GNSS Time-Series Prediction Using Moving Average Filter and Multi-Layer Perceptron Neuron Network

Tuan Minh DO[1)], Huynh Dinh Quoc NGUYEN[1)], Quang Ngoc PHAM[2,3)], Duc Tinh LE[2)], Long Quoc NGUYEN[2,4)], Van Anh TRAN[2,4)], Trong Nguyen GIA[2,3)]*

[1)] Ho Chi Minh city of Natural Resources and Environment, Hochiminh city, Vietnam; ORCID https://orcid.org/0000-0003-3172-5330; ORCID https://orcid.org/0009-0007-8447-9045;
[2)] Faculty of Geomatics and Land Administration, Hanoi University of Mining and Geology, Hanoi, Vietnam; ORCID https://orcid.org/0009-0006-0765-245X; ORCID https://orcid.org/0000-0003-2311-7351; ORCID https://orcid.org/0000-0002-0022-3453; ORCID https://orcid.org/0000-0002-4792-3684; ORCID https://orcid.org/0009-0003-1616-8625;
[3)] Geodesy and Environment research group, Hanoi University of Mining and Geology, Hanoi, Vietnam
[4)] Innovations for Sustainable and Responsible Mining (ISRM) Research Group, Hanoi University of Mining and Geology, Hanoi, Vietnam
* Corresponding author: nguyengiatrong@humg.edu.vn

http://doi.org/10.29227/IM-2024-02-98

Submission date: 02-11-2024 | Review date: 03-12-2024

Abstract

The Mekong Delta and Ho Chi Minh City in Vietnam are recognized as areas significantly impacted by land subsidence. This phenomenon has led to notable consequences, including increased vulnerability to issues such as saline intrusion and tidal flooding. GNSS-CORS technology, known for its capability to provide continuous time-series data, plays a crucial role in accurately monitoring changes in the land surface. Despite the existence of traditional algorithms for analyzing continuous measurement data collected through GNSS-CORS technology, their effectiveness is constrained by challenges in handling diverse input data and limitations in forecasting future displacements. Consequently, there is a growing trend towards the adoption of artificial intelligence techniques, particularly artificial neural networks (ANN), for predicting Up component in GNSS time-serries daily solution. This study leverages data from the CTHO GNSS CORS station located in the Mekong Delta to evaluate proposed models. An innovative hybrid approach, which integrates the Moving Average Filter (MAF) and Multilayer Perceptron Neural Network (MLPNN), is introduced to enhance the accuracy of forecasting. Performance evaluation metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) are utilized to assess the effectiveness of the models. Results demonstrate the superior performance of the MLPNN model, achieving high prediction accuracy with metrics including MAE = 0.001, MSE = 0.000, and RMSE = 0.002. This research underscores the robustness of the proposed model in forecasting GNSS time-serries daily solution, highlighting its potential for practical applications in geodetic research.

Keywords: *land vertical movement, plate tectonic, Gamit/Globk, GNSS data analysis, machine learning, GNSS time-serries*

1. Introduction and literature review

Vertical displacement of the land surface is the change in elevation of the ground due to the tectonic activities of the Earth's crust and surface subsidence due to causes such as mining, groundwater extraction or sediment compaction, etc. For areas with low terrain such as the Mekong Delta, Ho Chi Minh City – Vietnam, vertical displacement of the ground exacerbates the inundation of the city during high tides, heavy rain or saline intrusion (Construction 2019, Wang, Jiang et al. 2021). To determine the vertical displacement, some methods can be used such as GNSS, InSAR, PInSAR, geometric levelling or using gravity gradient data.

GNSS allows to determine the movement of coordinate components with high accuracy by processing data with some specialized software packages such as Bernese, Gamit/Globk Based on the GNSS data from thousands of tide gauges, (Hammond, Blewitt et al. 2021) created the vertical displacement map on a global scale with precision to mm. In addition, the earth's crust movement can be computed by using data measured in discrete cycle (Abidin, Andreas et al. 2013, Duong, Sagiya et al. 2013, Trần, Nguyễn et al. 2013) or continuous observation series (Kiani 2020, Lau, Coleman et al. 2021, Trọng, Nghĩa et al. 2022, Dinh, Nguyen et al. 2023).

It can be seen that the advantage of GNSS is determining the crustal displacement with high accuracy, but the disadvantage is that it only can observe the displacement at the point where the GNSS receiver is installed but not present for the entire surface. To overcome the aforementioned disadvantage, some different techniques are combined with GNSS to monitor the movement of the land surface. (Abidin, Andreas et al. 2011, Abidin, Andreas et al. 2015) combined levelling data, GPS and InSAR data to observe the subsidence in Jakarta city with the highest subsidence value up to 16cm/year and the displacement in 10 years from 1997 to 2007 up to 90cm. Through analyzing the GNSS and InSAR data, (Catalao, Raju et al. 2013) computed the highest vertical displacement at some places in Singapore up to 7mm/year, and up to 70mm/year in Karrapina – Turkey (Orhan, Oliver-Cabrera et al. 2021). Moreover, by using a combination of different data types such as SAR, levelling, hydrology, InSAR, (Chen, Gong et al. 2017) observed the maximum subsidence value up to 100mm/year and displacement features at some places in Beijing city – China.

