Establishment of a corrective geoid surface by spline approximation of Iranian GNSS/levelling network

Hosseini-Asl, M.; Amiri-Simkooei, A. R.; Safari, A.

DOI
10.1016/j.measurement.2022.111341

Publication date
2022

Document Version
Final published version

Published in
Measurement: Journal of the International Measurement Confederation

Measurement 197 (2022) 111341

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Establishment of a corrective geoid surface by spline approximation of Iranian GNSS/levelling network

M. Hosseini-Asl [a,*], A.R. Amiri-Simkooei [b,c], A. Safari [a]

[a] School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Iran
[b] Department of Geoscience and Remote Sensing, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Netherlands
[c] Department of Geomatics Engineering, Faculty of Civil Engineering and Transportation, University of Isfahan, 81746-73441 Isfahan, Iran

ARTICLE INFO

Keywords:
Geoid height approximation
GNSS/leveling network
Height reference
Systematic bias
Corrector surface
B-spline

ABSTRACT

The performance of the gravimetric geoid models is usually evaluated by comparison of geoid heights with the GNSS/levelling derived geoid. But the GNSS/levelling network can be infected by significant systematic biases and random errors, especially in large and uneven areas. This contribution addresses the challenging problem of the corrector surface development along with the elimination of biases. To this end, fitting an appropriate geometric surface to the GNSS/levelling geoid heights is required, which is accomplished by applying the least squares B-spline approximation theory to the GNSS/levelling data. In addition, the 3D affine transformation is used to detect systematic effects of the GNSS/levelling network compared to a global geoid model. This strategy is applied to the adjustment of the Iranian GNSS/levelling network. A significant tilt is observed across the country ranging from −0.35 m to 0.04 m. The entire study area is divided into four zones and the corrector surfaces are obtained in each zone. The accuracy of three global geoid models, EGM2008, EIGEN-6C4 and SGG-UGM-1, along with that of the Iranian regional geoid model IRG2016, are then investigated based on the raw GNSS/levelling heights and the corrector surfaces. The evaluations show the reliable results regarding the corrector surfaces against the raw data set. All geoid models show their maximum RMSE values of discrepancies in the mountainous zones and their minimum RMSEs in the zones having lower variety in topography. This indicates that the irregular topographies are not well detected by the global models over the study area.

1. Introduction

The accuracy of the gravimetric geoid models is usually assessed by comparing them with the GNSS/levelling network. It is therefore assumed that the geometric geoid derived as $N = h - H$ provides the reference surface, where h is the ellipsoidal height measured by the Global Navigation Satellite System (GNSS) receivers and H is the orthometric height measured by levelling methods related to the geoid [1]. But this simple mathematical expression cannot be easily established in practice and there might be significant systematic biases and random errors in the GNSS/levelling network.

The GNSS/levelling networks offer the short and ultra-short wavelength components of geoid, but there may be a systematic bias and a noticeable tilt across the study area due to the levelling network. One way to detect this systematic effect is to compare them with the global geoid models. This is often the case in larger regions such as the accumulated systematic effects in the levelling network of the North American Vertical Datum of 1988, NAVD88 in the US [2]. The biases can also occur due to the theoretical approximations, datum inconsistencies inherent among the height types, incorrect height corrections and changes in station coordinates over time [3]. The elimination of the systematic effects can lead to a more reliable reference surface. On the other hand the observational errors, the uneven distribution of the GNSS/levelling control points, the lack of simultaneous observations, land subsidence and earthquakes can reduce the GNSS/levelling accuracy. Therefore, the development of a corrector surface has become of great importance in dealing with this noisy data. We need an appropriate functional model to establish such an optimal corrector surface. Moreover, the unknown heights of new points can be computed using this corrector surface. On the other hand, the accurate GNSS/levelling network, enables the users to replace the traditional height determination techniques by faster and more cost-effective technique as visibility between stations are not required, and it can be operated in all-weather conditions [4]. It therefore plays an essential role in the geodetic

* Corresponding author.
E-mail addresses: mahin.hosseiniasl@ut.ac.ir (M. Hosseini-Asl), a.amirisimkooei@tudelft.nl (A.R. Amiri-Simkooei), asafari@ut.ac.ir (A. Safari).

https://doi.org/10.1016/j.measurement.2022.111341
Received 20 October 2021; Received in revised form 30 April 2022; Accepted 9 May 2022
Available online 13 May 2022

infrastructure where geoid-related heights are required in geodesy, geophysics, oceanography and engineering.

Many studies have used the polynomial surface fitting problem to model the geoid heights. But a unique method cannot be established to model the geoid surfaces in different areas. The degree of the polynomial depends on the size of the study area and the variation of the geoid heights. A plane or a low order polynomial is usually applied to model the normal variations of geoid surface in a small area of interest. But different classifications of polynomials such as bi-quadratic, bi-cubic, bi-quartic, and bi-quintic surfaces are applied for relatively large or large area [5]. Fotopoulos (2005) applied various polynomial functions to adjustment of geoid models and showed that the type of functional model affects the estimated values of standard deviations [6]. Khazraei et al. (2017) used the bi-linear and the bi-quadratic polynomials to approximate and analyze the accuracy of the geoid models in a small area (~100 km^2). The results prove that the bi-quadratic polynomial was the best model fitted in this study [7]. Das et al. (2017) developed the geoid surface using second, third and fourth degree polynomials [8]. The study shows that the third degree polynomial provided the best accuracy for the corrective surface in a small size (~25 km^2) and nearly flat area. In an attempt to improve the accuracy of the corrector surface, Eteje et al. (2019) applied the five degree, bi-quintic polynomial surface over the study area [5]. Erol and Erol (2021) also applied the finite elements based bivariate (BIVAR) interpolation method in local geoid modeling, in an area of approximately 44 × 77 km^2 in the west of Turkey. The BIVAR algorithm divides the area into triangles, and expresses each sub-element (finite elements cells consisting of three data points) with an individual polynomial model, and employs some continuity and differentiability conditions along the boundaries of each geometrical element [9].

