https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

Geoscientific Model Development Discussions

Novel Deep Learning Approaches for Mapping Variation of Ground Level from Spirit Level Measurements

Fawzi Zarzoura [1], Mosbeh R. Kaloop[1,2,3,4], Pijush Samui[5], Jong Wan Hu[2,3*], Md Shayan Sabri[5], Tamer ElGharbawi[6]

[1]Public Works Engineering Department, Mansoura University, Mansoura, Egypt
[2]Department of Civil and Environmental Engineering, Incheon National University, Incheon, Korea
[3]Incheon Disaster Prevention Research Center, Incheon National University, Incheon, Korea
[4] DigInnoCent s.r.o., Liberec, Czech Republic
[5]Department of Civil Engineering, National Institute of Technology Patna, Patna, India
[6]Civil Engineering Department, Suez Canal University, Ismailia, Egypt
*Correspondence to: Jong Wan Hu (jongp24@inu.ac.kr)
(F.Zarzoura: fawzihamed@mans.edu.eg ; M.Kaloop:mosbeh@mans.edu.eg, P.Samui: pijush@nitp.ac.in ; M.Sabri: mds.pg21.ce@nitp.ac.in; T.ElGharabawi: tgh@eng.suez.edu.eg)

Abstract: This study investigates the use of new machine learning techniques in mapping variation in ground levels based on ordinary spirit levelling (SL) measurements. Convolution Neural Network (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and bi-directional LSTM (BI-LSTM) were developed and compared in the current study to estimate the leveling through SL measurements. SL measurements of the Manzalla region, Egypt, were used in the current study. 3253 datasets of SL observation points, including 229 benchmarks of precise levelling (PL), were used to design and verify the proposed model's results. The results show the developed LSTM model outperforms CNN, RNN, and BI-LSTM in modeling ground leveling in the training and testing stages. The root mean square error and correlation determination of the LSTM model are 7.4 cm and 0.99, respectively, in the testing stage. The accuracy of mapping ground levelling through the developed LSTM model is close to 99% in terms of model error.

Keywords: Spirit levelling, Deep learning, CNN, LSTM, Fitting

1. Introduction

Modeling the variation of the earth's surface is one of the essential requirements in engineering applications. Traditional leveling methods are commonly used in small scale engineering projects; however, satellite systems, e.g., satellite images and coordinate systems such as global positioning systems (GPS), are commonly used in large-scale projects (Ahmed EL-Mowafy 2004; IHO 2011; CDT 2012; Shanker and Acharya 2022). Ordinary spirit levelling (SP) is a lower-cost method compared to other surveying methods, and it is almost always used to cover a wide area of construction projects. However, satellite systems and precise leveling (PL) are almost as costly and used in special survey engineering networks (Karila et al. 2013; Kemboi 2016; Janos et al. 2022). In order to decrease the cost and time of survey engineering works, this study aims to develop a soft computing technique that can be used to map variations of the earth's surface through SP measurements in construction and infrastructure projects.

SP, or levelling, is a process to estimate the land elevation of a measured point based on the known elevation of another point with a level instrument and an ordinary vertical staff. It is known as a relative measurement of leveling with low accuracy. However, it is widely used in construction projects. The details of SP can be

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

found in (Kemboi 2016; LSC 2018). Machine learning was applied for modeling the geoid undulation (Yılmaz et al. 2006; Kaloop et al. 2020b; Tütüncü et al. 2021; Asenso-Gyambibi et al. 2022). However, mapping variation in earth surface or surface elevation through SP is still limited based on our literature. Latitude and longitude are commonly used in leveling modeling (Veronez et al. 2011; Erol and Erol 2013). Erol and Erol (Erol and Erol 2013) applied multivariable polynomial regression equations (MPRE), artificial neural networks (ANNs), adaptive network-based fuzzy inference system (ANFIS) and especially wavelet neural networks (WNNs) to interpolate the geoid surface; ANFIS and WNN outperformed other models. In addition, the ANN model was tested to estimate the geoid height in Brazil, and the results found it was efficient compared to the Brazilian geoid model (MAPGEO2004) (Veronez et al. 2011). Kernel Ridge Regression (KRR) was applied to estimate the Kuwait geoid model based on GPS/Levelling measurements, with the results that its performance is better than that of least squares support vector regression (LSSVR), gaussian process regression (GPR), and multivariate adaptive regression splines (MARS) in modeling the geoid (Kaloop et al. 2019). More studies can be found in (Zhong 1997; Veronez et al. 2011; Rabah and Kaloop 2013; Sorkhabi et al. 2015; Kaloop et al. 2018; Tütüncü et al. 2021) for modeling the geoid. However, due to the limitations of the data used in modeling the geoid, the use of deep learning in geoid modeling is still limited.

Nowadays, Deep learning techniques, such as convolutional neural networks (CNN), recurrent neural networks (RNN), long short-term memory (LSTM), have been used in modeling and classifying land use/land cover (LULC) based on satellite images. Rußwurm and Korner (Rußwurm and Körner 2017) found LSTM to be more efficient in LULC classification. Sun et al. (Sun et al. 2019) evaluated LSTM-RNN, RCNN, and CNN in mapping and classification LULC; their results found the LSTM-RNN model can be precisely used in LULC. Modeling of Land-use and land-cover change (LULCC) through machine learning techniques was collected and discussed in Wang et al. (Wang et al. 2022); the review summarized that machine and deep learning may be limited in "(i) describing occurrence, transition, and spatial patterns of changes ; (ii) unavailability of training data for all the change drivers, particularly sequence data, and (iii) lack of inclusion of local ecological, hydrological, and social-economic drivers when addressing the spectral feature change". Bi-directional long short-term memory (Bi-LSTM) was integrated with the optimal guidance-whale optimization algorithm (OG-WOA) technique to classify and map the LULC (Vinaykumar et al. 2023). The accuracy of Bi-LSTM was found to be better than that of CNN and RNN in LULC classification. Furthermore, CNN, Deep Boltzmann Machines (DBM), Deep Belief Net (DBN), and RNN are used to estimate and evaluate the geodetic velocity, and the results showed that the CNN was better than other deep learning models (Sorkhabi et al. 2022). Other applications of soft computing techniques in water leveling and wave height modeling can be found in (Kaloop et al. 2016, 2020a; Miky et al. 2021; Sinha and Abernathey 2021; Minuzzi and Farina 2023).

Therefore, this study aims to use deep learning techniques, such as CNN, RNN, LSTM, and BI-LSTM, in the elevation interpolation of grid points. With modeling the ground leveling, even in the relative measurements, the time and cost of traditional leveling fieldworks should be decreased. To train the proposed models, 3253 datasets leveling points of SL, including 229 benchmarks of precise levelling (PL), were used. These measurements were collected from a project in the Manzalla region, Egypt. The accuracy of the proposed models was evaluated based on the collected datasets; this means the accuracy should be changed based on the volume of data used and the topography of the study regions. However, the concept of the proposed models can be used in a similar area.

2. Methods

In this study, four input-output deep learning techniques, CNN, RNN, LSTM, and BI-LSTM, are applied and compared to map the leveling based on Latitude and longitude measurements. The following is a theory summary of the proposed models.

