Mansoura University
Faculty of Engineering
Mansoura Engineering Journal

GNSS Cycle Slip Detection and Estimation Using LSTM-based Autoencoder Model

Ahmed Ragheb, *Ahmed Zekry* and Mohamed Elhabiby*

KEYWORDS:

GNSS; Cycle Slip; Slip Net; LSTM; Autoencoder

***Abstract*— Global Navigation Satellite Systems (GNSSs) are used in many navigation and positioning applications. Unfortunately, a GNSS signal may suffer from some errors, such as cycle slips, which deteriorate the positioning solution. A cycle slip is defined as a sudden jump by an integer number of cycles in the GNSS carrier phase observations. Signal blockage or/and high troposphere activities are the most common causes for GNSSs' cycle slips. Therefore, cycle slips should be detected and corrected to determine reliable positioning estimations. A new approach for cycle slip detection and repair is proposed based on a master-rover phase-difference with a deep Long Short-Term Memory (LSTM) neural network model; our SlipNet model can classify defective data where a cycle slip has occurred and then predict the exact epoch where the cycle slip(s) occurred. The proposed SlipNet network would be the first end-to-end learning framework to solve the integer ambiguity problem in GNSS measurements with high performance results, %99.7 detection and localization accuracy, and 0.045 MAE for slip estimation and recovery. These results are on par with the latest classical cycle slip detection methods of cycle slip detection and correction.**

I. INTRODUCTION

A Global Navigation Satellite System (GNSS) is an absolute positioning technique where the user receives transmitted data from at least four GNSS satellites to determine the position related to a fixed coordinate frame [1]. However, GNSS signals may be subjected to different error types that must be eliminated or modeled to achieve a reliable solution. The error sources are satellite, receivers, and propagation media biases [2]. The GNSS carrier phase observation is described as the fractional beat phase (the difference between the phase of the satellite transmitted carrier wave and the phase of the receiver-generated replica signal), and an integer counter is then initialized [2]. During the tracking of the satellite, the counter is incremented by one, whenever the accumulated phase changes from zero to 2π where at a given epoch, the observed phase is the sum of the fractional phase and the previously mentioned integer counter [3]. The initial integer number of complete cycles N (integer ambiguity) between the satellite and the receiver is unknown and remains constant, as long as no loss of the signal lock occurs, i.e., no cycle slips occur. The integer counter is re-initialized when cycle slips that cause a sudden jump (change) in the instantaneous accumulated phase by an integer number of cycles occur; this sudden jump is called a cycle slip. When plotting the measured phase versus time, a smooth curve with some noise should be obtained unless a cycle slip occurs (i.e., there is a sudden jump in the

Received: (30 April, 2021) - Revised: (06 June, 2021) - Accepted: (12 June, 2021)

Ahmed E. Ragheb, Associate professor of Surveying and Geodesy, Public Works Department, Ain Shams University, Faculty of Engineering, Cairo, Egypt. (Email: aragheb@eng.asu.edu.eg)

Ahmed Zekry, is with Department of Electrical and Computer Engineering, Queen's University, Kingston, Canada (Email: amar@queensu.ca)

***Corresponding Author:** Mohamed Elhabiby, Associate professor at Public Works Department, Faculty of Engineering, Ain Shams University (Email: mmelhabiby@eng.asu.edu.eg)*

(Ser. No. BFEMU-2104-1111)

observations at the epoch of cycle slip occurrence), as shown in Fig. 1.

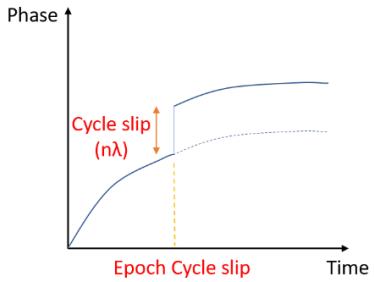

Fig. 1. GNSS cycle slip

A cycle slip affects the range between the satellite and the user receiver; thus, it affects the accuracy of the derived user position and, therefore, it should be detected and estimated. The first stage of cycle slip detection and correction is that of creating a test function, a slowly time-varying function that combines the code and/or phase observations at a single receiver or between different receivers. Techniques were applied to avoid cycle slip occurrence, such as increasing the satellites mask angle. For low satellite mask angle, the GNSS signal path is longer through the atmosphere. This leads to a low Signal to Noise Ratio (SNR), which may cause multipath or/and cycle slip. Detecting of the real cycle slip in such situation is very hard because of the contamination of the signal with noise [4].

II. Related Works

There are several conventional approaches for GNSS cycle slip detection and repair as shown in Fig. 2

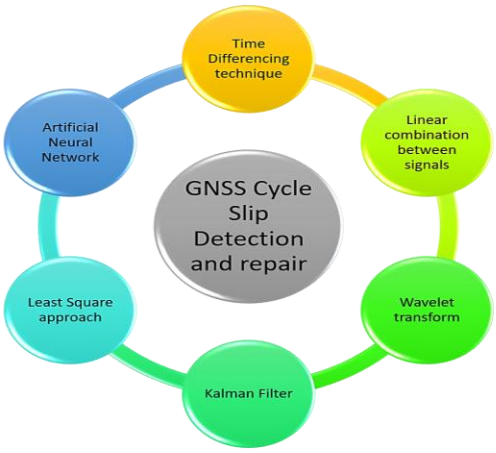

Fig. 2. Different approaches for GNSS cycle slip detection and repair

Time differencing technique is used for GNSS cycle slip detection and repair, where GNSS Observations are subtracted between two successive epochs to reduce or eliminate some biases [5]; any data discontinuities (i.e., a sudden jump for the test function with time, due to the existence of the cycle slips) will be amplified in higher order differences and, thus, enhance the cycle slip detection process [2]. This approach acts as a subtractive filter that passes the low frequencies (the ordinary GNSS signal) while it amplifies the high frequencies represented by the slipped cycles [2]. The second approach is based on a low-degree polynomial fitting depending on different test functions, such as ionospheric residual [6] and range residual, which depend on the combination of the carrier phase and code observations [7]. The Melbourne–Wübbena (M–W) linear combination has been widely applied to cycle slip detection [8]. Some researchers, such as [9], used the double differences phase observation as a test function and modeled by low degree polynomial functions. These approaches necessitate the user's intervention for tuning purposes.