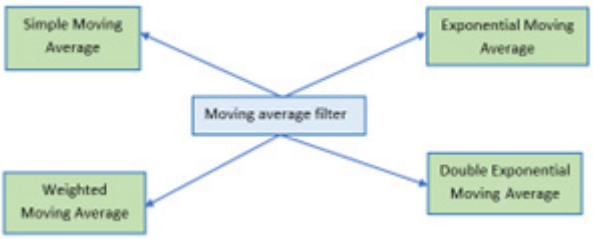

Fig. 1. Types of moving average filters
Rys. 1. Rodzaje filtrów średniej ruchomej

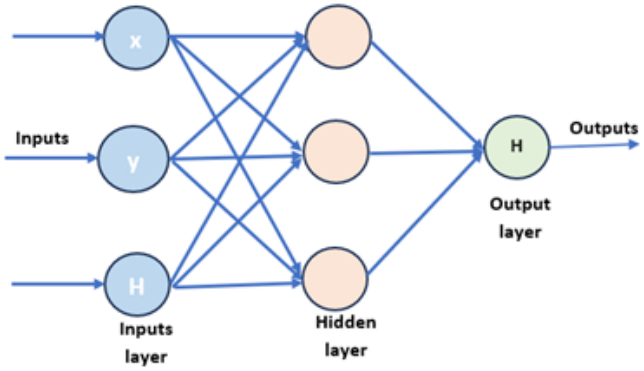

Fig. 2. An example structure of the MLP model for forecasting GNSS time-serries daily solution
Rys. 2. Przykładowa struktura modelu MLP do prognozowania dziennego rozwiązania dla serii czasowych GNSS

GNSS CORS (Continuously Operating Reference Station) is a new solution of GNSS technology in recent years. With the advantage is working continuously over a long period, data from CORS not only can monitor the displacement but also determine the shifting rules as well as the related effects (Wang, Jiang et al. 2021). According to this, the coordinate or displacement values can be computed accurately through the analysis of GNSS data at each station by using professional software packages such as Gamit/Globk, Bernese (Trọng, Nghĩa et al. 2022), Bernese (Kall, Oja et al. 2019). Moreover, the time series of coordinate values obtained from GNSS data can also be analyzed by using other methods such as physical models and digital models. Especially, some components such as linear velocity, annual and semi-annual amplitudes of sine and cosine functions can be determined by using the traditional analytical method.

Artificial Neuron Network (ANN) is widely used in GNSS data processing such as correcting compensation error when integrating GNSS with inertial positioning (INS) (Al Bitar and Gavrilov 2021), evaluating the performance of TEC forecasting models over equatorial low-latitude GNSS station (Sivavaraprasad, Deepika et al. 2020) and prediting the Earth's crust vertical displacement with short time series data (Duong Van Phong 2023). In parallel with the development of modern geodetic techniques, information technology and especially artificial intelligence (AI) are being applied more and more to research the Earth's crust displacement. AI including artificial neuron network, machine learning and deep learning has significantly contributed to improving the accuracy of forecasting, interpolating missing data and analyzing the time series of GNSS data gain (Wang, Jiang et al. 2021, Gao, Li et al. 2022). Many machine learning models have been used in analyzing the time series of GNSS data such as long short-term memory (LSTM), convolutional neural network long short-term memory (CNN-LSTM), artificial neural network (ANN) (Chen, Lu et al. 2023).

Among them, LSTM is a type of recurrent neural network (RNN) architecture designed to address the limitations of traditional RNNs in capturing and remembering long-range dependencies in sequential data. Wenzong Gao and others evaluated the effectiveness of the LSTM technique in forecasting GNSS data compared with other techniques such as gradient boosting decision tree (GBDT) and support vector machine (SVM) whereby the forecasting accuracy increased 30% to using the traditional least square(LS) method (Gao, Li et al. 2022). (Wang, Jiang et al. 2021) used MSSW-LSTM for GNSS time-serries forecasting to improve the precision of the LSTM method with the results reducing RMSE value up to 23,7% and MEA value to 22,2%. By combining Variational Mode Decomposition (VMD) and LSTM, (Chen, Lu et al. 2023) built the DVMD-LSTM model to analyze the GNSS series data in 22 years, the results showed the improvement in forecasting accuracy up to 36,5% and average velocity up to 33,02%.

The forecasting results of vertical displacement of the Earth's crust in (Duong Van Phong 2023) did not compare the effectiveness of the ANN model to the traditional methods and none of solution was proposed in case there is noise in the input data series. This article's target is evaluating the effectiveness of ANN model to the forecasting result of using the Moving Average Filter as well as verify it's ability in case of noise existing in the time series of GNSS data. Another traditional approach, namely Autoregressive Intergrated Moving Average (ARIMA) was also developed to forecast vertical land movement based on GNSS CORS data and compared to the ANN model to highlight the superiority of the ANN model.

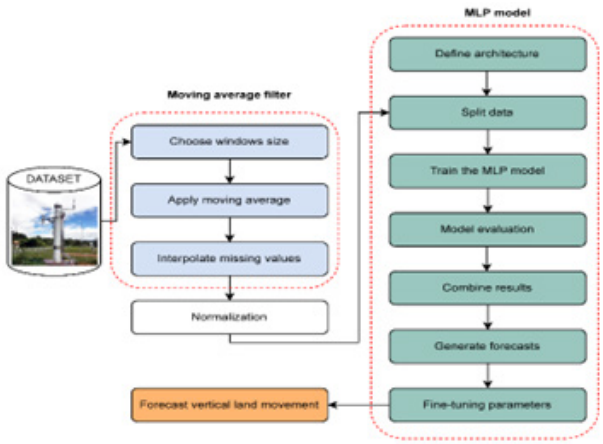

Fig. 3. Proposed framework of the MAF-MLPNN model for forecasting GNSS time-series daily solution
Rys. 3. Proponowana struktura modelu MAF-MLPNN do prognozowania dziennych rozwiązań szeregów czasowych GNSS

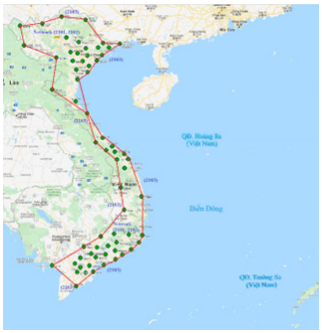

Fig. 4. VNGEONET network
Rys. 4. Sieć VNGEONET

Fig. 5. Structure of VNGEONET station
Rys. 5. Struktura stacji VNGEONET

Tab. 1. Data information of CTHO station
Tab. 1. Informacje o danych stacji CTHO

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

2. Methodology

2.1 Moving average filter (MAF)

The Moving Average Filter, a Finite Impulse Response (FIR) smoothing filter, is employed to mitigate short-term overshoots or noisy fluctuations in a signal. It aids in preserving the true signal representation and retaining a sharp step response. This statistical tool is simple yet elegant, serving as an effective means for denoising signals in the time domain.