2.1. Convolution neural network (CNN)

CNN is one of the most essential deep learning methods, in which multiple layers are powerfully trained. The convolutional layer is the center of the CNN and is the cause of its name. This layer receives and processes data by performing a convolution function. The CNN is made up of multiple convolutional layers that can be integrated or completely connected, as well as multilayer perceptron's (Oh et al. 2019; Wang et al. 2020; Zhang et al. 2022; Sorkhabi et al. 2022). This approach is very effective and one of the most popular approaches in a variety of computer vision applications. The convolution layer, the pooling layer, and the fully connected layer are the three major layers that make up a CNN network.

Different layers carry out different tasks (Jang et al. 2019). The architecture of 1D conventional Neural Network (1D-CNN) for layer-by-layer leveling is shown in Figure 1.a. In general, there are two training phases in each CNN: the feedforward phase and the backpropagation phase. The input signal is fed into the network in the first stage, which consists of multiplying the input by the parameters of each neuron and then performing convolution in each layer to produce the network output (Sorkhabi et al. 2022). In this case, the network parameters are modified, or to put it another way, the network is trained, and the output product is used to figure out how much network error there is. To do this, compute the error rate by comparing the network output with an accurate response using an error function (loss function). Based on the calculated error rate, the backpropagation process starts in the following step. The gradient of each parameter is determined in this phase using the chain rule, and all parameters are changed based on their impact on the error introduced into the network (Sorkhabi et al. 2022). Following the updating of the parameters, the following feed-forward process starts. The network training comes to a conclusion after repeating a significant number of these steps (Sorkhabi et al. 2022). In general, a convolutional neural network is a hierarchical neural network that has a number of completely connected layers after the pooling and convolutional layers.

2.2. Recurrent neural networks (RNN)

The convolutional model works with a fixed number of inputs and produces a fixed vector as an output with a predefined number of steps. We can manipulate vector sequences at both the input and the output thanks to return grids (Hang et al. 2019; Sorkhabi et al. 2022; Amalou et al. 2022). In the case of the RNN, the link between the units forms a direct cycle. The inputs and outputs of a recursive neural network are connected rather than independent, in contrast to conventional neural networks. Additionally, each layer of the RNN uses the same standard settings. RNN design is shown in Figure 1.b. The backpropagation method can be used to train the return network to mimic a conventional neural network (Hang et al. 2019). Here, the flow step is only one factor that is considered in the computation of the gradient. The two-way neural network takes into account both the expected future output and the prior output. Deep learning can be achieved in two-way and direct RNN by adding numerous hidden levels. With a lot of learning data, these deep networks have a greater learning capacity (Sorkhabi et al. 2022).

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

Geoscientific Model Development
Discussions
Open Access

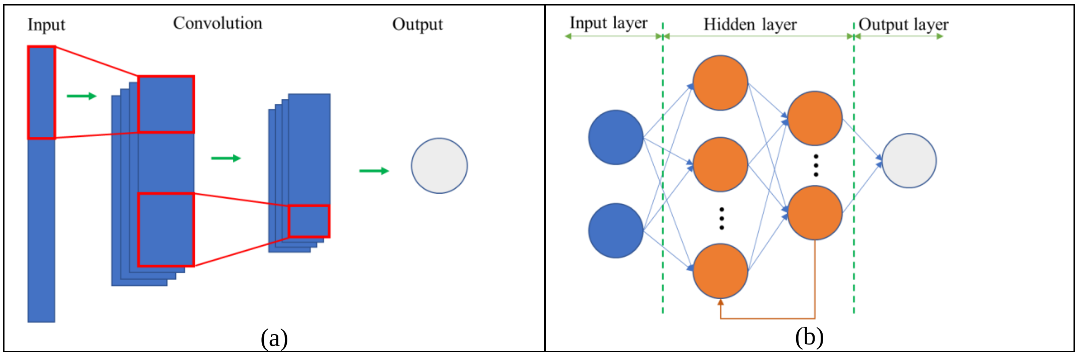

Figure 1. Diagram of Leveling estimation using (a) 1D convolution CNN and (b) Simple RNN

2.3. Long short-term memory (LSTM)

LSTM is a model or structure for sequential data created by (Hochreiter and Schmidhuber 1997) for the advancement of RNN. It employs a unique combination of hidden units, elementwise products, and sums between units to create gates that control "memory cells." These cells are intended to store information without modification for extended periods of time (Apaydin et al. 2020). The most important feature of LSTM is its capacity to learn long-term dependency, which RNNs cannot do. To anticipate the next step, the weight values on the network must be updated, which necessitates the preservation of information from the previous steps. RNN can only learn a finite number of short-term relationships and cannot learn long-term series. However, because LSTM has three gates—input, forget, and output—it can effectively learn these long-term relationships. (Figure 2a). To show how much of the prior memory is remembered and how much of it has been lost, the forget gate is embedded. The concealed state h_t for LSTM is calculated as follows:

$$i_t = \sigma(w_i X_t + u_i h_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(w_f X_t + u_f h_{t-1} + b_f) \quad (2)$$

$$O_t = \sigma(w_O X_t + u_O h_{t-1} + b_O) \quad (3)$$

$$\tilde{C}_t = tanh(w_C X_t + u_C h_{t-1} + b_C) \quad (4)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (5)$$

$$h_t = \tanh\,(C_t) \times O_t \quad (6)$$

Where, $i_t, f_t, and\ O_t$ are the input, forget, and output gates at time t, respectively; $w_i, w_f, w_O, and\ w_C$ are weights that map the hidden layer input to the three gates of input, forget, and output while $u_i, u_f, u_O, and\ u_C$ weights matrices map the hidden layer output to gates; $b_i, b_f, b_O, and\ b_C$ are vectors. Additionally, $C_t\ and\ h_t$ represent the results of the cell and stratum, respectively (Apaydin et al. 2020).

2.4. Bi-directional long short-term memory (Bi-LSTM)

In fact, the network is processed in two directions rather than just one: backward and forward, with two distinct hidden levels. Bidirectional networks performed better than unidirectional networks in situations like phonemic grouping, as shown by Graves and Schmidhuber (Graves and Schmidhuber 2005). Figure 2b shows the bidirectional network's layout. These networks have a structure with a forward and backward LSTM layer based on this image. The forward layer output order, $\vec{h}_t$, is computed repeatedly from time $t - n$ to time $t - 1$ using positive order inputs, whereas the backward layer outcome order, $\overleftarrow{h}_t$, is computed

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

repeatedly using inverted inputs [36]. Both the forward and backward layers outputs are calculated similarly to the unidirectional LSTM. In the Bi-LSTM layer, Yt is computed from Equation (7):

$$Y_t = \sigma(\vec{h}_t, \overleftarrow{h}_t) \tag{7}$$

where σ function is used to combine the two output sequences.

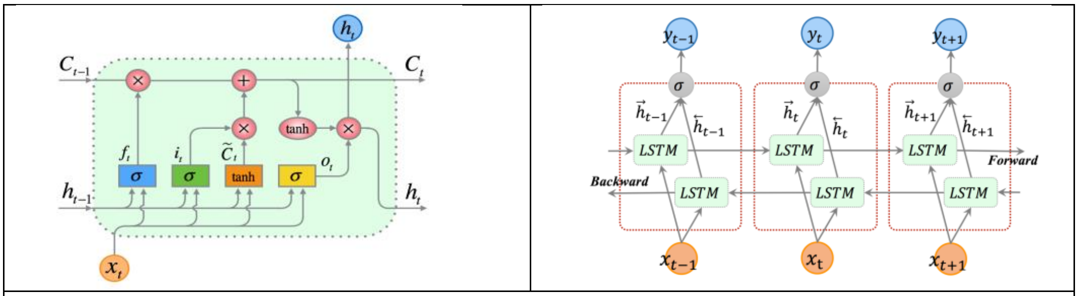

Figure 2. LSTM (a) and BI-LSTM (b) diagrams for levelling prediction (Apaydin et al. 2020)

2.5. Design models and evaluation

To design the proposed models, CNN, RNN, LSTM, and BI-LSTM, trial-and-error runs were performed. Table 1 proposes the details of the model`s configuration. The optimum CNN consists of two input layers, one convolutional layer with 32 neurons followed by a flat layer, a dense layer with 64 neurons, and an output layer. For the best model fit and optimization, mean square error and Adam were used. For a simple RNN, the number of hidden neurons in the hidden layers was investigated in the range of 2 to 40 using trial and error approach, the appropriate value of hidden neuron was determined to be 20 and 8 for the two hidden layers, respectively. The optimum RNN model consists of two input layers, a simple RNN hidden layer with 20 neurons followed by a dense layer with 8 neurons, and an output layer. In each hidden layer, ReLU (rectified linear unit) was used as the activation function. In particular, "mean square error" and "rmsprop" were utilized in the processes of model fitness and optimization, respectively. To achieve optimal performance, LSTM neural networks are structured into the following layers: two input layers, one LSTM hidden layers consisting of 50 neurons each, and an output layer. Similarly, the optimal BI-LSTM network consist of a two-input layer, one hidden BI-LSTM layer with 64 neurons, and one output layer. Notably, the 'sigmoid' function was utilized in each hidden layer, while the Adam and mean squared errors were used for model fitness and optimization, respectively, in both of the models. It is also noted that, different epoch numbers of 100, 300, and 500 were tested in each run of the LSTM and BI-LSTM models, with the optimum epoch found to be 500.

Table 1. Parametric configuration of the developed models

Model	Optimum model configuration
CNN	1D convolutional layer, flatten layer, dense layer, learning rate = 0.0001 batch size = 10, epoch = 500
RNN	20 hidden neurons in simple RNN layer, 8 hidden neurons in dense layer, input shape = (2,1), learning rate = 0.0001, batch size = 16, epoch = 500
LSTM	100 hidden neurons, input shape = (2,1), learning rate = 0.01, batch size = 15, epoch = 500
BI-LSTM	64 hidden neurons, input shape = (2,1), learning rate = 0.01, batch size = 15, epoch = 500

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

To assess the accuracy of the proposed models, different statistical indices were used. The accuracy of the proposed models in term of the correlation between measured and estimated values was evaluated using the coefficient of determination (R^2), where 1 is the best, and Nash-Sutcliffe efficiency (NAF), where 100 is the best. In addition, the model errors were evaluated using root mean square error (RMSE); where 0 is the best, mean absolute error (MAE), where 0 is the best, and mean bias error (MBE), where 0 is the best. Furthermore, percentage error (PE) is applied to measure the accuracy of the proposed models in error terms, 0 is the best, and the overall performance of models is tested using the performance index (PI), where 2 is the best. These indices are presented as follows:

$$R^2 = \frac{\sum_{i=1}^{N}(Z_i - Z_{mean})^2 - \sum_{i=1}^{N}(Z_i - Z_{pi})^2}{\sum_{i=1}^{N}(Z_i - Z_{mean})^2} \quad (8)$$

$$VAF = 100 \times (1 - \frac{var(Z_i - Z_{pi})}{var(Z_i)}) \quad (9)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Z_i - Z_{pi})^2}{N}} \quad (10)$$

$$MAE = \frac{\sum_{i=1}^{N}|(Z_i - Z_{pi})|}{N} \quad (11)$$

$$MBE = \frac{1}{N}\sum_{i=1}^{N}(Z_i - Z_{pi}) \quad (12)$$

$$PI = adj.R^2 + (0.01 \times VAF) - RMSE \quad (13)$$

$$PE = 100 \times \frac{RMSE}{Z_{max} - Z_{min}} \quad (14)$$

Where, Z_i and Z_{pi} represent the measured and predicted levelling, $z_{mean}, z_{max}, and\ z_{min}$ are the average, maximum and minimum, respectively, of measured values, $adj.R^2$ is the adjustment R^2, and N is the number of the data sample.

Figure 3 proposes the data processing and mapping in three stages. Data collection, adjustment, and improvement are implemented in the first stage using the least squares method and PL benchmarks. E and N are collected using GPS observation networks. Z was calculated at observed points using SL equipment. The data were divided into training and testing stages. The training datasets were used to design the proposed models in the second stage. In the last stage, training and testing datasets were used to assess the performance of the proposed models. In addition, the whole datasets were used to validate the best-fit model. Also, grids of 500 m were generated to map the ground level of the study area.