Hosseini-Asl et al. (2021) employed the least squares 2D bi-cubic spline approximation (LS-BICSA) theory to the combined adjustment of the high-degree global gravitational model, European Improved Gravity model of the Earth by New techniques (EIGEN-6C4) and the Iranian regional geoid model IRG2016 in a large-scale and complex terrain study area [10]. The LS-BICSA method was proposed by Amiri-Simkooei et al. (2018) to estimate a 2D smooth surface of an irregularly distributed data set [11]. The region must be divided into some patches and the patches must be smoothly connected together by some continuity and differentiability conditions at their boundaries. We apply such a strategy to the adjustment of the Iranian geometric geoid derived from GNSS/levelling network. The method can alternatively be formulated based on the B-spline method, addressed for example by De Boor (1962) and Hayes and Halliday (1974) [12,13]. B-spline method formulates a spline as a linear combination of the piecewise basis functions [14]. Therefore, for the cubic splines, the first and second derivatives are automatically continuous along and across the borders.

The Iranian GNSS/levelling network consists of 1288 stations over the large-scale area (~1632000 km^2) with a variety of different characteristics of topography. The GNSS/levelling network accuracy is not completely known in Iran, especially for the levelling observations. Therefore, detailed study of this network is essential in terms of systematic and random errors. First the existence of possible biases in the network is discussed and then a corrector surface is obtained using the B-spline model. The entire study area is then divided into four zones and the accuracy of the data is obtained in each zone. The results are expected to vary from complex terrain zones to the other zones. After removing the biases and obtaining the corrector surfaces, the accuracy of three global geopotential models is evaluated based on the corrector surface in each zone. Three high-degree global geoid models, namely EIGEN-6C4, Earth Gravitational Model 2008 (EGM2008) and Satellite Gravity Gradiometry (SGG-UGM-1) are selected to this evaluation. In addition to the global models, the Iranian regional geoid IRG2016 has also been used in this evaluation.

The remaining parts of this paper are organized as follows. In the following section, the functional model formulation based on B-spline is explained. We apply the method to adjustments of the geometric geoid derived from GNSS/levelling network covering Iran. The global geoid models, the regional geoid model and GNSS/levelling network of Iran are then introduced in a subsequent section. The systematic effects in the GNSS/levelling network are also investigated. We develop the geometric surface by employing B-spline model and thereafter analyze the accuracy of the model in the study area. The global and regional geoid models are evaluated based on the corrector surface. Finally, some conclusions are made in the last section.

2. Functional model of B-spline

To approximate the 2D function values using a smooth spline surface, B-spline can be introduced as a reliable method. The method can formulate a spline function as a linear combination of basis functions, called B-spline [12,14]. B-spline is one of the most popular methods due to its universality, local control and optimal continuity [15]. Consider a pth-degree B-spline curve $f(u) = \sum_{i=1}^{h} \gamma_i M_{i,p}(u)$ with unknown control points γ_i for a sequence of given data points. B-spline basis function $M_{i,p}(u)$ is a piecewise function of pth-degree polynomials on the knot vector. The knot vector includes the number and placement of joining points of subsequent piecewises,$\lambda_1, \lambda_2, \ldots, \lambda_h$. If we have a set of basis functions in u-axis with respect to the knots $\lambda_1, \lambda_2, \ldots, \lambda_h$ and a set of basis functions in v-axis with respect to the knots$\mu_1, \mu_2, \ldots, \mu_k$, the 2D B-spline is then represented in the form [13].

$$f(u,v) = \sum_{i=1}^{h} \sum_{j=1}^{k} \gamma_{i,j} M_{i,p}(u) N_{j,q}(v) \tag{1}$$

where $M_{i,p}(u)$ and $N_{j,q}(v)$ are the B-spline basis functions of pth and qth-degree piecewise polynomials along u and v axes, respectively. Eq. (1) can be considered as the following representation of the functional model.

$$E(y) = Ax\ ;\ D(y) = Q_y \tag{2}$$

where E and D denote the mathematical expectation and dispersion operators, respectively. With the m-dimensional vector of observations, the design matrix A has m rows and $h \times k$ columns, and Q_y is the $m \times m$ covariance matrix of observations. The least squares estimate $\hat{x}$ of the unknown parameters is.

$$\hat{x} = (A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1} y \tag{3}$$

and the least squares estimate of the observations and residuals follows as $\hat{y} = A\hat{x}$ and $\hat{e} = P_A^{\perp} y$ respectively. The matrix $P_A^{\perp} = I_m - A(A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1}$ is an orthogonal projector, with I_m an identity matrix of size m. The knot placement, choice of control points and the best parameterization of the data points play obviously a key role in this approximation, see [15,16].

There are also alternative functional models for B-spline. Zangeneh-Nejad et al. (2017) introduced the least squares cubic spline approximation (LS-CSA) to approximate the 1D data in the least squares sense [17]. The functional model is a piecewise cubic curve fitted to a number of consecutive data under some continuity conditions. Amiri-Simkooei et al. (2018) proposed the least squares 2D bi-cubic spline approximation (LS-BICSA) method to estimate a smooth surface fitted to 2D data set [11]. LS-BICSA is formulated by the pure bi-cubic spline functions, but in the least squares sense and the theory of constrained least squares is used to impose the continuity conditions at the boundaries. When considering LS-BICSA and B-spline methods, it follows that B-spline automatically applies all of the continuity constraints, whereas LS-BICSA imposes only the user-specified constraints, along and across the borders. Here we consider all constraints up to and including the second order, and hence LS-BICSA and B-spline provide identical results.

Table 1
Global geopotential models.

Models	Year	Deg.	Data Sources
EGM2008	2008	2190	GRACE, Gravity, Altimetry [20]
EIGEN-6C4	2014	2190	GRACE, GOCE, LAGEOS, Gravity, Altimetry [21]
SGG-UGM-1	2018	2159	EGM2008, GOCE [22]

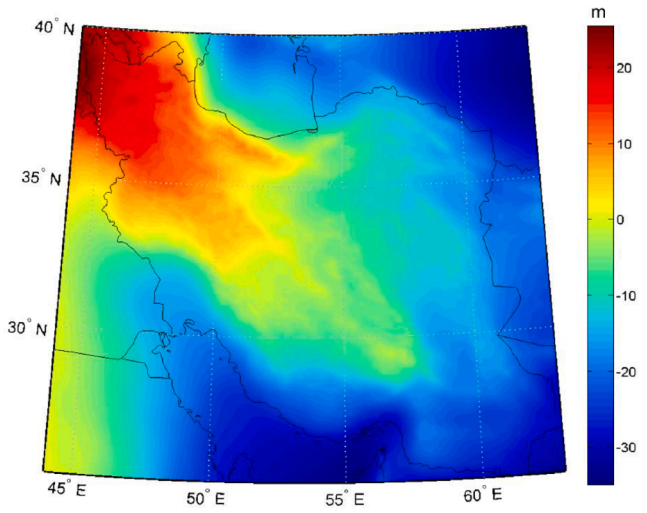

Fig. 1. EIGEN-6C4 geoid heights over Iran.