Other approaches propose automated/semi-automated techniques for cycle slip detection, such as [10] who developed an automated cycle slip detection and repair technique based on the Chebyshev polynomial and least-square combination scheme.

Wavelet techniques are used in detecting the GNSS cycle slip in frequency domain; [4] proposed a wavelet based function to detect cycle slip in the details' coefficient, where the cycle slip amplitude was determined in addition to a comparison between the Kalman filter and the wavelet approach, which was done based on the phase linear combination..

Many researchers proposed integrated techniques [11]; [12]; [13]; [14]; and [15], such as neural networks and/or aiding statistical approaches for cycle slip detection and repair. However, these researchers did not focus on the severe noise effect on their approaches.

Kalman filtering is used in different research methods to detect cycle slips where the filter predicts the carrier phase observations and compare them with the actual observed data. A cycle slip is detected when a huge difference between the predicted and actual observations is indicated [16].

There are other approaches for cycle slip detection that are based on Global Positioning System/Inertial Navigation System (GPS/INS) integration [17]. Unfortunately, this approach is difficult in many applications due to the INS installation cost and complexity.

[18] proposed a geometric free linear combination for detecting and fixing the GNSS cycle slips for both Beidou and GPS signals, where an improved adaptive Particle Swarm Optimization (PSO) algorithm is used for cycle slip fixation. A cycle slip detection method for Beidou3 was developed by [19] using an elevation-based model to assist the code measurement noise, where code-phased combination is used as a test quantity.

An enhanced cycle slip repair approach was proposed by [20] using Kalman filter for dual and triple differencing under different satellites elevation and ionospheric conditions. [21] developed a cycle slip detection based on comparison between geometric range and satellite navigation message position, through a moving average technique. A new cycle slip detection and repair method was proposed using Total Electron Content Rate (TECR) through Sudden Increase of Pseudorange Error (SIPE) methods [22].

[23] developed a real-time cycle slip detection and repair approach under high ionospheric activity for undifferenced GPS/BeiDou Satellite System triple-frequency observations using a single receiver, while the second order time difference approach was used for the code/phase linear combination to reduce the ionospheric effect. Different GNSS signal linear combination were investigated as test quantities by [24] for cycle slip detection and repair.

A multi-scale singularity detection was developed, evaluated, and tested by [25] over GPS Code minus Carrier (CmC) and Phase1 minus Phase2 measurements to detect and remove, with a good degree of accuracy, cycle slip errors. This dataset was collected under two conditions; it was assumed that the rover was not in Kinematic mode and that all the data were collected with a clear-view during one session only (without any obstructions) [25].

Another approach used based on modeling the positional errors resulting from cycle slips using Artificial Neural Networks (ANN) was proposed by [26]. Here, it is concluded that the constructed ANNs are capable of modeling the positional errors with the five selected parameters. It has been clarified that for any of the constructed ANN (Back-Propagation (BP)) there were about 160,000 different outputs available that were divided into two groups. The first group (representing 90% of the available data) was used in establishing the ANN, whereas the remaining 10% (about 16,000 outputs) were used to check the reliability of the established ANN. The predicted results have been compared between experimental (an experiment to illustrate the effectiveness of the proposed algorithm) and the BP model. This research work proved that the ANN training model was in agreement with the experimental results. However, the amount of testing data has been relatively small, and there were no measures of accuracy [26].

[27] developed an INS/GPS integration method based on ANN to fuse uncompensated INS and differential GPS (DGPS) positioning and navigation (POS/NAV) measurements. Both "Position Update Architecture (PUA)" and "Position and Velocity Update Architecture (PVUA)" mechanization have shown a superior performance over the conventional INS/DGPS integration techniques. In addition, the PUA and the PVUA have shown the ability to provide the most stable and accurate INS/DGPS solution if compared to other techniques, without any time slots prediction of the occurred cycle slips.

A cycle slip detection and repair technique based on dual-frequency GPS data observed from a single receiver under high ionospheric activity was proposed by [28], where a Forward and Backward Moving Window Averaging (FBMWA) algorithm was used, as well as the Second-Order Time-Difference Phase Ionospheric Residual (STPIR) algorithm to precisely detect cycle slips with the use of only carrier phase observations. Unfortunately, STRIP is sensitive to ionospheric disturbances and, therefore, the integration of the FBMWA and STPIR algorithms allow the cycle slips to be uniquely detected and determined, even under high ionospheric activities. Although, the occurrence of cycle slips won't be signified when the cycle slips are on L_1 and L_2 and have the same size and same sign, thus, cancelling one another [28]. [29] developed a real-time algorithm to detect, determine, and validate (one that has been resolved correctly) the cycle slips for triple-frequency GPS. The cycle-slip detection was implemented by simultaneously applying two geometry-free phase combinations to detect more insensitive cycle slips, this is applicable for high data rate applications. The cycle-slip determination adaptively used the predicted phase data and the code data. The LAMBDA technique was applied to search for the cycle-slip candidates. This technique includes the following attributes:

1. Cycle-slip detection, which checks the occurrence of cycle slips.
2. Cycle-slip determination, which quantifies the sizes of cycle slips.
3. Cycle-slip validation, which tests whether the cycle slips are correctly resolved.
4. Cycle-slip removal, which removes the cycle slips from the phase measurement.

However, the results indicate that the proper performance of the cycle-slip detection relies on the slight change of ionospheric delay between two adjacent epochs; hence, this approach is only valid for the applications with a high data rate. In some extreme cases, (for example, during a magnetic storm) the detection approach may provide unexpected results. Multipath errors will not significantly affect the cycle-slip detection, but if the phase noise is high, the sensitivity of the cycle-slip detection will be degraded [29].

