

Hossam Talaat Elshambaky\*

# Using direct transformation approach as an alternative technique to fuse global digital elevation models with GPS/levelling measurements in Egypt

<https://doi.org/10.1515/jag-2018-0050>

Received December 25, 2018; accepted February 10, 2019

**Abstract:** Open global digital elevation models (GDEMs) represent a free and important source of information that is available to any country. Fusion processing between global and national digital elevation models is neither easy nor inexpensive. Hence, an alternative solution to fuse a GDEM (GTOPO30 or SRTM 1) with national GPS/levelling measurements is adopted. Herein, a transformation process between the GDEMs and national GPS/levelling measurements is applied using parametric and non-parametric equations. Two solutions are implemented before and after the filtration of raw data from outliers to assess the ability of the generated corrector surface model to absorb the effect of the outliers' existence. In addition, a reliability analysis is conducted to select the most suitable transformation technique. We found that when both the fitting and prediction properties have equal priority, least-squares collocation integrated with a least-squares support vector machine inherited with a linear or polynomial kernel function exhibits the most accurate behavior. For the GTOPO30 model, before filtration of the raw data, there is an improvement in the mean and root mean square of errors by 39.31% and 68.67%, respectively. For the SRTM 1 model, the improvement in mean and root mean square values reached 86.88% and 75.55%, respectively. Subsequently, after the filtration process, these values became 3.48% and 36.53% for GTOPO30 and 85.18% and 47.90% for SRTM 1. Furthermore, it is found that using a suitable mathematical transformation technique can help increase the precision of classic GDEMs, such as GTOPO30, making them to be equal or more accurate than newer models, such as SRTM 1, which are supported by more advanced technologies. This can help overcome the limitation of shortage of technology or restricted data, particularly in developed countries. Henceforth, the proposed direct transformation technique represents an alternative faster and more economical way to utilize unfiltered measurements of GDEMs to estimate national digital elevations in areas with limited data.

**Keywords:** Least-Squares Collocation, Least-Squares Support Vector Machines, Artificial Neural Network, Global Digital Elevation Models, Support Vector Machines, Bursa–Wolf, Molodensky, SRTM, GTOPO30

## 1 Introduction

The digital elevation model (DEM) is a computer representation of the Earth's surface; it provides a base dataset from which topographic parameters can be digitally generated. Since global digital elevation models (GDEMs) were generated, many applications were developed based on them, such as continental-scale hydrologic modeling; radiometric and geometric correction of satellite image data; global land cover mapping; climate modeling, visualization and animation; contributions in establishing gravity field modeling; determining radio wave line-of-sight for telecommunications; remote sensing; development of geopotential global models; use in navigation systems for commercial aviation; geomorphology; and archeology [13, 43]. It is known that different GDEMs were generated from different sources of data, different technologies, and different mathematical manipulation processes [50, 89]. As a consequence, the final products have different accuracies [3, 6]. The factors affecting their accuracies can be listed as, for example, choice of vertical datum system; position accuracy; frequency and distribution of the source data; topography characteristics; type of land cover; agreement between the models; and DEM generation (interpolation) methods [3, 27]. Thus, GDEMs cannot be used directly in precise engineering applications because they are typically used without an estimation of their accuracies or reliabilities and without considering errors that can affect their results [91]. In addition, based on the GDEMs' actual accuracy values, they should not be the first option in a high-precision geomatics' activity [20]. Based on previously mentioned shortcomings, many studies worldwide have tried first to assess the quality of different GDEMs relative to many parameters such as the study area, release versions, sources that they are compared with, vegetation canopy corrections, horizontal shift corrections, outlier removal, and various statistical terms used to assess the models [4, 62]. The results of the quality assessment process were classified into two broad categories based

\*Corresponding author: Hossam Talaat Elshambaky, Civil Dept., Misr Higher Institute for Engineering and Technology in Mansoura, Mansoura, Egypt, e-mail: [hossam4000talat@gmail.com](mailto:hossam4000talat@gmail.com)

on quantitative and qualitative assessments, where, in the former category, only statistical analyses with internal and external validations were considered [6]. After the assessment process, the GDEMs should be rectified to be suitable entirely with national height data for any country; in this research, Egypt is taken as a case study. In general, attempts have been made to improve the accuracies of GDEMs by a fusion process, where two GDEMs are fused in a new model. The process includes many operations such as treatment of grid spacing, datum transformation, residuals analysis, vertical shift manipulation, interpolation, weights' calculations for every model to be fused, and finally assessment of the new model's quality [6, 39].

Egypt is a developed country, and it is suffering from a shortage of topographic map production, particularly for the western desert region [42]. Additionally, it has no official national digital elevation model [28]. Thus, it will be of great significance if GDEMs can be used as base models to cover this part of the country. It is clear that using open sources of GDEMs represents an easy way to establish and update the topographic maps because of the economic savings, time savings, and the ability to work under inclement atmospheric conditions compared with aerial photogrammetry and classical ground survey [6, 11]. If an acceptable level of accuracy can be secured to achieve this goal, it will be a tremendously helpful tool for developed countries, such as Egypt. As a consequence, many Egyptian researchers have attempted to fuse national data or national DEMs with open-source GDEMs to establish more accurate national DEMs and overcome shortage of data in uncovered regions [2]. The main objective of this study is to find an alternative solution for the fusion process. This alternative solution comprises a mathematical transformation process of GDEM data and different types of national data (of a country such as Egypt) such as ground control points (GCPs), national DEMs, or topographic maps. The transformation process aims to generate a corrector surface model (CSM) between the two sources of data. A CSM can absorb the errors that might appear between the two sources of data due to datum inconsistencies and systematic distortions that usually exist in the height datasets [41]. The most important advantage of this alternative solution is to deal directly with the original data generated from GDEMs, which exhibit different types of errors such as gross errors, systematic errors, and random errors [38]. In this study, two open sources of GDEMs (GTOPO30 and SRTM 1) were chosen to be applied for the following reasons: (i) both models represent an old and new generation of GDEMs, (ii) it is known that SRTM is an update of the GTOPO30 model by averaging the data of GTOPO30 to 30 seconds resolution and replacing the data of GTOPO30 between the

latitudes 60° N and 58° S [57], (iii) both models use different vertical datums, different sources of data, and different technologies to collect and compile the data [24, 89, 92], and (iv) from the available Egyptian studies for different GDEMs, there is a consensus to consider SRTM as the most suitable model to represent the topography of Egypt with sufficient accuracy to update the topographic maps with 1:50,000 scale [8, 11, 20, 42]. To execute the transformation process, classical geodetic transformation equations such as Bursa–Wolf and Molodensky–Badekas were used [23], and due to the wide range of soft computing techniques' applications in geodesy [5, 16, 31, 79, 94], they are also used here to tackle the transformation problem. Soft computing techniques that were utilized include artificial neural network (ANN), support vector machine (SVM), least-squares-support vector machine (LS-SVM), linear interpolation (LI) and cubic interpolation (CI), and polynomial regression (PR) of different degrees. Moreover, the least-squares collocation technique was applied along with previous established equations and techniques to carry out the same task. Then, statistical parameters i. e., key performance indices (KPIs) such as range, root mean square error (RMSE), mean, standard deviation (S.D.), and correlation factor (R) were applied to assess the transformation process [29, 30]. Both external and internal assessments were used to evaluate the fitting and prediction abilities of generated CSMs. At the final stage, a reliability index (RI) based on the weighted linear combination method is used to make the final decision of choosing the most suitable transformation technique. This study is divided into six parts: (i) background about commonly used GDEMs, (ii) summary of available data, (iii) explanation of mathematical transformation techniques, (iv) narration of research methodology, (v) analysis of results, and (vi) presentation of conclusions and recommendations.

## 2 Background of GDEMs

As mentioned before, two open sources of GDEMs (GTOPO30, and SRTM 1) are used in this study. The reasons to use these particular models have been explained in the introduction. Henceforth, a brief background based on different assessment studies, describing their origin and accuracy, is presented.

### 2.1 GTOPO30 model

The GTOPO30 model is an open-source GDEM with a horizontal grid spacing of 30 arcseconds (approximately

1.0 km). The model was completed in the second half of 1996 through the contribution of many agencies. It covers the full latitude from 90° S to 90° N, and from 180° W to 1180° E. The horizontal coordinate system is referenced to the World Geodetic System 1984 (WGS84) ellipsoid, while the vertical datum is the mean sea level. The model is generated from eight different sources of data, and it is worthy to note that approximately 56 % of the Africa DEM is based on “Digital Terrain Elevation Data,” and 44 % is based on the “Digital Chart of the World.” [44, 86]. The model describes the physical surface of the earth (elevation of the bare ground surface) excluding objects such as vegetation or buildings. Hence, the model is called the digital terrain model (DTM) [62]. The absolute vertical accuracy of the model varies from 3 m to 500 m based on a linear error with a confidence interval of 90 %. This variation is attributed to the various types of data sources, the unknown accuracy of maps used or their degree of compliance with the standard accuracy of the cartographic sources, and the unknown errors in the digitizing and elevation surface interpolation processes. [43, 45]. Although the model is rarely assessed [51, 57], there are some studies that have determined its accuracy level globally, continentally, nationally, and at last locally (in Egypt). Globally, the model was assessed against Shuttle Laser Altimeter data (SLA-01) as an error-free source, and the overall accuracy was 70.3 m RMSE [43]. Furthermore, the assessment was made globally against a database of the altitude of mountain peaks with ultra-topographic prominence, and the result was 383.09 m RMSE [46]. The level of accuracy for the African continent was estimated to be better than 100 m RMSE during the generation process [44]. Then, this level is ascertained to be 44.75 m against SLA-01 data [48]. For different countries such as Germany and Croatia in the European continent, the average accuracy was 27 m and 21.6 m, respectively [24, 89], while in Turkey and Iran in the Eurasia region it approached 16.46 and 19.03 m, respectively [57, 92]. Finally, in Egypt the model was assessed partially by two studies through the Delta region with a flat terrain topography, the raw data was filtered from outliers, and the accuracy was in the range from 6.8 m to 7.32 m RMSE [20, 28]. Two studies have attempted to utilize the available GCPs distributed through the Egyptian territory and have determined their accuracy to be in the range of 14.22 m to 25.20 m RMSE [6, 8]. From the above studies, we can conclude that the accuracy of GTOPO30 in the African continent is better than 50 m, while in Egypt it is better than 30 m RMSE. This accuracy can be less in the mountainous areas like the Red Sea mountain chain in the east, and the south west of Egypt.

## 2.2 Shuttle radar topography mission model

The 11-day STS-99 Shuttle Endeavor mission period, February 11–22, 2000, completed its objective and produced the final shuttle radar topography mission (SRTM) product through the cooperation between four agencies: National Aeronautics and Space Administration (NASA), National Geospatial Intelligence Agency, German space agency (DLR), and the Italian Space Agency. Through this mission, two-radar interferometry was used to give 3D measurements of the earth’s land surface [57]. C-band (5.6 cm) and X-band (3.1 cm) radar were utilized to gather the data required between the latitudes 60° N and 58° S to cover nearly 80 % of the earth’s land surface. NASA’s Jet Propulsion Laboratory (JPL) was responsible for C-band, and DLR was responsible for X-band; more details about the mission can be found in [84]. The horizontal datum is WGS84 ellipsoid, and the Earth Gravitational Model 1996 (EGM96) is the vertical datum. The spatial resolution is three arcseconds (90 m) and one arcsecond (30 m) coverage, the projection is geographic, and the raster size is 1-degree tiles [87]. On September 23, 2014, the decision was made by the US government to release the highest-resolution topographic data generated from NASA’s SRTM in the year 2000 globally by late 2015. The void-filled version of SRTM is called SRTM Plus or SRTM NASA Version 3. The model is designed to have linear vertical absolute height error less than 16 m, linear vertical relative height error less than 10 m, circular absolute geolocation error less than 20 m, and circular relative geolocation error less than 15 m. All the above requirements will represent 90 % of the data [55]. The model is assessed globally relative to each continent against GCPs of the kinematic global positioning system. In the African continent, it has an accuracy of 11.9 m absolute geolocation error, 5.6 m absolute height error, 9.8 m relative height error, and 3.1 m as long wavelength height error [72, 73]. To estimate the accuracy level of the SRTM model inside Egypt, many assessment studies for both versions of SRTM were carried out. A summary of those studies is listed in Table 1.