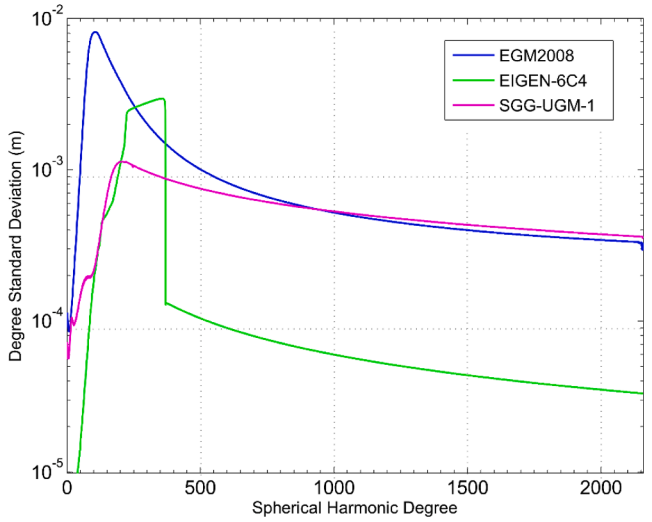

Fig. 2. Degree standard deviations of EGM2008, EIGEN-6C4 and SGG-UGM-1.

3. Data description

3.1. Global geopotential model

The Global Geopotential Models (GGMs) are composed of the spherical harmonic coefficients, derived from the satellite gravity missions, Challenging Mini-satellite Payload (CHAMP), Gravity Recovery and Climate Experiment (GRACE) and Gravity and steady-state Ocean Circulation Explorer (GOCE) and terrestrial gravity and satellite altimeter data. Therefore, the geoid height in terms of the spherical harmonics can be represented at any point of interest (φ, λ) as follows [18].

$$N(\varphi,\lambda) = N_0 + \frac{GM}{R\gamma}\sum_{n=2}^{N_{\max}}\left(\frac{R}{r}\right)^{n+1}\sum_{m=0}^{n}\left[\overline{C}_{nm}\cos m\lambda + \overline{S}_{nm}\sin m\lambda\right]\overline{P}_{nm}(\sin\varphi) \quad (4)$$

where GM is the geocentric gravitational constant, R is the reference radius, γ is the mean gravity of the reference ellipsoid. $\overline{C}_{nm}$ and $\overline{S}_{nm}$ are the fully normalized spherical geopotential coefficients of degree n and order m. $\overline{P}_{nm}(\sin\varphi)$ are the fully normalized associated Legendre functions, r is the geocentric radius of the computation point and $N_{\max}$ is the maximum degree of the model. The spherical harmonic coefficients are online available on https://icgem.gfz-potsdam.de, operated by GFZ Potsdam [19]. The zero-degree term of geoid heightN_0, as an additive constant value, can also be computed based on the mass and potential differences between the model and the reference ellipsoid. Providing the standard deviation of the spherical harmonic coefficients by GGMs, the error degree variances can then be evaluated by the following expression [20].

$$\sigma_n^2 = \left(\frac{GM}{R\gamma}\right)^2\sum_{m=0}^{n}\left[\left(\sigma_{\overline{C}_{nm}}\right)^2 + \left(\sigma_{\overline{S}_{nm}}\right)^2\right] \quad (5)$$

where $\sigma_{\overline{C}_{nm}}$ and $\sigma_{\overline{S}_{nm}}$ are the standard deviation of the spherical harmonic coefficients $\overline{C}_{nm}$ and$\overline{S}_{nm}$, respectively.

Three extra-high degree spherical harmonic models, EGM2008, EIGEN-6C4 and the recent high-degree GGM model, SGG-UGM-1 are selected in this study. The brief description of these global models is provided in Table 1. Fig. 1 shows the study area along with EIGEN-6C4 geoid heights computed in terms of a resolution 5-arcmin grid, covering the study area, with reference to the WGS84 ellipsoid.

The error degree variances in geoid heights of EGM2008, EIGEN-6C4 and SGG-UGM-1 up to degree 2159 is computed from Eq. (5), and shown in Fig. 2. The differences among the error degree variances are significant. The values for EIGEN-6C4 are lower than EGM2008 and SGG-UGM-1, especially in higher degrees. This makes sense because EIGEN-6C4 is computed from more data sources than the other models, according to Table 1. Not only GRACE but also the GOCE and Laser Geodynamics Satellite (LAGEOS) data along with more terrestrial gravity data were integrated to develop the EIGEN-6C4. On the other hand, the error degree variances of EGM2008 and SGG-UGM-1 are very close to each other in higher degrees, but lower for SGG-UGM-1 model up to the degree 900. The reason is probably due to the use of the GOCE data in the SGG-UGM-1 model.

3.2. GNSS/levelling network

The global and regional geoid models are usually evaluated based on the geometric geoid derived from the GNSS/levelling network. Using precise levelling networks, the orthometric height of the origin station is spread throughout the country. The GNSS/levelling network of Iran consists of 1288 stations where their orthometric heights are connected to the Iranian first order precise levelling network. This network is tied to the reference benchmark DN-G1001, located in the southern part of the country. DN-G1001 has been connected to the zero point, which is a reference tide-gauge station. The levelling network provides a reference surface with respect to the Mean Sea Level (MSL) and not to the geoid [23]. MSL is not an equipotential surface because of the existence of the Sea Surface Topography (SST) due to the diverse dynamic phenomena in oceans [24]. The Iranian tide gauge data have not been corrected for an SST model. Thus orthometric heights derived by this levelling network cannot be used to assess the discrepancies among the geoid models. But a few studies have defined the Iranian height datum offset with respect to the geoid at the reference benchmark [25,26]. Ebadi et al. (2019) showed the offset of −0.254 m compared to a geoid with the potential value of $W_0 = 62636853.4\ m^2/s^2$ [23].