This paper introduces a new end-to-end learning approach based on phase-difference with LSTM. This new method of architecture has been designed to classify the epoch where a cycle slip could occur based on phase measurements, which have their own benefits in providing accurate results. Subsequently, this new method has the ability to predict the cycle slip epoch. Moreover, the actual classified (detected) cycle slip(s) phase-difference will be corrected. The simulated cycle slips will be added to the original dataset phase-difference in order to test the classification, prediction, and correction models' accuracy, with the goal of producing noise-free output measurements without any cycle slip. In this paper, two main phase-difference measurements sources are presented: one for the master and the other for the rover for each satellite. The updated measurements will be evaluated based on the simulated cycle slips that will be added to the original data which have been obtained from all satellites.

III. LSTM MODEL ARCHITECTURE

Cycle slip is defined as a sudden discontinuity in the carrier phase observation that follows a polynomial trend. The carrier phase observation mathematical model [30] is described in equation (1), where carrier phase ($\lambda\phi$) and pseudo-range (ρ) measurements, geometrical range (p), d_ϕ, and $d\rho$ represent the noise in carrier and code measurements, tropospheric error

(d_{trop}), orbital error (d_{Ephem}), satellite and receiver clock errors (*cdt*,*cdT*), and ionospheric error (d_{ion}).

$$\lambda\Phi = p + cdt - cdT - d_{ion} + d_{trop} + d_{EPHEM} + d_{\Phi} + \lambda N \quad (1)$$

First, a test quantity is selected for detecting and estimating GNSS cycle slip to omit the geometrical range (p) and to decrease the time-varying biases through the carrier phase and the code observations linear combination at the same frequency (e.g., C/A code and phase in L_1), or with the carrier phase measurements linear combination at a different frequency (e.g., carrier phase measurements in both L_1 and L_2 frequencies). In this case cycle slip could be easily detected where the test quantity has the advantage of being a slow time variance [25]. Equation (2) describes the code observation mathematical model.

$$\rho = p + cdt - cdT + d_{ION} + d_{TROP} + d_{EPHEM} + d_{\rho} \quad (2)$$

The proposed method uses phase-difference as a test quantity; however, LSTM is deployed because of its memory advantage. LSTM has, by design, the ability to deal with correlated sequences of data using memory-keeping gating mechanism [31].

The constructed deep network architecture consists of three separate steps. Therefore, three neural network models with three different outputs at each step are determined:

1. Classification model:

This model binary-classifies each segment of epochs: five epochs per segment are used and determined whether this segment contains a cycle slip or not.

2. Localization model:

This model determines epochs, those which have been affected by the cycle slips and classified by the previous model. This model outputs another binary classification, but for each single epoch. In case it has been fed by the segments, which have been classified by the previous model, this model can act as a further filtration for the output from the previous model.

3. Cycle slip estimation model:

This model estimates a correct phase-difference before the cycle slip(s) is (are) added. The output from the localization model can be used in conjunction with this model to correct the cycle slips in real-time (epoch by epoch).

Each of these models has some similarities, such as each model is based on 1D LSTM network in the famous encoder-decoder setup that signifies the autoencoder architecture. The autoencoder architecture is adopted for its proven usage in noise reduction and producing noise-free data.

As for the two datasets that have been used for training, the data have been prepared as n satellites, with five epochs, and two channels (L_1 and L_2), for each master and rover phase difference. So, two channels are used: the first channel (L_1) represents the L_1 phase-difference between the master and the rover; and the second channel (L_2) represents the L_2 phase-difference between the master and the rover. Hence, the LSTM model input will be the time-difference of the receivers' phase-differences. According to the previous description, the dataset will be in the following matrix form as shown in equation (3).

$$D = [S \times EP] \quad (3)$$

Where *S* represents the satellites number and *EP* is the number of epochs. For the n satellites, each one will be passed individually. So, by substitution in equation (4)

$$D = [5 \times 1] \quad (4)$$

Before adding the noisy data (time-difference measurements with simulated cycle slips), it is necessary to assure that the original data are noise-free (without any cycle slip). So, "Novatel's Waypoint Precise Positioning Processing GrafNav" software has been used to check whether the measurements are cycle slip free or not. After that, cycle slips will be added among the epochs according to a fixed sequence and with a specific range of values, which guarantee that the proposed model can generalize a wide range of integer ambiguity and sparse distribution of cycle slips (as they represent a rare event anomaly). After cycle slip addition, the datasets are shuffled and separated into %70 training data and 30% testing data. Finally, the confusion matrix (precision, recall, and f1-score) is used to evaluate the performance of the developed model.

The base model takes the time-difference of the master-rover phase-difference as an input. This is fed by five epoch segments (four time-differenced master-rover phase-difference L_1/L_2), so the classification model can classify and detect the cycle slips. Then the classified cycle slips are passed to the prediction model to pinpoint the cycle slip epoch(s). In parallel, the classified segments with cycle slips are passed to cycle slip error estimation model to recover the original time-differenced master-rover phase-difference values.

The simulated cycle slips have two important factors:

1. CS_num: The number of added (simulated) cycle slips.
2. CS_factor: The added cycle slips values.

So, after testing and experiments, the results show that higher CS_num in the training dataset achieves a better accuracy, precision, and recall results in the testing stage.

After cycle slip classification (detection), a pipeline technique is applied to enhance the prediction (localization) model. The classification output is fed to the prediction model, so the cycle slip distribution will be evaluated according to the classification results (precision). Now, it is time to correct the produced cycle slip(s) and estimate the produced error.

As mentioned before, the confusion matrix (precision, recall, and f1score) is employed to better represent the model results' effect. As cycle slips are rare events, accuracy metric from the TensorFlow library would not be very representative

of the true value S in our model. The classification, localization, and error estimation models are depicted in Fig. 3, 4, and 5.

