 155–184.
3. Srinivasan, M.; Tsontos, V. Satellite altimetry for ocean and coastal applications: A review. *Remote Sens.* **2023**, *15(16)*, 3939. https://doi.org/10.3390/rs15163939.
4. Gómez, C.; White, J.C.; Wulder, M.A. Optical remotely sensed time series data for land cover classification: A review. *ISPRS J. Photogramm. Remote Sens.* **2016**, *116*, 55–72.
5. Banskota, A.; Kayastha N.; Falkowski M.J.; Wulder M.A.; Froese R.E.; White J.C. Forest monitoring using Landsat time series data: A review. *Can. J. Remote Sens.* **2014**, *40(5)*, 362–384.
6. Vrieling, A.J.C. Satellite remote sensing for water erosion assessment: A review. *Catena* **2006**, *65(1)*, 2–18.
7. Li, S.; Dragicevic, S.; Castro, F.A.; Sester, M.; Winter, S.; Coltekin, A.; Pettit, A.; Jiang, B.; Haworth, J.; Stein, A.; Cheng, T. Geospatial big data handling theory and methods: A review and research challenges. *ISPPS J. Photogramm. Remote Sens.* **2016**, *115*, 119–133.
8. Trong, N.G.; Tinh, L.Đ.; Cuong, N.V.; Quang, P.N. GNSS Data Processing: Theory, Software, and Applications. Transport Publishing House, 2023, pp. 300.
9. Li, Y. Analysis of GAMIT/GLOBK in high-precision GNSS data processing for crustal deformation. *Earthquake Res. Adv.* **2021**, *1(8-11)*, 100028. https://doi.org/10.1016/j.eqrea.2021.100028.
10. Cetin, S.; Aydin, C.; Dogan, U. Comparing GPS positioning errors derived from GAMIT/GLOBK and Bernese GNSS software packages: A case study in CORS-TR in Turkey. *Surv. Rev.* **2019**, *51(369)*, 533–543.
11. Klos, A.; Bogusz, J.; Bos, M.S.; Gruszczynska, M. Modelling the GNSS time series: different approaches to extract seasonal signals. *Geodetic Time Ser. Anal. Earth Sci.* 2020, pp. 211–237.
12. Wang, J.; Nie, G.; Gao, S.; Wu, S.; Li, H.; Ren, X. Landslide deformation prediction based on a GNSS time series analysis and recurrent neural network model. *Remote Sens.* **2021**, *13(6)*, 1055. https://doi.org/10.3390/rs13061055.

13. Gao, W.; Li, Z.; Chen, Q.; Jiang, W.; Feng, Y. Modelling and prediction of GNSS time series using GBDT, LSTM and SVM machine learning approaches. *J. Geod.* **2022**, *96(10)*, 71. https://doi.org/10.1007/s00190-022-01662-5.
14. Li, Z.; Lu, T.; Yu, K.; Wang, J. Interpolation of GNSS position time series using GBDT, XGBoost, and RF machine learning algorithms and models error analysis. *Remote Sens.* **2023**, *15(18)*, 4374. https://doi.org/10.3390/rs15184374.
15. Hoffer, E.; Hubara, I.; Soudry, D. Train longer, generalize better: Closing the generalization gap in large batch training of neural networks. *Adv. Neural Inf. Process. Syst.* **2017**, 1–13.
16. McCandlish, S.; Kaplan, J.; Amodei, D.; Team, O.D. An empirical model of large-batch training. *ArXiv* **2018**, 1–35. https://doi.org/10.48550/arXiv.1812.06162.
17. Masters, D.; Luschi, C. Revisiting small batch training for deep neural networks. *Comput. Sci. Mach. Learn.* **2018**, 1–18. https://doi.org/10.48550/arXiv.1804.07612.
18. Thai, T.H.; Khiem, M.V.; Thuy, N.B.; Ha, B.M.; Ngoc, P.K. Building a regression neural network model to predict significant wave heights at Con Co station, Quang Tri, Vietnam. *J. Hydro-Meteorol.* **2022**, *EME4*, 73–84.
19. Phong, D.V.; Trong, N.G.; Chien, N.V.; Thanh, N.H.; Ha, L.L.; Quan, N.V.; Quang, P.N. Analysis of land vertical movement using ANN function from the results of processing GNSS time series data. *J. Hydro-Meteorol.* **2023**, *752*, 41–50.
20. Phong, N.D.; Duong, H.H. Application of deep learning models in forecasting surface water quantity of Bac Hung Hai irrigation system. *Water Resour. Mag.* **2023**, *1*, 61–72.
21. Available online: https://geoweb.mit.edu/gg/.
22. Trong, N.G.; Nghia, N.V.; Khai, P.C.; Thanh, N.H.; Ha, L.L.; Dung, V.T.; Quan, N.V.; Quang, P.N. Determination of tectonic velocities in Vietnam territory based on data of CORS stations of VNGEONET network. *J. Hydro-Meteorol.* **2022**, *739*, 59–66.
23. Available online: https://www.python.org/.
24. Available online: https://anaconda.org/anaconda/pandas.
25. Savchuk, S.; Doskich, S.; Golda, P.; Rurak, A. The Seasonal Variations Analysis of Permanent GNSS Station Time Series in the Central-East of Europe. *Remote Sens.* **2023**, *15(15)*, 3858. https://doi.org/10.3390/rs15153858.
26. Carbonari, R.; Riccardi, U.; Martino, P.D.; Cecere, G.; Maio, R.D. Wavelet-like denoising of GNSS data through machine learning. Application to the time series of the Campi Flegrei volcanic area (Southern Italy). *Geomatics Nat. Hazards Risk* **2023**, *14(1)*, 2187271. https://doi.org/10.1080/19475705.2023.2187271.