The precise geodetic height and horizontal position of the stations was measured by the dual-frequency GNSS receivers and the observations were transformed to the Iranian Geodetic Datum IRGD2010. IRGD2010 is the original (or Doppler) version of WGS84 coordinate system. Therefore, it is necessary to transform the ellipsoidal

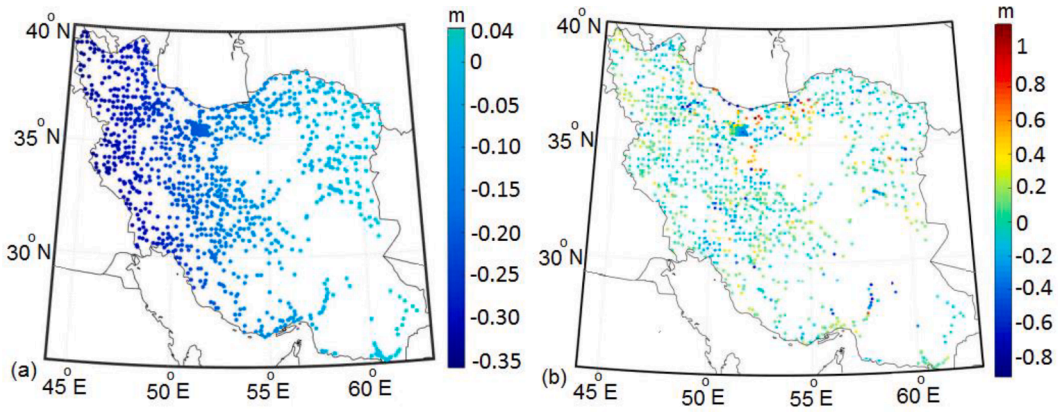

Fig. 3. The modeled systematic effects of GNSS/levelling network with respect to EIGEN-6C4 (a). The difference in geoid heights between EIGEN-6C4 and GNSS/levelling network (b).

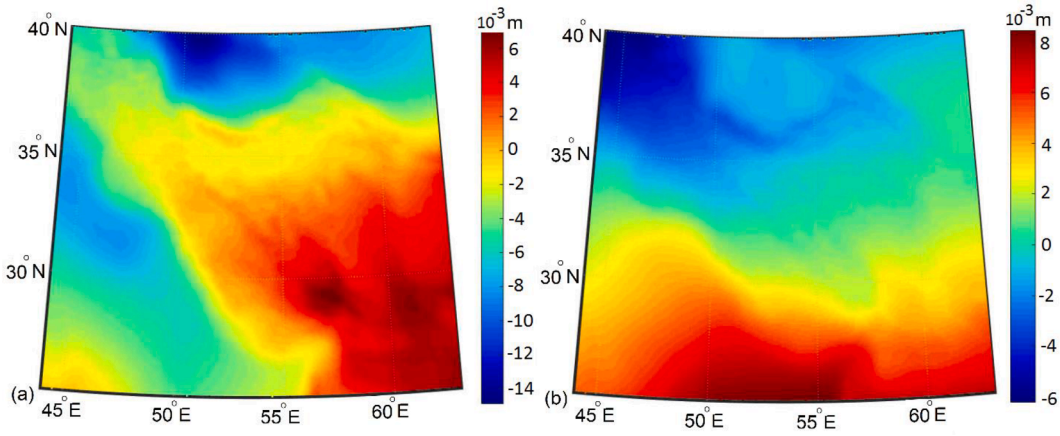

Fig. 4. The modeled systematic effects of (a) EGM2008, and (b) SGG-UGM-1 with respect to EIGEN-6C4.

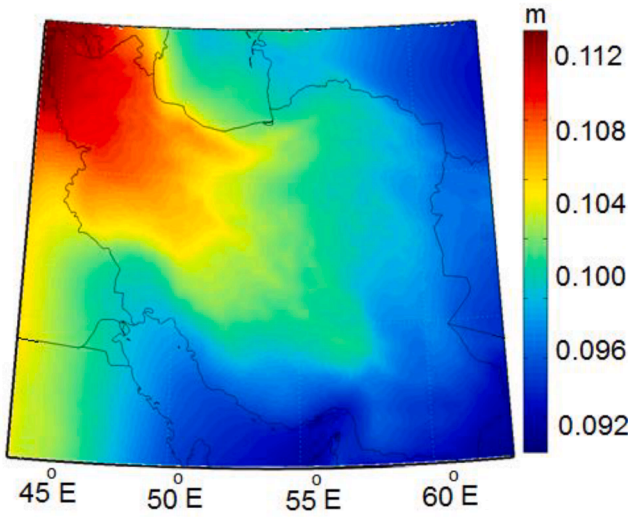

Fig. 5. The modeled systematic effects of IRG2016 with respect to EIGEN-6C4.

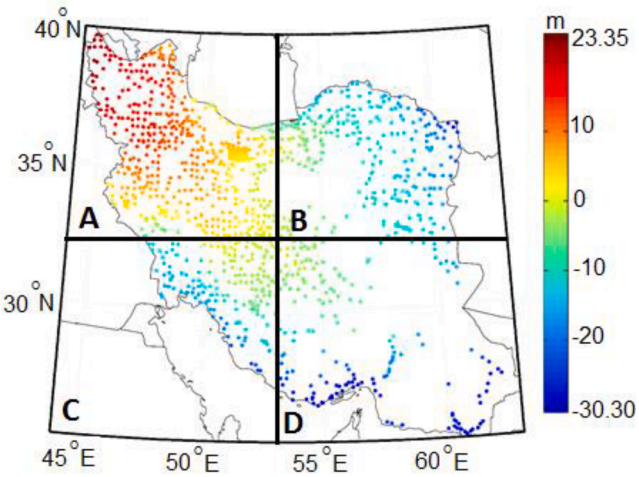

Fig. 6. Distribution of 1288 GNSS/levelling control points over the divided zones.

coordinates from IRGD2010 to the realization of WGS84, used in the recent global geoid models.