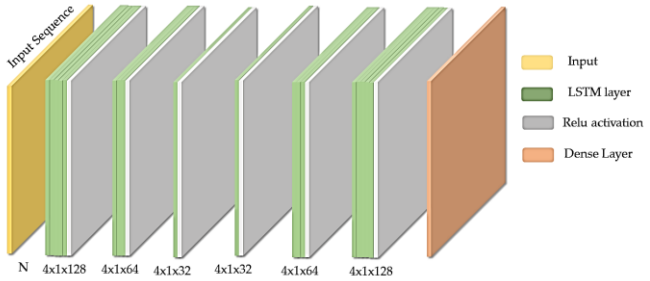

Fig. 3. Classification Model

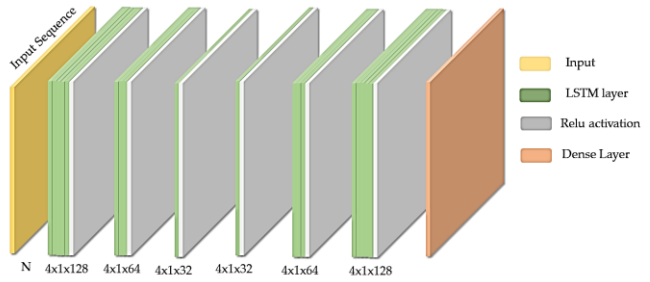

Fig. 4. Localization Model

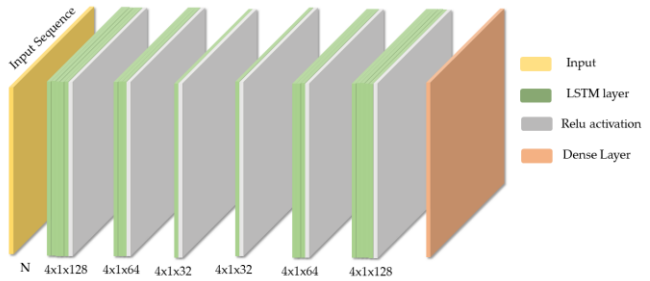

Fig. 5. Error Estimation Model

For example, if a single cycle slip per 1,000 epochs occurred and the model fails to predict such cycle slip, then the accuracy will be equal to "0.999," when the model in that case has no real value to the problem at hand and the recall value of "0.0" would be more descriptive of the actual value of the model.

Such scenarios are faced while training and testing processes, such as in SlipNet-3 to 5 and SlipNet-35, where very few cycle slips in the training dataset have occurred, the accuracy metric was not less than "0.999," where zero true positive has been indicated and, subsequently, no value would come from the resulting trained model.

So, in order to have an accuracy metric that represents the actual performance of the model, the confusion matrix is used: more specifically the true positive, false negative, true negative, and false positive quantities for each of the trained and tested models. As these quantities might be confusing, the description of each test case in the context of cycle slip detection is as follows:

1. True-Negative: The model claimed it is **slip-free** while it actually was **slip-free**.
2. False-Positive: The model claimed it has **cycle slip** while it actually was **slip-free**.
3. True-Positive: The model claimed it has **cycle slip** while it actually had **cycle slip**.
4. False-Negative: The model claimed it **slip-free** while it actually had **cycle slip**.

As observed, the true negative is the most common case as the natural, slip-free epochs usually dominate the phase measurements. However, true positives are the most important. The reason for this is that true positives represent the actual cycle slips detected by the model. To complete the picture of how accurate the model is, the ratio between true positives and false negatives is used (which are cycle slips the model failed to detect) in order to have what is called the recall value. The recall metric represents how many cycle slips the model can correctly detect of all cycle slips in the test dataset. The false positives are also rather important, as they represent how many cycle slips are falsely detected by the model. Together with the true positives, they represent the expected precision of the model. In other words, how much of the model detected cycle slips that will go forward into the correction step is actually a cycle slip that needs correction. The precision value was used to determine the cycle slip ratio in the training and testing datasets for subsequent steps, i.e., the localization and cycle slip estimation models.

When dealing with regression network models, like the proposed LSTM cycle slip estimation model, the sigmoid non-linear activation function is removed at the last layer of the model, because its purpose is to limit the output in the classification models. This is unlike the binary cross-entropy loss function used in training the two classification models, which minimizes the Mean Absolute Error (MAE) during the cycle slip estimation training.

During training, the hyper-parameters of the training procedure were set as follows:

1. Number of epochs per segment:

(a) Five epochs (fixed).

Number of satellites with simulated cycle slips:

(a) All satellites.
(b) Decreasingly, from five to one satellite.

2. Number of epochs between cycle slips:

(a) 250 (fixed).

3. Batch size:

(a) 500
(b) 100 (in cycle slip estimation model training).

4. *Epochs*:
   (a) 10
   (b) 200 (in cycle slip estimation model training).

5. Learning rate:
   (a) 0.001(fixed).

6. Train-Test-split:
   (a) 0.3 (fixed).

Different batch size and epochs are selected for the cycle slip estimation model training, to address the regression problem and preferred method of training longer on smaller batch sizes. Hence, in this way the model can generalize better on data it has never seen before and achieve a better MAE.

IV. EXPERIMENTAL RESULTS

Two datasets of different sources are used to train the proposed network; both datasets are similar in collection conditions. These were collected where the rover was fixed with a clear sky and a moderate baseline. The first dataset comprised more satellite phase measurements over the second dataset, which comprised more overall epochs for the satellite readings. However, the first dataset was collected in a different globe hemisphere than the second dataset, this group had a different set of satellite measurements, which was considered a benefit for the developed network architecture, as these measurements could easily be used for other datasets.

As shown in Tables 1 and 2, CS_factor within the range of 1, 6 has the least f1-score in both datasets. While training on the first dataset within this CS_factor range may get better accuracy results than training on the second dataset, this may be a result of the magnitude of CS_factor compared to the actual phase-difference values in both datasets. This might be leading to a contrast in the second dataset to that of first dataset.

TABLE I
CLASSIFICATION MODEL RESULTS WHEN TRAINED AND TESTED ON THE FIRST DATASET

Model version	CS factor	Precision	Recall	F1 Score	True neg	False pos	True pos	False neg
SlipNet-26	100-200	1	0.852	0.92	48186	0	109	19
SlipNet-27	50-100	0.991	0.819	0.897	48175	1	113	25
SlipNet-28	25-50	1	0.861	0.925	48177	0	118	19
SlipNet-29	10-20	1	0.881	0.937	48188	0	111	15
SlipNet-30	1-6	0.911	0.713	0.8	48176	9	92	37
SlipNet-36	10-200	0.991	0.85	0.915	48187	1	107	19

TABLE II
CLASSIFICATION MODEL RESULTS WHEN TRAINED AND TESTED ON THE SECOND DATASET

Model version	CS factor	Precision	Recall	F1 Score	True neg	False pos	True pos	False neg
SlipNet-6	100-200	0.972	0.998	0.984	334092	26	887	2
SlipNet-7	50-100	0.978	0.999	0.988	334054	21	931	1
SlipNet-8	25-50	0.995	0.999	0.997	334119	4	883	1
SlipNet-9	10-20	0.973	0.98	0.977	334071	25	894	18
SlipNet-10	1-6	0.876	0.193	0.317	334104	24	170	709
SlipNet-17	10-200	0.97	0.838	0.899	334072	24	764	148

TABLE III
CLASSIFICATION MODEL RESULTS WHEN TRAINED AND TESTED ON THE FIRST DATASET (NUMBER OF SATELLITES)

Model version	Sats	Precision	Recall	F1 Score	True neg	False pos	True pos	False neg
SlipNet-31	5	1	0.951	0.975	48253	0	58	3
SlipNet-32	4	0.794	0.947	0.864	48243	14	54	3
SlipNet-33	3	1	0.886	0.940	48270	0	39	5
SlipNet-34	2	1	0.500	0.667	48290	0	12	12
SlipNet-35	1	0.000	0.000	0.000	48292	0	0	22

TABLE IV
CLASSIFICATION MODEL RESULTS WHEN TRAINED AND TESTED ON THE SECOND DATASET (NUMBER OF SATELLITES)

Model version	Sats	Precision	Recall	F1 Score	True neg	False pos	True pos	False neg
SlipNet-11	5	0.966	0.941	0.953	334606	13	365	23
SlipNet-13	4	0.923	0.951	0.937	334696	23	274	14
SlipNet-14	3	0.860	0.939	0.897	334743	35	215	14
SlipNet-15	2	0.954	0.954	0.961	334849	7	146	5
SlipNet-16	1	0.000	0.000	0.000	334946	0	0	61

Table 1 and 2 show the results with respect to the fixed CS_factor range of 10-20, these results show the effect of the cycle slip value (in cycles) on the model precision and recall metrics. The last four columns are the number of epochs where the SlipNet model correctly or mistakenly predicts cycle slips; this can then be tested regarding the true negative, false positive, true positive, and false negative (true_neg, false_pos, true_pos, and false_neg, respectively).

Tables 3 and 4 show the confusion matrix error metrics of accepted classification results regarding the low number of satellites, which decreases the cycle slip numbers, CS_num, and achieves higher confusion matrix values, but with larger number of satellites. In addition, environmental factors have to be taken into consideration in order to estimate and analyze these results.

Fig. 6 depicts the difference in precision metric results and the magnitude ranges of simulated cycle slips between the trained models on both datasets at different CS factor ranges. Fig. 7 shows the recall metric results vs. the magnitude ranges of simulated cycle slips for the proposed classification model. Finally, Fig. 8 exhibits the f1 score and the magnitude ranges for both datasets.

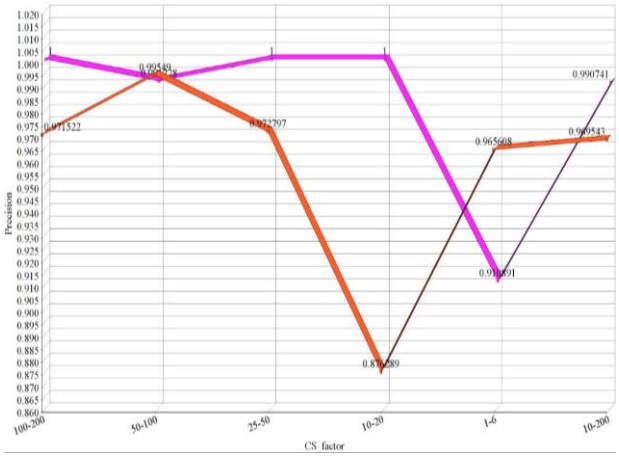

Fig. 6. Precision metric results vs. magnitude ranges of simulated cycle slips for the proposed classification model on the two datasets (orange: first dataset; red: second dataset).

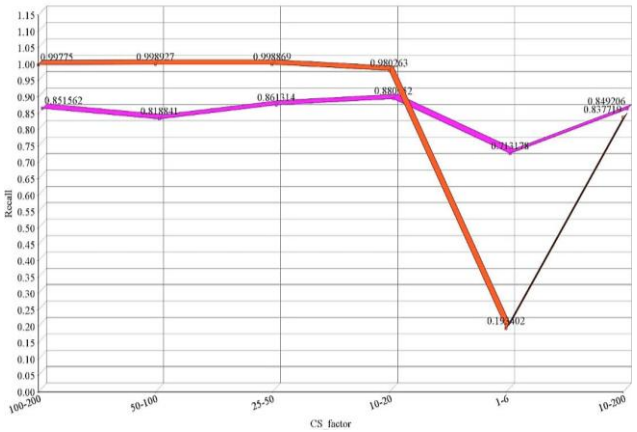

Fig. 7. Recall metric results vs. magnitude ranges of simulated cycle slips for the proposed classification model on the two datasets (orange: first dataset; red: second dataset).