Preferred in many Digital Signal Processing (DSP) applications involving time-series data, the Moving Average Filter stands out for its simplicity, speed, and remarkable noise suppression capabilities while retaining a sharp step response. Consequently, it is considered an optimal choice for signals encoded in the time domain.

While the Moving Average filter excels as a smoothing filter in the time domain, it performs poorly in the frequency domain. Its effectiveness shines in applications solely reliant on time-domain processing, but in scenarios where information is encoded in both time and frequency or exclusively in the frequency domain, it may prove to be an unsuitable option.

Various types of moving average filters exist, with simple, cumulative moving average, weighted moving average, and exponentially weighted average filters forming the fundamental basis for most other variants. Although numerous filter variants exist, the core structure generally boils down to four types, as illustrated in Figure 1. In this study, a simple moving average filter was applied to reduce dataset noise.

The Simple Moving Average (SMA) represents one of the simplest forms of moving average filters, known for its ease of understanding and application. Its main advantage lies in its straightforward formula, requiring no complex mathematics for interpretation.

However, the drawback of SMA is that it assigns equal weightage to all samples, resulting in less effective suppression of noisy signals. The mathematical model of SMA is described as follows:

Suppose that we have a given dataset [a1, a2, …, an]. If we consider the periodicity or window length as 'k,' (assume k = 3), then the average of 'k' elements would be:

$$SMA_1 = \frac{a_1 + a_2 + a_3}{3} \tag{1}$$

$$SMA_2 = \frac{a_2 + a_3 + a_4}{3} \tag{2}$$

$$SMA_{N-k+1} = \frac{a_n + a_{n-1} + a_{n-2}}{3} \tag{3}$$

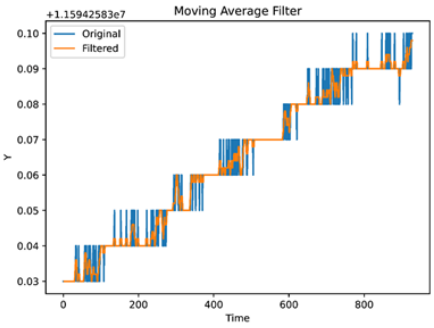

Fig. 6. MAF results of the dataset used
Rys. 6. Wyniki MAF dla użytego zbioru danych

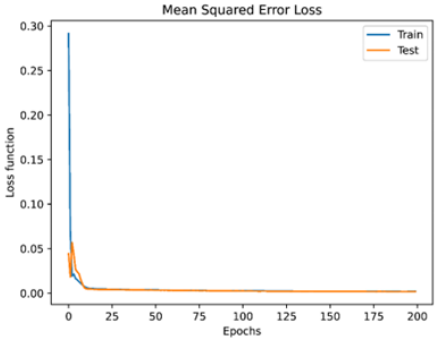

Fig. 7. Training performance of the MAF-MLPNN model for forecasting Up component
Rys. 7. Wydajność treningu modelu MAF-MLPNN dla prognozowania składnika Up

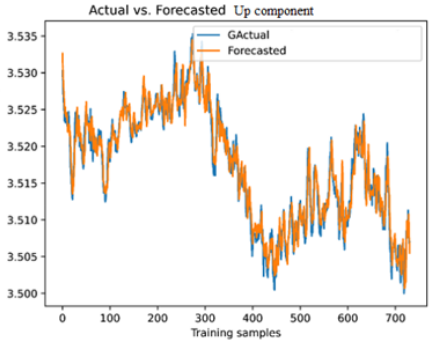

Fig. 8. Comparison of actual and forecasted Up component using the MAF-MLPNN model on the training samples
Rys. 8. Porównanie rzeczywistej i prognozowanej składowej Up przy użyciu modelu MAF-MLPNN na próbkach treningowych

Finnally, simple moving average result will be calculated as follows:

$$SMA = [SMA_1, SMA_2, \ldots SMA_{N-k+1}] \quad (4)$$

2.2 Multi-layer perceptron neural network (MLPNN)

The Multilayer Perceptron (MLP) is a straightforward and efficient multi-layer feedforward neural network with distinct input, hidden, and output layers. The hidden layer comprises one or more layers, each hosting a specific number of neurons responsible for summing and activating weighted inputs to generate outputs. The primary algorithm employed is the Backpropagation Neural Network (BPNN), where signals propagate forward during the learning process, and errors propagate backward. In a magnified MLP neural network, each node is composed of an adder and an activation function.

The MLP model follows four essential steps:

1. Variable selection: Decide on the variables that will be used in the model.
2. Methodologies for instruction, evaluation, and formula verification: Establish methodologies for training, evaluation, and verifying the model's formulas.
3. Buildings and other structures: Develop the network architecture, including input, hidden, and output layers.
4. Model validation and forecasting: Validate the model and use it for forecasting.