The levelling network can contain a significant systematic bias and a noticeable tilt across the country. This is often the case in larger regions such as the accumulated systematic effects in the levelling network related to the North American Vertical Datum of 1988, NAVD88 in the US. This systematic effect can be detected from geoid models derived from satellite means. Detailed error analysis has been performed on NAVD88 by Li (2018) and showed a significant diagonal tilt in NAVD88 ranging from −0.20 m to 1.30 m [2]. The simplest approach to model the systematic effects is to use the regression analysis to the difference

Table 2
Estimated standard deviation of corrector surfaces based on 2D polynomials.

Zone	Bi-linear		Bi-quadratic		Bi-cubic	
	Max err. [m]	STD [m]	Max err. [m]	STD [m]	Max err. [m]	STD [m]
A	−9.8	2.7	−6.5	1.5	−5.1	1.3
B	−6.6	1.8	−4.5	1.2	3.2	0.9
C	−7.4	2.7	5.1	1.8	−4.3	1.0
D	10.5	3.1	7.1	2.2	6.7	1.5

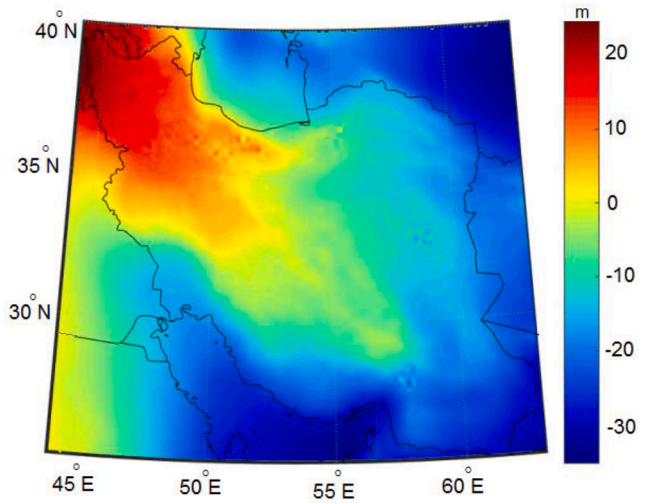

Fig. 7. The estimated surface over the study area.

Table 3
Maximum error and estimated standard deviation of B-spline corrector surface, before and after data snooping.

Zone	No. of observations	Before data snooping		No. of blunders	After data snooping	
		Max err. [m]	STD [m]		Max err. [m]	STD [m]
A	647	−0.902	0.254	4	−0.645	0.246
B	255	−0.962	0.312	2	0.597	0.239
C	208	0.227	0.150	0	0.227	0.150
D	178	−0.377	0.165	0	−0.377	0.165

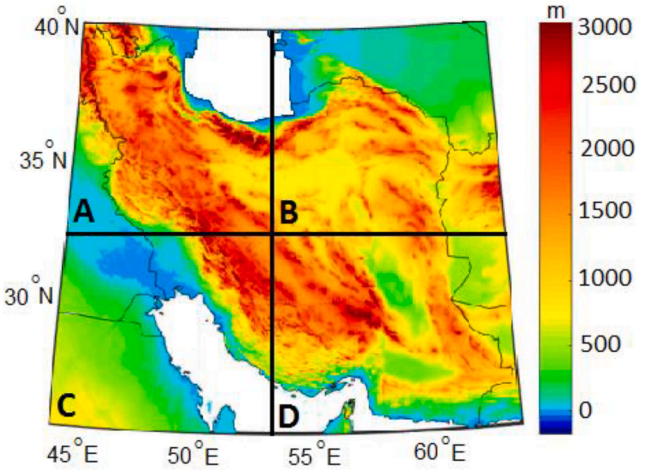

Fig. 8. The digital elevation model generated from SRTM 90 m over Iran.

between the heights derived from a referenced geoid model and the GNSS/levelling data [2,27]. We have used the 3D affine transformation to detect the systematic effects of the GNSS/levelling network compared to the EIGEN-6C4. If we have the coordinates of the points in two different systems, any combination of translation, rotation, scaling and shears can be obtained through 12 unknown coefficients by the 3D affine transformation as.

$$\begin{aligned} x' &= a_{xx}x + a_{xy}y + a_{xz}z + b_x \\ y' &= a_{yx}x + a_{yy}y + a_{yz}z + b_y \\ z' &= a_{zx}x + a_{zy}y + a_{zz}z + b_z \end{aligned} \tag{6}$$

The coefficients a's were estimated for 1288 GNSS/levelling control points compared to their corresponding points in EIGEN-6C4, and the computed errors are shown in Fig. 3a. The results indicate a significant tilt in the Iranian GNSS/levelling network ranging from 4 cm (south-east) to −35 cm (north-west). This is somehow the dominant direction for which the topography changes from plain and desert regions to the mountainous area. This is likely due to the different unmodelled error sources in the regions with high topography changes. The Iranian height datum is only tied to a single tide-gauge station DN-G1001. Increasing the errors from South to North can also indicate the accumulated errors in the levelling network. These biases can also occur for a variety of reasons, such as theoretical approximations made in processing observed data and network adjustments, approximate or inexact normal or orthometric height corrections, and instability of reference station monuments over time, etc [3]. The first order levelling network was measured for the second time from 2001 to 2009. Therefore, changes in station coordinates over the past years, due to tectonic movements of the crust, earthquakes and subsidence phenomenon can reduce the assessment accuracy. Fig. 3b also shows the residuals of GNSS/levelling geoid height from EIGEN-6C4 after the elimination of the systematic biases.

We have used the 3D affine transformation to detect the possible systematic effect of the global model, EGM2008, compared to the EIGEN-6C4 (see Fig. 4a). The results indicate an error ranging from −15 mm to 7 mm. Fig. 4b also shows the modeled systematic effects of SGG-UGM-1with respect to EIGEN-6C4. The results indicate an error ranging from −6 mm to 8 mm. These do not seem to be significant for both models.

3.3. Regional geopotential model

The Iranian regional geoid model, IRG2016, uses the data sets of 21,525 gravity data provided by National Cartographic Center (NCC) of Iran and is based on the radial basis functions (RBFs) [28]. The residual gravity disturbances were computed by subtracting the EIGEN-6C4 model up to and including degree 360 and applied to determine the unknown RBF parameters by the stabilized orthogonal matching pursuit (SOMP) algorithm. The model was fitted to 1288 GNSS/levelling control points by applying the six parameter polynomial surface. The results show an RMSE value of approximately 0.23 m for the difference in geoid height. The 2.5-arcmin gridded model is available at the International Service for the Geoid (ISG) website, while the model interpolation can be performed by https://irg2016.ncc.gov.ir/.