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

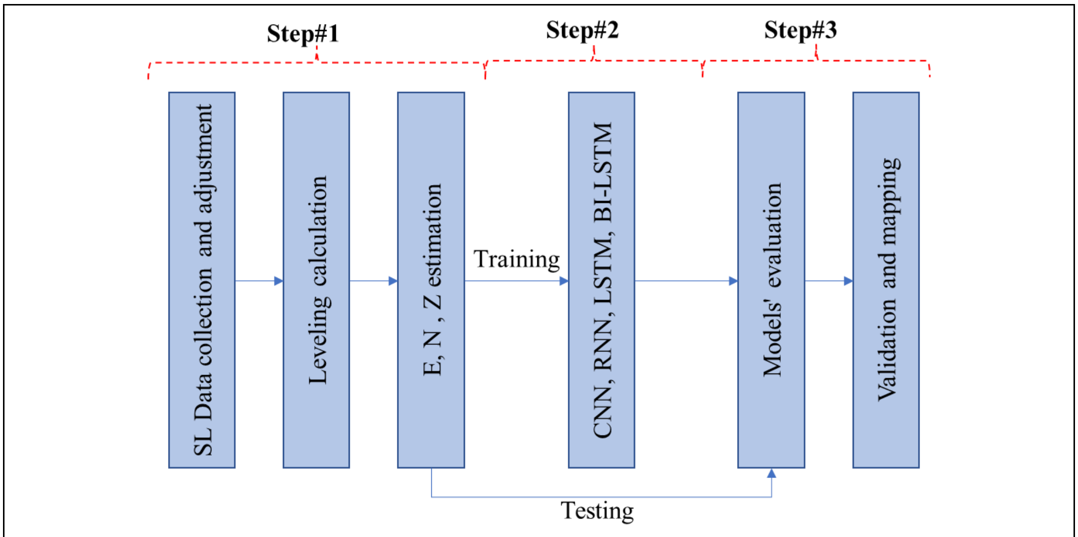

Figure 3. Data processing and modeling flowchart

3. Data collection

The ground level observations were collected in 2019 for a national Egyptian project that aims at the full rehabilitation of the sanitary networks of the Egyptian small cities. The Sokkia B40 automatic level device was utilized to measure elevations, and the benchmarks were connected by paths to form a closed network. The corrections for the measurements were estimated using the Least Squares method. The observations were carried out by a professional team of surveyors under the management of a consulting engineering team. Figure 4 shows the study area, Manzalla region, and distribution of levelling points. The project started by fixing more than 250 benchmarks covering the whole study region. Then these benchmarks were connected using a full, precise leveling network, which was observed with great care by professional surveyors and calibrated spirit level instruments. The network leveling observations were analyzed, filtered, and corrected using the main principles of random error theory and rejection of outlier observations. Then, the precise leveling network was corrected using the least squares method, which presented an estimated standard deviation of nearly 6 millimeters for the estimated observations of the benchmarks. All the benchmarks that presented a standard deviation of more than 1 cm were removed from the network, leaving 229 benchmarks used in the ground level observations.

After estimating the corrected reduced levels of the selected benchmarks, they were used as a reference for ground leveling observations, which were conducted along the longitudinal center of every street in the study region with spacing ranging between 10 and 20 meters. The ground level observations were collected, analyzed, and filtered to remove any blunder observations.

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

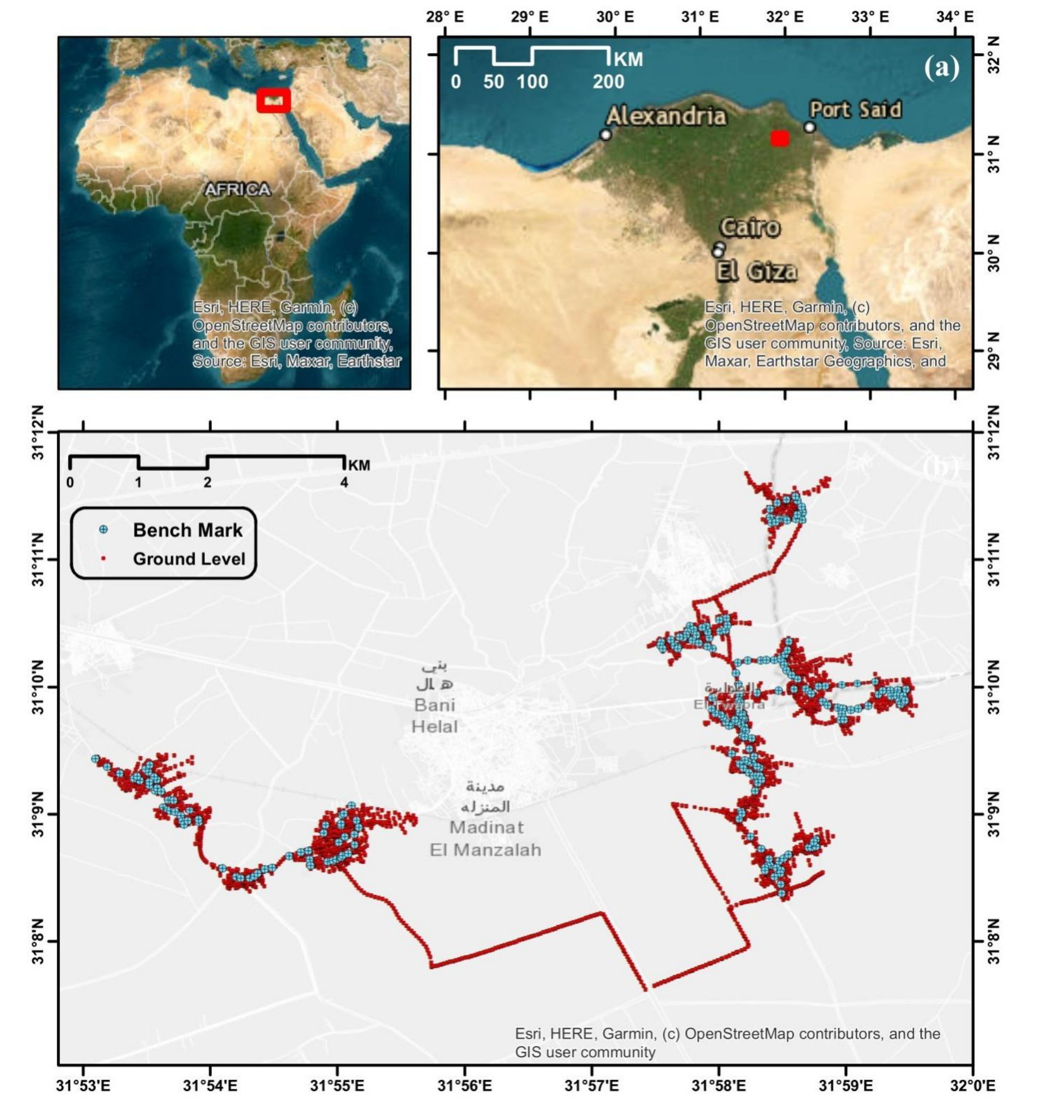

Figure 4. Study area and levelling points (a) study location, (b) leveling points and routes (*Images resources from Esri*)

Figure 5 and Table 1 present the data collection evaluation. histogram of trend data used, and the normal distribution of data used. The mean (M), maximum (MX), minimum (MN), standard deviation (STDEV), and correlation (Corr) between input variables (E and N) and output variables (Z) are presented in Table 1. The data distribution is shown to be non- normal. Negative and positive correlations between E and N and Z, respectively, are observed. The statistical evaluation and data distribution show there is a nonlinear correlation between Z and E,N. This indicates that a non-linear relationship between the input and output

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

variables can be detected, which is advantageous for using deep learning approaches in modeling the ground levels.

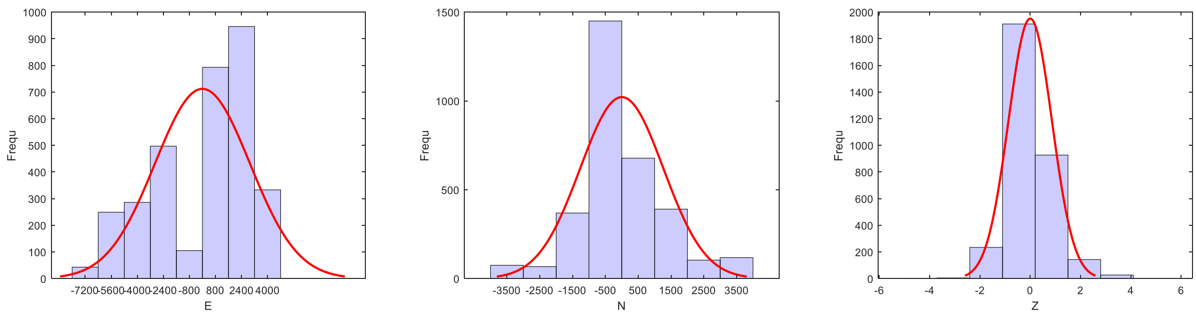

Figure 5. Histogram and normal distribution of data used

Table 1. Statistical evaluation of measured datasets

Variables	E	N	Z
M	399952.067	3447484.489	2.023
MX	403979.133	3451633.725	6.640
MN	393723.542	3444154.308	-2.939
STDEV	3089.666	1558.468	0.876
Corr	-0.394	0.218	1.000

4. Results and Discussion

Figure 6 and Table 2 present the performance evaluations of the proposed models. The rank of the proposed models is presented in Table 2. A high score value indicates the best performance. In the training stage, LSTM outperformed other deep learning techniques. The correlation between measured and predicted levels for LSTM is shown to be high, R^2 and VAF are 0.99 and 99%, respectively. The estimation error of the LSTM model is shown to be low: RMSE = 8.0 cm, MAE = 7.2 cm, and MBE = -2 cm. The overall performance of LSTM is the best compared to other models PI = 1.9, and the percentage of the model error is 0.84%. The rank of LSTM is 26, followed by BI-LSTM, rank = 22. The model followed in modeling the leveling is BI-LSMT with PI=1.67 and PE = 2.18%; while the worst model is shown to be the CNN model. From Figure 6, it can be seen that the scatter plot of LSTM is very close to the best fitting line (dashed line), while the variation around the best fitting line is high for CNN and RNN.

As a result, in the testing stage, the LSTM model outperformed other models in all statistical indices. The rank evaluation of the proposed models showed that LSTM has a high rank compared to other models. CNN was shown to be better than RNN at this stage. BI-LSTM still followed LSTM in the testing stage to model the ground levels. From Figure 6, it can be seen that the scatter plot of LSTM is very close to the best fitting line, while the variation around the best fitting line is high for CNN and RNN. The performance of BI-LSTM is shown to be acceptable in the testing stage, rank = 21; however, LSTM performance achieved a high rank (28) in estimating ground levels.

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

Geoscientific Model Development
Discussions
Open Access
EGU

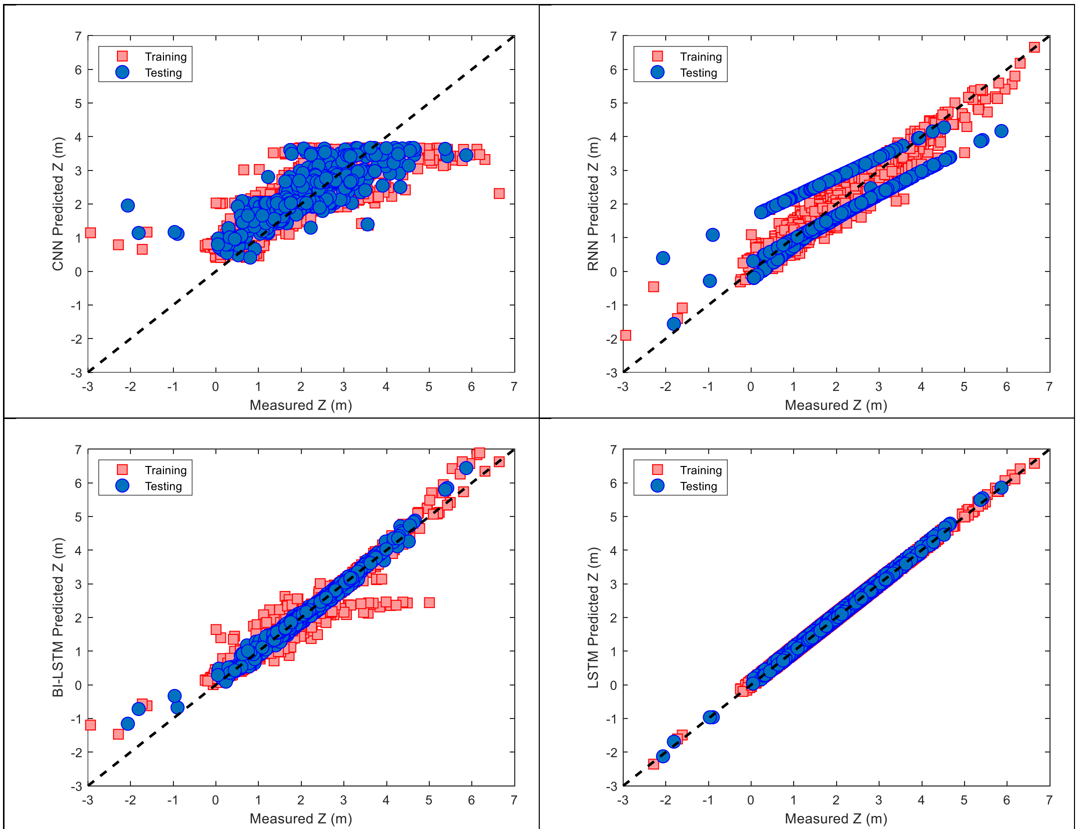

Figure 6. Training and testing scatter plot of predicted and measured leveling

Table 2. Models' performance assessments in the training and testing stages

Training		R^2	VAF	PI	RMSE	MAE	MBE	PE	Total score
CNN	Value	0.614	60.807	0.840	0.382	0.426	-0.189	3.986	7
	Ranke	1	1	1	1	1	1	1	
RNN	Value	0.890	89.003	1.510	0.270	0.239	-0.015	2.819	15
	Ranke	2	2	2	2	2	3	2	
LSTM	Value	**0.991**	**99.145**	**1.903**	**0.080**	**0.072**	-0.020	**0.839**	**26**
	Ranke	**4**	**4**	**4**	**4**	**4**	2	**4**	
BI-LSTM	Value	0.940	93.961	1.670	0.209	0.126	**-0.001**	2.180	22
	Ranke	3	3	3	3	3	**4**	3	
Testing									
CNN	Value	0.645	63.232	0.905	0.372	0.429	-0.186	4.685	13
	Ranke	2	2	2	2	2	1	2	
RNN	Value	0.516	48.580	0.439	0.562	0.557	0.053	7.082	8
	Ranke	1	1	1	1	1	2	1	
LSTM	Value	**0.993**	**99.326**	**1.912**	**0.074**	**0.062**	**-0.012**	**0.936**	**28**
	Ranke	**4**	**4**	**4**	**4**	**4**	**4**	**4**	
BI-LSTM	Value	0.978	97.840	1.825	0.132	0.100	-0.014	1.663	21
	Ranke	3	3	3	3	3	3	3	

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

Geoscientific Model Development
Discussions

In addition, the performances, of LSTM and BI-LSTM models in modeling ground variation levels are presented in Figure 7 in the testing phase. From the figure, it can be seen that the variation of errors for both models are small. The error ranges of LSTM and BI-LSTM are (-0.13 to 0.07) and (-1.09 to 0.26) m, respectively. Thus, both models can be used to estimate levelling.