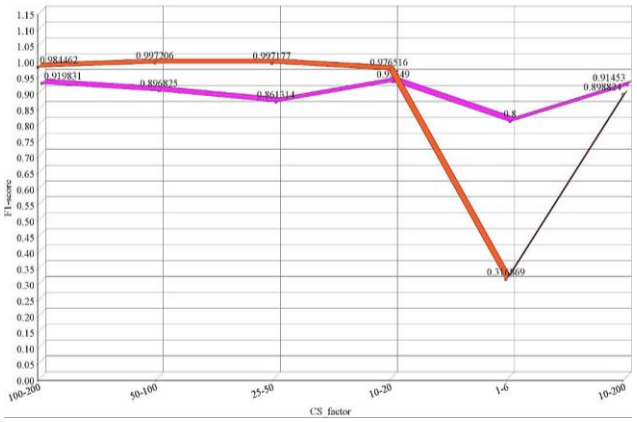

Fig. 8. F1 score metric results vs. magnitude ranges of simulated cycle slips for the proposed classification model on the two datasets (orange: first dataset, red: second dataset).

It was observed from the previous figures that at higher ranges of CS factor, model precision is considerably lower when trained on the second dataset than precision when trained on the first dataset. This is a direct result of using higher cycle slip magnitude in epochs with simulated cycle slips in the training dataset. This results in fewer overall false

positives, (i.e., may even reach zero) contributing to a higher precision value, though this will not generalize well in a broad sense.

On the other hand, the recall value and the f1-score prove higher cycle slip detection accuracy at CS factor ranges higher than the 1, 6 range, when the model is trained on the second dataset compared to models trained on the first dataset; while models in the range of 1, 6 trained on the first dataset perform better than the ones trained on the second dataset.

It was concluded that in a dataset with higher numbers of epochs (as in the second dataset), the model can achieve great recall results, meaning it can detect more cycle slips out of the total simulated cycle slips, but it still performs well even at low number of available epochs with simulated cycle slips (as in the first dataset). Additionally, it should be noted that zero recall is reached when a single satellite phase-difference measurement is used when the simulated cycle slips are added. This means it is less sensitive to single satellite cycle slip occurrences and it is unlikely to detect such occurrences.

Fig. 9, 10, and 11 depict the difference between the trained models on the first and the second datasets at a different number of satellites with added cycle slips for the precision, recall, and f1 score, respectively.

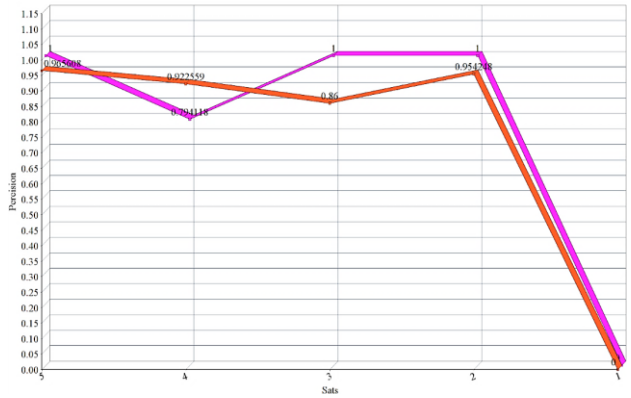

Fig. 9. Precision metric results vs. number of satellites with the simulated cycle slips for the proposed classification model on the two datasets (orange: first dataset; red: second dataset).

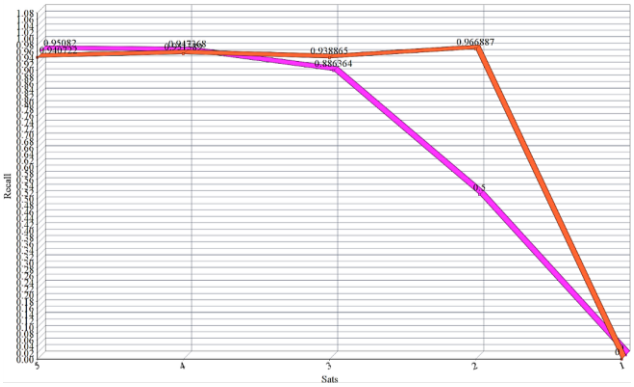

Fig. 10. Recall metric results vs. number of satellites with the simulated cycle slips for the proposed classification model on the two datasets (orange: first dataset; red: second dataset).

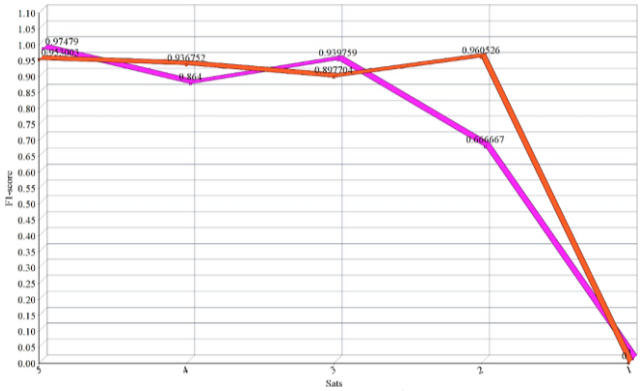

Fig. 11. F1 score metric results vs. number of satellites with the simulated cycle slips for the proposed classification model on the two datasets (orange: first dataset; red: second dataset).

For the localization model results, it is clear from Table 5 that the proposed pipelined network technique has its positive impact on the outcome results. Using the output of the classification model as an input for the localization model proved beneficial in terms of overall precision and recall metrics.