We have also used the 3D affine transformation to detect the possible systematic effect of the regional model, IRG2016, compared to the EIGEN-6C4 (see Fig. 5). The results indicate an error ranging from 9 cm to 11 cm, indicating mostly a shift.

4. Numerical results and discussions

The aim of this study is to develop a geometric geoid model based on the B-spline least-squares adjustment of the functional model, and thereafter to analyse the accuracy of the model. The study area is bounded by 25° N to 40° N in latitude and 44° E to 63° E in longitude. The Iranian GNSS/levelling network consists of 1288 stations over this

Table 4
Statistics of geoid height differences at GNSS/levelling control points. Units in meters.

Zone	Geoid model	Raw GNSS/levelling data				Approximated GNSS/levelling			
		Max	Mean	Min	RMS	Max	Mean	Min	RMS
A	EGM	1.603	0.086	−1.082	0.288	1.629	0	−1.155	0.274
	EIGEN	1.220	0.038	−0.889	0.252	1.296	0	−0.825	0.245
	SGG	1.316	0.048	−0.875	0.258	1.366	0	−0.872	0.252
	IRG2016	2.235	0.299	−2.570	0.708	1.245	0	−0.834	0.236
B	EGM	0.891	−0.074	−1.093	0.325	0.786	0	−0.992	0.263
	EIGEN	0.818	−0.074	−0.750	0.280	0.718	0	−0.683	0.215
	SGG	0.752	−0.074	−0.758	0.283	0.631	0	−0.965	0.219
	IRG2016	1.300	−0.217	−2.221	0.601	0.712	0	−0.676	0.210
C	EGM	1.131	0.032	−0.761	0.229	1.143	0	−0.680	0.215
	EIGEN	0.616	0.055	−0.735	0.179	0.593	0	−0.559	0.166
	SGG	0.730	0.056	−0.726	0.183	0.728	0	−0.569	0.167
	IRG2016	1.729	0.603	−0.088	0.693	0.569	0	−0.602	0.160
D	EGM	0.563	−0.150	−1.182	0.333	0.465	0	−0.927	0.244
	EIGEN	0.601	−0.139	−0.950	0.295	0.768	0	−0.717	0.207
	SGG	0.568	−0.131	−1.042	0.287	0.671	0	−0.817	0.210
	IRG2016	1.638	0.293	−0.830	0.554	0.723	0	−0.711	0.200

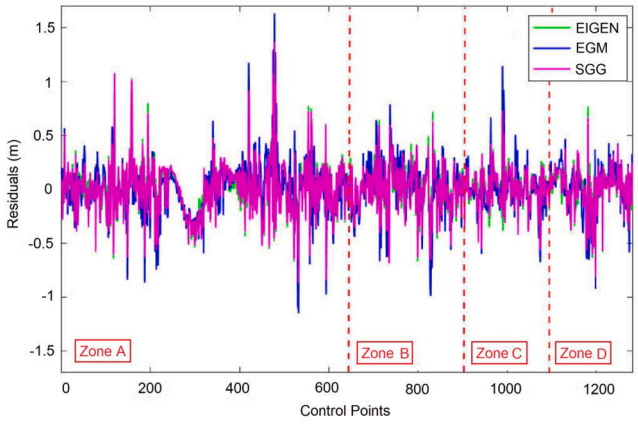

Fig. 9. Global geoid height residuals at GNSS/levelling control points in four zones.

large-scale area. The inhomogeneous distribution of the GNSS/levelling control points along with their geoid heights after removing the biases are shown in Fig. 6. As can be seen, the GNSS/levelling network has the large gaps located in the south-east of the country and the Lut and Kavir central desert of Iran. On the other hand, Iran is characterized by complex terrain. Therefore, for a more detailed study of the corrector surface, the region is divided into four zones, A, B, C and D (see Fig. 6). Zone A contains the highest data density with 647 control points. The zones B, C and D contain 255, 208 and 178 data, respectively.

The efficiency of the proposed B-spline functional model was evaluated in comparison with a few 2D polynomials commonly used in similar studies. Solving for the 2D polynomials, the estimated surfaces are obtained in each zone and the accuracy of the corrector surfaces are obtained. Each zone is separately approximated by the first order bi-linear, second-order bi-quadratic and third-order bi-cubic polynomials with four, nine and 16 unknown coefficients, respectively. The difference between the actual and approximated function values (error values) has been obtained in each zone. Table 2 presents the maximum error value and the estimated standard deviation for each approximation method over the study area. The results show the high amounts of errors and standard deviations in the range of 1 to 3 m in these zones. This indicates that the commonly used 2D polynomials cannot appropriately approximate such a large region.

Solving for the proposed B-spline functional model, the estimated geometric surface is obtained in each zone. The splines are considered with 30-arcmin patch sizes over the zones. It is also essential to eliminate the possible blunders that can cause large errors in the final results. We set the criterion of 3 times of the standard deviation to identify the mistake data, corresponding to the 99.7% confidence interval. GNSS/levelling points that have residuals larger than 3σ are defined and eliminated from the final adjustment. The lack of GNSS/levelling data in some patches causes the singularity in the functional model. The global geoid model, EIGEN-6C4 has been used to fill in empty patches and to eliminate such singularity. For this purpose, EIGEN-6C4 geoid heights computed in terms of a resolution 5-arcmin grid and a weight of 100 times smaller than its actual weight has been used. Fig. 7 shows the estimated surface over the study area.

Table 3 presents the maximum error value and the estimated standard deviation of B-spline method before and after data snooping. This table also shows the number of blunders detected in each zone. As indicated removing the blunders has affected the accuracy of the approximation.

The accuracy of the approximated surface differs from 0.15 m in zone C to 0.246 m in zone A. This indicates that the accuracy of the GNSS/levelling network is not the same throughout the study region. This can be due to uneven data density as well as uneven data accuracy over the study area. Zone A has the maximum error of −0.645 m and the standard deviation of 0.246 m. This zone has the maximum amount of error despite the highest data density. Zone B also has an accuracy close to the zone A with the maximum error and the standard deviation of 0.597 m and 0.239 m, respectively. This makes sense because these zones have the high variation of topography. Fig. 8 shows the digital elevation

Table 5
Correlation among different geoid models.