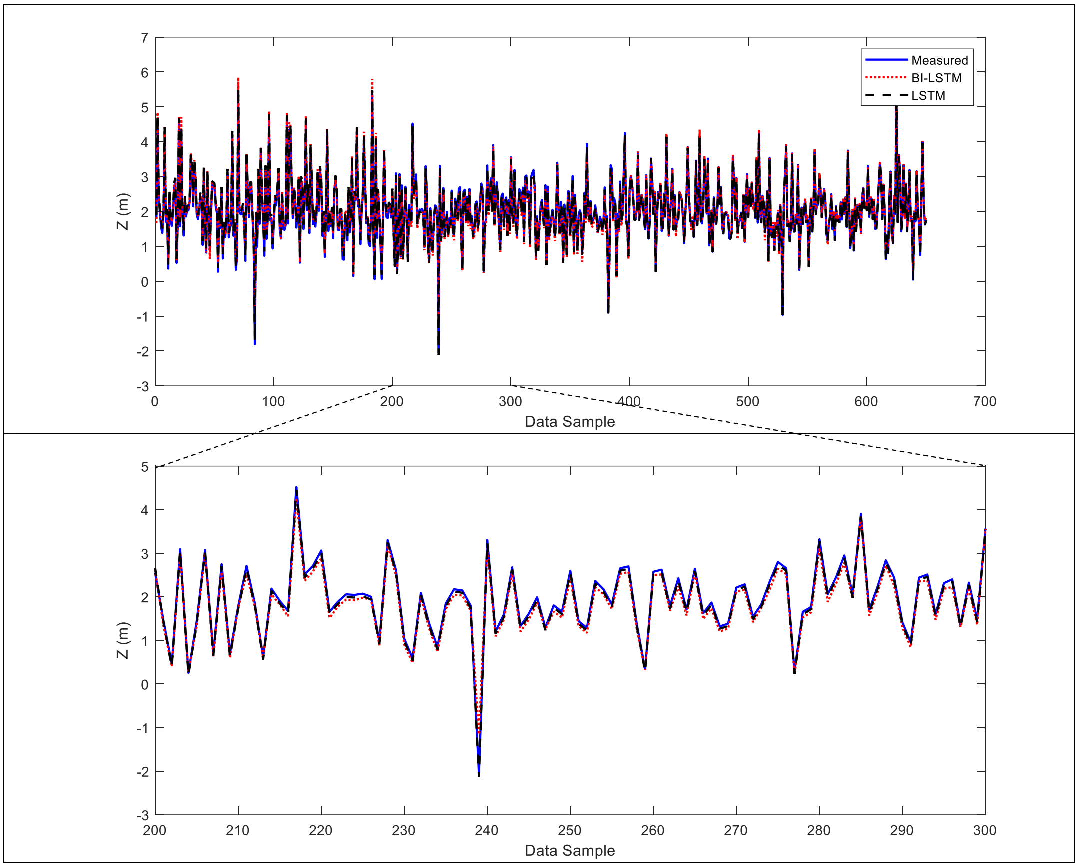

Figure 7. Modeling performance of LSTM and BI-LSTM models in the testing stage

For more investigation, a visualization plot is used to assess the performance of the proposed models. The Taylor diagram (Figure 8.a) and boxplot (Figure 8.b) are presented to evaluate overall performance and model errors, respectively. Taylor's diagram is a two-dimensional diagram that provides a comparative review of models in terms of R-value, root mean square deviation (RMSD), and ratio of standard deviation between measured and predicted values. The best model is the one closest to the reference point. The details of the Taylor diagram can be found in (Taylor 2005). Here, test datasets are used to assess the proposed models based on untuned datasets of models. Taylor's diagram shows the overall performance of LSTM in modeling ground level is better than that of BI-LSTM, CNN, and RNN, respectively. The accuracy of BI-LSTM is close to that of LSTM, and it can be used for ground level estimation. However, boxplots show there are outliers that can be observed with BI-LSTM. In addition, the model error range of LSTM is very

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

Geoscientific Model Development
Discussions
Open Access

small, and the interquartile range of LSTM is very small compared to other models. Although the CNN model outperforms the RNN model in terms of overall performance, it has a high number of outliers. Boxplot obviously shows the LSTM model can be accurately used in ground level estimation.

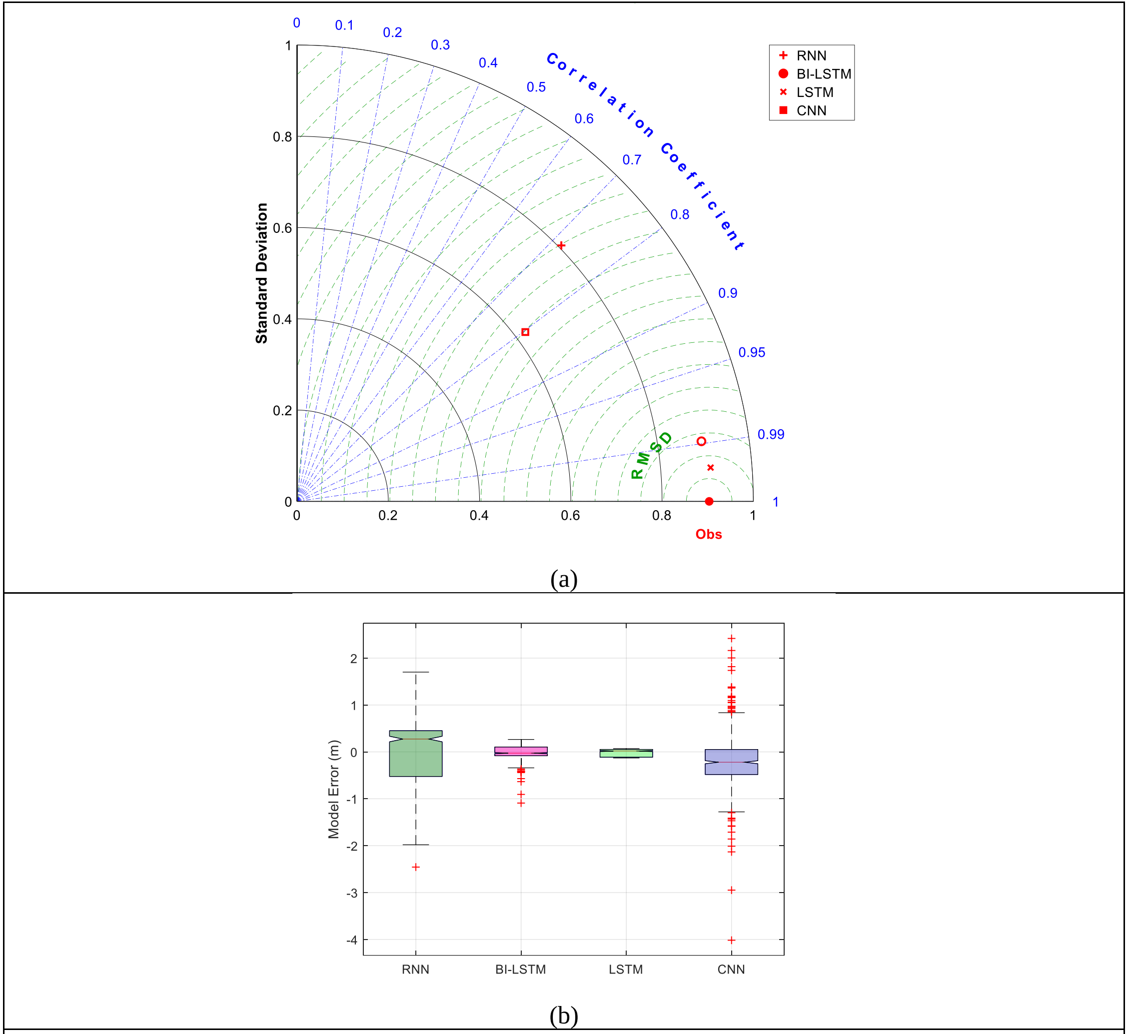

Figure 8. Visualization analysis of the proposed models (a) Taylor diagram (b) boxplot

Figure 9 shows the map estimation of the study area for ground levels and the error of the estimation levels. The error is the difference between levels of mapping and measurements. From the map, it can be seen that the ground level is smooth and slopes from 0 to 6 m in one direction. From the measured errors, it can be seen that the absolute mean error of the estimated ground levels is 0.187 cm, and the standard deviation of error is 0.666 m. The error distribution is roughly normal, and the majority of the confidence in the model error falls within the 95% confidence interval. This indicates that the estimated levels are acceptable, and that LSTM can be accurately applied to estimate the ground level of the study area. These results reveal the

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

Geoscientific Model Development Discussions
Open Access
EGU

proposed model is accurate in estimating ground level, and LSTM can be applied in similar areas to decrease the cost and time of SL field works.

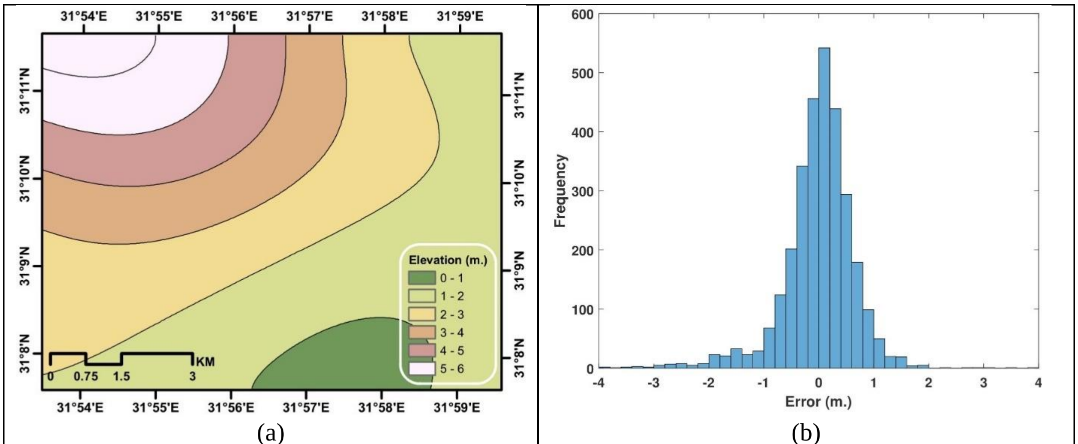

Figure 9. Mapping of the ground level, (a) contour map and (b) model error

5. Conclusions

In the current study, the applicability of using deep learning techniques for mapping ground relative levelling from spirit leveling (SL) measurements was investigated. Convolution Neural Network (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and bi-directional LSTM (BI-LSTM) were developed and compared to estimate the leveling through SL measurements of Manzalla region, Egypt. 3253 datasets leveling points of SL including 229 benchmark points of precise levelling (PL) were used to map an area of about 77 Km2 and to verify the proposed models.

In a comparative study, the proposed models showed overall performances of RNN, BI-LSTM, and LSTM models of 1.51, 1.67, and 1.90, respectively, in the training stage. The overall performance of CNN is 0.91, while the overall performance (PI) of the BI-LSTM and LSTM models is 1.82 and 1.91, respectively, in the testing stage. The accuracy of BI-LSTM and LSTM models in estimating ground level reaches up to 98.5% and 99% in terms of model error (PE). The visualization evaluation of the proposed models showed LSTM outperformed other models in terms of the Taylor diagram and box plot. Thus, the LSTM model can be considered an accurate soft computing model that can be used to estimate the ground level of the study area. With the same concepts, it can be applied in the same regions. LSTM is applied to map the ground level of the study area, and the results show that the estimated accuracy of the ground level of the study area is 0.187 cm + 0.666 m. The error distribution of the model error is significantly within the 95% interval. These results reveal the proposed model is accurate in estimating ground level, and LSTM can be applied in similar areas to decrease the cost and time of SL field works.

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

Geoscientific Model Development
Discussions
Open Access

Author contributions: FZ, MRK, and TE conceptualized the study, collected, and analyze the data, evaluated the results, and wrote the manuscript. **MRK, PS, and MSS** visualize the data, modeling designed and evaluated, and revised the manuscript. **MRK and JWH** revised the final form.

Competing interests: The contact author has declared that none of the authors has any competing interests.

Acknowledgements: Authors are grateful for the Sustainable Rural Sanitation Services Program (SRSSP) in Dakahlia Governorate for providing ground level observations.

Financial support: This work is supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2022-00143541).

Data availability: The used data in this study can be found in the supplementary materials.

References

Ahmed EL-Mowafy (2004) Surveying with GPS for Construction Works Using the National RTK Reference Network and Precise Geoid Models. In: 1st FIG International Symposium on Engineering Surveys for Construction Works and Structural Engineering

Amalou I, Mouhni N, Abdali A (2022) Multivariate time series prediction by RNN architectures for energy consumption forecasting. Energy Reports 8:1084–1091. https://doi.org/10.1016/j.egyr.2022.07.139

Apaydin H, Feizi H, Sattari MT, et al (2020) Comparative analysis of recurrent neural network architectures for reservoir inflow forecasting. Water (Switzerland) 12:. https://doi.org/10.3390/w12051500

Asenso-Gyambibi D, Lamkai N, Peprah M, et al (2022) Novel Ellipsoidal Heights Predictive Models Based on Artificial Intelligence Training Algorithms and Classical Regression Models Techniques: A Case Study in the Greater Kumasi Metropolitan Area Local Geodetic Reference Network, Kumasi, Ghana. International Journal of Earth Sciences Knowledge and Applications 4:493–515

CDT (2012) Global Positioning System (GPS) Survey Specifications. California

Erol B, Erol S (2013) Learning-based computing techniques in geoid modeling for precise height transformation. Comput Geosci 52:95–107. https://doi.org/https://doi.org/10.1016/j.cageo.2012.09.010

Graves A, Schmidhuber J (2005) Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Networks 18:602–610. https://doi.org/https://doi.org/10.1016/j.neunet.2005.06.042

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

Hang R, Liu Q, Hong D, Ghamisi P (2019) Cascaded Recurrent Neural Networks for Hyperspectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing 57:5384–5394. https://doi.org/10.1109/TGRS.2019.2899129

Hochreiter S, Schmidhuber J (1997) Long Short-Term Memory. Neural Comput 9:1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

IHO (2011) Positioning. https://iho.int/uploads/user/pubs/cb/c-13/english/C-13_Chapter_2.pdf. Accessed 19 Feb 2023