Based on the above table, it is clear that both versions of the SRTM model have better accuracy than GTOPO30. For flat topography in the Delta region, the accuracy is within a range of 2.6 m to 8.28 m. Moreover, for all Egyptian territorial areas including mountainous and western desert, the accuracy is within 4.3 m to 12.68 m. It is also important to note that the general accuracies of both SRTM versions in Egypt are below the designed level, and they are consistent with the global accuracy assessment of the African continent.

**Table 1:** Egyptian assessment studies for the SRTM model.

Model	No.	Comparison data	Study Area	Tested Height	RMSE (m)	Reference
SRTM 3"	1	National DEM + 950GCPs	Delta	Orthometric	4.72	[11]
	2	116 GPS/GCPs	Delta, West desert	ellipsoidal	12.68	[27]
	3	1227 GPS/GCPs	All Egypt	Orthometric	4.30	[8]
	4	601 GPS/GCPs	All Egypt	ellipsoidal	8.28	[71]
	5	416 GPS/GCPs	Delta	ellipsoidal	2.80	[20]
	6	359 GPS/GCPs	All Egypt	Orthometric	10.07	[6]
	7	National DEM + 625 GCPs	Delta	Orthometric	5.71	[42]
SRTM 1	1	Digitized map	Delta	ellipsoidal	8.56	[65]
		389 GPS/GCPs	All Egypt		7.04	
	2	601 GPS/GCPs	All Egypt	ellipsoidal	7.53	[71]
	3	416 GPS/GCPs	Delta	ellipsoidal	2.60	[20]
	4	359 GPS/GCPs	All Egypt	Orthometric	6.23	[6]
	5	National DEM + 625 GCPs	Delta	Orthometric	5.35	[42]

### 3 Available data

There are two packages of data used in this study. The first belongs to the available heights downloaded for both models (GTOPO30 and SRTM 1) related to the study area. It is within 22° N–32° N and 25° E–37° E. Those data were downloaded from the following sites <http://lta.cr.usgs.gov/GTOPO30>, and <http://www2.jpl.nasa.gov/srtm>, respectively. The second package was the available collection of 311 GCPs distributed through the study area within the same previous window limits. All GCPs have their coordinates in 3D relative to WGS84, and their GPS measurements were carried out relative to the Egyptian Surveying Authority (ESA) national geodetic reference framework of Egypt. The accuracies of the ellipsoidal heights were in a range of 3 mm to 6 mm. In addition, their orthometric heights were observed relative to the Egyptian vertical datum. The origin of the vertical datum is based on the mean sea level (MSL) at the Alexandria tide gauge of 1906. The orthometric heights were a final product of spirit leveling observations' adjustment. The observations were performed as closed loops that run between known high-precision benchmarks established by the ESA. The accuracies of those orthometric heights are not available because ESA did not publish this information. The spatial distribution of the GCPs is scattered and concentrated at the north coast and Nile valley where the most activities exist, although there were a small number of points in the Sinai Peninsula, Western and Eastern deserts [29, 30, 70]. The distribution of GCPs is shown in Fig. 1. For the first data package, the elevations were extracted from both GDEMs at the location of the GCPs using Global Mapper 19 software. The extracted elevations will be used in the transformation and analysis processes, and GCPs are classified to common and checkpoints. The number of checkpoints

is 70 points distributed through Delta region, Nile Valley, North cost, Suez Canal, and Western desert. They were chosen to have no outliers as will be shown later. Both types of points are used in the internal and external quality evaluation process as will be shown in the next sections.

## 4 Mathematical structure of the transformation techniques

Herein, the concept of corrector surface model generated from the transformation technique will be explained. In addition, different mathematical structures of various transformation techniques will be illustrated next. The transformation techniques can be classified as: (i) classical geodetic transformation equations such as Bursa–Wolf, Molodensky–Badekas, (ii) soft computing techniques like ANN, SVM, LS-SVM, (iii) LI and CI, (iv) polynomial regression (PR) of different degrees, and (v) least-squares collocation.

### 4.1 Transformation technique and corrector surface model concept

This part describes the interrelation between the transformation process and the concept of CSM, where the relation between two sources of data can be represented as in Eq. (1) or Eq. (2) below. Eq (1) describes the direct mathematical transformation function  $f(\phi, \lambda, H^{DEM})$  between the two height systems: (i) the national height system  $H^L$  and (ii) heights extracted from the GDEM  $H^{DEM}$  at the same locations of the GCPs representing the common points with geodetic coordinates  $(\phi, \lambda)$ . In Eq (2) the relation shows

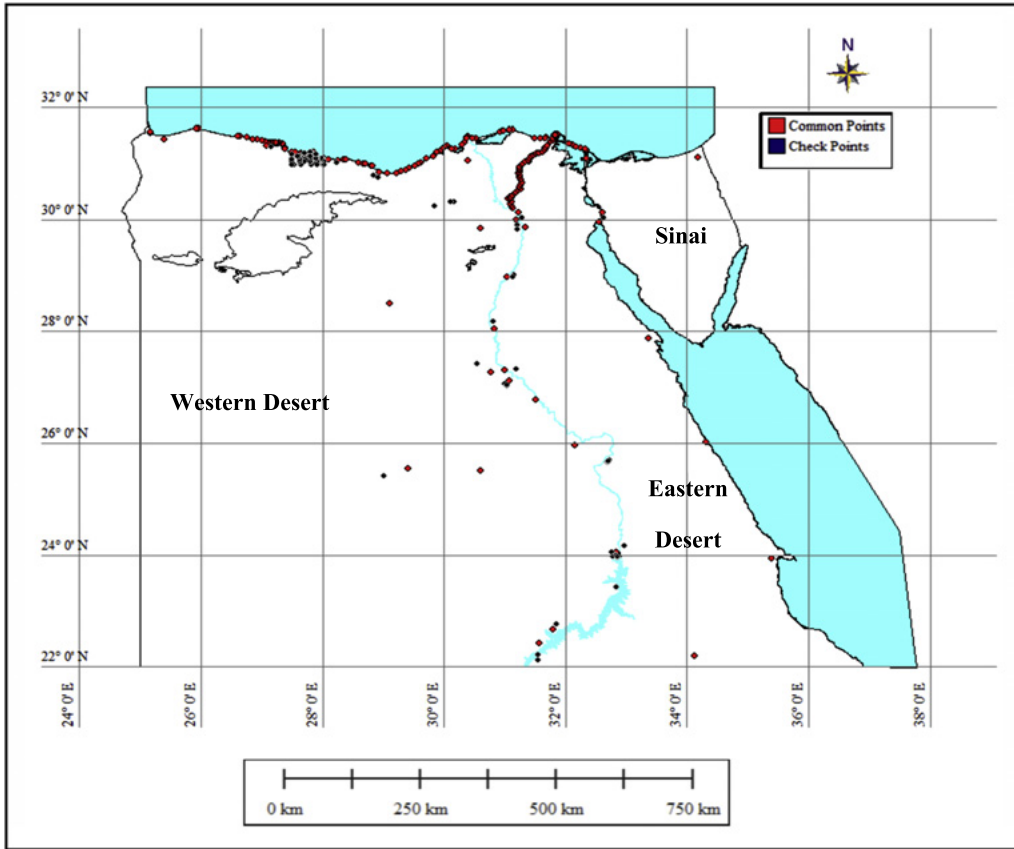


Figure 1: Distribution of common and checkpoints.

a simple transformation mathematical model relating the same previous datasets through a shift vector  $\Delta H$  between the two geodetic surfaces, i. e., it represents the vector of errors between the two systems.

$$H^L = f(\phi, \lambda, H^{DEM}) \quad (1)$$

$$H^L = H^{DEM} + \Delta H \quad (2)$$

Vector  $\Delta H$  can be represented by a CSM through which this surface can describe all possible datum inconsistencies and other systematic effects in the datasets by minimizing and mitigating those errors [78]. The corrector surface can be represented according to Eq. (3) where  $f(\phi, \lambda)$  is any mathematical function expression in the geodetic coordinates' latitude and longitude, respectively, of the common points [95].

$$\Delta H = f(\phi, \lambda) \quad (3)$$

To facilitate the mathematical solution based on least-squares adjustment, both Eqs. (1) and (3) can be linearized for every individual observation as in Eq. (4), where  $\alpha_i$  represents a vector of known coefficients with dimension  $n \times 1$ ,

$x$  is a vector of unknown parameters with a dimension of  $n \times 1$ , and vector  $v$  denotes a residual random noise term.

$$\Delta H_i = \alpha_i^T x + v_i \quad (4)$$

The main disadvantage of the previous equation is how to find an excellent parametric model of  $\alpha_i^T x$  to comprise all possible systematic inconsistencies in both national and global DEM datasets [58]. Based on the previous statement, many types of mathematical models can be used, as mentioned above, to find the most suitable one that represents the transformation process between the national and global DEM datasets.

## 4.2 Bursa–Wolf and Molodensky–Badekas models

For Eq. (1), the direct transformation between the two datasets will be implemented using the two famous classical geodetic datum transformation models: Bursa–Wolf, and Molodensky–Badekas. Herein, CSM is generated through the least-squares adjustment process to obtain

transformation parameters [63]. Seven transformation parameters are determined, including scale, three components of rotation, and three components of the shift. The formulation of both models is illustrated in Eqs. (5) and (6) where  $\omega$  is the scale parameter,  $R$  represents the rotation matrix between the two datums, and  $[\phi \ \lambda \ H]_L^T$ ,  $[\phi \ \lambda \ H]_{GDEM}^T$  are the two vectors of geodetic coordinates at common points with radian units for latitudes and longitudes. Meanwhile,  $[\Delta\phi \ \Delta\lambda \ \Delta H]^T$  represents the shift vector between the two datums and  $[\hat{\phi} \ \hat{\lambda} \ \hat{H}]^T$  are the coordinates of the average point at the GDEM's datum; this point is also referred to the initial point.

$$\begin{bmatrix} \phi \\ \lambda \\ H \end{bmatrix}_L = \omega R \begin{bmatrix} \phi \\ \lambda \\ H \end{bmatrix}_{GDEM} + \begin{bmatrix} \Delta\phi \\ \Delta\lambda \\ \Delta H \end{bmatrix} \quad (5)$$

$$\begin{bmatrix} \phi \\ \lambda \\ H \end{bmatrix}_L = \omega R \begin{bmatrix} \phi - \hat{\phi} \\ \lambda - \hat{\lambda} \\ H - \hat{H} \end{bmatrix}_{GDEM} + \begin{bmatrix} \Delta\phi \\ \Delta\lambda \\ \Delta H \end{bmatrix} + \begin{bmatrix} \hat{\phi} \\ \hat{\lambda} \\ \hat{H} \end{bmatrix} \quad (6)$$

More details about those models' applications, advantages, and disadvantages can be found in the following studies [77, 96]. Because both vectors of elevations ( $H^L, H^{DEM}$ ) from Eq. (1) belong to the same horizontal datum (WGS84) while having a different vertical datum, a geodetic transformation in 3D can be performed with added conditions as shown in Eq. (7), at which  $(\phi, \lambda)$  will have the same values for every common point in both datasets, i. e., it can be considered as a transformation in 1D [37, 60].

$$\begin{bmatrix} \phi \\ \lambda \end{bmatrix}_L = \begin{bmatrix} \phi \\ \lambda \end{bmatrix}_{GDEM} \quad (7)$$

It is clear from Eqs. (5) and (6) that, when the CSM is determined, the heights from the GDEM can be transferred to the national height datum directly and enrich the national dataset, particularly in the bare regions. Moreover, based on Eq. (7), the degree of freedom will increase resulting in a more reliable solution.

### 4.3 Soft computing models

In this study, soft computing techniques were chosen for their advantages such as their ability to deal with the transformation problem with few observations, nonlinearities, high dimensions, using few parameters, fast training, and representing a powerful nonlinear regression method due to inherent kernel functions. On the other hand, they suffer from over-fitting problems, time-consuming processing

with high-dimensional kernels, unexpected behavior, in-apprehensible learned function, and they depend mainly on trial and error [33, 83, 90].

#### 4.3.1 Artificial neural network

An ANN is a universal approximator function. This function, with a suitable number of neurons and hidden layers, can approximate any integrable function from one finite dimensional space to another finite dimensional space [52]. From Eq. (3), the CSM can be achieved through three steps. First, an intermediate output function  $S_k^l$  is related to the input geodetic coordinates of the common GCPs within the hidden layer, as in Eq. (8). Here,  $(w_i^l, b_0^l)$  represent the connection weight vector from the previous layer of neurons, and the bias weight vector that corresponds to an additional independent input,  $n$  is the number of input signal sources to a definite destination neuron,  $l$  is the number of hidden layers, and  $k$  is the number of destination neuron artificial neural networks.

$$S_k^l = \sum_{i=1}^n [\phi \ \lambda]_i w_i^l + b_0^l \quad (8)$$

In the second step, the intermediate function  $S_k^l$  will be transformed to an output  $a_k^l$  by applying a pre-determined transfer function such as a hyperbolic tangent function which is depicted in Eq. (9).

$$a_k^l = f(S_k^l) = \frac{e^{S_k^l} - e^{-S_k^l}}{e^{S_k^l} + e^{-S_k^l}} \quad (9)$$

At the third step, the CSM in Eq. (3) will be related to the transferred output  $a_k^l$  as in Eq. (10).

$$\Delta H = \sum_{i=1}^n a_k^l w_i^l + b_0^l \quad (10)$$

Now, the issue is turned to an optimization problem. The objective of ANN algorithm is to find the optimum  $(w_i^l, b_0^l)$  to minimize the errors between the estimated shift vector  $\Delta\tilde{H}$  and the input shift vector  $\Delta H$ . This objective can be achieved by applying the Levenberg–Marquardt algorithm, which adaptively varies the unknown parameters  $(w_i^l, b_0^l)$  between the gradient descent update and the Gauss–Newton update according to Eq. (11), until the algorithm reaches a stable state. At this point, the algorithm adopts the latest estimated parameters of the model.  $J^T$  is the transpose Jacobian matrix for the unknown parameters  $(w_i^l, b_0^l)$ , and  $\mu_k$  is a scalar value used to help the Hessian matrix ( $J^T J$ ) to be invertible [14].

$$\min(w_i^l, b_0^l) = -[J^T J + \mu_k I]^{-1} J^T (\Delta H - \Delta\tilde{H}) \quad (11)$$

In this study, two ANN techniques were used. The first was multi-feed forward neural network (MFFNN) as described above, and the second was a radial base neural network (RBNN). In the second technique, only two differences exist; instead of applying Eq. (8), Eq. (12) will be used, and the Gaussian transfer function will replace the hyperbolic tangent in Eq. (9). More details about the design process, and the advantages and disadvantages of ANN can be found in [47, 49].

$$S_k^l = \sum_{i=1}^n \|w_i^l - [\phi \quad \lambda]_i\| b_0^l \quad (12)$$

#### 4.3.2 Support vector machine and least squares-support vector machine models

The SVM was introduced by Vapnik within the area of statistical learning theory [80, 81, 88]. It is an extension of the support vector classifier [53], while LS-SVM is a derivation from SVM with a slight difference in the mathematical structure as will be shown next [30]. Again, to relate the CSM with the mathematical model of SVM, the input vector  $[\phi \quad \lambda]_i$  of coordinates for  $N$  common GCPs will be transferred from its primal space  $n$  to other higher space  $n_h$  or feature space  $(\varphi([\phi \quad \lambda]_i) : \mathbb{R}^n \rightarrow \mathbb{R}^{n_h})$  by using an unknown mapping function  $\varphi([\phi \quad \lambda]_i)$  to facilitate the linear inequality as illustrated in Eq. (13).  $\Delta H_i$  is as defined in Eq. (4),  $(w^T, b)$  is the unknown transpose weight vector associated with the input coordinate vector, and the bias of the linear model, respectively, where  $(b \in \mathbb{R})$ . The term  $\epsilon$  is the Vapnik-insensitive loss function.  $(\xi_i, \xi_i^*)$  represent the slack variables.

$$\begin{aligned} \Delta H_i - w^T \varphi([\phi \quad \lambda]_i) - b &\leq \epsilon + \xi_i - \Delta H_i + w^T \varphi([\phi \quad \lambda]_i) + b \\ &\leq \epsilon + \xi_i^*, \xi_i^* \geq 0 \end{aligned} \quad (13)$$

The system of inequality in Eq. (13) will be changed to a linear system of equations in LS-SVM as in Eq. (14). Here,  $v_i$  is defined in Eq. (4). The similarity between Eqs. (4) and (14) is clear, as the unknown CSM is equal to the first and second parts of the right-hand side of Eq. (14).

$$\Delta H_i = w^T \varphi([\phi \quad \lambda]_i) + b + v_i \quad (14)$$

Again, the issue is turned into an optimization problem as in Eq. (15) for SVM, and as depicted in Eq. (16) for LS-SVM.

$$\min(w, b, \xi_i, \xi_i^*) = \frac{1}{2} w^T w + \frac{1}{2} c \sum_{i=1}^N (\xi_i + \xi_i^*) \quad (15)$$

Here,  $c$  is a constant  $> 0$  and determines the amount up to which deviations from the desired insensitivity  $\epsilon$  accuracy is tolerated.

$$\min(w, b, e) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^N e_i^2 \quad (16)$$

Furthermore, the term  $\gamma$  is a regularization parameter, and it will be calculated by the cross-validation method [82]. Next, the Lagrange formulation is applied to Eqs. (15) and (16) in order for them to have the optimality conditions for SVM (as represented in Eq. (17)) or be solved as a linear system in dual space under a least-squares cost function based on a least-squares quadratic loss function for LS-SVM (as shown in Eq. (18)) [83, 93].

$$\begin{aligned} \mathcal{L}(w, b, \alpha, \alpha^*, \eta, \eta^*) &= \frac{1}{2} w^T w + \frac{1}{2} c \sum_{i=1}^N (\xi_i + \xi_i^*) \\ &\quad - \sum_{i=1}^N \alpha_i (\epsilon + \xi_i - \Delta H_i + w^T \varphi([\phi \quad \lambda]_i) + b) \\ &\quad - \sum_{i=1}^N \alpha_i^* (\epsilon + \xi_i^* + \Delta H_i - w^T \varphi([\phi \quad \lambda]_i) - b) \\ &\quad - \sum_{i=1}^N (\eta_i \xi_i + \eta_i^* \xi_i^*) \end{aligned} \quad (17)$$

$$\begin{aligned} \mathcal{L}(w, b, e, \alpha) &= \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^N v_i^2 \\ &\quad - \sum_{i=1}^N \alpha_i (w^T \varphi([\phi \quad \lambda]_i) + b + v_i - \Delta H_i) \end{aligned} \quad (18)$$

Here,  $(\alpha_i, \alpha_i^*, \eta_i, \eta_i^*) \in \mathbb{R}$  represent the Lagrange multipliers  $> 0$ . Eqs. (17) and (18) will be partially differentiated to the unknowns  $(w, b, \alpha, \alpha^*, \eta, \eta^*, v)$  and equated to zero. The sequential minimal optimization algorithm is used to solve the system of SVM [34, 35, 69], and for LS-SVM, Eq. (19) is produced.

$$\begin{aligned} w &= \sum_{i=1}^N \alpha_i \varphi([\phi \quad \lambda]_i); \quad \sum_{i=1}^N \alpha_i = 0; \quad \alpha_i = v_i v_i; \\ w^T \varphi([\phi \quad \lambda]_i) + b + v_i - \Delta H_i &= 0 \end{aligned} \quad (19)$$

By elimination both  $(w, v)$  and solving for  $(b, \alpha)$ , the following solution is obtained as in Eq. (20).

$$\begin{bmatrix} 0 & 1_v^T \\ 1_v & \Omega + I/\gamma \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \Delta H_i \end{bmatrix} \quad (20)$$

$1_v = [1, \dots, 1]$  is a vector with dimension  $(N, 1)$ ,  $\alpha = [\alpha_1, \dots, \alpha_N]^T$ , and  $\Omega$  is a matrix representing the effect of

the nonlinear mapping function by applying Mercer's theory for the kernel, as it is calculated implicitly instead of obtaining the mapping function explicitly, as in Eq. (21).

$$\Omega_{ij} = K(x_i, x_j) = \varphi([\phi \quad \lambda]_i)^T \varphi([\phi \quad \lambda]_j) \quad (21)$$

Here,  $K(x_i, x_j)$  represents the positive definite kernel function, this function can take many forms such as linear, polynomial, and radial base kernels. All kernel functions used in this study are listed in Table 2 [22, 33, 53, 81, 93]. After the unknowns are determined, the CSM can be calculated to any GCP as in Eq. (22)

$$\Delta H = \sum_{i=1}^N \alpha_i K(x_i, x_j) + b \quad (22)$$

**Table 2:** Kernel functions.

Kernel Functions	Formula
Linear	$K(x_i, x_j) = x_i^T x_j$
Polynomial of degree ( $d$ )	$K(x_i, x_j) = (t + x_i^T x_j)^d, t \geq 0$
Radial base function (RBF)	$K(x_i, x_j) = e^{-\frac{\ x_i - x_j\ ^2}{2\sigma^2}}$ , where $\sigma^2$ the variance of Gaussian Kernel

### 4.3.3 Linear and cubic interpolation models

In this study, we chose non-parametric methods to generate the corrector surface with LI and CI models. The interpolation is a method for estimating the value at a query location within the domain of a set of sample data points [59]. The interpolation process can be classified as being performed either globally or locally. For global interpolation, the estimating value depends on the values of all of the data points, while in the local interpolation, the estimating value depends mainly on the nearby data points [75]. The Delaunay triangulation method is used to achieve the locality characteristic, where the triangulation is obtained by connecting all the neighboring points in the Voronoi diagram of the given point set [10]. In LI, the interpolated value at any location on the corrector surface  $\Delta H$  belonging to any specified Delaunay triangle can be calculated by applying the barycentric interpolation as depicted in Eq. (23), where  $\beta_i$  is the barycentric weights, and  $\Delta H_i$  is the CSM value at the common GCPs. The barycentric weights for every individual Delaunay triangle have to fulfill the condition shown in Eq. (24). More information about barycentric interpolation can be found in a previous study [15, 36].

$$\Delta H = \sum_{i=1}^3 \beta_i \Delta H_i \quad (23)$$

$$\sum_{i=1}^3 \beta_i = 1 \quad (24)$$

Because the LI scheme is not smooth across the boundaries between Delaunay triangles, CI will be used to overcome this drawback. CI, or Hermite cubic interpolant, is a piecewise interpolant with a continuous derivative [56]. Herein, the MATLAB application for curve fitting is used to apply Hermite interpolation within the vector of coordinates and the corresponding values of their CSM at the known GCPs to calculate the interpolated values at different locations within the study area [61].

### 4.3.4 Polynomial regression

It is known from interpolation techniques that the interpolated surface will pass through the data points. This is adequate when the data is assumed to have zero noise. On the other hand, if the data is considered to have noise, it is better to use the PR with different degrees to minimize the error between the estimated and true value at the common GCPs [56].

PR is considered as the most popular method used to describe a complex surface based on pre-assumed coefficients [1, 66]. Herein, PR can be used to represent the CSM illustrated in Eq. (25) where  $x_{ij}$  represents the elements of the unknown parameters in the vector  $x$  of Eq. (4), and the rest of the right-hand side of Eq. (25) are the elements of the coefficients  $\alpha_i^T$  in Eq. (4). Through the least-squares adjustment, the CSM describing the surface  $\Delta H$  will be generated. The degree of PR will be determined based on the available common GCPs, and the importance of the parameters used, based on the statistical hypothesis tests [63, 64].

$$\Delta H = \sum_{i=0}^k \sum_{j=0}^i x_{ij} \lambda^{i-j} \phi^j \quad (25)$$

### 4.3.5 Least-squares collocation

Least-squares collocation (LSC) is a package that includes the three processes of adjustment, filtering, and prediction within one algorithm [26]. The concept of LSC is to have unbiased minimum error variance prediction when the underlying function is considered as a second-order stochastic process with a known or estimated covariance function [25]. LSC has many applications to solve different



geodetic problems [12, 85]. Based on this, LSC is chosen to be one of the mathematical techniques to represent the CSM. The LSC algorithm is subdivided into five steps as follows: First, the raw data of  $\Delta H$  calculated from Eq. (2) using both sources of data ( $H^L, H^{DEM}$ ) is reduced to a zero-mean value through subtracting the average value of the vector  $\Delta H$  from its entire values to guarantee the zero mean, stationary, and isotropic-two-dimensional model [18, 54]. In the second step, the empirical function is calculated based on the type of the spatial distribution of the common GCPs. Based on the scatter distribution of the common GCPs, Eq. (26) will be used to generate the empirical function as a function of a specified interval distance [67, 68, 76].  $C_r$  is covariance between points  $i$  and  $j$  at the Euclidian distance less than or equal to interval distance  $r$ ,  $N_r$  is the total number of pair points, and  $(P_i P_j)$  is the product of the signals at common points location  $i$  and  $j$ .

$$C_r = \frac{1}{N_r} \sum P_i P_j \quad (26)$$

Next, in the third step, the autocovariance function (ACF) will be calculated to fit with the empirical function using its variance of the noise and correlation distance as a fitting parameter. In this study, a classical geodetic ACF, such as Cauchy, Gauss, Markov 2D, and Markov 3D, is used. The mathematical formulas for those ACF are illustrated in Table 3 where  $C_0$  represents the variance value at the zero-correlation distance,  $D$  is the optimum correlation distance, and  $d$  is any distance at which the covariance value  $C(d)$  is required. The soft computing equations like ANN, SVM, LS-SVM, LI, and CI are also used to calculate the ACF.

**Table 3:** Classical geodetic ACF.

Name	Formula
Cauchy (Hirvonen)	$C(d) = C_0 / (1 + (d/D)^2)$
Gaussian 1	$C(d) = C_0 e^{-(d^2/D^2)}$
Gaussian 2	$C(d) = C_0 e^{-d^2/D^2}$
Second Markov	$C(d) = C_0 (1 + d/D) e^{-(d/D)}$
Third Markov	$C(d) = C_0 (1 + d/D + d^2/3D^2) e^{-(d/D)}$

In the fourth step, the variance–covariance matrix is calculated for both common and checkpoints [17]. At the final step, the CSM can be related to the LSC as in Eq. (27). This equation is a particular case of LSC adjustment as it represents only a prediction of signals at the computational (check) points, where  $C_{qp}$  is the cross-covariance matrix between both common and computational points,

$C_{pp}$  is the autocovariance matrix for the common points' signals, and  $P$  represent the signals at common points. More details about the steps to establish the LSC algorithm can be found in [30].

$$\Delta H = C_{qp}(C_{pp})^{-1}P \quad (27)$$

## 5 Research methodology

In this part of the study, we will describe the research methodology and process as depicted in Fig. 2. This includes the steps of preparing the raw data and generating different CSMs, the definition of different KPIs used to validate the various transformation techniques, and the description of the reliability index used to make the final decision.

### 5.1 Preparing raw data

In this step, all elevations of the study area related to both the GTOPO30 and SRTM 1 GDEMs are downloaded from the previously mentioned official sites. Then, the elevations related to both common and checkpoints are extracted from both models using Global Mapper 19 software. Next, the vector  $\Delta H$  is calculated for both common and checkpoints as in Eq. (2). To evaluate the capability of the transformation techniques to deal with the inherent errors in the raw data, we will deal with the common points'  $\Delta H$  values by two solutions: First, the elevations' differences will not be filtered against outliers. They will be considered to be errors-free data. Then, in the second solution,  $\Delta H$  will be filtered from the previously mentioned error through applying a  $3\sigma$  test routine to remove any value greater than  $3\sigma$ , where  $\sigma$  represents the standard deviation of the data [40]. After that, the mean value is subtracted to give the final white noise [18, 54]. More details about preparing the raw data steps can be found in [30]. Table 4 shows the effect of the filtration process. The following points are evident from this table, the number of detected outliers was 17 points, which represent nearly 5 % of the total points. The total points were subdivided to the common and checkpoints with a ratio of 77 % and 23 %, respectively. Based on the raw data before the filtration process, SRTM 1 is more accurate than the GTOPO30 model where the RMSE and S.D. are 11.485 m and 11.275 m, compared to 17.752 m, 17.673 m, respectively. These values are consistent with the previous studies of accuracy in Egypt.

It is noticed from the mean value that SRTM 1 suffers from systematic errors more than the GTOPO30 model

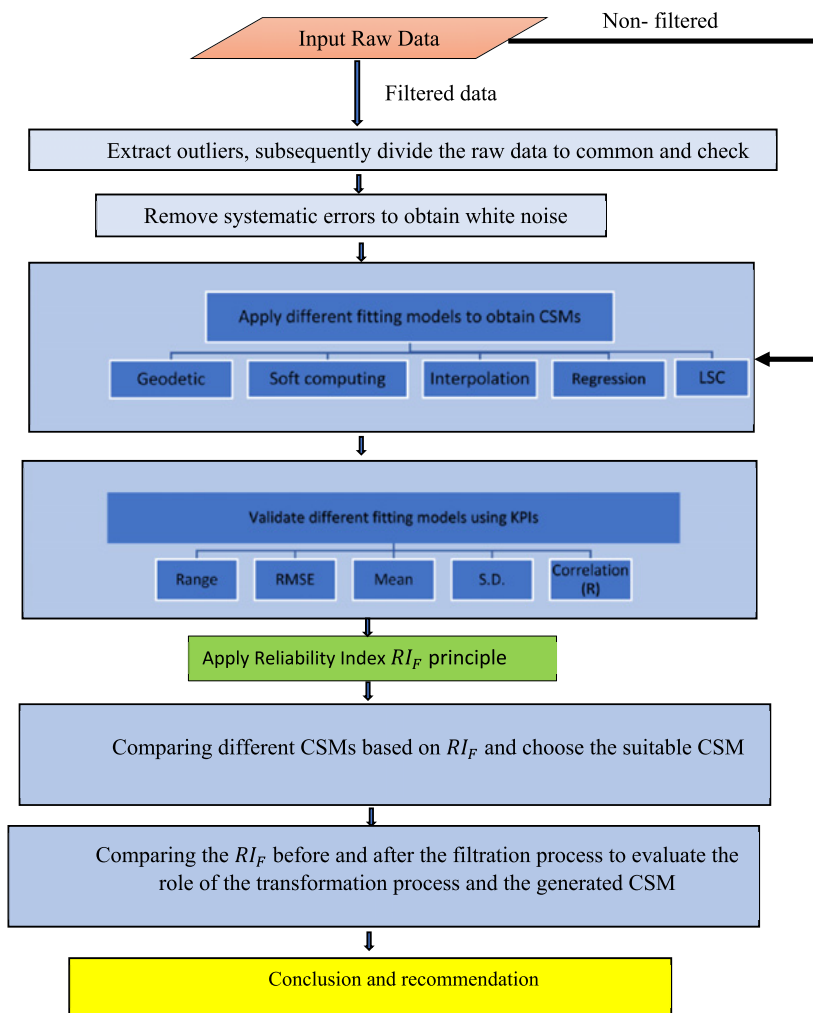


Figure 2: Flowchart of the study methodology.

Table 4: The effect of outliers' filtration process.

Model	Before filtration						After filtration					
	GCPs	Range (m)	RMSE (m)	Mean (m)	S.D. (m)	R	GCPs	Range (m)	RMSE (m)	Mean (m)	S.D. (m)	R
GTOPO30	311	275.622	17.752	1.947	17.673	0.982	294	55.128	8.984	1.420	8.995	0.995
SRTM 1	311	196.281	11.485	2.274	11.275	0.993	294	23.733	4.007	1.833	3.57	0.999

even after removing the outliers. This may be attributed to two factors: SRTM 1 is a GDSM, not a global DTM, and its vertical datum is EGM96, not MSL as in GTOPO30 model. After filtration, both GDEMs have improved to 49 % and 67 % in their accuracy based on RMSE and S.D., respectively. The improvement in their systematic error reached 27 % and 19 % for GTOPO30 and SRTM 1, respectively. This improvement is just for the study area as the available point cannot represent the whole study area; furthermore,

this improvement can be attributed to the cleaning of the utilized ground points.

## 5.2 Generating corrector surface model

In this step, all predefined mathematical techniques in part 4 of this study are used to generate different CSMs. There will be two solutions, one for non-filtered and one

for filtered common points. Then, the estimated values of both common and checkpoints from different CSMs will be used to assess the generated models as will be shown in the next step.

### 5.3 Definition of used key performance indices

As mentioned before, different CSMs need to be evaluated. Henceforth, the KPIs will be used to assess their behavior. The error vector ( $e$ ) between the measured national elevation vector  $H^L$  at the GCPs and the estimated national elevation vector  $\hat{H}^L$ , which is extracted from different CSMs as defined in Eq. (28) below, will be calculated as in Eq. (29). The Different KPIs are illustrated in Table 5;  $e_i$  is any element in the vector of errors [7, 9, 74]. Parameters  $(\overline{H^L}, \overline{\hat{H}^L})$  are the average values for both vectors  $H^L$  and  $\hat{H}^L$ , respectively. A high KPI implies a low error generated by a high accuracy model, whereas a low KPI implies a high error generated by a low accuracy model. KPIs will be used for common and checkpoints as internal and external quality assessments [30].

$$\hat{H}^L = H^{GDEM} + CSM \quad (28)$$

$$e = H^L - \hat{H}^L \quad (29)$$

### 5.4 Validation of corrector surface models using final reliability index

As a result of using many KPIs to evaluate the behavior of different CSMs, the final reliability index ( $RI_F$ ) will collect them under its umbrella as one unit for every individual CSM to facilitate the comparison process. Henceforth, the weighted linear combination method will be

used [19, 20, 21, 32]. The steps to calculate the  $RI_F$  are as follows: first, the KPIs including range, RMSE, absolute of mean, S.D., and R will be scaled to be  $KPI_{Sc}$  and to be ranked from 0 to any pre-chosen constant number ( $N$ ) ascendingly as in Eq. (30) where  $KPI_{max}$  and  $KPI_{min}$  are the maximum and minimum values, respectively, and  $KPI_i$  is any value needed to be scaled and ranked. Second, the weights are assigned values as; 3 for the mean; 2 for RMSE, S.D., and R; and 1 for the range. It is worthy to note that these chosen values represent our interest to monitor the behavior of different transformation processes in reducing the systematic errors under the existence of outliers to get high accuracy and nearly normal distribution of residuals. As shown in Eq. (31),  $RI_w$  is the weighted internal or external reliability index and representing the weighted average of  $KPI_{Sc}$ .

$$KPI_{Sc} = \frac{(KPI_i - KPI_{min})}{(KPI_{max} - KPI_{min})} * N \quad (30)$$

$$RI_w = \frac{\sum KPI_{Sc} W_j}{\sum W_j} \quad (31)$$

Third, the internal and external reliability indices are combined to get the total reliability index  $RI_T$  as in Eq. (32) where ( $\alpha, \beta$ ) are control factors. Here, we used three cases; (1,0) if we consider the fitting to the common points is more important than the prediction of checkpoints; (0.5,0.5) if we need a balanced state between the fitting and prediction, and using (0,1) is considered if the prediction is more important than the fitting.

$$RI_T = \alpha * RI_{in} + \beta * RI_{ex} \quad (32)$$

Finally, the scaled and ranked final reliability index  $RI_F$  is calculated as in Eq. (33) where ( $RI_{T_{max}}, RI_{T_{min}}$ ) are the maximum and minimum values of  $RI_T$ , and  $RI_i$  is any value needed to be scaled and ranked. Based on Eq. (33), the values of  $RI_F \in [0, 100]$ .

$$RI_F = \frac{(RI_i - RI_{T_{min}})}{(RI_{T_{max}} - RI_{T_{min}})} * 100 \quad (33)$$

### 5.5 Comparison between non-filtered and filtered solutions

At the final step of the research methodology of Fig. 2, there will be a comparison between the two solutions to investigate the role of the generated CSM in reducing systematic error and increasing the accuracy of the GDEMs.

**Table 5:** Key performance indices.

Key Performance Indices (KPIs)	Mathematical Expression
Maximum Error	$\text{Max}(e)$
Minimum Error	$\text{Min}(e)$
Range	$\text{Max}(e) - \text{Min}(e)$
Mean Error ( $\bar{e}$ )	$\sum_i^n e_i / n$
Standard Deviation ( $\sigma_e$ )	$\sqrt{\frac{\sum_i^n (e_i - \bar{e})^2}{n - 1}}$
Root Mean Square Error (RMSE)	$e^T e / n$
Correlation (R)	$\frac{\sum_i^n (H_i^L - \overline{H^L})(\hat{H}_i^L - \overline{\hat{H}^L})}{\sqrt{\sum_i^n (H_i^L - \overline{H^L})^2} \times \sqrt{\sum_i^n (\hat{H}_i^L - \overline{\hat{H}^L})^2}}$

## 6 Result analysis

In this section, the analysis of the results from applying different transformation techniques will be illustrated. It is worthy to note that the gained results are restricted to the study area due to the uneven distribution of the available points. Tables 6 and 7 show the resulted resulting KPIs internally and externally for both GDEMs before the filtration process, while Tables 8 and 9 show the KPIs for both models in the second solution after the filtration process. In these tables, the abbreviated names of the built-in kernel functions are appended to the soft computing technique names LS-SVM and SVM to further delineate the models. In an ANN, the most optimized number of neurons is attached to the name of ANN. Only the optimized degree of the PR is recorded. In the LSC technique, Tables 6 and 7

show the optimized interval distance in radians and the type of soft computing technique associated with its kernel function to generate ACF. Every group of the transformation techniques is shown with a different color to be easily tracked and be differentiated from the other techniques.

To assess the behavior of every transformation technique, a reliability analysis will be used as mentioned before. Eqs. (30)–(33) are applied in Tables 6–9 with different ratios of  $\alpha$ ,  $\beta$ , as explained previously. The results are listed in Tables 10–11. Based on  $RI_F$  values, Tables 12 and 13 show the rank of every individual transformation technique according to the three different conditions (fitting, balance, and prediction). We will enclose our discussion on the balance condition because, if we consider fitting only, the overfitting problem can arise and affect the prediction of

**Table 6:** GTOPO30 KPIs before the filtration process (first solution).

KPIs	Internal Quality					External Quality				
Model	Range	RMSE	Mean	S.D.	R	Range	RMSE	Mean	S.D.	R
Bursa–Wolf	273.127	19.546	0.000	19.587	1.000	56.560	11.876	3.793	11.331	0.996
Molodensky–Badekas	273.127	19.546	0.000	19.587	1.000	56.490	11.876	3.793	11.331	0.996
LS-SVM-RBF	245.692	16.925	0.000	16.961	0.992	57.227	11.549	3.812	10.978	0.996
LS-SVM-linear	265.177	19.052	0.000	19.093	0.990	62.301	12.563	3.895	12.027	0.995
LS-SVM-poly	260.215	18.475	0.000	18.514	0.991	58.336	11.830	4.336	11.083	0.996
SVM-linear	271.455	19.381	1.614	19.354	0.990	56.723	12.611	5.374	11.488	0.996
SVM-quadratic	278.572	19.290	1.462	19.275	0.990	52.964	11.732	5.885	10.219	0.997
SVM-cubic	253.794	30.974	3.688	30.818	0.982	102.268	28.252	11.151	26.138	0.987
SVM-fine Gaussian	270.367	18.280	0.651	18.307	0.991	54.976	11.452	4.524	10.593	0.996
SVM-medium Gaussian	274.068	18.933	0.965	18.948	0.990	50.192	11.689	5.496	10.388	0.996
SVM-coarse Gaussian	272.493	19.330	1.519	19.311	0.990	55.308	12.343	5.262	11.243	0.996
MFFNN (07) neurons	199.379	14.197	0.001	14.227	0.995	62.188	11.821	5.029	10.772	0.996
RBNN (200) neurons	262.749	18.235	0.000	18.273	0.991	59.503	11.701	4.543	10.857	0.997
LI	73.859	3.570	0.000	3.578	1.000	54.862	11.439	5.357	10.177	0.997
CI	73.859	3.570	0.000	3.578	1.000	58.892	11.997	5.120	10.925	0.996
PR (3 <sup>rd</sup> ) degree	259.736	18.355	0.000	18.393	0.991	57.481	11.598	4.637	10.704	0.997
LSC-rad002-LS-SVM-poly	0.001	0.000	0.000	0.000	1.000	48.406	11.479	5.034	10.387	0.996

**Table 7:** SRTM 1 KPIs before the filtration process (first solution).

KPIs	Internal Quality					External Quality				
Model	Range	RMSE	Mean	S.D.	R	Range	RMSE	Mean	S.D.	R
Bursa–Wolf	194.856	12.494	0.000	12.520	1.000	23.657	4.348	-1.409	4.143	0.999
Molodensky–Badekas	194.856	12.494	0.000	12.520	1.000	23.657	4.348	-1.409	4.143	0.999
LS-SVM-RBF	190.508	12.279	0.000	12.305	0.996	23.628	4.686	-1.430	4.495	0.999
LS-SVM-linear	190.339	12.276	0.000	12.302	0.996	23.771	4.711	-1.420	4.524	0.999
LS-SVM-poly	190.441	12.278	0.000	12.303	0.996	23.686	4.696	-1.426	4.507	0.999
SVM-linear	194.061	12.402	0.126	12.428	0.996	22.363	4.260	-1.332	4.076	0.999
SVM-quadratic	194.523	12.458	0.202	12.482	0.996	22.921	4.442	-1.245	4.295	0.999
SVM-cubic	192.988	12.762	2.008	12.630	0.996	21.626	4.179	0.851	4.121	0.999
SVM-fine Gaussian	193.937	12.387	0.357	12.408	0.996	23.507	4.203	-1.227	4.049	0.999
SVM-medium Gaussian	193.895	12.326	0.230	12.349	0.996	21.923	4.300	-1.328	4.120	0.999
SVM-coarse Gaussian	193.938	12.398	0.000	12.424	0.996	22.231	4.371	-1.504	4.134	0.999
MFFNN (12) neurons	81.040	5.936	0.000	5.948	0.999	28.545	5.386	-1.618	5.174	0.999
RBNN (200) neurons	188.548	11.919	0.000	11.944	0.996	24.786	5.092	-1.273	4.966	0.999
LI	81.043	3.919	0.000	3.927	1.000	42.685	5.454	-1.002	5.400	0.999
CI	81.043	3.919	0.000	3.927	1.000	45.923	5.808	-0.928	5.775	0.999
PR (2 <sup>nd</sup> ) degree	188.090	12.212	0.000	12.238	0.996	22.295	4.738	-1.448	4.543	0.999
LSC-rad002-SVM-linear	0.001	0.000	0.000	0.000	1.000	36.241	5.917	-1.325	5.808	0.999

**Table 8:** GTOPO30 KPIs after the filtration process (second solution).

KPIs	Internal Quality					External Quality				
Model	Range	RMSE	Mean	S.D.	R	Range	RMSE	Mean	S.D.	R
Bursa–Wolf	55.104	9.242	0.000	9.262	1.000	53.577	12.544	5.748	11.227	0.996
Molodensky–Badekas	55.104	9.242	0.000	9.262	1.000	53.577	12.544	5.748	11.227	0.996
LS-SVM-RBF	13.162	1.690	0.000	1.693	1.000	46.764	12.311	6.510	10.521	0.996
LS-SVM-linear	54.828	9.220	0.000	9.241	0.997	53.906	12.559	5.746	11.245	0.996
LS-SVM-poly	54.781	7.882	0.000	7.900	0.998	51.699	11.927	6.334	10.177	0.997
SVM-linear	54.611	9.206	0.000	9.227	0.997	54.258	12.605	5.763	11.288	0.996
SVM-quadratic	57.947	8.067	0.423	8.074	0.998	48.485	11.982	6.758	9.963	0.997
SVM-cubic	129.644	21.292	0.543	21.333	0.986	89.696	20.333	7.835	18.892	0.991
SVM-fine Gaussian	59.065	7.421	-0.379	7.429	0.998	49.867	11.536	5.465	10.230	0.997
SVM-medium Gaussian	59.240	7.953	-0.130	7.970	0.998	48.482	11.884	6.251	10.178	0.997
SVM-coarse Gaussian	54.006	8.959	0.075	8.979	0.997	52.722	12.443	5.843	11.062	0.996
MFFNN (08) neurons	52.713	6.920	0.000	6.936	0.998	46.704	11.645	5.358	10.410	0.996
RBNN (200) neurons	51.181	7.653	0.000	7.670	0.998	51.511	11.826	6.349	10.046	0.997
LI	9.996	0.496	0.000	0.497	1.000	40.279	11.349	6.411	9.430	0.997
CI	9.996	0.496	0.000	0.497	1.000	41.528	11.471	6.132	9.762	0.997
PR (5 <sup>th</sup> ) degree	52.320	7.206	0.000	7.223	0.998	49.418	11.674	6.071	10.039	0.997
LSC-rad002-LS-SVM-linear	0.027	0.006	-0.002	0.006	1.000	51.535	11.445	5.549	10.078	0.997

**Table 9:** SRTM 1 KPIs after the filtration process (second solution).

KPIs	Internal Quality					External Quality				
Model	Range	RMSE	Mean	S.D.	R	Range	RMSE	Mean	S.D.	R
Bursa–Wolf	21.962	3.233	0.000	3.241	1.000	22.506	4.185	-1.019	4.089	0.999
Molodensky–Badekas	21.962	3.233	0.000	3.241	1.000	22.506	4.185	-1.019	4.089	0.999
LS-SVM-RBF	16.790	2.761	0.000	2.767	1.000	22.567	4.198	-1.118	4.076	0.999
LS-SVM-linear	21.603	3.346	0.000	3.354	1.000	21.430	4.208	-1.087	4.094	0.999
LS-SVM-poly	18.342	3.112	0.000	3.119	1.000	21.658	4.249	-1.047	4.148	0.999
SVM-linear	25.204	3.352	-0.069	3.359	1.000	22.004	4.145	-1.085	4.029	0.999
SVM-quadratic	25.492	3.310	0.039	3.317	1.000	22.621	4.353	-0.979	4.272	0.999
SVM-cubic	30.547	4.606	1.497	4.366	0.999	20.899	4.333	0.820	4.285	0.999
SVM-fine Gaussian	19.531	2.781	0.054	2.786	1.000	22.856	4.140	-1.015	4.043	0.999
SVM-medium Gaussian	22.935	3.135	-0.041	3.142	1.000	21.777	4.228	-1.127	4.105	0.999
SVM-coarse Gaussian	24.025	3.306	-0.124	3.311	1.000	21.977	4.219	-1.160	4.085	0.999
MFFNN (12) neurons	15.451	2.413	0.000	2.418	1.000	21.235	4.101	-1.397	3.883	0.999
RBNN (200) neurons	17.589	3.056	0.000	3.063	1.000	21.978	4.378	-1.006	4.292	0.999
LI	4.723	0.264	0.000	0.265	1.000	24.707	4.494	-1.016	4.410	0.999
CI	4.723	0.264	0.000	0.265	1.000	24.322	4.620	-0.973	4.549	0.999
PR (5 <sup>th</sup> ) degree	16.634	2.867	0.000	2.873	1.000	22.013	4.239	-1.017	4.145	0.999
LSC-rad002-LS-SVM-poly	0.001	0.000	0.000	0.000	1.000	20.284	4.278	-1.140	4.153	0.999

checkpoints. On the other hand, if we consider prediction only, the fitted surfaces can deviate from the common points. In general, it is clear that the internal quality for LSC integrated with LS-SVM always scores high rank. The explanation for this result can be found in Eq. (27), as the cross-covariance  $C_{qp}$  matrix will be replaced by the auto-covariance  $C_{pp}$  matrix, and the final result will identical to the common points' values.

For the GTOPO30 model in Table 12, it is evident that SVM with the cubic kernel is the worst technique of all. Its behavior did not change even when the outliers were removed from the raw data. On the other hand, the LSC integrated with LS-SVM inherent with the poly kernel functions (LSC-rad002-LS-SVM-poly and LSC-rad002-LS-SVM-linear) shows the best behavior before and after

the filtration process. Its CSM can overcome the existence of outliers in the raw data and predict the checkpoint values with high reliability compared to the other transformation techniques. The most important note is that the non-parametric equations (LSC and interpolation) scored a higher ranking than the parametric equations at the fitting and balance conditions. The majority of the SVM group comes at the end of the table, meaning that this technique cannot deal smoothly with high inconsistency data. The remaining soft computing techniques like LS-SVM, ANN, and the geodetic equations come in the middle between the highest and lowest ranking techniques.

In Table 13, SVM with cubic kernel proves again its inability to represent variation between the two groups

**Table 10:** Reliability index for GTOPO30.

$RI_F$	Before filtration			After filtration		
Model	Fitting	Balance	Prediction	Fitting	Balance	Prediction
Bursa–Wolf	51.768	75.073	98.879	69.765	78.727	88.058
Molodensky–Badekas	51.768	75.079	98.893	69.765	78.727	88.058
LS-SVM-RBF	53.418	76.466	100.000	94.310	90.480	86.189
LS-SVM-linear	47.418	70.343	93.778	67.703	77.582	87.891
LS-SVM-poly	49.291	73.162	97.564	72.196	81.377	90.934
SVM-linear	37.782	64.883	92.673	67.759	77.354	87.361
SVM-quadratic	38.522	67.455	97.131	55.854	71.772	88.512
SVM-cubic	0.000	0.000	0.000	0.000	0.000	0.000
SVM-fine Gaussian	45.755	72.237	99.351	59.209	79.079	100.000
SVM-medium Gaussian	42.075	69.438	97.477	66.861	79.380	92.489
SVM-coarse Gaussian	38.356	65.951	94.248	65.738	76.788	88.341
MFNN	62.123	78.680	95.474	75.069	86.693	98.833
RBNN (200) neurons	49.671	73.806	98.480	73.121	82.200	91.646
LI	90.322	94.748	98.893	97.922	97.737	97.302
CI	90.322	92.944	95.226	97.922	98.016	97.876
PR	49.545	74.046	99.098	74.292	84.455	95.047
LSC-rad002-LS-SVM-poly	100.000	100.000	99.531	100.000	100.000	99.755

**Table 11:** Reliability index for SRTM 1.

$RI_F$	Before filtration			After filtration		
Model	Fitting	Balance	Prediction	Fitting	Balance	Prediction
Bursa–Wolf	30.825	85.270	96.084	49.483	62.266	71.162
Molodensky–Badekas	30.825	85.270	96.084	49.483	62.266	71.162
LS-SVM-RBF	22.056	46.561	79.140	57.508	68.986	71.995
LS-SVM-linear	22.079	44.583	77.591	48.088	62.342	73.269
LS-SVM-poly	22.066	45.759	78.510	52.295	61.193	65.248
SVM-linear	20.035	69.706	99.821	45.914	64.744	80.689
SVM-quadratic	18.994	52.923	88.466	46.783	45.671	44.928
SVM-cubic	0.000	0.000	75.473	0.000	0.000	28.311
SVM-fine Gaussian	17.821	65.746	100.000	55.628	70.060	76.594
SVM-medium Gaussian	19.374	66.003	97.943	49.936	62.164	70.338
SVM-coarse Gaussian	21.317	69.369	97.702	46.195	60.234	72.127
MFNN (12) neurons	65.226	82.516	44.013	62.611	88.434	100.000
RBNN (200) neurons	23.862	18.149	54.733	53.291	50.241	44.012
LI	77.279	75.319	20.989	94.914	68.590	18.448
CI	77.279	47.945	0.000	94.914	58.401	0.000
PR	22.497	45.459	77.655	56.140	63.912	64.740
LSC-rad002-LS-SVM-poly	100.000	100.000	6.915	100.000	100.000	68.135

of data (SRTM 1 and national GPS/leveling). The LSC integrated with LS-SVM inherent with linear kernel and polynomial function (LSC-rad002-LS-SVM-linear and LSC-rad002-LS-SVM-poly) shows the best behavior before and after the filtration process. Based on previous analysis, the decision has been made to choose LSC-rad002-LS-SVM-poly and LSC-rad002-LS-SVM-linear as the most suitable transformation technique at the balance condition to fuse GTOPO30 and SRTM 1 with the national GPS/leveling data. Table 14 shows the final KPIs related to the two GDEMs after applying the suitable transformation techniques before and after the filtration process for all GCPs. The results

show an improvement in all KPIs relative to the original values in Table 4.

Table 15 lists the improvement percentage for every KPI. It is clear that the generated CSM reduced the inherent systematic errors and increased the accuracy for both GDEMs. Moreover, for GTOPO30 before filtration, its RMSE and S.D. values equal to 5.561 m and 5.443 m, respectively, while these values in SRTM 1 in Table 4 before filtration reached 11.485 m and 11.275 m, respectively. This means that, when a suitable transformation technique is applied to the old model GTOPO30, its KPIs give better results than the SRTM 1 model. This proves the idea of this re-

**Table 12:** The rank of transformation techniques related to GTOPO30.

Rank	Before filtration			After filtration		
	Fitting	Balance	Prediction	Fitting	Balance	Prediction
1	LSC-rad002-LS-SVM-poly	LSC-rad002-LS-SVM-poly	LS-SVM-RBF	LSC-rad002-LS-SVM-linear	LSC-rad002-LS-SVM-linear	SVM-fine Gaussian
2	LI	LI	LSC-rad002-LS-SVM-poly	LI	CI	LSC-rad002-LS-SVM-linear
3	CI	CI	SVM-fine Gaussian	CI	LI	MFFNN (8) neurons
4	MFFNN (7) neurons	MFFNN (7) neurons	PR (3) degree	LS-SVM-RBF	LS-SVM-RBF	CI
5	LS-SVM-RBF	LS-SVM-RBF	Molodensky-Badekas	MFFNN (8) neurons	MFFNN (8) neurons	LI
6	Bursa-Wolf	Molodensky-Badekas	LI	PR (5) degree	PR (5) degree	PR (5) degree
7	Molodensky-Badekas	Bursa-Wolf	Bursa-Wolf	RBNN (200) neurons	RBNN (200) neurons	SVM-medium Gaussian
8	RBNN (200) neurons	PR (3) degree	RBNN (200) neurons	LS-SVM-poly	LS-SVM-poly	RBNN (200) neurons
9	PR (3) degree	RBNN (200) neurons	LS-SVM-poly	Bursa-Wolf	SVM-medium Gaussian	LS-SVM-poly
10	LS-SVM-poly	LS-SVM-poly	SVM-medium Gaussian	Molodensky-Badekas	SVM-fine Gaussian	SVM-quadratic
11	LS-SVM-linear	SVM-fine Gaussian	SVM-quadratic	SVM-linear	Bursa-Wolf	SVM-coarse Gaussian
12	SVM-fine Gaussian	LS-SVM-linear	MFFNN (7) neurons	LS-SVM-linear	Molodensky-Badekas	Bursa-Wolf
13	SVM-medium Gaussian	SVM-medium Gaussian	CI	SVM-medium Gaussian	LS-SVM-linear	Molodensky-Badekas
14	SVM-quadratic	SVM-quadratic	SVM-coarse Gaussian	SVM-coarse Gaussian	SVM-linear	LS-SVM-linear
15	SVM-coarse Gaussian	SVM-coarse Gaussian	LS-SVM-linear	SVM-fine Gaussian	SVM-coarse Gaussian	SVM-linear
16	SVM-linear	SVM-linear	SVM-linear	SVM-quadratic	SVM-quadratic	LS-SVM-RBF
17	SVM-cubic	SVM-cubic	SVM-cubic	SVM-cubic	SVM-cubic	SVM-cubic
Legend						
	Geodetic	LS-SVM	SVM	ANN	Interpolation	LSC

**Table 13:** The rank of transformation techniques related to SRTM 1.

Rank	Before filtration			After filtration		
	Fitting	Balance	Prediction	Fitting	Balance	Prediction
1	LSC-rad002-LS-SVM-linear	LSC-rad002-LS-SVM-linear	SVM-fine Gaussian	LSC-rad002-LS-SVM-poly	LSC-rad002-LS-SVM-poly	MFFNN (12) neurons
2	LI	Bursa-Wolf	SVM-linear	LI	MFFNN (12) neurons	SVM-linear
3	CI	Molodensky-Badekas	SVM-medium Gaussian	CI	SVM-fine Gaussian	SVM-fine Gaussian
4	MFFNN (12) neurons	MFFNN (12) neurons	SVM-coarse Gaussian	MFFNN (12) neurons	LS-SVM-RBF	LS-SVM-linear
5	Bursa-Wolf	LI	Bursa-Wolf	LS-SVM-RBF	LI	SVM-coarse Gaussian
6	Molodensky-Badekas	SVM-linear	Molodensky-Badekas	PR (5) degree	SVM-linear	LS-SVM-RBF
7	RBNN (200) neurons	SVM-coarse Gaussian	SVM-quadratic	SVM-fine Gaussian	PR (5) degree	Bursa-Wolf
8	PR (2) degree	SVM-medium Gaussian	LS-SVM-RBF	RBNN (200) neurons	LS-SVM-linear	Molodensky-Badekas
9	LS-SVM-linear	SVM-fine Gaussian	LS-SVM-poly	LS-SVM-poly	Bursa-Wolf	SVM-medium Gaussian
10	LS-SVM-poly	SVM-quadratic	PR (2) degree	SVM-medium Gaussian	Molodensky-Badekas	LSC-rad002-LS-SVM-poly
11	LS-SVM-RBF	CI	LS-SVM-linear	Bursa-Wolf	SVM-medium Gaussian	LS-SVM-poly
12	SVM-coarse Gaussian	LS-SVM-RBF	SVM-cubic	Molodensky-Badekas	LS-SVM-poly	PR (5) degree
13	SVM-linear	LS-SVM-poly	RBNN (200) neurons	LS-SVM-linear	SVM-coarse Gaussian	SVM-quadratic
14	SVM-medium Gaussian	PR (2) degree	MFFNN (12) neurons	SVM-quadratic	CI	RBNN (200) neurons
15	SVM-quadratic	LS-SVM-linear	LI	SVM-coarse Gaussian	RBNN (200) neurons	SVM-cubic
16	SVM-fine Gaussian	RBNN (200) neurons	LSC-rad002-LS-SVM-linear	SVM-linear	SVM-quadratic	LI
17	SVM-cubic	SVM-cubic	CI	SVM-cubic	SVM-cubic	CI
Legend						
	Geodetic	LS-SVM	SVM	ANN	Interpolation	LSC

search that, with suitable direct mathematical transformation technique, the old GDEMs can give results with equal or better precision than the outcomes from the new GDEMs which are supported by new technology. This concept can

help developed countries where there is a shortage in the latest technology, or when the new data is blocked or not available. Furthermore, the direct transformation process can help in utilizing the unfiltered measurements that



**Table 14:** The resultant KPIs after applying the transformation process.

Model	Before filtration						After filtration					
	GCPs	Range (m)	RMSE (m)	Mean (m)	S.D. (m)	R	GCPs	Range (m)	RMSE (m)	Mean (m)	S.D. (m)	R
GTOPO30	311	48.406	5.561	1.182	5.443	0.999	294	51.543	5.702	1.370	5.544	0.999
SRTM 1	311	36.242	2.807	−0.298	2.796	0.999	294	20.284	2.088	−0.272	2.073	0.999

**Table 15:** The improvement percentage of KPIs relative to the original KPIs.

Model	Before filtration						After filtration					
	GCPs	Range (%)	RMSE (%)	Mean (%)	S.D. (%)	R (%)	GCPs	Range (%)	RMSE (%)	Mean (%)	S.D. (%)	R (%)
GTOPO30	311	82.43	68.67	39.31	69.20	1.73	294	6.50	36.53	3.48	38.36	0.40
SRTM 1	311	81.53	75.55	86.88	75.20	0.60	294	14.53	47.90	85.18	41.92	0.00

come from the GDEMs in the fusion process with the national measurements to densify or to fill the gap areas.

## 7 Conclusion

From this study, we reached to the following conclusions: The GDEMs cannot be used directly in various applications, or fused with national DEMs, due to many causes; the existence of outliers; different vertical datums; and the different nature of GDEMs. The filtration of raw data to reduce the effect of these causes should be considered first. The direct transformation technique between the GDEMs and the national data shows an acceptable result based on adopted key performance indices. It is found that at balance state where the fitting and prediction have equal priority, least-squares collocation integrated with least-squares support vector machine inherited with linear or polynomial kernel function exhibits the most accurate behavior among the different transformation techniques. The resulting KPIs for this model are shown in Table 14 for both GTOPO30 and SRTM 1. For the GTOPO30 model, before filtration of the raw data, there is an improvement in the mean and root mean square of errors by 39.31% and 68.67%, respectively. In the SRTM 1 model, these values reached 86.88% and 75.55%, respectively. On the other hand, after the filtration process, these values became 3.48% and 36.53% for GTOPO30, and 85.18% and 47.90% for SRTM 1, respectively. The improvement percentage of all KPIs are listed in Table 15. The most important note is, using a suitable direct mathematical transformation technique can help to increase the precision of classic GDEMs such as GTOPO30 to have equal or better

accuracy than the new SRTM 1 which is supported by advanced technology. This can help to overcome the problem of the shortage of supporting technology or blocked data, particularly in developed countries. Henceforth, the direct transformation technique represents a fast and economical way to utilize the unfiltered measurements that come from the GDEMs in the fusion process with the national data to densify or to fill the gap areas.

## References

- [1] Abd-Elmotaal, H. A., (1994), "Comparison of polynomial and similarity transformation-based datum-shifts for Egypt", *Bulletin Geodesique*, vol. 68, pp. 168–172.
- [2] Abd-Elmotaal, H. A., (2011), "The new Egyptian height models EGH10", *NRIAG journal of astronomy and geophysics*, Special issue, pp. 249–261.
- [3] Aguilar, F. J., Agüera, F., Aguilar, M. A., and Carvajal, F., (2005), "Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy", *Photogrammetric Engineering and Remote Sensing*, vol. 71, no. 7, pp. 805–816, doi: 10.14358/PERS.71.7.805.
- [4] Aguilar, F. J., Aguilar, M. A., and Agüera, F., (2007), "Accuracy assessment of digital elevation models using a non-parametric approach", *International Journal of Geographic Information Science*, vol. 21, no. 6, pp. 667–686, doi: 10.1080/13658810601079783.
- [5] Akyilmaz, O., Özlüdemir, M. T., Ayan, T., and Çelik, R. N., (2009), "Soft computing methods for geoidal height transformation", *Earth Planets and Space*, vol. 61, no. 7, pp. 825–833.
- [6] Ali, M. E. O., Shaker, I. F. M., and Saba, N. M., (2017), "A perspective of reliable and accurate DEM using world DEMs data fusion", *ISER*, vol. 8, no. 6. ISSN 2229-5518.
- [7] Ali, M. H. and Abustan, I., (2014), "A new novel index for evaluating model performance", *Journal of Natural*



- Resources and Development, vol. 2014, no. 04, pp. 1–9, doi: 10.5027/jnrd.v4i0.01.
- [8] Al-Karagy, E. M., Hosny, M. M., and Dawod, G. M., (2015), “Investigation the precision of recent global geoid models and global digital elevation models for geoid modeling in Egypt”, Regional Conference on Surveying and Development, Sharm El-Sheikh, Egypt, 3–6 October 2015.
- [9] Al-Krargy, E. M., Doma, M. I., and Dawod, G. M., (2014), “Towards an Accurate Definition of the Local Geoid Model in Egypt using GPS/Leveling Data: A Case Study at Rosetta Zone”, International Journal of Innovative Science and Modern Engineering (IJISME), vol. 2, no. 11. ISSN: 2319-6386.
- [10] Amidror, I., (2002), “Scattered data interpolation methods for electronic imaging systems: a survey”, Journal of Electronic Imaging, vol. 11, no. 2, pp. 157–176.
- [11] Amin, M. M., El-fatrai, S. M., and Saba, N. M., (2013), “Accuracy assessment of world DEMs versus local DEM in Egypt”, Civil Engineering Research Magazine CERM, vol. 35, no. 3. Published by Faculty of Engineering, Al-Azhar University, Cairo, Egypt.
- [12] Arabelos, A. and Tziavos, I. N., (1983), “Determination of Deflection of the vertical using a combination of spherical harmonics and gravimetric data for the area of Greece”, Bull. Géod., vol. 57, pp. 240–256.
- [13] Arabelos, D., (2000), “Inter comparisons of the global DTMs ETOPO5, Terrain Base and JGP95E”, Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy, vol. 25, no. 1, pp. 89–93.
- [14] Beale, M. H., Hagan, M. T., and Demuth, H. B., (2015), “Neural network toolbox user’s guide”, The Math Works, Inc.
- [15] Böhm, W., Farin, G., and Kahmann, J., (1984), “A survey of curve and surface methods in CAGD”, Comput. Aided Des., vol. 1, pp. 1–60.
- [16] Cakir, L. and Yilmaz, N., (2014), “Polynomial, radial basis functions and multilayer perceptron neural network methods in local geoid determination with GPS/levelling”, J. Measurements, vol. 57, pp. 48–153.
- [17] Cross, P. A., (1983), “Advanced least squares applied to position fixing”, North East London Polytechnic Department of Land Surveying.
- [18] Darbeheshti, N. and Featherstone, W. E., (2009), “Non-stationary covariance function modeling in 2D least-squares collocation”, J. Geod., vol. 83, pp. 495–508, doi: 10.1007/s00190-008-0267-0.
- [19] Dawod, G., (2013), “Suitability analysis for tourist infrastructures utilizing multi-criteria GIS: A case study in Al-Hada city, Saudi Arabia”, International journal of geomatics and geosciences, vol. 4, no. 2, pp. 313–324.
- [20] Dawod, G. and Al-Ghamdy, K., (2017), “Reliability of recent global digital elevation models for geomatics application in Egypt and Saudi Arabia”, Journal of Geographic Information System, vol. 9, pp. 685–698, doi: 10.4236/jgis.2017.96043.
- [21] Dawod, G. and Mandoer, M. S., (2016), “Optimum Sites for Solar Energy Harvesting in Egypt Based on Multi-Criteria GIS”, The 11th The First Future University International Conference on New Energy and Environmental Engineering Cairo, Egypt. April 11–14, 2016.
- [22] De Brabanter, K., Karsmakers, F., Ojeda, C., Alzate, J., De Brabanter, J., Pelckmans, K., De Moor, B., Vandewalle, J., and Suykens, J. A. K., (2011), “LS-SVMLab Toolbox User’s Guide Version 1.8”, ESAT-SISTA Technical report 10-146, Katholieke Universiteit Leuven, Belgium, [http://www.esat.kuleuven.be/\\_sista/lssvmlab/](http://www.esat.kuleuven.be/_sista/lssvmlab/).
- [23] Deakin, R. E., (2006), “A Note on the Bursa-Wolf and Molodensky-Badekas Transformations”, School of mathematical and geospatial sciences, RMIT University, Australia.
- [24] Denker, H., (2005), “Evaluation of SRTM3 and GTOPO30 Terrain Data in Germany”. In: Jekeli C., Bastos L., Fernandes J. (eds.) Gravity, Geoid and Space Missions. International Association of Geodesy Symposia, vol. 129, pp. 218–223, Springer, Berlin, Heidelberg, doi: 10.1007/3-540-26932-0\_38.
- [25] Dermanis, A., (1984), “Kriging and Collocation – A comparison”, Manuscript Geodaetica, vol. 9, pp. 159–167.
- [26] Doganalp S., (2016), “Geoid height computation in strip-area project by using least-squares collocation”, Acta Geodyn. Geomater., vol. 13, no. 2 (182), pp. 167–176, doi: 10.13168/AGG.2015.0054.
- [27] Ebaid, H., (2014), “Accuracy enhancement of SRTM and ASTER DEMs using weight estimation regression model”, IJRET, vol. 3, no. 08, ISSN: 2319-1163 (online), ISSN: 2321-7308 (print).
- [28] El-quilish, M., El-ashquer M., Dawod, G., and El fiky, G., (2018), “Development and accuracy assessment of high-resolution digital elevation model using GIS approaches for the Nile delta region, Egypt”, American Journal of Geographic Information System, vol. 7, no. 4, pp. 107–117, doi: 10.5923/j.ajgis.20180704.02.
- [29] Elshambaky, H. T., (2017), “Application of neural network to determine a corrector surface for global geopotential model using GPS/levelling measurements in Egypt”, J. Appl. Geodesy, vol. 12, no. 1, pp. 29–44, doi: 10.1515/jag-2017-0017.
- [30] Elshambaky, H. T., (2018), “Enhancing the predictability of least-squares collocation through the integration with least-squares-support vector machine”, J. Appl. Geodesy, doi: 10.1515/jag-2018-0017.
- [31] Elshambaky, H. T., Kaloop, M. R., and Hu, J. W., (2018), “A novel three-direction datum transformation of geodetic coordinates for Egypt using artificial neural network approach”, Arabian Journal of Geoscience, vol. 11, pp. 110, doi: 10.1007/s12517-018-3441-6.
- [32] El shouny, A., Al-karagy, E. M., Mohamed, H. F., and Dawod, G. M., (2018), “GIS-based accuracy assessment of global geopotential models: a case study of Egypt”, American Journal of Geographic Information System, vol. 7, no. 4, pp. 118–124, doi: 10.5923/j.ajgis.20180704.03.
- [33] Espinoza, M., Suykens J. A. K., and De Moor, B., (2005), “Load forecasting using fixed-size least squares support vector machines”. In: Cabestany, J., Prieto, A., Sandoval, F. (eds.) “Computational Intelligence and Bioinspired Systems”, IWANN 2005. Lecture notes in computer science, vol. 3512, Springer, Berlin, Heidelberg, doi: 10.1007/11494669\_125.
- [34] Fan, R. E., Chen, P. H., and Lin, C. J., (2005), “Working set selection using second order information for training support vector machines”, Journal of Machine Learning Research, vol. 6, pp. 1871–1918.
- [35] Fan, R. E., Chen, P. H., and Lin, C. J., (2006), “A study on SMO-type decomposition methods for support vector machines”, IEEE Transactions on Neural Networks, vol. 17, pp. 893–908.

- [36] Farin, G., (1997), "Curves and Surfaces for Computer Aided Geometric Design", 4th edn., Academic, San Diego.
- [37] Fazilova, D., (2017), "The review and development of a modern GNSS network and datum in Uzbekistan", *Geodesy and Geodynamics*, pp. 2–7.
- [38] Fisher, P. F. and Tate, N. J., (2006), "Causes and consequences of error in digital elevation models", *Progress in Physical Geography*, vol. 30, no. 4, pp. 467–489, doi: 10.1191/0309133306pp492ra.
- [39] Florinsky, I. V., (2012), "Digital terrain analysis in soil science and geology", Academic Press.
- [40] Fotopoulos, G., Featherstone, W. E., and Sideris, M. G., (2002), "Fitting a gravimetric geoid model to the Australian height datum via GPS data", IAG Third Meeting of the International Gravity and Geoid Commission, Thessaloniki, Greece, Aug. 26–30, 2002.
- [41] Fotopoulos, G., Kotsakis, C., and Sideris, G., (2003), "How accurately can we determine Orthometric height differences from GPS and geoid data", *J. Surv. Eng.*, vol. 129, no. 1, pp. 1–10, doi: 10.1061/(ASCE)0733-9453(2003)129:1(1).
- [42] Gad, M. A., Odalović, O. R., and Zaky, K. M., (2018), "Case study – Accuracy assessment of SRTM 1,3-arcsec by using topographic DEM over limited area of Egypt territory", *IJSER*, vol. 9, no. 8.
- [43] Gesch, D. B., (1998), "Accuracy assessment of a global elevation model using shuttle laser altimeter data", *IGARSS '98. Sensing and Managing the Environment. 1998 IEEE International Geoscience and Remote Sensing Symposium Proceedings*. (Cat. No. 98CH36174), Seattle, USA, doi: 10.1109/IGARSS.1998.699601.
- [44] Gesch, D. B. and Larson, K. S., (1996), "Techniques for development of global 1-kilometer digital elevation models", in *Proceedings, Pecora Thirteen Symposium*, Sioux Falls, South Dakota, August 20–22, 1996 (CD-ROM), Am. Soc. for Photogrammetry and Remote Sens., Bethesda, Md., 1998.
- [45] Gesch, D. B., Verdin, K. L., and Greenlee, S. K., (1999), *New land surface digital elevation model covers the Earth*. Eos, Transactions American Geophysical Union, vol. 80, no. 6, pp. 69–70.
- [46] Grohmann, C. H., (2016), "Comparative analysis of global digital elevation models and ultra-prominent mountain peaks", *ISPRS Annals of the photogrammetry, Remote Sensing and Spatial Information Science*, vol. III-4, 2016, XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic.
- [47] Hagan, M. T., Demuth, H. B., and Beale, M. H., (1996), "Neural Network Design", Boston, MA PWS Publishing.
- [48] Hardin, D. J., Gesch, D. B., Carabajal, C. C., and Luthcke, S. B., (1998), "Application of the shuttle laser altimeter in an accuracy assessment of GTOPO30, a global 1-kilometer digital elevation model", *ISPRS Mapping surface structure and topography by airborne and spaceborne lasers*, vol. WG III-5, XXXII-3/W14 ISPRS Congress, 9–11 Nov. 1999, La Jolla, USA, <http://www.isprs.org/PROCEEDINGS/XXXII/3-W14/default.aspx>.
- [49] Haykin, S., (2001), "Neural Network: A Comprehensive Foundation", 2nd edn., Hamilton, Ontario, Canada.
- [50] Hengl, T. and Evans, I. S., (2009), "Mathematical and digital models of the land surface", *Geomorphometry concepts, software, applications*, T. Hengl and H. I. Reuter, pp. 31–63.
- [51] Hilton, R. D., Featherstone, W. E., Berry, P. A. M., Johnson, C. P. D., and Kirby, J. F., (2003), "Comparison of digital elevation models over Australia and external validation using ERS-1 satellite radar altimetry", *Australian Journal of Earth Science*, vol. 50, pp. 157–168, doi: 10.1046/j.1440-0952.2003.00982.x.
- [52] Hornik, K. M., Stinchcombe, M., and White, H., (1989), "Multilayer feedforward Networks are universal approximators", *Neural Networks*, vol. 2, no. 5, pp. 359–366.
- [53] James, G., Witten, D., Hastie, T., and Tibshirani, R., (2013), "An introduction to statistical learning: with application in R", *Springer Texts in statistics*, vol. 103, doi: 10.1007/978-1-4614-7138-7\_9.
- [54] Jordan, S. K., (1972), "Self-consistent statistical models for the gravity anomaly, vertical deflection, and undulation of the geoid", *J. Geophys. Res.*, vol. 77, no. 20, pp. 3660–3670.
- [55] JPL-Shuttle Radar Topography Mission, (2015), "U.S. Releases Enhanced Shuttle Land Elevation Data" <https://www2.jpl.nasa.gov/srtm/>.
- [56] Kahaner, D., Cleve, M., and Stephen, N., (1988), "Numerical Methods and Software", Upper Saddle River, NJ: Prentice Hall.
- [57] Kimehr, R. and Sjöberg, L. E., (2005), "Effect of the SRTM global DEM on the determination of high-resolution geoid model: a case study in Iran", *J. Geod.*, vol. 2005, no. 79, pp. 540–551, doi: 10.1007/s00190-005-0006-8.
- [58] Kotsakis, C. and Sideris, M. G., (1999), "On the adjustment of combined GPS/levelling/geoid networks", *Journal of Geodesy*, vol. 73, pp. 412–421, doi: 10.1007/s001900050261.
- [59] Kreyszig, E., Kreyszig, H. and Norminton, E. J., (2011), "Advanced engineering mathematics", 10th edn., John Wiley & Sons, Inc., USA.
- [60] Mataija, M., Pogarčič, M., and Pogarčič, I., (2014), "Helmert transformation of reference coordinating systems for geodesic purposes in local frames", *Procedia Engineering*, vol. 69, pp. 168–176.
- [61] Math Works Inc., (2015), "Curve fitting toolbox user's guide", R2018b, Math Work Inc., pp. 6-2–6-47.
- [62] Maune, D. F., (2007), "Digital Elevation Model Technologies and Applications, the DEM User's Manual", Bethesda, MD: American Society for Photogrammetry and Remote Sensing.
- [63] Mikhail, E. M. and Ackermann, F., (1976), "Observations and Least Squares". Dun Donnelly, New York.
- [64] Milton, J. S. and Arnold, J. C., (1995), "Introduction to Probability and Statistics Principals and Applications for Engineering and the Computing Science", 3rd edn., McGraw-Hill Book Company, New York, USA.
- [65] Mohamed, M. H. and Saleh, S. S., (2018), "Fusion of SRTM and ASTER GDEM2 DEMs based on height error weighted average technique", *AJBAS*, vol. 12, no. 6, pp. 23–29, doi: 10.22587/ajbas.2018.12.6.5.
- [66] Moore, T. and Smith, M. J., (1998), "Back to basics geodetic transformations", *The University of Nottingham. Survey Review*, vol. 34, p. 270.
- [67] Moritz, H., (1978), "Least – Squares Collocation", *Review of Geophysics and Space Physics*, vol. 16, pp. 421–430.
- [68] Moritz, H., (1980), "Advanced physical geodesy", Abacus, Tunbridge Wells Kent.
- [69] Platt, J., (1998), "Sequential minimal optimization: A fast algorithm for training support vector machines", *Technical report, MSR-TR-98-14*.

- [70] Powell, S. M. (1997). "Results of the Final Adjustment of the New National Geodetic Network", Technical report, Egyptian Surveying Authority, Egypt.
- [71] Rabah, M., El-Hattab, A., and Abdallah, M., (2017), "Assessment of the most recent satellite based digital elevation models of Egypt", *NRIAG journal of astronomy and geophysics*, vol. 6, no. 2017, pp. 326–335, doi: 10.1016/j.nrag.2017.10.006.
- [72] Rodriguez, E., Morris, C. S., Belz, J. E., (2006), "A global assessment of the SRTM performance", *Photogrammetric Engineering and Remote Sensing*, vol. 72, no. 3, pp. 249–260. doi: 10.14358/PERS.72.3.249V.
- [73] Rodriguez, E., Morris, C. S., Belz, J. E., Chapin, E. C., Martin, J. M., Daffer, W., and Hensley, S. (2005), "An assessment of the SRTM topographic products", *JPL Publ.*, D31639.
- [74] Samui, P., Kim, D., and Aiyer, B. G., (2015), "Pullout capacity of small ground anchor: a least square support vector machine approach", *Journal of Zhejiang University-Science A (Applied Physics & Engineering)*, ISSN 1673-565X (print), ISSN 1862-1775 (online), [www.zju.edu.cn/jzus](http://www.zju.edu.cn/jzus); [www.springerlink.com](http://www.springerlink.com).
- [75] Schumaker, L. L., (1982), "Fitting surfaces to scattered data", No. 19830007490, Conference Paper, Proc. of the NASA Workshop on Surface Fitting, pp. 27–94.
- [76] Schwarz, K. P. and Lachapelle, G., (1980), "Local characteristics of the gravity anomaly covariance function", *Bull. Géod.*, vol. 54, pp. 21–36.
- [77] Shen, Y. Z., Chen, Y., and Zheng, D. H., (2006), "A quaternion-based geodetic datum transformation algorithm", *J Geod.*, vol. 80, no. 5, pp. 233–239.
- [78] Shuanggen, J., (2012), "Global navigation satellite system: Signals, Theory, and Application", ISBN 978-953-307-843-4, In Tech Europe, University Campus SteP Ri, Slavaka Krautzeka 83/A, 51000 Rijeka, Croatia, [www.intechopen.com](http://www.intechopen.com).
- [79] Stopar, B., Ambrozic, T., Kuhar, M., and Turk, G., (2006), "GPS-derived geoid using artificial neural network and least squares collocation", *Survey Review*, vol. 38, p. 300.
- [80] Suykens J. A. K., (2001), "Support vector machines: A nonlinear modelling and control perspective", *European Journal of Control*, vol. 7, pp. 311–327.
- [81] Suykens J. A. K. and Vandewalle J., (1999), "Least squares support vector machine classifiers", *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300.
- [82] Suykens J. A. K., Van Gestel T., De Brabanter J., De Moor, B., and Vandewalle, J., (2002), "Least Squares Support Vector Machines", World Scientific, Singapore.
- [83] Szu-Pyng, K., Chao-Nan, C., Hui-Chi, H., Yu-Ting, S., (2014), "Using a least squares support vector machine to estimate a local geoid model", *Bol. Ciênc. Geod.*, vol. 20, no. 2, doi: 10.1590/S1982-21702014000200025.
- [84] Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D., (2007), "The Shuttle Radar Topography Mission", *Rev. Geophys.*, vol. 45, RG2004, doi: 10.1029/2005RG00183.
- [85] Tscheerning, C. C., (2010), "The use of Least-Squares Collocation for the processing of GOCE data", *Vermessung & Geoinformation*, vol. 1, pp. 21–26.
- [86] U.S. Geological Survey, (1996), "GTOPO30 Documentation", <http://lta.cr.usgs.gov/GTOPO30>.
- [87] U.S. Geological Survey, (2015), "Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global", <https://lta.cr.usgs.gov/SRTM1Arc>.
- [88] Vapnik, N. V., (1998), "Statistical Learning Theory", John Wiley & Sons, New York.
- [89] Varga, M. and Tomislav Bašić, (2015), "Accuracy validation and comparison of global digital elevation models over Croatia", *International Journal of Remote Sensing*, vol. 36, no. 1, pp. 170–189, doi: 10.1080/01431161.2014.997420.
- [90] Wang, J., Y. Hu, and Zhou, J., (2009), "Combining model for regional GPS height conversion based on least squares support vector machines", *Proceedings – 2009 International Conference on Environmental Science and Information Application Technology*, ESIAT 2009 vol. 2, no. 2, pp. 639–641. doi: 10.1109/ESIAT.2009.182.
- [91] Wechsler, S. P., (2003), "Perceptions of Digital Elevation Model uncertainty by DEM users." *URISA Journal*, vol. 15, no. 2, pp. 57–64.
- [92] Yastikli, N., Koçak, G., and Büyüksalih, G., (2006), "Accuracy and morphological analysis of GTOPO30 and SRTM X-C band DEMs in the test area Istanbul", *ISPRS Topographic mapping from space with special emphasis on small satellite*, Vol. XXXVI-1/W41, 2006, ISPRS Congress, 14–16 Feb. 2006, Ankara, Turkey, 6th session.
- [93] Ye, J. and Xiong, T., (2007), "SVM versus least squares SVM", *Proc. 7th artificial intelligence and statistics*, 21–24 March 2007, San Juan, Puerto Rico, Vol. 2, pp. 644–651.
- [94] Zaletnyik, P., Völguesi, L., and Paláncz, B., (2008), "Modelling local GPS/levelling geoid undulations using Support Vector Machine", *Periodica Polytechnica, Civil Engineering*, vol. 52, no. 1, pp. 39–43, web: <http://www.pp.bme.hu/ci>.
- [95] Zaletnyik, P., Völguesi, L., Kirchner, I., and Paláncz, B., (2007), "Combination of GPS/Leveling and gravimetric geoid by using the thin plate spline interpolation technique via finite element method", *Journal of Applied Geodesy*, vol. 1, no. 2007, pp. 233–239 doi: 10.1515/JAG.2007.025.
- [96] Závoti, J. and Kalmár, J., (2016), "A comparison of different solutions of the Bursa–Wolf model and of the 3D, 7-parameter datum transformation", *Acta Geodaetica et Geophysica*, vol. 51, no. 2, pp. 245–256.