Zone	Correlation among the models					
	EGM2008 and EIGEN	EGM2008 and SGG	EIGEN and SGG	IRG2016 and EIGEN	IRG2016 and EGM2008	IRG2016 and SGG
A	0.876	0.897	0.989	0.994	0.871	0.983
B	0.784	0.858	0.982	0.996	0.779	0.977
C	0.744	0.861	0.971	0.989	0.734	0.961
D	0.817	0.886	0.982	0.996	0.814	0.979

model (DEM) generated from Shuttle Radar Topography Mission (SRTM) over Iran. The realization of a uniform network over the large regions with high mountains and also large deserts is highly challenging and expensive.

After removing the biases with 3D affine transformation according to Eq. (6) in Figs. 3 to 5, removing the blunders, and then obtaining the final corrector surfaces according to Table 3, the accuracy of the three global geopotential models, EGM2008, EIGEN-6C4 and SGG-UGM-1 and Iranian regional geoid model, IRG2016, is evaluated based on the approximated surface in each zone. This evaluation is performed by comparing the approximated GNSS/levelling geoid heights at 1282 stations with their corresponding values calculated from global geoid models. The statistics of this comparisons, maximum, minimum, mean and the root mean square error (RMSE) values are given in Table 4. Also the comparison of the geoid models with the raw GNSS/levelling data is summarized in this table. The geoid models are usually evaluated based on the geoid derived from the raw GNSS/levelling network. Such a network is expected to be a valid reference for evaluating the accuracy of the geoid models. The results show that the global geoid models have the maximum RMSE in zone D and the regional model IRG2016 has the maximum RMSE in zone A. But the results are quite different if the approximated GNSS/levelling network is used, a corrector surface with possible biases and gross errors removed. This reliable result shows the maximum RMSE value of each geoid model in zone A. This is in agreement with the results provided in Table 3, which provides the maximum standard deviation value of the corrector surface in zone A. On the other hand use of the global geoid models may not guarantee high accurate heights, especially in mountainous areas over Iran. The global geoid models usually use the terrestrial data to provide the short and medium wavelengths of geoid. Due to the lack of Iranian proprietary gravity and GNSS/levelling data in GGM computations, irregular topographies are not well detected by these models. All three global geoid models and the regional geoid model have also the minimum RMSE values in zone C where the topographic variations are relatively smooth compared to the other zones.

Comparing the accuracy of the three global models shows that EIGEN-6C4 and SGG-UGM-1 have a higher accuracy than the model EGM2008 in all zones. The reason is probably that EIGEN-6C4 and SGG-UGM-1 use more data sources, especially the GOCE data. The geoid models that use GOCE observations are sensitive to the short wavelengths of the gravity field due to the gravitational gradiometry [29]. The results also indicate that EIGEN-6C4 gives at least slightly better results than SGG-UGM-1 in terms of RMSE value in each zone. Therefore, EIGEN-6C4 and SGG-UGM-1 can be used in parallel over the Iranian territory. The regional geoid model IRG2016 also gives better results than the global geoid models in each zone. Fig. 9 shows the global geoid height differences at GNSS/levelling control points (residuals) in each zone. According to the figure, the residuals of all three models show a roughly similar behavior. Also Table 5 shows high correlation coefficients among the residuals. These high correlations may be seen again in Fig. 2 in which the behavior of degree variances of global models are very close to each other. EIGEN-6C4 and SGG-UGM-1 have a very strong correlation, more than 0.97 in all zones. On the other hand the model EGM2008 has the maximum correlation with SGG-UGM-1. The regional geoid model IRG2016 shows the maximum correlation with EIGEN-6C4 in each zone, because of the use of this global geoid model in its generation.

To further demonstrate the performance of the proposed method, the methodology was also applied to approximate and correct the GNSS/levelling network of Sweden. The results of this GNSS/levelling network approximation have a higher accuracy than the Iranian GNSS/levelling network. Further presentation and analysis of the results is beyond the scope of the present contribution, and can likely be considered in future studies.

5. Conclusions

The optimal approximation of the Iranian GNSS/levelling heights using the B-spline function along with the elimination of systematic effects using the 3D affine transformation of the GNSS/levelling network, compared to the EIGEN-6C4, can lead to a more reliable reference surface. We observed a significant systematic effect and tilt across the country ranging from −0.35 m to 0.04 m due to the Iranian levelling network. Then an adjustment procedure was required to eliminate the possible blunders and handle random errors. In this contribution, B-spline method was introduced as the functional model that can estimate a smooth 2D corrector surface of an irregularly distributed data set. Further, the region is divided into four zones due to the large and complex terrain study area. Therefore each zone was separately approximated by B-spline method and the accuracy of the corrector surface was obtained. Also the possible blunders were identified and eliminated from final adjustment of each zone.

The efficiency of the proposed functional model modelling was evaluated in comparison with a few commonly used 2D polynomials in similar studies. Solving for the 2D polynomials, the estimated surfaces were obtained in each zone and the accuracy of the corrector surfaces were obtained. Each zone was separately approximated by the first order bi-linear, the second-order bi-quadratic and the third-order bi-cubic polynomials with four, nine and 16 unknown coefficients, respectively. The results of Table 2 showed the high amounts of errors and standard deviations in the range of 1 to 3 m in each zone, indicating that the commonly used 2D polynomials cannot appropriately approximate such a large region. But the results of the B-spline approximation indicated that the accuracy of the corrector surface differs from 0.15 m in zone C to 0.25 m in the highly variable topographic zone A.

After obtaining the corrector surfaces, the accuracy of three global geopotential models, EGM2008, EIGEN-6C4 and SGG-UGM-1 and the Iranian regional geoid model IRG2016 was evaluated in each zone. This evaluation was obtained by comparing the approximated GNSS/levelling geoid heights with their corresponding values calculated from geoid models. The results showed the maximum RMSE value of each geoid model in mountainous zone A. All geoid models have also the minimum RMSE values in zone C where the topographic variations are relatively smooth compared to the other zones. Comparing the accuracy of the three models, EIGEN-6C4 and SGG-UGM-1 have a higher accuracy than the model EGM2008 in all zones. Also EIGEN-6C4 and SGG-UGM-1 have a very strong correlation (more than 0.97) within all zones. Therefore, EIGEN-6C4 and SGG-UGM-1 can be used in parallel, over the Iranian territory. The regional geoid model, IRG2016 also provides better results than the global geoid models in each zone. IRG2016 shows the maximum correlations of 0.99 with EIGEN-6C4 in each zone, because this global geoid model has been used to generate IRG2016.

In conclusion, the biases such as the theoretical approximations, datum inconsistencies incorrect height corrections, changes in station coordinates over time, systematic effects of levelling network and also the observational errors, the uneven distribution of the GNSS/levelling control points, the lack of simultaneous observations, land subsidence and earthquakes can lead to have a suboptimal reference geoid. Therefore, the proposed methods can deal with a large part of these errors to have a corrected reference surface, in large areas in general and having irregular topography in particular.

CRediT authorship contribution statement

M. Hosseini-Asl: Conceptualization, Methodology, Software, Validation, Writing – original draft, Investigation, Formal analysis. **A.R. Amiri-Simkooei:** Supervision, Conceptualization, Investigation, Formal analysis, Writing – review & editing. **A. Safari:** Visualization, Writing – review & editing, Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to acknowledge the National Cartographic Center of Iran for providing us with the GNSS/levelling control stations used in this research.

References

[1] A. Abdalla, M. Mustafa, Horizontal displacement of control points using GNSS differential positioning and network adjustment, Measurement 174 (2021) 108965, https://doi.org/10.1016/j.measurement.2021.108965.
[2] X. Li, Modeling the North American vertical datum of errors 1988 in the conterminous United States, J. Geod. Sci. 8 (2018) 1–13.
[3] G. Fotopoulos, C. Kotsakis, M.G. Sideris, How accurately can we determine orthometric height differences from GPS and geoid data? J. Surv. Eng. 129 (1) (2003) 1–10.
[4] J. Zhou, C. Luo, W. Jiang, X. Yu, P. Wang, Using UAVs and robotic total stations in determining height differences when crossing obstacles, Measurement 188 (2022) 110372, https://doi.org/10.1016/j.measurement.2021.110372.
[5] S.O. Eteje, O.F. Oduyebo, P.D. Oluyori, Relationship between Polynomial Geometric Surfaces Terms and Observation Points Numbers and Effect in the Accuracy of Geometric Geoid Models, International Journal of Environment, Agric. Biotechnol. (IJEAB) 4 (4) (2019) 1181–1194.
[6] G. Fotopoulos, Calibration of geoid error models via a combined adjustment of ellipsoidal, orthometric and gravimetric geoid height data, J. Geod. 79 (1-3) (2005) 111–123.
[7] S.M. Khazraei, V. Nafisi, A.R. Amiri-Simkooei, J. Asgari, Combination of GPS and Levelling Observations and Geoid Models Using Least-Squares Variance Component Estimation, J. Surv. Eng. 143 (2) (2017) 04016023.
[8] R.K. Das, S. Samanta, S.K. Jana, R. Rosa, Polynomial interpolation methods in development of local geoid model, Egypt. J. Remote Sens. Space. Sci. 21 (3) (2018) 265–271.
[9] S. Erol, B. Erol, A comparative assessment of different interpolation algorithms for prediction of GNSS/levelling geoid surface using scattered control data, Measurement 173 (2021) 108623.
[10] M. Hosseini-Asl, A.R. Amiri-Simkooei, A. Safari, Combination of regional and global geoid models at continental scale: Application to Iranian geoid, Ann. Geophys. 64 (4) (2021).
[11] A.R. Amiri-Simkooei, M. Hosseini-Asl, A. Safari, Least squares 2D bi-cubic spline approximation: Theory and Applications, Measurement 127 (2018) 366–378.
[12] C. De Boor, Bicubic spline interpolation, J. Math. Phys. 41 (1962) 212–218.
[13] J.G. Hayes, J. Halliday, The least-squares fitting of cubic spline surfaces to general data sets, J. Inst. Math 14 (1) (1974) 89–103.
[14] M.G. Cox, The numerical evaluation of B-splines, J. Appl. Math. 10 (1972) 134–149.
[15] Y. Zhang, J. Cao, Z. Chen, X. Li, X.M. Zeng, B-spline surface fitting with knot position optimization, Comput. Graph. 58 (2016) 73–83.
[16] J.-J. Fang, C.-L. Hung, An improved parameterization method for B-spline curve and surface interpolation, Comput. Aided Des. 45 (6) (2013) 1005–1028.
[17] F. Zangeneh-Nejad, A.R. Amiri-Simkooei, M.A. Sharifi, J. Asgari, Cycle slip detection and repair of undifferenced single-frequency GPS carrier phase observations, GPS Solut. 21 (4) (2017) 1593–1603.
[18] W.A. Heiskanen, H. Moritz, Physical Geodesy, Bull. Geodesique 86 (1) (1967) 491–492.
[19] International Centre for Global Earth Models (ICGEM). [On Line]. Available From: http://icgem.gfz-potsdam.de/tom_longtime (2022) (accessed 30 April 2022).
[20] N.K. Pavlis, S.A. Holmes, S.C. Kenyon, J.K. Factor, The development and evaluation of the Earth Gravitational Model 2008 (EGM2008), J. Geophys. Res 117 (2012) B04406.
[21] C.H. Förste, S.L. Bruinsma, O. Abrikosov, J.M. Lemoine, T. Schaller, H.J. Götze, J. Ebbing, J.C. Marty, F. Flechtner, G. Balmino, R. Biancale, EIGEN-6C4 The latest combined global gravity field model including GOCE data up to degree and order 2190 of GFZ Potsdam and GRGS Toulous, in: 5th GOCE User Workshop, Paris, 2014, pp. 25–28.
[22] W. Liang, X. Xu, J. Li, G. Zhu, The determination of an ultra-high gravity field model SGG-UGM-1 by combining EGM2008 gravity anomaly and GOCE observation data, Acta Geodaetica et Cartographica Sinica 47 (4) (2018) 425–434.
[23] A. Ebadi, A.A. Ardalan, R. