

A PROPOSED MERGING METHODS OF DIGITAL ELEVATION MODEL BASED ON ARTIFICIAL NEURAL NETWORK AND INTERPOLATION TECHNIQUES FOR IMPROVED ACCURACY

Mustafa K. ALEMAM^{1,2}, Bin YONG^{1,3}, Abubakar SANI-MOHAMMED⁴

 ¹ School of Earth Sciences and Engineering, Hohai University, Nanjing, China
 ² Faculty of Engineering, Assiut University, Assiut, Egypt
 ³ State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, China
 ⁴ Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

e-mail: mustafaalemam2014@gmail.com

ABSTRACT. The digital elevation model (DEM) is one of the most critical sources of terrain elevations, which are essential in various geoscience applications. Most of these applications need precise elevations, which are available at a high cost. Thus, sources like the Shuttle Radar Topography Mission (SRTM) DEM are frequently accessible to all users but with low accuracy. Consequently, many studies have tried to improve the accuracy of DEMs acquired from these free sources. Importantly, using the SRTM DEM is not recommended for an area that partly contains high-accuracy data. Thus, there is a need for a merging technique to produce a merged DEM of the whole area with improved accuracy. In recent years, advancements in geographic information systems (GIS) have improved data analysis by providing tools for applying merging techniques (like the minimum, maximum, last, first, mean, and blend (conventional methods)) to improve DEMs. In this article, DEM merging methods based on artificial neural network (ANN) and interpolation techniques are proposed. The methods are compared with other existing methods in commercial GIS software. The kriging, inverse distance weighted (IDW), and spline interpolation methods were considered for this investigation. The essential step for achieving the merging stage is the correction surface generation, which is used for modifying the SRTM DEM. Moreover, two cases were taken into consideration, i.e., the zeros border and the H border. The findings show that the proposed DEM merging methods (PDMMs) improved the accuracy of the SRTM DEM more than the conventional methods (CDMMs). The findings further show that the PDMMs of the H border achieved higher accuracy than the PDMMs of the zeros border, while kriging outperformed the other interpolation methods in both cases. The ANN outperformed all methods with the highest accuracy. Its improvements in the zeros and H border respectively reached 22.38% and 75.73% in elevation, 34.67% and 54.83% in the slope, and 40.28% and 52.22% in the aspect. Therefore, this approach would be cost-effective, especially in critical engineering projects.

Keywords: digital elevation model, GIS, artificial neural network, interpolation methods, SRTM



1. INTRODUCTION

1.1. Background and related work

Accurate terrain elevation is crucial for numerous geoscientific engineering applications. digital elevation model (DEM) is one of the most important sources of elevation (Hawker, Neal, and Bates 2019). In geostatistical analysis, DEM is used to quantitatively represent the bare Earth's surface, which is used to describe the spatial patterns of various surfaces, e.g., surface water, ground surface, and canopy (Mukherjee et al. 2013). This can be represented in a raster or vector model in geographical information system (GIS) (Habib et al. 2020, Huang and Yang 2011). Aside the DEM, digital surface model (DSM) and digital terrain model (DTM) are other elevation surfaces that are frequently used for the ground terrain (Kim et al. 2020). However, DTM represents the Earth's terrain, i.e., bare ground (Xiong, Wang, and Wessel 2017), while DSM includes objects on the ground, such as buildings, vegetation, and human-made structures (Nwilo et al. 2022). Apparently, most data providers and professional users use the term "digital elevation model" as a general term for digital topography (Badura and Przybylski 2005), (Varga and Bašić 2015). It is important to note that there are four different representations of DEM. They are regular grid, outlines, triangulated grid, and triangulated irregular network (TIN) (Khemiri et al. 2013, Leitão, Prodanović, and Maksimović 2016). Generally, DEM can be created using remote sensing (Habib et al. 2020), field surveying (Mesa-Mingorance and Ariza-López 2020), aerial photography (Arun 2013), satellite images (Yap et al. 2019), interferometry, and laser surveying. This can also be achieved through scanning and digitizing topographic maps (Khemiri et al. 2013) and field surveying, e.g., the leveling, total station, and global navigation satellite system (GNSS) survey (Nwilo et al. 2022).

Recently, Light Detection and Ranging (LIDAR), Radio Detection and Ranging (RADAR) altimetry (Yap et al. 2019), and Interferometric Synthetic Aperture Radar (InSAR) (Zhou et al. 2012) have been the most advanced and widely used technologies for acquiring threedimensional information over large areas with high horizontal resolution and high vertical accuracy (Jain et al. 2007). These techniques are expensive, i.e., high instrument and survey costs (Shaikh, Yadav and Manekar 2021, Capolupo 2021). Thus, they are not well integrated in developing countries (Jakovljevic et al. 2019, Kulp and Strauss 2018). Consequently, most users in the developing countries resort to freely accessible open-source products.

There are several global DEMs (GDEMs), such as the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and the Shuttle Radar Topography Mission (SRTM), created for free and commercial use (Reuter, Strobl and Mehl 2011). These GDEMs represent the new-generation models, which have been significantly improved in comparison with the first released models like the global 30 arc-second elevation (GTOPO30) (Mukherjee et al. 2013) and one arc-minute global relief model (ETOPO1) (Varga and Bašić 2015).

In this study, the SRTM DEM (1 arc-second) is considered in the investigation. It is a joint international project of the National Aeronautics and Space Administration (NASA), the National Geospatial-Intelligence Agency, and the German and Italian Space Agencies to collect three-dimensional digital mapping (Yang, Meng and Zhang 2011). The SRTM DEM is mostly used as the essential alternative data when there is no DEM availability, or there is budget constraint to purchase costly high-resolution satellite data (Wendi et al. 2016). Moreover, the GDEMs are vital sources of topographic information for many geoscience studies and applications (Mirza, Dawod and Al-Ghamdi 2011). Some examples include, the geodetic survey (Yamazaki et al. 2017), geological survey, water resources and hydrology (Hawker et al. 2019), evaluation of natural hazards and vegetation survey (Wendi et al. 2016), and mineral and petroleum industries (Badura and Przybylski 2005). In addition, they are applied in engineering works, e.g., highway construction and wind turbine location optimization (Akturk

and Altunel 2019); land surface modeling, such as management of flood risk (Jakovljevic et al. 2019), volcanology, ecology, and glaciology (Kim et al. 2020, Mukherjee et al. 2013). Despite these achievements, their accuracy, still, does not meet the requirements for precise applications due to DEM errors (Yamazaki et al. 2017). Therefore, the accuracy assessment of DEMs is critical in these applications (Varga and Bašić 2015).

DEM error sources include data collection error (gross error), lacking orientation of stereoimages (systematic error), data processing methodology (Yue et al. 2007), georeferencing of stereo-images, grid spacing, interpolation techniques for DEM generation, characteristics of the terrain surface (Chen and Yue 2010), and random error, which cannot be avoided (Varga and Bašić 2015, Mukherjee et al. 2013). As a result, several investigations on DEM have been conducted by researchers leading to many methodical propositions to reduce errors and increase accuracy. For example, Yue et al. (2007), Cook et al. (2012), Wilson (2012), Sari et al. (2019), Hugenholtz et al. (2013), Chen and Yue (2010), Warriner (2005), and Zhou et al. (2012) suggested different methods for generating improved DEM. Further, DEM accuracy was evaluated by Hawker et al. (2019) over Croatia. Varga and Bašić (2015) investigated floodplain sites, while Akturk and Altunel (2019) considered a study area that was highly broken and vegetated terrain. Furthermore, other studies evaluated the vertical accuracy of open-source DEM, i.e., SRTM and ASTER (Yap et al. (2019), Miliaresis and Paraschou (2005), Suwandana et al. (2012), Mirza et al. (2011), Mukherjee et al. (2013), Falorni et al. (2005), and Shaikh et al. (2021)). To improve the accuracy of the DEM, Yamazaki et al. (2017) produced a highaccurate map of global terrain elevations, Capolupo (2021) enhanced the accuracy of GDEM of differences in the Google Earth Engine for 3D change detection analysis, Muhadi, Mohd Kassim, and Abdullah (2019) Improved DEM using data fusion technique for oil palm replanting phase. In addition, Xu and Zhou (2016) suggested recovering distorted DEMs of regular terrain while Deilami et al. (2012) recommended reducing pit and flat error for improving DEM accuracy.

In constructing DEMs, interpolation methods make use of an effective factor. Consequently, many investigations have been achieved for selecting the suitable method for enhancing DEM accuracy. Such as, Habib et al. (2020) studied the impact of interpolation techniques on the accuracy of the large-scale DEM. Rishikeshan, Katiyar and Mahesh (2014) also evaluated DEM interpolation methods in GIS. Moreover, Jana (2011), Arun (2013), and Schwendel, Fuller, and Death (2012) established the appropriate interpolation methods for creating DEMs.

In the past few years, artificial neural networks (ANNs) have received increased attention for achieving unprecedented results in many complicated geomatics problems. ANN has been applied in the creation of DEMs, leading to some contributions like the SRTM DEM improvement (Kim et al. 2020, Wendi et al. 2016, Kulp and Strauss 2018), landslide detection (Gorsevski et al. 2016, Kawabata and Bandibas 2009), calibration and verification of DEM parameters (Ye et al. 2019, Benvenuti, Kloss and Pirker 2016), DTM creation and flood risk mapping (Jakovljevic et al. 2019), and the analysis of remotely sensed data for producing DEM (Mas and Flores 2008).

Recently, an effective merging technique was applied to the elevation data sources. This has been suggested due to the availability of elevation data sets for more than one area. In some cases, these datasets have low accuracy in some areas and high accuracy levels in other areas. Therefore, using a low-accuracy dataset for the whole area when parts of the area have high-accuracy data is not recommended. Consequently, merging the most accurate data sources (e.g., DEMs) to produce a single DEM that covers the whole area of interest with the highest possible resolution and accuracy was proposed. For example, Gruber et al. (2013), Choussiafis, Karathanassi and Nikolakopoulos (2012), and Jain et al. (2007) applied mosaic methods to improve DEM accuracy.

In this study, we hypothesize that, if there is available data with high accuracy (i.e., GNSS observations or high-accuracy DEM) for a small area S located in a large area B, a highly accurate DEM for the large area B can be created. This can be achieved by proposing DEM merging methods based on ANN and interpolation techniques. These techniques are employed for generating a correction surface (Vs) for the large area. The SRTM DEM is adjusted by adding the correction surface and merging with the small area S. Note that the small area S is the overlap area between the SRTM DEM and the large area B. Therefore, the main objectives of this study are to (i) generate a highly accurate DEM for a large area based on GNSS observations for a small portion of that area, (ii) mutually evaluate the proposed DEM merging methods (PDMMs) in comparison with the conventional DEM merging methods (CDMMs) available in ArcGIS software (in terms of elevations, slopes, and aspects), and (iii) compare and analyze the accuracy of the merged DEMs with SRTM DEM.

1.2. Interpolation methods

Interpolation is the process of predicting the values of unsampled locations based on collected samples, i.e., neighboring measurements (Sacchi, Ulrych and Walker 1998). The GIS software presents powerful analytical tools for various interpolation techniques (Habib et al. 2020). Some examples include kriging, inverse distance weighted (IDW), natural neighbor, triangulation with linear interpolation (TLI), spline, spline with barriers, topo to raster, and trend. However, the most widely used interpolation methods are Kriging, IDW, and spline (Yang et al. 2004). The IDW and spline are classified as deterministic interpolation methods. They are directly based on the surrounding measurements or specified mathematical formulas that determine the smoothness of the resulting surface. Kriging is classified as a geostatistical method (i.e., based on statistical models). This includes the statistical relationships among the measured points. Geostatistical techniques create a prediction surface and evaluate it at appreciable accuracy (Heine 1986). These methods are illustrated in the following sections.

1.2.1. Kriging

Kriging is an advanced geostatistical method that generates an estimated surface from scattered points with z-values (Rishikeshan et al. 2014). It depends mainly on a variogram that displays the variability between data points as a function of distance instead of the actual data values (Arun 2013). If a good variogram model is available, kriging provides the best estimation of the input data (Kim et al. 2020). Also, kriging assumes that the distance between sample points reflects a spatial correlation that can be used to explain the variations in the surface. This method fits a mathematical function to a specified number of points, or all points within a specified radius, to determine the surface elevations (Royle and AG 1981). Kriging follows multistages: exploratory statistical analysis of the dataset, variogram modeling, creating the surface, and (optionally) exploring a variance surface (Oliver and Webster 1990). The mathematical formula of this interpolator is formed as a weighted sum of the data (Lenda et al. 2016):

$$z(s_0) = \sum_{i=1}^n \lambda_i z(s_i) \tag{1}$$

where $z(s_i)$ is the measured value at the *i*th location, λ_i is an unknown weight for the measured value at the *i*th location, s_0 is the prediction location, and *n* is the number of measured values.

In other words, kriging achieves the prediction process through two steps. The first is the creation of the variograms and covariance functions to estimate the statistical dependence (called spatial autocorrelation) values that depend on the model of autocorrelation (fitting a model). The second is the prediction of the unknown values.

Semivariogram modeling is a key step between spatial description and spatial prediction. The empirical semivariogram provides information on the spatial autocorrelation of datasets. Nevertheless, this information is not for all directions and distances. Thus, fitting the model is necessary for this reason and to ensure that kriging predictions have positive kriging variances (Meng, Liu and Borders 2013).

The kriging tool provides several functions for modeling the empirical semivariogram (Oliver and Webster 1990): circular, spherical, exponential, Gaussian, and linear. Significantly, the spherical model is one of the most commonly used models.

1.2.2. IDW

This technique estimates values at an unsampled location through a weighted average of data points within a specified radius generated around each grid cell (Arun 2013). Consequently, the points closer to the prediction location will influence the estimated value more than those farther away (Rishikeshan et al. 2014, Lu and Wong 2008). A cell's value of the interpolated surface is calculated using Equation 2 (Habib et al. 2020):

$$z_g = \frac{\sum_{i=1}^{n} z_i w_i}{\sum_{i=1}^{n} w_i}, \qquad w_i = \frac{1}{d_i^k}$$
(2)

where z_g is the predicted value at the unsampled location, n is the number of measured points used for the interpolation, z_i is the known value, w_i is the weight, d_i is the separation distance between the grid node and the data point, and k is a smoothing parameter of the estimated surface.

IDW principally relies on the inverse of the distance raised to mathematical power. The power parameter controls the significance of known points on the interpolated values based on their distance from the output point. It is a positive real number whose default value is 2 (Arun 2013). By defining a higher power value, more emphasis can be put on the nearest points resulting in nearby data being influenced most while the surface has more detail (be less smooth). As the power increases, the interpolations begin to approach the value of the nearest sample point. Conversely, specifying low power turn to influence surrounding points farther away, resulting in a smoother surface (Lu and Wong 2008).

1.2.3. Spline

This interpolation method estimates points using a mathematical function that minimizes the overall surface curvature, resulting in a smooth surface (Rishikeshan et al. 2014). The spline interpolation requires that the surface pass precisely through the data points and have minimum curvature (Houdek, Verlinden and Hajžman 2022). There are two types of spline interpolation: regularized and tension (Childs 2004). While the regularized approach creates a smooth, gradually changing surface with points that may lie outside the sample data range, the tension approach controls the stiffness of the surface according to the character of the modeled phenomenon (Song et al. 2022). The spline algorithm for the surface interpolation is expressed in Equation 3:

$$S(x, y) = T(x, y) + \sum_{j=1}^{n} \lambda_j R(r_j)$$
(3)

where j = 1, 2, ..., n, *n* is the number of points, λ_j are coefficients found by the solution of a system of linear equations, and r_j is the distance from the point (x, y) to the *j*th point. T(x, y) and R(r) are defined differently, depending on the selected option (Mitáš and Mitášová 1988).

It is important to note that this research is not focusing on studying the different parameters of the interpolation methods mentioned. Nonetheless, the commonly used parameters according to the previous investigations are selected for the trial tests during the development of the proposed methods. Consequently, the spherical model, the power parameter equal 2, and the regularized approach are considered for the kriging, IDW, and spline interpolation methods, respectively.

1.3. Artificial Neural Networks

Artificial neural network (ANN) is one type of machine learning algorithm that has, lately, received increased attention for achieving promising results in different geoscience applications (Marmanis et al. 2015, Mas and Flores 2008). This algorithm is implemented as a system of interconnected processing elements that are functionally analogous to biological neurons (Kim et al. 2020). ANN consists of three layers: the input, hidden, and output layers (Figure 1).

As shown in Figure 1, the input layer includes input P, which has a single vector of R elements that relate to the hidden layer by synaptic weights, $W_{N^h,R}^h$. The synaptic weights W_{N^o,N^h}^o make connections between the hidden and output layers (Gurney 2018). The superscripts h and o represent the hidden and output layer, respectively. The first index indicates the specific neuron destination for that weight; the second index denotes the source of the signal fed to the neuron; while N^h, N^o represent the number of neurons (NON) in the hidden and output layers, respectively. These neurons represent the processing units in ANN, where each neuron has a bias b, summation S, activation function f, and output a. The hidden layer's output a^h is the input for the output layer, which produces an output a^o . The bias b is like a weight, but it has a constant input of 1. The summation S can be calculated in the hidden and output layers using Equations (4) and (5) (Hagan, Demuth and Beale 1996):

$$S_{j}^{h} = \sum_{i=1}^{R} W_{ji}^{h} P_{i} + b_{j}^{h}$$
(4)

$$S_{k}^{o} = \sum_{j=1}^{N^{h}} W_{kj}^{o} a_{j}^{h} + b_{k}^{o}$$
(5)

where S_j^h and S_k^o represent the weighted inputs summed with the bias at the hidden neuron *j* and output neuron *k*, respectively.

This summation is fed to the activation function (i.e., transfer function), which is transformed to specific values depending on the type of activation function considered. There are three activation functions that are most commonly used (Han and Moraga 1995, Haykin 2009). The first function (Equation 6) is a linear transfer function (Purelin), which produces output equal to its input, where:

$$f(S) = S \tag{6}$$

The second function (Equation 7) is the log sigmoid (Logsig); this transfer function takes the input, which lies between plus and minus infinity, and transforms the output into the range 0-1, using the expression:

$$f(S) = \frac{1}{1 + e^{-S}}$$
(7)



Figure 1. ANN architecture

Equation (8) presents the third function, which is the hyperbolic tangent sigmoid (Tansig); this function is like the log sigmoid function, but the output here is transformed into the range -1 to 1 using the expression:

$$f(S) = \frac{e^{S} - e^{-S}}{e^{S} + e^{-S}} f(S) = \frac{1}{1 + e^{-S}}$$
(8)

Although there are different types of ANNs, the feed-forward neural network, the most straightforward type of ANN, is frequently considered. It generates a relationship between the inputs and outputs using a backpropagation (BP) algorithm (Seiffert 2001, Schmidhuber 2015). The Levenberg–Marquardt (LM) algorithm, often the fastest BP algorithm, is applied for training datasets in the network to minimize errors (cost function) using the mean squared error (Levenberg 1944, Marquardt 1963). The training is continued until the stopping criteria meet the requirements. The dataset, in most cases, is divided into 70% for training, 15% for validation test, and 15% for independent testing of the network's performance.

1.4. Conventional DEMs merging methods

GIS software provides functions for merging two or more grid-based (raster) datasets. These methods include the first, the last, the minimum, the maximum, the mean, and the blend. They are based on the assumption that grid-based DEMs have the same spatial resolutions (cell size) as well as a coordinate system.

The First method determines the pixel value from the first raster dataset, i.e., high accuracy DEM, in the mosaic list. It does not operate any elevation adjustment on the DEMs; the DEMs are just superimposed. Then, the existing raster dataset is considered first. Figure 2 shows the results when there are four overlapping pixels, and the First option is selected. The values of the first raster dataset (on the left) are superior to another mosaicked dataset. Therefore, the result is the same as the first set of overlapping pixels.

The Last method determines the pixel values from the overlapping last raster dataset. Figure 3 shows the result of a mosaic where the Last option is selected. When the two rasters are mosaicked, the overlapping values from the second raster dataset become the output mosaic.



Figure 2. First method



Figure 3. Last method

Figure 4 shows the Minimum method; it determines the lower pixel value from the two overlapping raster datasets. The illustration below shows the result of a mosaic where the Minimum option is selected. When the two rasters are mosaicked, the output mosaic results from the minimum values of the two raster datasets.



Figure 4. Minimum method

The Maximum method determines the higher pixel values from the two overlapping raster datasets. Figure 5 shows the result of a mosaic where the Maximum option is selected. When the two rasters are mosaicked, the output mosaic results from the maximum values of the two raster datasets.

(DEM _h	igh accuracy)	(DEM _{lov}	w accuracy)		
4	5		6	6	6	6
3	4		1	3	3	4

Figure 5. Maximum method

The Mean method determines the average pixel values from the two overlapping raster datasets. As shown in Figure 6, the result of a mosaic resulting from the Mean option is shown; the output pixel type is float. When the two rasters are mosaicked, the output results from averaging the two overlapping values. If many raster datasets are overlapping, only two raster datasets are processed at one time. Also, if the output pixel type is integer, then the values are rounded.

Figure 7 displays the Blend method; it uses a distance-weighted algorithm to determine the value of overlapping pixels. The output cell is a blend of the overlapping areas. This method is the most computationally intensive option for mosaicking.



Figure 6. Mean method

In the illustration in Figure 7, the diagram shows two overlapping raster datasets. The cell where the x is located has two values: the value of the pixel in dataset R1 (outlined red) and the value of the pixel in dataset R2 (outlined blue). Since the x is closer to dataset R2, the value of the R2 pixel is more heavily weighted in the output. For more details, see ESRI (2020).



Figure 7. Blend method

As was previously stated, the resulting DEM by min., max., first, and last methods may generate terrain surface discontinuities (cliffs) along the boundaries. These discontinuities are created due to the elevation differences in (variation in accuracies) DEMs, leading to errors in the slope and aspect values (Leitão et al. 2016). On the contrary, Mean and Blend methods determine the average elevation value within the overlapping area of the two DEMs. Thus, a smooth transition between the merged DEMs is achieved. This results in new elevations.

The drawbacks identified in the conventional DEM merging methods reveal the need for new methods. These new methods can retain high-accuracy DEM data while creating smooth transitions between the two merged DEMs.

2. MATERIALS AND METHODS

2.1. Study area and data sources

The study area is a region with an area equal to 3276400 m2 located in the Assiut governorate desert, Egypt. The elevations of the area's topography vary between 88 m and 134 m. A reference station (R), fixed by the Egyptian survey authority, was selected for the GNSS survey. Figure 8 and Table 1 show the coordinates of the corner points of the study area and the reference point R.

Two types of data were collected.

- I. DEM file (SRTM 1 sec.) was downloaded from the U.S. Geological Survey (USGS) website (<u>https://pubs.er.usgs.gov</u>).
- II. Field observations were collected according to the DGNSS technique, i.e., two units of GNSS receivers (TRIMBLE R8s). The horizontal and vertical accuracies of this GNSS receiver type are shown in Table 2. The first unit (base station) was set up at reference station R, and the second unit (rover) was used for recording data by applying

postprocessed kinematic (PPK) mode. The GNSS receivers' parameters were set at a sampling time = 1 second, occupation time = 1 minute, and elevation mask = 12° .

	U	TM Coordinat	tes	Geodetic Coordinates						
Point	Easting	Northing	Orthometric height (H)		Longitude			Lati	Ellipsoid height (h)	
	m	m	m	0	'	"	0	,	"	m
P1	327256.448	3021899.541	86.214	31	15	15.321	27	18	33.418	98.977
P2	328756.137	3023120.616	102.135	31	16	9.246	27	19	13.765	114.913
P3	330325.359	3022440.421	127.317	31	17	6.662	27	18	52.371	140.085
P4	328521.254	3020944.203	96.403	31	16	1.802	27	18	2.955	109.153
R	329450.670	3019874.311	109.047	31	16	36.137	27	17	28.615	121.783

Table 1. Corners' coordinates of the study area and the reference point R

UTM – Universal Traverse Mercator (WGS 84 – UTM ZONE 36 N)

 Table 2. Root Mean Square Error (RMSE) of the different GNSS surveying types for TRIMBLE receivers (R8s)

GNSS surveying type	RMSE				
GNSS surveying type	Horizontal	Vertical			
Static	2 3 mm + 0.1 ppm 3.5 mm + 0				
РРК	8 mm + 1 ppm	15 mm + 1 ppm			
RTK	8 mm +1 ppm	15 mm + 1 ppm			

PPK – Postprocessing Kinematic; RTK – Real-Time Kinematic; ppm – parts per million



Figure 8. Study area and the reference point R

2.2. Methodology

This study is based much on statistical comparisons. Thus, relevant statistical and engineering software and programs were employed for highly accurate results. These programs are:

- Trimble Business Centre (TBC),
- Autodesk civil 3d vers.2021,
- AutoCAD vers. 2021,
- Microsoft Excel,
- MATLAB R2021a,
- ArcGIS 10.8.

The investigation was performed in five stages, which are shown in Figure 9 and discussed in the following sections.

2.2.1. Data processing and preparation

The GNSS observations were processed using Trimble Business Center (TBC) software based on the known station (R). The processed data was in the geodetic coordinate system (Latitude, Longitude, Ellipsoidal height); thus, the coordinates were converted to projected coordinates (Easting, Northing, Orthometric Height) according to the Universal Transverse Mercator (UTM) ZONE 36N.

The TBC outputs were used (in Autodesk Civil 3d and Excel programs) to produce grid data (5 m x 5 m). Also, the border coordinates (E, N) of the whole area with zeros and H elevations were assigned. Then, the overlap area coordinates (E, N, H) were extracted from the grid. These

overlap coordinates were added to the border coordinates in two cases (Figure 10): the border with zeros elevations (zeros border) and the border with H elevations (H border). The grid files of the overlap and the whole area were converted to DEM files in the ArcGIS environment. For achieving a high level of accuracy in the statistical stage, especially in comparisons, the horizontal coordinates of the whole area grid (E, N) were used for extracting the elevations of the DEMs under test, which means the same horizontal coordinates were applied for all DEMs. These coordinates were extracted in an Excel file and converted to a shape file named Extraction File (E.F.) in ArcGIS. Accordingly, the E.F. was used to extract the elevations from the DEM_{SRTM}, thereby converting DEM_{SRTM} (raster data) to vector data. After that, the following files were retrieved:

- DEM file for the whole area (DEM_{GNSS}), which was used as a reference surface in the statistics stage (Figure 11),
- DEM file for the overlap area (DEM Overlap),
- DEM file for the whole area from the USGS website, i.e., DEM_{SRTM} (Figure 12),
- Grid file for the whole area (GNSS processed data),
- Grid file for the whole area (extracted from DEM_{SRTM} data),
- Grid file for the overlap area,
- Grid files for the overlap area with the two cases: zeros and H borders.

The DEM $_{Overlap}$ and its location on the DEM $_{SRTM}$ are shown in Figure 13 and Figure 14, respectively.





Figure 9. Methodology flow chart



Case 1: Zeros border



Case 2: GNSS Points border





Figure 11. DEMGNSS for the study area



Figure 12. SRTM DEM for the study area



Figure 13. DEM of the overlap area



Figure 14. The position of overlap area on SRTM DEM

2.2.2. Correction surface generation

This is the basic stage of the study. As mentioned in Section 1.1, a small area of the overlapping DEM_{GNSS} and DEM_{SRTM} was selected from the study area. The differences (corrections) between the DEM_{SRTM} and $DEM_{Overlap}$ were calculated based on two techniques: (1) interpolation technique and (2) ANN technique. This was used as the correction surface, which was extended to the entire area.

2.2.2.1. Interpolation techniques

The most popular interpolation methods, considered, were kriging, IDW, and spline. It is important to note that the external borders do not have data to interpolate. Thus, additional data is required at the borders of the whole area to possibly execute the interpolation operation. Consequently, the two scenarios mentioned in *Section 2.2.1* were applied. The first scenario assumed that the differences between the elevations of DEM_{SRTM} and the border elevations are zeros (zeros borders) (Warriner 2005). The second scenario was proposed to test the effect of the same boundary with H elevations (H border). In this case, the differences between the elevations of DEM_{SRTM} and that of the borders were estimated. After that, the two scenario outputs were added to the overlap differences.

2.2.2.2. ANN technique

Also, the ANN algorithm was formulated in MATLAB software for generating the correction surface (Figure 15 9). According to the ANN components, the horizontal coordinates of the overlap area (E_o , N_o) were used as inputs, while the elevation differences between the overlap area and DEM_{SRTM} were the target of outputs. The horizontal coordinates of DEM_{SRTM} (E_D , N_D) were the sample data.

It is important to mention that ANN initialization produces different results. Thus, the ANN initializations are often repeated several times to guarantee more accurate results. Consequently, the number of initializations was set at 100; this iterated automatically. One hidden layer with ten neurons was selected, according to the recommendation of Alemam, Yong, and Mohammed (2022). During the development of this ANN algorithm, the results indicated that the linear activation function (pure-line), in the hidden and output layers, is suitable for this network and data compared to the other activation functions. The Train L.M. was selected as the training function. The involved training parameters according to the Train L.M. function and learning rate are shown in Table 3. Unlike the interpolation method, the ANN does not need a zeros border. Apparently, two scenarios were applied; ANN only and ANN with H elevations (H border) boundary. This network produced the DEM_{SRTM} corrections (V_D) as the output. The average value of these outputs was estimated by applying the least squares method (LSM) to minimize the weighted residuals (ν) by applying the normal equation:

$$L + v = AX \tag{9}$$

and the solution model:

$$X = (A^{T}A)^{-1}A^{T}L (10)$$

in the output data, *L* is the output matrix, *A* is the design matrix, and *X* is the unknown matrix.

Epoch	Goal	Max_fail	Min_fail	Mu	Learning rate
1000	0	6	1e-7	0.001	0.01

Table 3. The parameters involved in the ANN algorithm

Epoch – maximum number of epochs to train; Goal – performance goal; Max_fail – maximum validation failrues; Min_fail – minimum validation failures; Mu – adaptive value

This resulted in eight types of corrections surfaces (V_S). Consequently, DEMSRTM elevations were modified according to the equation:

$$H_{SRTM_{modified_i}} = H_{SRTM_{USGS}} + V_{S_i} \tag{11}$$

where V_{S_i} is the *i*th corrections surface.

The eight modified HSRTM were exported to ArcGIS software to generate DEM files.

2.2.3. DEMs generation

The imported data in ArcGIS software was converted to shape files, then to triangular irregular network (TIN) files, which were further transformed into raster data (DEM files).

The two scenarios (zeros and H border) were applied for creating Vs by interpolation and ANN methods. According to the first scenario, the generated DEMs were $DEM^0_{Kriging}$, DEM^0_{IDW} , and DEM^0_{Spline} . Then, $DEM^P_{Kriging}$, DEM^P_{IDW} , and DEM^P_{Spline} were also generated in the second scenario. The ANN outputs presented DEM_{ANN} and DEM^P_{ANN} , where the superscripts o and p represent the zeros and H border, respectively.

2.2.4. DEMs merging

The first, last, minimum, maximum, mean, and blend methods, in ArcGIS software, are the Conventional DEMs Merging Methods (CDMMs). These methods were applied for merging DEMSRTM (lower surface) and DEM Overlap (upper surface). The resulting merged DEMs (MDEM) were MDEMFirst, MDEMLast, MDEMMin, MDEMMax, MDEMMean, and MDEMBlend. Cognizant of the fact that the first method retains the upper surface properties, it was used for merging the DEMs produced in the previous section with the DEM Overlap. Consequently, the Proposed DEM Merging Methods (PDMMs) produced eight merged DEMs, including MDEM0 Kriging, MDEM0 IDW, MDEM0 Spline, MDEMANN, MDEMP Kriging, MDEMP IDW, MDEMPSpline, and MDEMPANN.

2.2.5. Statistical analysis

The processed data of the whole area (study area), whether the vector (i.e., E, N, and H coordinates) or raster (DEM_{GNSS}), represent the reference (comparison) surface for the merged DEMs. The results were evaluated according to the differences in elevations, slopes, and aspects between the merged DEMs and the comparison surface. Likewise, the differences between the DEM_{SRTM} and DEM_{GNSS} were estimated. These differences were analyzed in the case of the absolute value and considering the signs. The absolute values statistics were achieved in terms of mean absolute error (MAE), Maximum value (Max.), Minimum value (Min.), and root mean square error (RMSE) as expressed in Equations 12 and 13:

$$MAE = \frac{\sum_{i=1}^{n} |v_i|}{n} \tag{12}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (v_i)^2}{n}}$$
(13)

where v_i is the *i*th value of v, which represents the elevation differences between the GNSS processed data and the different DEMs produced; n is the number of observations.





 E_0 – Easting coordinate of the overlap area; N_0 – Northing coordinate of the overlap area; V_0 – the corrections (elevations differences) at the overlap area; E_D – Easting coordinate of the SRTM DEM; N_D – Northing coordinate of the SRTM DEM; V_D – the output corrections that generate the corrections surface (Vc); Nt – number of initializations

Figure 15. Algorithm steps of the ANN stage

In the case of considering the sign, the data were analyzed based on the box and whisker plot (Figure 16). It is very significant when a large number of datasets are analyzed. This is a standardized way of displaying the distribution of the dataset based on the five-number summary: the sample median, the lower and upper quartiles, any outliers (computed using the interquartile range), and the minimum and maximum values that are not outliers. Each box's top and bottom edges are the upper and lower quartiles. The distance between the top and bottom edges is the interquartile range (IQR). The upper quartile corresponds to the 0.75 quantile, and the lower quartile corresponds to the 0.25 quantile. Outliers are the values more than 1.5 IQR away from the top or bottom of the box. The whiskers are lines that extend above and below each box. One whisker connects the upper quartile to the nonoutlier maximum value, and the other connects the lower quartile to the nonoutlier minimum value. For more information on this, see Wickham and Stryjewski (2011), Thomas (2010), and Dekking et al. (2005). To clarify, assume the following elevations samples: 54, 66, 69, 70, 71, 73, 73, 80, 84, 85, 86, 88, 90, 91, and 95 m. Accordingly:

- Min. = 54;
- Max. = 91;
- Lower Quartile = 70;
- Upper Quartile = 88;
- Median = 80.



Figure 16. Box and whisker plot elements

3. RESULTS AND DISCUSSIONS

The DEM_{GNSS} , which was selected as a comparison surface for the statistical analysis, was assumed to have no error. The following sections present the results and discussions for the elevation, slope, and aspects differences between the DEM_{GNSS} and all DEMs merging methods, i.e., the CDMMs and the PDMMs.

3.1. Elevation

The elevation differences between DEM_{GNSS} and DEM_{SRTM} are shown in Figure 17 and illustrated in Table 4. It can be observed that, for the data distribution, the median is closer to the top of the box (upper quartile); thus, the distribution is negatively skewed (skewed left). In

contrast, Figure 18, which represents the elevation differences in the case of CDMMs, shows a positive skewed in all methods. Further, comparing the median value in this case (Table 5) with that obtained in DEM_{SRTM}, disregarding the sign, it shows that all the median values obtained by the CDMMs are less than the DEM_{SRTM} median. Furthermore, the upper and lower quartile values indicate that DEM_{SRTM} results are more dispersed than those from the CDMMs.

Considering the absolute value, the statistics demonstrate that the mean absolute error (MAE) for all CDMMs is better than the MAE in the case of DEM_{SRTM} . These results have been confirmed through the RMSEs, which are shown in Tables 4, 5, and Figure 19. Excluding the last method's results, whose RMSE is equal to that produced by DEM_{SRTM} , all the CDMMs achieved higher accuracy than DEM_{SRTM} . The first and blend methods have provided the best results in terms of the median, MAE, and RMSE. This is because the first method takes advantage of the highly accurate DEM through the merging operation and the distance-weighted algorithm that is applied in the blend method.



Figure 17. Elevation differences between DEM_{SRTM} and DEM_{GNSS}

	Absolute values statistics				Box chart statistics				
DEM type	MAE	Max.	Min.	RMSE	Median	Whisker (max.)	Whisker (min.)	25 th percentile	75 th percentile
DEM _{SRTM}	3.64	19.49	0.00	4.78	-2.47	3.37	-8.68	-4.16	-1.15

Table 4. The differences in elevations between the DEM_{SRTM} and DEM_{GNSS}

MAE – Mean Absolute Error

	Abso	olute val	ues sta	tistics	Box chart statistics						
DEMs	MAE	Max.	Min.	RMSE	Median	Whisker (max.)	Whisker (min.)	25 th percentile	75 th percentile		
MDEM _{First}	3.26	19.64	0.00	4.41	2.27	8.67	-4.07	0.71	3.89		
MDEM _{Last}	3.63	19.64	0.00	4.78	2.46	8.74	-3.44	1.13	4.17		
MDEM _{Min}	3.59	19.64	0.00	4.75	2.46	8.74	-3.44	1.13	4.17		
MDEM _{Max}	3.30	19.64	0.00	4.44	2.27	8.68	-4.08	0.71	3.89		
MDEM _{Mean}	3.38	19.64	0.00	4.45	2.36	8.43	-3.47	0.99	3.97		
MDEM _{Blend}	3.26	19.64	0.00	4.41	2.27	8.67	-4.07	0.71	3.89		

Table 5. The elevation differences between the conventional DEMs merging methods and the DEM_{GNSS}



Figure 18. Elevation differences between the conventional DEMs merging methods and DEM_{GNSS}



Figure 19. Root mean square error of the elevation's differences between DEM_{GNSS} and the merged DEMs in the case of conventional DEMs merging methods

On the other hand, Figure 20 a–c shows the results of the PDMMs in the case of the zeros border. The results reveal a negative skew (skewed left), but the results of ANN in the same case (Figure 20 d) show that the median is closer to the bottom of the box, which means the distribution is positively skewed (skewed right). Besides, comparing these results (Table 6 a) and the best results of the CDMMs demonstrate that the data dispersion from the lower and upper quartile is significantly less than those of the CDMMs. The median, MAE, and RMSE show improvement in the merged DEM than the CDMMs. The MAE of the spline method is less than that produced by the IDW and kriging methods. However, the RMSE of the kriging method is better than both spline and IDW. Moreover, the ANN achieved the best results in the median, MAE, and RMSE.

Manad	Ab	solute v	alues stati	stics	Box chart statistics							
DEMs	MAE	Max.	Min.	RMSE	Median	Whisker (max.)	Whisker (min.)	25 th percentile	75 th percentile			
a. Zeros border												
MDEM ⁰ Kriging	2.86	19.97	0.00	4.14	-1.71	5.01	-8.63	-3.52	-0.11			
MDEM ⁰ IDW	2.98	22.17	0.00	4.30	-1.87	4.91	-8.75	-3.63	-0.22			
MDEM ⁰ Spline	2.85	20.17	0.00	4.20	-1.60	5.12	-8.55	-3.43	-0.01			
MDEM _{ANN}	2.19	22.01	0.00	3.71	0.37	5.53	-4.69	-0.86	1.70			
				b.	. H borde	er						
MDEM ^P _{Kriging}	1.99	19.54	0.00	3.23	-0.01	4.46	-4.79	-1.32	0.99			
MDEM ^P IDW	2.12	21.93	0.00	3.51	-0.06	4.48	-5.11	-1.51	0.88			
MDEM ^P Spline	2.15	20.25	0.00	3.48	-0.02	4.89	-5.38	-1.53	1.04			
MDEM ^P _{ANN}	0.85	4.85	0.00	1.16	0.18	2.77	-2.34	-0.43	0.85			

Table 6. The statistics of elevations differences between the proposed DEM merging methodsand the DEM_{GNSS}

Table 6 b and Figure 21 display the results of the PDMMs in the case of the H border. The results show that the lower and upper quartile as well as the median values indicate fewer scatter than in all previous cases. Also, while the PDMMs related to the interpolation methods show a negative skew, the ANN shows a positive skew. According to the absolute values (Table 6 b), it can be concluded that the PDMMs, by the interpolation methods, achieved higher accuracy for the merged DEMs than that produced using the zeros border. Additionally, the kriging method achieved better MAE and RMSE than the spline method, which improved the IDW results except for MAE. Generally, the ANN proved its superiority as it outperformed all in terms of MAE and RMSE. The RMSEs of the PDMMs are illustrated in Figure 22.



Figure 20. Elevation differences according to DEM_{GNSS} for the proposed DEM merging methods in the case of zero border and the ANN



Figure 21. Errors in elevation between the DEM_{GNSS} and the proposed DEM merging methods in the case of additional observations at the study area border



Figure 22. Root mean square error of the elevations differences between DEM_{GNSS} and the merged DEMs in the case of PDMMs

To confirm the previous results, the elevation differences for all applied methods were classified into four categories. These categories were between less than one meter (highly accurate) and more than five meters (least accurate).

Table 7 and Figure 23 display the classification results of the DEM_{SRTM}. In addition, the results of the CDMMs are shown in Figure 24 and illustrated in Table 8. The CDMMs achieved higher accuracy (diff. ≤ 1 m) than the DEM_{SRTM}, which achieved a high percentage that is equal to the mean method in the medium accuracy category (i.e., $1 \text{ m} < \text{diff.} \leq 3 \text{ m}$). The results also show that the first and blend methods achieved the highest percentage in the first category, while the last and minimum methods achieved a high percentage in the fourth category (very low accuracy).

DEM type	1 st category high accuracy	2 nd category medium accuracy	3 rd category low accuracy	4 th category very low accuracy	
	diff. ≤ 1 m [%]	1 m < diff. ≤ 3 m [%]	3 m< diff. ≤ 5 m [%]	diff. > 5 m [%]	
DEM _{SRTM}	14	41	24	21	

 Table 7. Classification of the difference between DEM_{GNSS} and DEM_{SRTM} into four categories of accuracy



diff. – elevation differences

Figure 23. Classification of the difference between DEM_{GNSS} and DEM_{SRTM} into four categories of accuracy

The PDMMs classifications are shown in Table 9 and presented in Figs. 25 and 26. The results for the zeros border (Table 9 a) show a higher percentage in the first category than the CDMMs, which produced a high percentage in the fourth category. Further, the spline method provided a better percentage in the first category than the kriging and IDW methods. The IDW produced the highest percentage in the second category (medium accuracy). The ANN outperformed all the other methods; it achieved the highest and lowest percentages in the first and fourth categories, respectively.

Merged DEMs	1 st category high accuracy	2 nd category medium accuracy	3 rd category low accuracy	4 th category very low accuracy
	diff. ≤ 1 m [%]	1 m < diff. ≤ 3 m [%]	3 m< diff. ≤ 5 m [%]	diff. > 5 m [%]
MDEM _{First}	21	38	23	18
MDEM _{Last}	15	40	24	21
MDEM _{Min}	15	40	24	21
MDEM _{Max}	20	38	23	19
MDEM _{Mean}	16	41	24	19
MDEM _{Blend}	21	38	23	18

Table 8. Four categories of the merged DEMs accuracy resulting from the differences
between DEM_{GNSS} and the merged DEMs in the case of CDMMs

Merged DEMs	1 st category high accuracy	2 nd category medium accuracy	3 rd category low accuracy	4 th category very low accuracy	
	diff. ≤ 1 m [%]	1 m < diff. ≤ 3 m [%]	3 m< diff. ≤ 5 m [%]	diff. > 5 m [%]	
		a. Zeros border			
MDEM ⁰ KRIGING	29	29 36		16	
MDEM ⁰ IDW	27	37	20	16	
MDEM ⁰ _{Spline}	31	36	17	16	
MDEM _{ANN}	41	28	19	12	
		b. H border			
MDEM ^P KRIGING	46	34	10	10	
MDEM ^P _{IDW}	45	34	9	12	
MDEM ^P _{Spline}	43	35	10	12	
MDEM ^P ANN	71	12	7	10	



Figure 24. Four categories of the merged DEMs accuracy resulting from the differences between DEM_{GNSS} and the merged DEMs in the case of CDMMs



Figure 25. Classification of the elevations differences between DEM_{GNSS} and the PDMMs in the case of zeros border and ANN

Referring to Figure 26 and Table 9 b, it can be noted that the interpolation methods, i.e., H border, achieved superiority over the case of the zeros border. Again, the spline method yielded a high percentage in the second group, but the kriging method was the best in the high accuracy and fourth categories. ANN and kriging achieved similar percentages in the fourth category, but ANN outperformed kriging in the high-accuracy category.

Generally, it can be concluded that, in interpolation methods, kriging performed best in the two applied cases (i.e., zeros border and H border). This could be due to kriging's ability to take into account the spatial structure of the data, as proven by Rishikeshan et al. (2014) and Arun (2013). Notwithstanding, ANN outperformed kriging in the merged DEMs.



Figure 26. Four categories of the merged DEMs accuracy resulting from the differences between DEM_{GNSS} and merged DEMs by PDMMs in the case of additional observations at the border

3.2. Slopes and aspects

The elevation accuracy directly impacts the slope and aspects accuracies. This was investigated to validate the claim. The differences in slope and aspect degrees according to DEM_{GNSS} were calculated based on RMSEs for comparison among the different DEMs merging methods.

Figs. 27 and 28 show the slope and aspects distributions of the DEM_{GNSS} and DEM_{SRTM}, respectively. Also, the slopes and aspects of the DEMs obtained by the CDMMs are displayed in Figs. 29 and 30, respectively. Tables 10 and 11 provide the slopes and aspects statistics of the DEM_{SRTM} and CDMMs. The results show that the slope accuracy resulting from DEM_{SRTM} is better than the last, minimum, and maximum methods. Also, the accuracy of the aspects for DEM_{SRTM} is better than the mean method. While the blend method achieved the best slope in terms of RMSE, the minimum method was the worst.

Similarly, the results of the aspect show that the last method achieved the best accuracy, while the largest error was seen in the mean method. Furthermore, the first and blend methods reached the same results of aspects. Consequently, it can be noted that the applied technique that achieved the best slopes may not necessarily achieve the best aspects. The RMSEs of the slopes and aspects are displayed in Figs. 31 and 32, respectively.



a. Slopes



Figure 27. DEMs of the slopes and aspects for the GNSS points



a. Slopes

b. Aspects

Figure 28. DEMs of the slopes and aspects for the SRTM data



 $\textbf{e.} \mathsf{MDEM}_{\mathsf{Mean}}$



Figure 29. Slopes distribution of the conventional DEMs merging methods





f. MDEM_{Blend}



DEM type	Slope [%]				Aspects [deg]			
DEM type	Min.	Max.	Mean	RMSE	Min.	Max.	Mean	RMSE
DEM _{SRTM}	-18.47	50.47	0.30	7.24	-351.69	359.07	32.67	112.65

Table 10. The summary of slope and aspects differences between DEM_{SRTM} and DEM_{GNSS}

Table 11. Statistics of the slope and aspects differences between DEM_{GNSS} and the merged DEMs in the case of conventional DEMs merging methods

Merged		Slope [%]				Aspects [deg]			
DEMs	Min.	Max.	Mean	RMSE	Min.	Max.	Mean	RMSE	
MDEM _{First}	-168.94	46.22	3.91	7.15	-352.30	360.97	219.42	89.99	
MDEM _{Last}	0.02	51.63	5.16	7.47	1.00	360.97	237.57	64.26	
MDEM _{Min}	-155.92	50.42	4.93	7.52	-351.15	360.97	234.90	69.68	
MDEM _{Max}	-204.99	50.37	4.33	7.50	-352.30	360.97	221.85	86.62	
MDEM _{Mean}	-123.96	46.22	4.15	6.75	-354.10	360.72	48.57	138.97	
MDEM _{Blend}	-134.84	41.07	3.51	6.43	-352.30	360.97	219.42	89.99	



Figure 31. Root mean square error of the slope differences between DEM_{GNSS} and the merged DEMs in the case of conventional DEMs merging methods



Figure 32. Root mean square error of the aspects differences between DEM_{GNSS} and the merged DEMs in the case of CDMMs

The slopes and aspects distribution of the PDMMs in the two cases (zeros and H border) are shown in Figs. 33, 34, 35 and 36, respectively. Table 12, Figure 37, and Figure 38 display the RMSEs of the slope and aspects for the PDMMs. As shown in Table 12 a, zeros border case, although the slope results establish that kriging and IDW are better than the CDMMs, the ANN achieved the best RMSE of the slope differences. The aspect results of the ANN were the best among the PDMMs and CDMMs, except for the last method, which achieved the minimum RMSE. In addition, kriging, IDW, and spline methods resulted in aspects like the first and blend methods, but better than the mean method. However, they were the worst compared to the remaining CDMMs.

Table 12 b, which represents the PDMMs (H border case), shows that the slope and aspects accuracies are the best. Although kriging achieved the highest accuracy among the selected interpolation techniques, the ANN achieved the best RMSE for the slope and aspect differences.

Merged DEMs	Slope [%]				Aspects [deg]				
	Min.	Max.	Mean	RMSE	Min.	Max.	Mean	RMSE	
a. Zeros border									
MDEM ⁰ KRIGING	-24.70	40.51	3.60	5.82	-352.30	360.97	219.35	89.99	
MDEM ⁰ _{IDW}	-27.28	45.06	3.95	6.41	-352.30	360.97	219.39	89.93	
MDEM ⁰ _{Spline}	-30.56	50.98	4.40	7.14	-352.30	360.97	219.40	89.92	
MDEMANN	-19.93	32.33	2.92	4.73	-264.23	270.73	164.67	67.28	
b. H border									
MDEM ⁰ KRIGING	-21.51	35.01	3.15	5.09	-253.66	259.90	157.93	64.79	
MDEM ⁰ IDW	-22.14	36.09	3.23	5.24	-257.18	263.51	160.15	65.65	
MDEM ⁰ Spline	-25.02	41.07	3.64	5.90	-260.00	266.39	161.92	66.36	
MDEMANN	-13.71	21.99	2.02	3.27	-211.38	216.58	131.73	53.82	

 Table 12. Characteristics of the slope and aspects differences between DEM_{GNSS} and the merged DEMs in the case of PDMMs



a. MDEM⁰KRIGING

b. MDEM⁰IDW



Figure 33. Slopes distribution of the proposed DEMs merging methods in the case of zeros border and the ANN



a. MDEM^PKRIGING

b. MDEM^P_{IDW}



Figure 34. Slopes distribution of the proposed DEMs merging methods in the case of additional observations at the study area border



a. MDEM⁰KRIGING

b. MDEM⁰IDW



Figure 35. Aspects distribution of the proposed DEMs merging methods in the case of zeros border and the ANN



a. MDEM^PKRIGING

b. MDEM^PIDW



c. $\mathsf{MDEM}^{\mathsf{P}}_{\mathsf{Spline}}$

d. MDEM^PANN

Figure 36. Aspects distribution of the proposed DEMs merging methods in the case of additional observations at the study area border



Figure 37. Root mean square error of the slope differences between DEM_{GNSS} and the merged DEMs in the case of PDMMs



Figure 38. Root mean square error of the aspects differences between DEM_{GNSS} and the merged DEMs in the case of PDMMs

The improvements in the elevations, slopes, and aspects were analyzed according to the DEM_{SRTM} for both CDMMs and PDMMs.

The improvements in the CDMMs are shown in Figure 39 and Table 13 a6. The results show no improvement in the elevation and slope for the last method. Also, some CDMMs show no improvements in accuracy, i.e., the minimum and maximum methods in the slopes and the mean method in the aspects. Nonetheless, some improvements are seen in other indicators. For example, the first and blend method in the elevations achieved the best improvement, while the blend improved in the slopes, and the last and minimum methods in the aspects.

Table 13 b and Figure 40 illustrate the percentages of improvement in the case of the PDMMs. The results show that all the proposed methods significantly improved in the three terms, i.e., the elevations, slopes, and aspects. The elevation and slope achieved more improvements than those of the CDMMs. kriging, IDW, and spline methods (H border case) improved more in the

elevations and aspects than for the ANN (zeros border case), which rather improved in the interpolation methods. Interestingly, kriging (in the two cases) improved than the IDW and spline methods.

It is also important to note that the ANN (zeros border case) achieved better results than the applied interpolation methods in the two cases (i.e., the zeros border and H border) for the PDMMs, for slope accuracy improvement. Noticeably, the ANN (H border case) outperformed all other methods in the elevations, slopes, and aspects, with improvements reaching 76%, 55%, and 52%, respectively.

Mangad DEMa	Improvement [%]								
Merged DEMS	Elevations	Slopes	Aspects						
a. Conventional DEMs merging methods									
MDEM _{First}	7.74	1.24	20.12						
MDEM _{Last}	<u>0.00</u>	<u>-3.18</u>	42.96						
MDEM _{Min}	0.63	<u>-3.87</u>	38.14						
MDEM _{Max}	7.11	<u>-3.59</u>	23.11						
$MDEM_{Mean}$	6.90	6.77	<u>-23.36</u>						
MDEM _{Blend}	7.74	11.19	20.12						
b. Proposed DEMs merging methods									
MDEM ⁰ KRIGING	13.39	19.61	20.12						
MDEM ⁰ IDW	10.04	11.46	20.17						
MDEM ⁰ _{Spline}	12.13	1.38	20.18						
MDEM _{ANN}	22.38	34.67	40.28						
MDEM ^P KRIGING	32.43	29.70	42.49						
MDEM ^P _{IDW}	26.57	27.62	41.72						
MDEM ^P Spline	27.20	18.51	41.09						
MDEM ^P ANN	75.73	54.83	52.22						

 Table 13. The improvement in elevations, slopes, and aspects by the merged DEMs compared to DEMsRTM based on RMSE



Figure 39. The improvement of elevations, slopes, and aspects by the conventional DEMs merging methods compared to DEM_{SRTM}



Figure 40. The improvement of elevations, slopes, and aspects by the proposed DEMs merging methods in comparison with DEM_{SRTM}

In view of the results, it can be noted that the PDMMs were better than the CDMMs in all statistical analyses. This could be due to the correction surface (Vs) applied for the PDMMs, which had a highly accurate overlapping zone before merging the DEMs. Thus, the CDMMs applied the merging technique only at the overlap area and retained the remaining zones without changes.

The ANN technique, in the case of the zeros border, performed the correction surface without additional points at the borders, but the interpolation methods cannot perform the interpolation properly without additional points at the border. This is because of the power of the

backpropagation algorithm and further the ANN mechanism that depends on establishing a relation (i.e., mapping) between the inputs and outputs data based on the activation function. This mapping presents an ANN simulation that can produce any data out of the input range.

In the case of the H border, the interpolation methods realized the superiority of the ANN without additional points at the border. This is because the interpolation technique depends, basically, on the surrounding points (Oliver and Webster 1990). So, the GNSS points at the borders increase the accuracy and the confidence level of the interpolated points more than in the case of zero points.

Overall, the ANN reached the best results with additional GNSS points at the border. This is because the ANN simulation performs best with a high volume of data (Alemam et al. 2022). Therefore, the ANN is preferred as its cost-effective. However, if there exist additional observations at the borders (GNSS points), the interpolation methods can equally provide reasonable results.

4. CONCLUSIONS

DEM applications in engineering projects are costly due to the need for high-precision instruments and a long processing time. In this article, a combination of highly accurate geodetic GNSS data (for just a small portion of an area) and a low-accurate SRTM DEM (of the whole area) improved the DEM (of the whole area), resulting in promising accuracy. This combination, known as merging technique, is applied in commercial GIS by several methods known as the conventional DEMs merging methods (CDMMs) (i.e., minimum, maximum, last, first, mean, and blend). However, PDMMs based on a correction surface generated by the interpolation methods and a designed ANN algorithm outperformed the existing CDMMs.

Kriging, IDW, and spline interpolation methods were applied considering two cases (zeros border and H border) for the whole border areas. Zero elevations values for the entire border area for the first case (zeros border), and, in the second case, GNSS points were added (H border). The GNSS observations of the whole area were collected for generating a high-accuracy DEM as a comparison surface (i.e., DEM_{GNSS}). The elevation differences were calculated and then represented according to the box and whisker chart, considering the signs. Moreover, the absolute statistics of elevations, slopes, and aspects differences were calculated. Additionally, the elevation differences were classified based on the accuracy into four categories, and further, the improvements by all DEMs merging methods in DEM_{SRTM} were determined.

According to the improvements for DEM_{SRTM}, it can be concluded that some CDMMs did not improve (e.g., the last, minimum, and maximum methods in the slopes, the mean method in the aspects, and the last method in the elevations). Nonetheless, the best improvements for the CDMMs were seen in the first and blend methods reaching 7.74% in elevations, blend in the slopes reaching 11.19%, and the last and minimum methods reaching 42.96% and 38.14%, respectively, in aspects. Apparently, the kriging method performed best in the two cases. Its improvements in the zeros and H border, respectively, reached 13.39% and 32.43% in elevation, 19.61% and 29.70% in the slope, and 20.12% and 42.49% in the aspect. Also, the ANN improved the interpolation methods for elevations, slopes, and aspects, reaching 22.38%, 34.67%, and 40.28%, respectively, in the zeros border case.

Notwithstanding, better improvements were seen in the H border case, except for the slope. Interestingly, the ANN for the H border case achieved the optimal improvements in the elevations, slopes, and aspects terms reaching 75.73%, 54.83%, and 52.22%, respectively. This indicates that the PDMMs retained the high accuracy of the overlap area and significantly

improved the output DEM without needing GNSS observations for the entire area, thus leveraging cost and time.

REFERENCES

Akturk, E. & A. O. Altunel (2019) Accuracy assessment of a low-cost UAV derived digital elevation model (DEM) in a highly broken and vegetated terrain. *Measurement*, 136, 382-386.

Alemam, M. K., B. Yong & A. S. Mohammed (2022) Integration of Artificial Neural Network and the Optimal GNSS Satellites' Configuration for Improving GNSS Positioning Techniques (A Case Study in Egypt). *Artificial Satellites*, 57, 18-46.

Arun, P. V. (2013) A comparative analysis of different DEM interpolation methods. *The Egyptian Journal of Remote Sensing and Space Science*, 16, 133-139.

Badura, J. & B. a. Przybylski (2005) Application of digital elevation models to geological and geomorphological studies-some examples. *Przegląd Geologiczny*, 53, 977-983.

Benvenuti, L., C. Kloss & S. Pirker (2016) Identification of DEM simulation parameters by Artificial Neural Networks and bulk experiments. *Powder technology*, 291, 456-465.

Capolupo, A. (2021) Improving the Accuracy of Global DEM of Differences (DoD) in Google Earth Engine for 3-D Change Detection Analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 12332-12347.

Chen, C. & T. Yue (2010) A method of DEM construction and related error analysis. *Computers & Geosciences*, 36, 717-725.

Childs, C. (2004) Interpolating surfaces in ArcGIS spatial analyst. *ArcUser, July-September*, 3235, 32-35.

Choussiafis, C., V. Karathanassi & K. Nikolakopoulos. 2012. Mosaic methods for improving the accuracy of interferometric based digital elevation models. In *Proceedings of 32nd EARSeL Symposium*, 552-560.

Cook, A. J., T. Murray, A. Luckman, D. G. Vaughan & N. E. Barrand (2012) A new 100-m Digital Elevation Model of the Antarctic Peninsula derived from ASTER Global DEM: methods and accuracy assessment. *Earth system science data*, 4, 129-142.

Deilami, B. R., A. Shikhi, M. Liaghat, H. Shahabi, M. Rashid & A. Al-saffar (2012) Improvement the Accuracy of Digital Elevation Model (DEM) with Reduction Pit and Flat Errors (without field work).

Dekking, F. M., C. Kraaikamp, H. P. Lopuhaä & L. E. Meester. 2005. *A Modern Introduction to Probability and Statistics: Understanding why and how*. Springer.

Falorni, G., V. Teles, E. R. Vivoni, R. L. Bras & K. S. Amaratunga (2005) Analysis and characterization of the vertical accuracy of digital elevation models from the Shuttle Radar Topography Mission. *Journal of Geophysical Research: Earth Surface*, 110.

Gorsevski, P. V., M. K. Brown, K. Panter, C. M. Onasch, A. Simic & J. Snyder (2016) Landslide detection and susceptibility mapping using LiDAR and an artificial neural network approach: a case study in the Cuyahoga Valley National Park, Ohio. *Landslides*, 13, 467-484.

Gruber, A., B. Wessel, M. Huber, M. Breunig & S. Wagenbrenner. 2013. The approach for combining DEM acquisitions for the TanDEM-X DEM mosaic. In 2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS, 2970-2973. IEEE.

Gurney, K. 2018. An introduction to neural networks. CRC press.

Habib, M., Y. Alzubi, A. Malkawi & M. Awwad (2020) Impact of interpolation techniques on the accuracy of large-scale digital elevation model. *Open Geosciences*, 12, 190-202.

Hagan, M., H. Demuth & M. Beale (1996) Neural Network Design (PWS, Boston, MA). *Google Scholar Google Scholar Digital Library Digital Library*.

Han, J. & C. Moraga. 1995. The influence of the sigmoid function parameters on the speed of backpropagation learning. In *International workshop on artificial neural networks*, 195-201. Springer.

Hawker, L., J. Neal & P. Bates (2019) Accuracy assessment of the TanDEM-X 90 Digital Elevation Model for selected floodplain sites. *Remote Sensing of Environment*, 232, 111319.

Haykin, S. 2009. Neural networks and learning machines, 3/E. Pearson Education India.

Heine, G. W. (1986) A controlled study of some two-dimensional interpolation methods. *COGS Computer Contributions*, 3, 60-72.

Houdek, V., O. Verlinden & M. Hajžman (2022) Non-uniform quaternion spline interpolation in vehicle kinematics.

Huang, Q. & C. Yang (2011) Optimizing grid computing configuration and scheduling for geospatial analysis: An example with interpolating DEM. *Computers & Geosciences*, 37, 165-176.

Hugenholtz, C. H., K. Whitehead, O. W. Brown, T. E. Barchyn, B. J. Moorman, A. LeClair, K. Riddell & T. Hamilton (2013) Geomorphological mapping with a small unmanned aircraft system (sUAS): Feature detection and accuracy assessment of a photogrammetrically-derived digital terrain model. *Geomorphology*, 194, 16-24.

Jain, D., C. Rao, S. R. Kumar & S. Suresh. 2007. Assessment of DEM mosaic accuracy. In SPIE Optics & Photonics 2007: Optical Engineering & Applications Conference.

Jakovljevic, G., M. Govedarica, F. Alvarez-Taboada & V. Pajic (2019) Accuracy assessment of deep learning based classification of LiDAR and UAV points clouds for DTM creation and flood risk mapping. *Geosciences*, 9, 323.

Jana, S. (2011) SELECTION OF APPROPRIATE INTERPOLATION METHODS FOR CREATION DEMS OF VARIOUS TYPES OF RELIEF BY COMPLEX APPROACH TO ASSESSMENT OF DEMS.

Kawabata, D. & J. Bandibas (2009) Landslide susceptibility mapping using geological data, a DEM from ASTER images and an Artificial Neural Network (ANN). *Geomorphology*, 113, 97-109.

Khemiri, S., I. Chenini, S. Saidi, B. Baamar, A. B. Mammou & F. Zargouni (2013) DEM-Based GIS Algorithms and 3D Geospatial Mapping for Creation of Hydrogeological Models Data in Foussana Basin (Central Tunisia).

Kim, D. E., S.-Y. Liong, P. Gourbesville, L. Andres & J. Liu (2020) Simple-yet-effective SRTM DEM improvement scheme for dense urban cities using ANN and remote sensing data: application to flood modeling. *Water*, 12, 816.

Kulp, S. A. & B. H. Strauss (2018) CoastalDEM: A global coastal digital elevation model improved from SRTM using a neural network. *Remote sensing of environment*, 206, 231-239.

Leitão, J. P., D. Prodanović & Č. Maksimović (2016) Improving merge methods for grid-based digital elevation models. *Computers & geosciences*, 88, 115-131.

Lenda, G., M. Ligas, P. Lewińska & A. Szafarczyk (2016) The use of surface interpolation methods for landslides monitoring. *KSCE Journal of Civil Engineering*, 20, 188-196.

Levenberg, K. (1944) A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics*, 2, 164-168.

Lu, G. Y. & D. W. Wong (2008) An adaptive inverse-distance weighting spatial interpolation technique. *Computers & geosciences*, 34, 1044-1055.

Marmanis, D., F. Adam, M. Datcu, T. Esch & U. Stilla (2015) Deep neural networks for aboveground detection in very high spatial resolution digital elevation models. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2, 103.

Marquardt, D. W. (1963) An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11, 431-441.

Mas, J. F. & J. J. Flores (2008) The application of artificial neural networks to the analysis of remotely sensed data. *International Journal of Remote Sensing*, 29, 617-663.

Meng, Q., Z. Liu & B. E. Borders (2013) Assessment of regression kriging for spatial interpolation–comparisons of seven GIS interpolation methods. *Cartography and geographic information science*, 40, 28-39.

Mesa-Mingorance, J. L. & F. J. Ariza-López (2020) Accuracy assessment of digital elevation models (DEMs): A critical review of practices of the past three decades. *Remote Sensing*, 12, 2630.

Miliaresis, G. C. & C. V. Paraschou (2005) Vertical accuracy of the SRTM DTED level 1 of Crete. *International Journal of Applied Earth Observation and Geoinformation*, 7, 49-59.

Mirza, M., G. Dawod & K. Al-Ghamdi (2011) Accuracy and relevance of digital elevation models for geomatics applications-A case study of Makkah municipality, Saudi Arabia. *International Journal of Geomatics and Geosciences*, 1, 803.

Mitáš, L. & H. Mitášová (1988) General variational approach to the interpolation problem. *Computers & Mathematics with Applications*, 16, 983-992.

Muhadi, N. A., M. S. Mohd Kassim & A. F. Abdullah (2019) Improvement of Digital Elevation Model (DEM) using data fusion technique for oil palm replanting phase. *International Journal of Image and Data Fusion*, 10, 232-243.

Mukherjee, S., P. K. Joshi, S. Mukherjee, A. Ghosh, R. Garg & A. Mukhopadhyay (2013) Evaluation of vertical accuracy of open source Digital Elevation Model (DEM). *International Journal of Applied Earth Observation and Geoinformation*, 21, 205-217.

Nwilo, P., J. Onyegbula, C. Okolie, O. Daramola, O. Abolaji & I. Arungwa (2022) Influence of land cover, slope, and aspect on the vertical accuracy of SPOT DEM at selected sites in Nigeria. *Applied Geomatics*, 14, 17-31.

Oliver, M. A. & R. Webster (1990) Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information System*, 4, 313-332.

Reuter, H., P. Strobl & W. Mehl. 2011. How to merge a DEM? In *Geomorphometry 2011*, 87-90.

Rishikeshan, C., S. Katiyar & V. V. Mahesh. 2014. Detailed evaluation of DEM interpolation methods in GIS using DGPS data. In *2014 International Conference on Computational Intelligence and Communication Networks*, 666-671. IEEE.

Royle, A. & R. AG (1981) PRACTICAL UNIVERSAL KRIGING AND AUTOMATIC CONTOURNING.

Sacchi, M. D., T. J. Ulrych & C. J. Walker (1998) Interpolation and extrapolation using a high-resolution discrete Fourier transform. *IEEE Transactions on Signal Processing*, 46, 31-38.

Sari, I. L., R. Maulana, H. S. Dyatmika, A. Suprijanto, R. Arief & S. Ali (2019) Study of Digital Elevation Model (DEM) Extraction using Stereo Radargrammetry TerraSAR-X in Madiun Area–Elevation Accuracy Improvement.

Schmidhuber, J. (2015) Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117.

Schwendel, A. C., I. C. Fuller & R. G. Death (2012) Assessing DEM interpolation methods for effective representation of upland stream morphology for rapid appraisal of bed stability. *River Research and Applications*, 28, 567-584.

Seiffert, U. 2001. Multiple layer perceptron training using genetic algorithms. In *ESANN*, 159-164. Citeseer.

Shaikh, M., S. Yadav & V. Manekar (2021) Accuracy assessment of different open-source digital elevation model through morphometric analysis for a semi-arid river basin in the western part of India. *Journal of Geovisualization and Spatial Analysis*, 5, 1-21.

Song, H., R. Yuan, Y. Lv, H. Liu & Y. Li (2022) Cubic spline interpolation-based refined composite multiscale dispersion entropy and its application to bearing fault identification. *Structural Health Monitoring*, 14759217221134050.

Suwandana, E., K. Kawamura, Y. Sakuno, E. Kustiyanto & B. Raharjo (2012) Evaluation of ASTER GDEM2 in comparison with GDEM1, SRTM DEM and topographic-map-derived DEM using inundation area analysis and RTK-dGPS data. *Remote Sensing*, 4, 2419-2431.

Thomas, K. 2010. Introductory business statistics.

Varga, M. & T. Bašić (2015) Accuracy validation and comparison of global digital elevation models over Croatia. *International journal of remote sensing*, 36, 170-189.

Warriner, T. 2005. Generating a new high resolution DTM product from various data sources. In *Proceedings of the 50th Photogrammetric Week*. Citeseer.

Wendi, D., S. Y. Liong, Y. Sun & C. D. Doan (2016) An innovative approach to improve SRTM DEM using multispectral imagery and artificial neural network. *Journal of Advances in Modeling Earth Systems*, 8, 691-702.

Wickham, H. & L. Stryjewski (2011) 40 years of boxplots. Am. Statistician, 2011.

Wilson, J. P. (2012) Digital terrain modeling. Geomorphology, 137, 107-121.

Xiong, L., G. Wang & P. Wessel (2017) Anti-aliasing filters for deriving high-accuracy DEMs from TLS data: A case study from Freeport, Texas. *Computers & Geosciences*, 100, 125-134.

Xu, J. & B. Zhou (2016) A Method of Recovering Distorted DEM of Regular Terrain.

Yamazaki, D., D. Ikeshima, R. Tawatari, T. Yamaguchi, F. O'Loughlin, J. C. Neal, C. C. Sampson, S. Kanae & P. D. Bates (2017) A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, 44, 5844-5853.

Yang, C.-S., S.-P. Kao, F.-B. Lee & P.-S. Hung. 2004. Twelve different interpolation methods: A case study of Surfer 8.0. In *Proceedings of the XXth ISPRS congress*, 778-785.

Yang, L., X. Meng & X. Zhang (2011) SRTM DEM and its application advances. *International Journal of Remote Sensing*, 32, 3875-3896.

Yap, L., L. H. Kandé, R. Nouayou, J. Kamguia, N. A. Ngouh & M. B. Makuate (2019) Vertical accuracy evaluation of freely available latest high-resolution (30 m) global digital elevation models over Cameroon (Central Africa) with GPS/leveling ground control points. *International Journal of Digital Earth*, 12, 500-524.

Ye, F., C. Wheeler, B. Chen, J. Hu, K. Chen & W. Chen (2019) Calibration and verification of DEM parameters for dynamic particle flow conditions using a backpropagation neural network. *Advanced Powder Technology*, 30, 292-301.

Yue, T.-X., Z.-P. Du, D.-J. Song & Y. Gong (2007) A new method of surface modeling and its application to DEM construction. *Geomorphology*, 91, 161-172.

Zhou, H., J. Zhang, L. Gong & X. Shang (2012) Comparison and validation of different DEM data derived from InSAR. *Procedia Environmental Sciences*, 12, 590-597.

Received: 2023-02-06

Reviewed: 2023-04-24 (A. Bhardwaj); 2023-05-25 (undisclosed name); 2023-06-06 (undisclosed name); 2023-08-27 (P. Kupidura)

Accepted: 2023-09-11