Jang Y, Ahn Y, Kim HY (2019) Estimating Compressive Strength of Concrete Using Deep Convolutional Neural Networks with Digital Microscope Images. Journal of Computing in Civil Engineering 33:1–11. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000837

Janos D, Kuras P, Ortyl Ł (2022) Evaluation of low-cost RTK GNSS receiver in motion under demanding conditions. Measurement 201:111647. https://doi.org/https://doi.org/10.1016/j.measurement.2022.111647

Kaloop MR, Kumar D, Zarzoura F, et al (2020a) A wavelet - Particle swarm optimization - Extreme learning machine hybrid modeling for significant wave height prediction. Ocean Engineering 213:107777. https://doi.org/https://doi.org/10.1016/j.oceaneng.2020.107777

Kaloop MR, Rabah M, Elnabwy M (2016) Sea Level Change Analysis and Models Identification Based on Short Tidal Gauge Measurements in Alexandria, Egypt. Marine Geodesy 39:. https://doi.org/10.1080/01490419.2015.1134735

Kaloop MR, Rabah M, Hu JW, Zaki A (2018) Using advanced soft computing techniques for regional shoreline geoid model estimation and evaluation. Marine Georesources & Geotechnology 36:688–697. https://doi.org/10.1080/1064119X.2017.1370622

Kaloop MR, Zaki A, Al-Ajami H, Rabah M (2019) Optimizing Local Geoid Undulation Model using GPS/Levelling Measurements and Heuristic Regression Approaches. Survey Review 1–11. https://doi.org/10.1080/00396265.2019.1665615

Kaloop MR, Zaki A, Al-Ajami H, Rabah M (2020b) Optimizing Local Geoid Undulation Model using GPS/Levelling Measurements and Heuristic Regression Approaches. Survey Review 52:. https://doi.org/10.1080/00396265.2019.1665615

Karila K, Karjalainen M, Hyyppä J, et al (2013) A comparison of precise leveling and Persistent Scatterer SAR Interferometry for building subsidence rate measurement. ISPRS Int J Geoinf 2:797–816. https://doi.org/10.3390/ijgi2030797

Kemboi KE (2016) ESTIMATION OF ORTHOMETRIC HEIGHT USING EGM2008 AND GPS OVER NAIROBI COUNTY AND ITS ENVIRONS Determination of an Optimal Trunk Sewer-line Route for Kikuyu Town Using Geospatial Technologies View project Geoid modelling and height systems View project

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

Geoscientific Model Development Discussions

LSC (2018) Spirit Leveling . In: USGS

Miky Y, Kaloop MR, Elnabwy MT, et al (2021) A Recurrent-Cascade-Neural network- nonlinear autoregressive networks with exogenous inputs (NARX) approach for long-term time-series prediction of wave height based on wave characteristics measurements. Ocean Engineering 240:109958. https://doi.org/https://doi.org/10.1016/j.oceaneng.2021.109958

Minuzzi FC, Farina L (2023) A deep learning approach to predict significant wave height using long short-term memory. Ocean Model (Oxf) 181:102151. https://doi.org/https://doi.org/10.1016/j.ocemod.2022.102151

Oh BK, Glisic B, Kim Y, Park HS (2019) Convolutional neural network-based wind-induced response estimation model for tall buildings. Computer-Aided Civil and Infrastructure Engineering 34:843–858. https://doi.org/10.1111/mice.12476

Rabah M, Kaloop M (2013) The use of minimum curvature surface technique in geoid computation processing of Egypt. Arabian Journal of Geosciences 6:. https://doi.org/10.1007/s12517-011-0418-0

Rußwurm M, Körner M (2017) Multi-temporal land cover classification with long short-term memory neural networks. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives. International Society for Photogrammetry and Remote Sensing, pp 551–558

Shanker KC, Acharya TD (2022) Advancements of Geodetic Activities in Nepal: A Review on Pre-and Post-2015 Gorkha Earthquake Eras with Future Directions. Remote Sens (Basel) 14

Sinha A, Abernathey R (2021) Estimating Ocean Surface Currents With Machine Learning. Front Mar Sci 8:. https://doi.org/10.3389/fmars.2021.672477

Sorkhabi OM, Milani M, Seyed Alizadeh SM (2022) Investigating the Efficiency of Deep Learning Methods in Estimating GPS Geodetic Velocity. Earth and Space Science 9:. https://doi.org/10.1029/2021EA002202

Sorkhabi OM, Propagation AB, Neural A, Bpann N (2015) Geoid Determination Based on Log Sigmoid Function of Artificial Neural Networks : (A case Study : Iran). Journal of Artificial Intelligence in Electrical Engineering 3:18–24

Sun Z, Di L, Fang H (2019) Using long short-term memory recurrent neural network in land cover classification on Landsat and Cropland data layer time series. Int J Remote Sens 40:593–614. https://doi.org/10.1080/01431161.2018.1516313

Taylor KE (2005) Taylor Diagram Primer

https://doi.org/10.5194/gmd-2023-62
Preprint. Discussion started: 12 April 2023

Tütüncü K, Şahman MA, Tuşat E (2021) A hybrid binary grey wolf optimizer for selection and reduction of reference points with extreme learning machine approach on local GNSS/leveling geoid determination. Appl Soft Comput 108:107444. https://doi.org/https://doi.org/10.1016/j.asoc.2021.107444

Veronez MR, de Souza SF, Matsuoka MT, et al (2011) Regional mapping of the geoid using GNSS (GPS) measurements and an artificial neural network. Remote Sens (Basel) 3:668–683. https://doi.org/10.3390/rs3040668

Vinaykumar VN, Babu JA, Frnda J (2023) Optimal guidance whale optimization algorithm and hybrid deep learning networks for land use land cover classification. EURASIP J Adv Signal Process 2023:13. https://doi.org/10.1186/s13634-023-00980-w

Wang J, Bretz M, Dewan MAA, Delavar MA (2022) Machine learning in modelling land-use and land cover-change (LULCC): Current status, challenges and prospects. Science of The Total Environment 822:153559. https://doi.org/https://doi.org/10.1016/j.scitotenv.2022.153559

Wang S, Zhou J, Lei T, et al (2020) Estimating Land Surface Temperature from Satellite Passive Microwave Observations with the Traditional Neural Network, Deep Belief Network, and Convolutional Neural Network. Remote Sens (Basel) 12:2691. https://doi.org/10.3390/rs12172691

Yılmaz M, Acar M, Ayan T, Arslan E (2006) Application of Fuzzy Logic Theory to Geoid Height Determination. In: Kłopotek MA, Wierzchoń ST, Trojanowski K (eds) Intelligent Information Processing and Web Mining. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 383–388

Zhang Y, Zhang C, Ma Q, et al (2022) Automatic prediction of shear wave velocity using convolutional neural networks for different reservoirs in Ordos Basin. J Pet Sci Eng 208:. https://doi.org/10.1016/j.petrol.2021.109252

Zhong D (1997) Robust estimation and optimal selection of polynomial parameters for the interpolation of GPS geoid heights. J Geod 71:552–561. https://doi.org/10.1007/s001900050123