TABLE V
LOCALIZATION MODEL RESULTS WHEN TRAINED AND TESTED ON BOTH DATASETS

Model version	Data	Precision	Recall	F1 Score	True neg	False pos	True pos	False neg
SlipNet-18	Dataset2	0.970	0.966	0.968	1337665	68	2221	78
SlipNet-22		0.998	0.997	0.998	1132462	1083	541021	1594
SlipNet-37	Dataset1	0.967	0.891	0.928	241229	10	295	36
SlipNet-38		0.927	0.955	0.941	156974	5942	75303	3541

Note that the ratio of cycle slips is not equivalent to the ratio of segments with simulated cycle slips declared in the CS_num column in Table 5. This is because the true_pos, for example, are not segments, but rather epochs that contains cycle slips, and this range may be within 1–4 epochs per segment. The same goes for other metrics, which are in epochs not segments. The reason is that a random generation process is used to generate the epoch at which the segment gets corrupted using the cycle slip integer. So, random epochs are generated within the range of 1–4 and a mean value is determined for the corrupted epochs of 2.5, which explains why the total number of cycle slips seem more like almost half of the total number of epochs in all segments, knowing that each segment has five epochs.

As discussed earlier, the cycle slip estimation model was treated differently in terms of loss function, batch size, and number of training epochs. This model was trained using a

batch size of 100 and for 200 epochs. The MAE was used as the loss function to be minimized to achieve a meaningful accuracy metric.

It can be shown in Table 6 that the model trained on the first dataset achieved less MAE than that trained on the second dataset. This is the result of having higher magnitude of phase-difference values in second dataset than in the first.

TABLE VI
CYCLE SLIP ESTIMATION MODEL RESULTS WHEN TRAINED AND TESTED

Model Version	Dataset	MAE
SlipNet-25	Second	0.459
SlipNet-39	First	0.045

V. CONCLUSION

Global Navigation Satellite Systems cycle slip affects the range computation between the GNSS satellites and receivers, which leads to a deteriorated positioning solution for the users. A new real-time deep neural network LSTM autoencoder model technique is proposed for cycle slip detection and repair that achieves comparable results to classical methods of cycle slip detection and estimation. Contrary to classical methods, the results were achieved using datasets with sparse cycle slip epochs, where cycle slips were treated as rare events. These results were achieved by using phase-difference techniques that solicited research for single receiver learning solutions. The results were compared to other methods with respect to the number of epochs with cycle slips and the magnitude of cycle slips (CS_num & CS_factor). This achieved rather promising results, albeit the first end-to-end learning framework to solve the cycle slip detection, localization, and correction problem.

Author's contributions

1- Conception or design of the work (A.R., A.Z., M.E.).
2- Data collection and tools (A.Z.)
3- Data analysis and interpretation (A.Z.)
4- Funding acquisition (M.E.)
5- Investigation (A.R., A.Z., M.E.).
6- Methodology (A.Z., M.E.).
7- Project administration (M.E.)
8- Resources (M.E.)
9- Software (A.Z.)
10- Supervision (M.E.)
11- Drafting the article (A.R., A.Z., and M.E.).
12- Critical revision of the article. (A.R., A.Z., M.E.).
13- Final approval of the version to be published (A.R., A.Z., and M.E.).

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] El-Rabbany, Ahmed. *Introduction to GPS: The Global Positioning System.* Boston, MA: Artech House, 2006

[2] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins, *Global Positioning System. Theory and practice.* Springer-Verlag Wien, 2001

[3] B. Remondi, "Global positioning system carrier phase: Description and use," *Bull. Geod. 59, 361–377*, 1985.

[4] F. and R. W. Collin, "Application of the Wavelet Transform for GPS Cycle Slip Correction and Comparison with Kalman Filter," *Manuscripta Geod. 20(3), 161-72.*, 1995.

[5] Z. Dai, "MATLAB software for GPS cycle-slip processing," *GPS Solut.*, 2012, doi: 10.1007/s10291-011-0249-1.

[6] G. Seeber, "Satellite Geodesy: Foundation, Methods and Applications," *Walter Gruyter, Berlin, New York*, 2003.

[7] P. Silva, "Cycle slip detection and correction using low cost IMU measurements," *Inst. to Super. tenico univerdate Tech. de, Lisboa.*, 2012.

[8] Z. Liu, "A new automated cycle slip detection and repair method for a single dual-frequency GPS receiver," *J. Geod.*, vol. 85, no. 3, pp. 171–183, 2011, doi: 10.1007/s00190-010-0426-y.

[9] U. H. Beutler, G., H. Bock, E. Brockmann, R. Dach, W. Gurtner, H. Habrich, P. D. Ineichen, A. Jaeggi, M. Meindl, L. Mervart, S. Rothacher, R. Schaer, T. Springer, and and R. W. Steigenberger, D. Svehla, D. Thaller, C. Urschl, "BERNESE GPS software," *Univ. Bern*, 2006.

[10] S. B. and R. B. L. Bisnath, "Automated cycle-slip correction of dualfrequency kinematic GPS data," *Proc. 47th Annu. Conf. Can. Aeronaut. Sp. Institute, 121-5.*, 2000.

[11] T. Yi, H. Li and G. Wang, "Cycle Slip Detection and Correction of GPS Carrier Phase Based on Wavelet Transform and Neural Network," *Sixth International Conference on Intelligent Systems Design and Applications*, 2006, pp. 46-50, doi: 10.1109/ISDA.2006.129

[12] B. Huang, L. Lintao and G. Guangxing. "Detection of Cycle-slip in the GPS Precise Point Positioning Based on Wavelet Transform." Geomatics and Information Science of Wuhan University (2006)

[13] X. Shi, X.Su, R. Sheng, "Cycle-slip detection of GPS carrier phase observable based on wavelet technology," *J. Nanjing Univ. Sci. Technol. 29(1), 105-22.*, 2005.

[14] Y. Haitao "An Cycle-Slip Correction Method for Real-Time Kinematic GPS Data Based on Triple Differences Observation." *University* (2007)*Geomatics Inf. Sci. Wuhan Univ. 32(8), 711-4.*, 2007.

[15] Z. Zuoya, L. Xiushan, W. Xinzhou, C. Chuanfa, "GPS phase measurement cycle-slip detection based on a new wavelet function," International Symposium on GPS/GNSS Vol.2 v.2 , 2006

[16] S Bisnath, D. Kim and R. Langley. , "A New Approach to an Old Problem Carrier-Phase Cycle Slips," *GPS World, May 2001, pp 46-51.*, 2001.

[17] H.-K. Lee, J. Wang, and C. Rizos, "Effective Cycle Slip Detection and Identification for High Precision GPS/INS Integrated Systems," *Journal of Navigation*, vol. 56, no. 3, pp. 475–486, 2003.

[18] X. Zhao, Z. Niu, G. Li, Q. Shuai, and B. Zhu, "A new cycle slip detection and repair method using a single receiver's single station B1 and L1 frequencies in ground-based positioning systems," *Sensors (Switzerland)*, vol. 20, no. 2, pp. 1–19, 2020, doi: 10.3390/s20020346.

[19] X. Fan, R. Tian, X. Dong, W. Shuai, and Y. Fan, "Cycle slip detection and repair for BeiDou-3 triple-frequency signals," *Int. J. Adv. Robot. Syst.*, vol. 17, no. 3, pp. 1–14, 2020, doi: 10.1177/1729881420926404.

[20] T. Li and S. Melachroinos, "An enhanced cycle slip repair algorithm for real-time multi-GNSS, multi-frequency data processing," *GPS Solut.*, vol. 23, no. 1, pp. 1–11, 2019, doi: 10.1007/s10291-018-0792-0.

[21] N. S. Kosarev, K. M. Antonovich, and L. A. Lipatnikov, "The method of cycle-slip detection and repair GNSS meaturements by using receiver with high stability frequency oscillator," *Contrib. to Geophys. Geod.*, vol. 49, no. 3, pp. 283–301, 2019, doi: 10.2478/congeo-2019-0015.

[22] Z. Liu, "A new approach for cycle slip detection and fix using single GPS receiver's single satellite dual frequency data containing arbitrarily large pseudorange errors," *J. Glob. Position. Syst.*, vol. 16, no. 1, 2018, doi: 10.1186/s41445-018-0013-8.

[23] W. Liu, X. Jin, M. Wu, J. Hu, and Y. Wu, "A new real-time cycle slip detection and repair method under high ionospheric activity for a triple-frequency GPS/BDS receiver," *Sensors (Switzerland)*, vol. 18, no. 2, 2018, doi: 10.3390/s18020427.

[24] M. El-Tokhey, T. F. Sorour, A. E. Ragheb, M. Moussa, "GPS cycle slips detection and repair through various signal combinations," *Int. J. Mod. Eng. Res.*, vol. 4, no. 12, pp. 01–10, 2014, [Online]. Available: http://www.ijmer.com/papers/Vol4_Issue11/Version-2/A04011_02-0110.pdf.

[25] Ahmed Adel El-Ghazouly, "Multi-Resolution Spectral Techniques for Static DGPS Error Analysis and Mitigation," University of Calgary, 2013.

[26] T. Yi, H. Li, and G. Wang, "Cycle slip detection and correction of GPS carrier phase based on wavelet transform and neural network," 2006, doi: 10.1109/ISDA.2006.129.

[27] N. El-Sheimy, K. W. Chiang, and A. Noureldin, "The utilization of artificial neural networks for multisensor system integration in navigation and positioning instruments," *IEEE Trans. Instrum. Meas.*, 2006, doi: 10.1109/TIM.2006.881033.

[28] C. Cai, Z. Liu, P. Xia, and W. Dai, "Cycle slip detection and repair for undifferenced GPS observations under high ionospheric activity," *GPS Solut.*, 2013, doi: 10.1007/s10291-012-0275-7.

[29] Z. Dai, S. Knedlik, and O. Loffeld, "Instantaneous Triple-Frequency GPS Cycle-Slip Detection and Repair," *Int. J. Navig. Obs.*, vol. 2009, pp. 1–15, 2009, doi: 10.1155/2009/407231.

[30] and A. P. A. Grewal, M. S., L. R. Weill, "GLOBAL POSITIONING SYSTES, INERTIAL NAVIGATION, AND INTEGRATION.," *Hoboken, New Jersey. John Wiley Sons, Inc.*, 2007.

[31] Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. *Neural computation*, *9*(8), pp.1735-1780.

Arabic Title

فحص الانقطاع الموجي لنظم الملاحة العالمي عبر الأقمار الإصطناعية وتقديره باستخدام نموذج التشفير التلقائي القائم على ذاكرة طويلة المدى.

Arabic Abstract

تُستخدم أنظمة الأقمار الإصطناعية للملاحة العالمية في العديد من تطبيقات الملاحة وتحديد المواقع. ولكن، إشارة تلك الأنظمة تعاني من بعض الأخطاء ومنها الانقطاع الموجي مما يؤدي إلى تدهور حل تحديد المواقع. يُعرَّف الانقطاع الموجي على أنها قفزة مفاجئة بعدد صحيح من موجات الطور الحاملة. يعتبر نشاط التروبوسفير العالي من أكثر الأسباب شيوعًا لحدوث تلك الظاهرة. لذلك يجب اكتشاف وتصحيح تلك الظاهرة للحصول على تقديرات موثوقة لتحديد المواقع.

تم اقتراح نهج جديد للكشف عن الانقطاع الموجي وإصلاحه بناءً على فرق الطور الرئيسي مع نموذج شبكة عصبية عميقة قصيرة المدى للذاكرة. يمكن لمثل هذا النموذج تصنيف البيانات المعيبة حيث حدث الانقطاع بين البيانات بأكملها ثم التنبؤ بالفترة الدقيقة التي حدثت فيها الانقطاع الموجي.

يعتبر النظام المقترح من أوائل الأساليب التي تستخدم اطر جديدة للتعليم لحل مشكلة الغموض الصحيح في القياسات بجودة عالية بالمقارنة مع أحدث طرق الكشف عن الانقطاع الموجي وتصحيحه