The MLPNN approach is a type of feedforward neural network that employs the Backpropagation technique for learning. It consists of input neurons, one or more hidden layers for computation and iterations, and an output layer for forecastings. Initially introduced in 1943 by McCulloch and Pitts, ANNs have gained popularity for their non-linear mapping capabilities in statistical analyses. MLP, being flexible and possessing proper representational capability, is widely used and regarded as the most practical type of ANN. The backpropagation algorithm is applied to train MLPs, which function as feedforward neural tools and general approximators. The overall structure

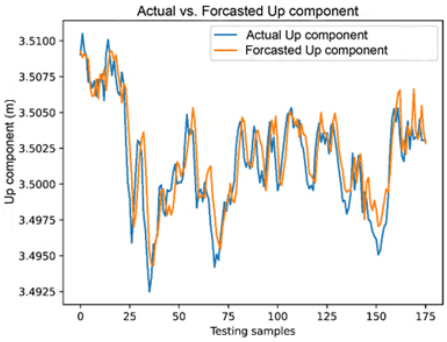

Fig. 9. Comparison of actual and forecasted Up component using the MAF-MLPNN model on the testing samples
Rys. 9. Porównanie rzeczywistej i prognozowanej składowej Up przy użyciu modelu MAF-MLPNN na próbkach testowych

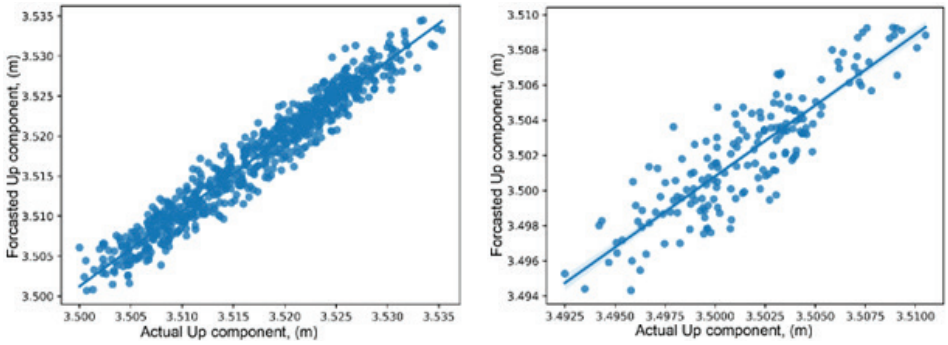

Fig. 10. Correlation between the actual and forecasted Up component based on the MAF-MLPNN model
Rys. 10. Korelacja między rzeczywistym i prognozowanym składnikiem Up na podstawie modelu MAF-MLPNN

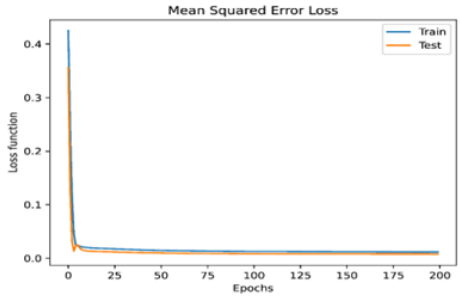

Fig. 11. Training performance of the MLPNN model for forecasting Up component
Rys. 11. Wydajność treningu modelu MLPNN dla prognozowania składnika Up

of the MLP considered in the study includes input and output layers, a hidden layer, and strong connections between adjacent layer neurons. Figure 2 shows the structure of MLP model for forecasting GNSS time-serries daily solution.

2.3 Multi-layer perceptron neural network (MLPNN)

In addressing the time series problem at hand, we propose a comprehensive framework that integrates the Moving Average Filter and Multilayer Perceptron Neural Network (MLPNN) model. This hybrid approach aims to leverage the smoothing capabilities of the Moving Average Filter alongside the forcastive power of the MLP model to enhance the accuracy and effectiveness of GNSS time-serries daily solution forecasting. Through this combined framework, we seek to capitalize on the strengths of both methods, providing a robust solution for capturing patterns, reducing noise, and making accurate forecasting s in time-dependent datasets. The framework of the MAF-MLPNN model for forecasting GNSS time-serries daily solution is proposed in Figure 3.

2.4 Evaluation methods

For evaluating the forecasting performance of the ANN and ARIMA models, three performance metrics, including mean absolute error (MAE), mean-squared error (MSE), and root-mean-squared error (RMSE) were applied, as calculated according to equations (5-7).

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{5}$$

Where:
n is the number of data points;
yi is the actual value at data point i;
$\hat{y}_i$ is the forecasted value at data point i.

$$\boldsymbol{MSE} = \frac{\mathbf{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}(\boldsymbol{y_i} - \hat{\boldsymbol{y}}_{\boldsymbol{i}})^{\mathbf{2}} \tag{6}$$

Where the expressions involving n, yi, $\hat{y}_i$ retain the same meaning as above.

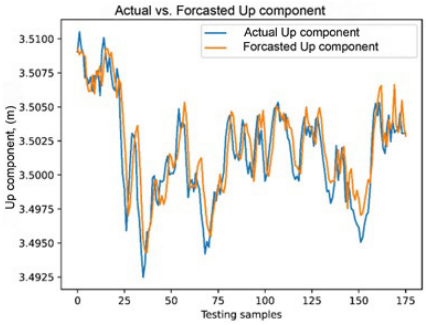

Fig. 12. Comparison of actual and forecasted Up component using the MLPNN model on the training samples
Rys. 12. Porównanie rzeczywistej i prognozowanej składowej Up przy użyciu modelu MLPNN na próbkach treningowych

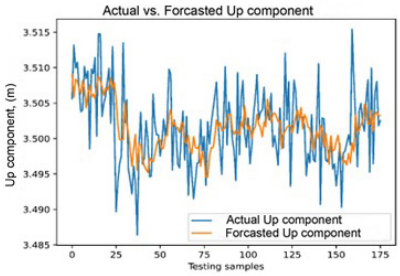

Fig. 13. Comparison of actual and forecasted Up component using the MLPNN model on the testing samples
Rys. 13. Porównanie rzeczywistej i prognozowanej składowej Up przy użyciu modelu MLPNN na próbkach testowych

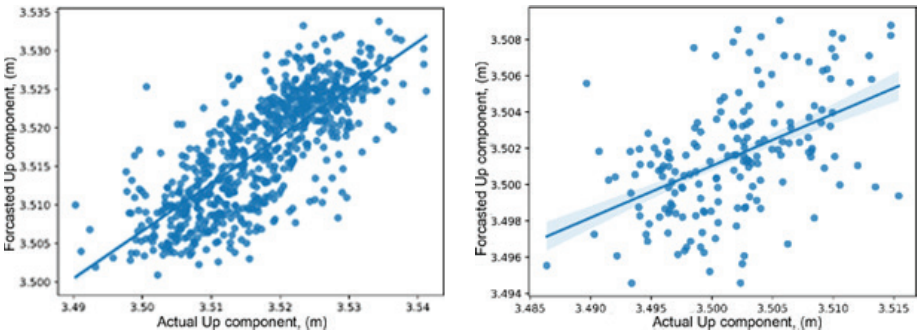

Fig. 14. Correlation between the actual and forecasted Up component based on the MLPNN model
Rys. 14. Korelacja między rzeczywistym i prognozowanym składnikiem Up w oparciu o model MLPNN

$$RMSE = \sqrt{MSE} \tag{7}$$

It is the square root of the Mean Squared Error and provides an assessment of the magnitude of forecasting errors in the units of the dependent variable.

3. GNSS Cors data and analysis

3.1 Data collection

The Vietnam's national network of satellite positioning stations (VNGEONET) is operated by the end of 2019 consists of 65 CORS stations (24 geodetic stations and 41 NRTK stations) (figure 4) for multiple purposes including the vertical displacement of the Earth's crust (Quân, Trung et al. 2021). To ensure the above tasks, the VNGEONET's stations are designed and buried deep to the bedrock layer (figure 5) (DOSM). From the data of this CORS network, some studies on the displacement of the earth's surface have been published (Nguyễn Gia Trọng 2021, Quân, Trung et al. 2021, Trọng, Nghĩa et al. 2022).

In Figure 4, the red line is used to indicate the boundary of the geographical location range within which the measurement solution with CORS stations can be applied, following the network solution - for the areas enclosed by the red line. The data information is presented in table 1.

3.2 Data collection

Firstly, the data as described in Table 1 was processed using the precise GNSS processing software. The results of this data analysis yielded the values of the three-dimensional spatial coordinates of the monitoring point over time. Subsequently, the time-series Up component was analyzed utilizing ARIMA and MLPNN models. There are many ways to process the time series of GNSS data such as using Bernese software (Zhao, Chen et al. 2023) or Gamit/Globk (Li 2021). In this article, the authors use Gamit/Globk software to analyze the GNSS data series with output results are point's coordinate components at each epoch.

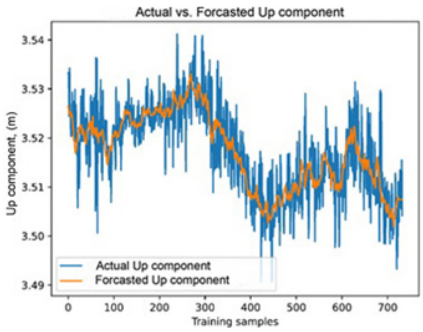

Fig. 15. Comparison of actual and forecasted Up component using the ARIMA (1,1,1) model on the training samples
Rys. 15. Porównanie rzeczywistego i prognozowanego składnika Up przy użyciu modelu ARIMA (1,1,1) na próbkach treningowych

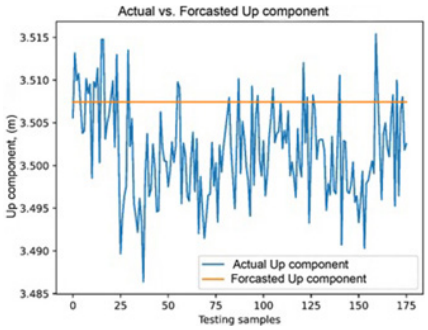

Fig. 16. Comparison of actual and forecasted Up component using the ARIMA (1,1,1) model on the testing samples
Rys. 16. Porównanie rzeczywistego i prognozowanego składnika Up przy użyciu modelu ARIMA (1,1,1) na próbkach testowych

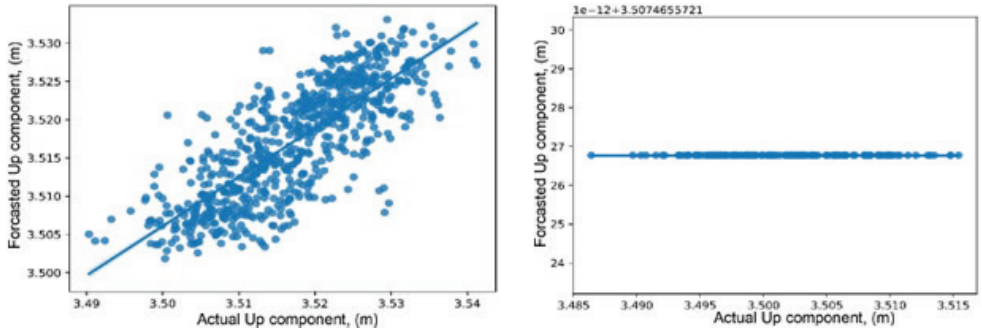

Fig. 17. Correlation between the actual and forecasted Up component based on the ARIMA (1,1,1) model
Rys. 17. Korelacja pomiędzy rzeczywistym i prognozowanym składnikiem Up na podstawie modelu ARIMA (1,1,1)

4. Results and discussion

The forecasting results using the MAF function are provided as shown in Figure 6. It can be observed that, during execution, the MAF function partitions the dataset into subsets based on which it forcasts step changes in values.

The MLPNN function is developed as follows: it divides the input data into a training dataset (80%) and a test dataset (20%). Using all variables (i.e., X, Y, H) to forecast Up component. The model training algorithm employed is the Adam algorithm. During model training, the epoch value is set to 10, with 01 hidden layer containing 10 hidden nodes. ReLU is the chosen activation function for the intermediate layer, and Sigmoid is the corresponding activation function for the output layer. To assess the loss value, the Mean Squared Error (MSE) function is selected with a total of 200 epochs. Figure 7 illustrates the correlation between the Loss function and Mean Squared Error Loss. It can be observed that, with the number of epochs reaching 25, the value of the Loss function has approached 0.

Subsequently, the performance and accuracy evaluation are conducted when employing the MAF-MLPNN function for forecasting GNSS time-serries daily solution. The evaluation results are depicted in Figure 8. Examining the performance and accuracy of the MAF-MLPNN model in forecasting GNSS time-serries daily solution with the training dataset reveals a remarkably high performance in the model's forecasting capabilities.

From the forecasted results of Up component for the training dataset (Figure 8) and the test dataset (Figure 9), it can be observed that the MAF-MLPNN model is entirely suitable for practical use. To further ascertain the suitability of the MAF-MLPNN function, regression and correlation levels are determined. The results indicating the regression and correlation levels are presented in Figure 10 for the training dataset (left) and the test dataset (right).

To compare the effectiveness of combining the MAF function with the MLPNN function, use the MLPNN function to forecast Up component with the same input dataset as before.

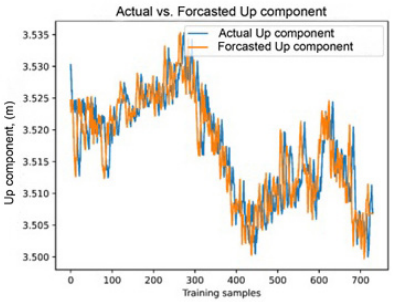

Fig. 18. Comparison of actual and forecasted Up component using the MAF-ARIMA (2,1,0) model on the training samples
Rys. 18. Porównanie rzeczywistego i prognozowanego składnika Up przy użyciu modelu MAF-ARIMA (2,1,0) na próbkach treningowych

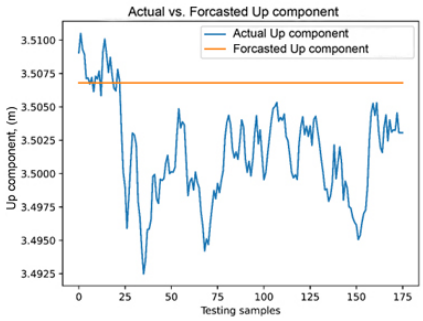

Fig. 19. Comparison of actual and forecasted Up component using the MAF-ARIMA (2,1,0) model on the testing samples
Rys. 19. Porównanie rzeczywistego i prognozowanego składnika Up przy użyciu modelu MAF-ARIMA (2,1,0) na próbkach testowych

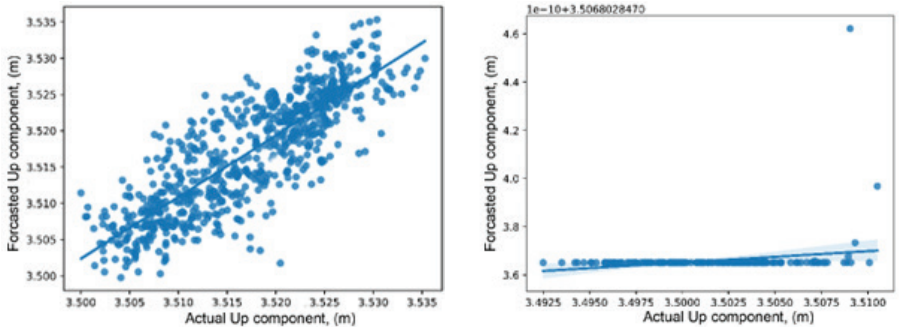

Fig. 20. Correlation between the actual and forecasted Up component based on the MAF-ARIMA (2,1,0) model
Rys. 20. Korelacja pomiędzy rzeczywistym i prognozowanym składnikiem Up na podstawie modelu MAF-ARIMA (2,1,0)

The results determining the loss function values for the training dataset using the MLPNN function are shown in Figure 11.

The accuracy of forecasting Up component for the training dataset and the test dataset is depicted in Figures 12 and Figures 13, respectively.

The correlation between the actual values of Up component and the corresponding forecasted values for both the training and test datasets is illustrated in Figure 14.

From the aforementioned analysis results, it can be observed that combining the MAF function with the MLPNN function yields significantly better forecasting results for Up component compared to using only the MLPNN function. Detailed results for model efficiency comparison will be provided in the following section.

To reassert the effectiveness of the proposed artificial intelligence model for forecasting Up component, a traditional forecasting function, specifically the ARIMA function, is chosen for comparison. To compare the effectiveness of using the combined MAF function, calculations have been performed for two scenarios: one using only the ARIMA function and the other using the combined MAF function with the ARIMA function.

When executing the two computation scenarios with the ARIMA function, the data is also divided into a training dataset (80%) and a test dataset (20%).

In the case of using the ARIMA function, the computational results show that after 10 iterations, the best ARIMA model has been identified with parameters for the training dataset: 1 autoregressive component (utilizing only the height component), 1 integration, and 1 moving average component – ARIMA(1,1,1). The total time for model fitting in this case is 8.745 seconds.

Figure 15 illustrates the comparison results between the actual of Up component values and their corresponding forecasted values for the training dataset.

The correlation between the actual values of Up component and their corresponding forecasted values for the test dataset using the ARIMA(1,1,1) function is shown in Figure 16.

Tab. 2. Performance of the MLPNN and ARIMA models with and without MAF

Tab. 2. Wydajność modeli MLPNN i ARIMA z i bez MAF

Forecast model	Training samples			Testing samples		
	MAE	MSE	RMSE	MAE	MSE	RMSE
MLPNN	0.005	0.000	0.006	0.004	0.000	0.005
MAF-MLPNN	0.001	0.000	0.002	0.001	0.000	0.002
ARIMA	0.004	0.000	0.006	0.007	0.000	0.008
MAF-ARIMA	0.003	0.000	0.004	0.005	0.000	0.006

It can be observed that the forecasted results of Up component using the ARIMA(1,1,1) function on the test dataset form a straight line and are highly inaccurate. The correlation between the actual values of Up component and their corresponding forecasted values for both the training and test datasets is illustrated in Figure 17.

The actual values of Up component and their corresponding forecasted values for the test dataset using the ARIMA(2,1,0) function are illustrated in Figure 19. It can be observed that while the forecasted values closely match the actual values on the training dataset, they are still highly inaccurate for the test dataset.

The regression and correlation of Up component for the training and test datasets when using the MAF-ARIMA function are shown in Figure 20. From Figure 20, it can be observed that the regression and correlation for the training dataset are very good, but conversely, for the test dataset, they are not as favorable.

The accuracy characteristics of the models, as mentioned in Formulas 5, 6, and 7 in Section 2.4, are summarized in Table 2. The corresponding values for MAE, MSE, and RMSE are 0.001, 0.000, and 0.002 for both the training and test datasets, demonstrating the superior effectiveness of the MAF-MLPNN model compared to other models.

From the values in Table 2, it can be observed that the prediction results of the Up component using the MAE-MLPNN model outperform those of previous publications. Specifically, in comparison to the findings reported by (Wang, Jiang et al., 2021) where the minimum MAE and RMSE values were 2.3864 and 3.1628 respectively, and the RMSE values published by (Gao, Li et al., 2022) ranged from 0.0027 to 0.0070.

5. Conclusion

This research proposed a hybrid model, the Moving Average Filter-Multilayer Perceptron Neural Network (MAF-MLPNN), to enhance the accuracy of forecasting GNSS time-serries daily solution. By integrating the noise-reducing capabilities of the Moving Average Filter with the predictive power of the MLPNN model, we aimed to provide a robust solution for addressing the challenges associated with predicting GNSS time-serries daily solution in dynamic environments.

Through rigorous data analysis and evaluation, we demonstrated the superior performance of the MAF-MLPNN model compared to traditional methods like ARIMA. The model exhibited exceptional accuracy metrics, even in the presence of noise in the dataset, showcasing its reliability and effectiveness for practical applications in geodetic studies. The results of this study contribute to advancing prediction models in current spatial data analysis research, presenting a new approach that combines AI techniques with established methods to achieve enhanced forecasting accuracy. The proposed MAF-MLPNN model holds significant potential in improving assessment strategies and minimizing risks in areas prone to GNSS time-serries daily solution, thereby supporting environmental management efforts and sustainable development endeavors.

Acknowledgment

The author group sincerely thanks the Vietnam Department of Survey, Mapping, and Geographic Information (DOSM) for providing the data collected by CORS stations in the VNGEONET network to conduct this paper.

Funding

The study was funded by the Ministry of Education and Training in Vietnam under grant number B2022-MDA-09.

Literatura – References

1. Abidin, H., et al. (2013). "Land subsidence in coastal city of Semarang (Indonesia): characteristics, impacts and causes." Geomatics, Natural Hazards and Risk 4(3): 226-240.
2. Abidin, H., et al. (2015). "On correlation between urban development, land subsidence and flooding phenomena in Jakarta." Proceedings of IAHS 370: 15-20.
3. Abidin, H. Z., et al. (2011). "Land subsidence of Jakarta (Indonesia) and its relation with urban development." Natural Hazards 59: 1753-1771.
4. Al Bitar, N. and A. Gavrilov (2021). "A new method for compensating the errors of integrated navigation systems using artificial neural networks." Measurement 168: 108391.
5. Catalao, J., et al. (2013). "Mapping vertical land movement in Singapore using InSAR GPS." ESA Spec. Publ 722: 54.
6. Chen, B., et al. (2017). "Characterization and causes of land subsidence in Beijing, China." International Journal of Remote Sensing 38(3): 808-826.
7. Chen, H., et al. (2023). "An Improved VMD-LSTM Model for Time-Varying GNSS Time Series Forcastion with Temporally Correlated Noise." Remote Sensing 15(14): 3694.
8. Construction, M. o. (2019). Trouble underground - Land Subsidence in the Mekong Delta. Vietnam.
9. Dinh, T. T., et al. (2023). "Crustal displacement in Vietnam using CORS data during 2018-2021." Earth Sciences Research Journal 27(1): 27-36.
10. Duong, N. A., et al. (2013). "Contemporary horizontal crustal movement estimation for northwestern Vietnam inferred from repeated GPS measurements." Earth, Planets and Space 65(12): 1399-1410.
11. Duong Van Phong, N. G. T., Nguyen Van Chien. Nguyen Ha Thanh, Ly Lam Ha, Nguyen Viet Quan, Pham Ngoc Quang (2023). "Analysis of land vertical movement using ANN function from the results of processing GNSS time series data." Vietnam Journal of Hydro-Meteorology: 41-50.
12. Ezquerro, P., et al. (2020). "Vulnerability assessment of buildings due to land subsidence using InSAR data in the ancient historical city of Pistoia (Italy)." Sensors 20(10): 2749.
13. Gao, W., et al. (2022). "Modelling and forcastion of GNSS time series using GBDT, LSTM and SVM machine learning approaches." Journal of Geodesy 96(10): 71.
14. Goudarzi, M. A. (2016). GPS inferred velocity and strain rate fields in eastern Canada, Université Laval.
15. Hammond, W. C., et al. (2021). "GPS imaging of global vertical land motion for studies of sea level rise." Journal of Geophysical Research: Solid Earth 126(7): e2021JB022355.
16. Kall, T., et al. (2019). "The noise properties and velocities from a time-series of Estonian permanent GNSS stations." Geosciences 9(5): 233.
17. Kiani, M. (2020). "Lateral land movement forcastion from GNSS position time series in a machine learning aided algorithm." arXiv preprint arXiv:2006.07891.
18. Lau, N. N., et al. (2021). "Determination of tectonic velocities of some continuously operating reference stations (CORS) in Vietnam 2016-2018 by using precise point positioning."
19. Li, Y. (2021). "Analysis of GAMIT/GLOBK in high-precision GNSS data processing for crustal deformation." Earthquake Research Advances 1(3): 100028.
20. Nguyễn Gia Trọng, L. T. T., Nguyễn Hà Thành, Phạm Ngọc Quang, Nguyễn Văn Cương (2021). First step determination the movement of some CORS stations in the Northern of Vietnam using Gamit/Globk software. National conference on Geospatial technology in Earth sicence and Environment, Hanoi University of Mining and Geology.
21. Orhan, O., et al. (2021). "Land subsidence and its relations with sinkhole activity in Karapınar region, Turkey: a multi-sensor InSAR time series study." Sensors 21(3): 774.
22. Quân, N. V., et al. (2021). "Ứng dụng mạng lưới trạm định vị vệ tinh quốc gia (VNGEONET) trong hoạt động đo đạc bản đồ, nghiên cứu khoa học Trái Đất và một số lĩnh vực khác trong thời kỳ chuyển đổi số." Hội nghị khoa học quốc gia về công nghệ địa không gian trong khoa học Trái Đất và môi trường, Hà Nội.
23. Shahvandi, M. K. and B. Soja (2021). Modified Deep Transformers for GNSS Time Series Forcastion. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS.
24. Sivavaraprasad, G., et al. (2020). "Performance evaluation of neural network TEC forecasting models over equatorial low-latitude Indian GNSS station." Geodesy and Geodynamics 11(3): 192-201.
25. Trần, Đ. T., et al. (2013). "Recent crustal movements of northern Vietnam from GPS data." Journal of Geodynamics 69: 5-10.

26. Trọng, N. G., et al. (2022). "Determination of tectonic velocities in Vietnam territory based on data of CORS stations of VNGEONET network." Vietnam Journal of Hydro-Meteorology: 59 - 66.

27. Wang, J., et al. (2021). "A new multi-scale sliding window LSTM framework (MSSW-LSTM): a case study for GNSS time-series forcastion." Remote Sensing 13(16): 3328.

28. Zhao, Q., et al. (2023). "The vertical velocity field of the Tibetan Plateau and its surrounding areas derived from GPS and surface mass loading models." Earth and Planetary Science Letters 609: 118107

Przewidywanie szeregów czasowych GNSS przy użyciu filtra średniej ruchomej i wielowarstwowej sieci neuronowej perceptronu

Delta Mekongu i Ho Chi Minh City w Wietnamie są uznawane za obszary w znacznym stopniu dotknięte osiadaniem gruntu. Zjawisko to doprowadziło do znaczących konsekwencji, w tym zwiększonej podatności na takie zjawiska, jak wnikanie soli i powodzie pływowe. Technologia GNSS-CORS, znana ze swojej zdolności do dostarczania ciągłych danych szeregów czasowych, odgrywa kluczową rolę w dokładnym monitorowaniu zmian powierzchni ziemi. Pomimo istnienia tradycyjnych algorytmów do analizy ciągłych danych pomiarowych zebranych za pomocą technologii GNSS-CORS, ich skuteczność jest ograniczona wyzwaniami związanymi z obsługą różnorodnych danych wejściowych i ograniczeniami w prognozowaniu przyszłych przemieszczeń. W związku z tym istnieje rosnąca tendencja do przyjmowania technik sztucznej inteligencji, w szczególności sztucznych sieci neuronowych (ANN), do przewidywania komponentu Up w codziennym rozwiązaniu GNSS. Niniejsze badanie wykorzystuje dane ze stacji CTHO GNSS CORS zlokalizowanej w delcie Mekongu do oceny proponowanych modeli. Innowacyjne podejście hybrydowe, które integruje filtr średniej ruchomej (MAF) i wielowarstwową perceptronową sieć neuronową (MLPNN), zostało wprowadzone w celu zwiększenia dokładności prognozowania. Do oceny skuteczności modeli wykorzystano wskaźniki oceny wydajności, takie jak średni błąd bezwzględny (MAE), średni błąd kwadratowy (MSE) i średni błąd kwadratowy (RMSE). Wyniki pokazują doskonałą wydajność modelu MLPNN, osiągając wysoką dokładność przewidywania dzięki wskaźnikom takim jak MAE

Słowa kluczowe: *ruch pionowy lądu, tektonika płyt, Gamit/Globk, analiza danych GNSS, uczenie maszynowe, serie czasowe GNSS*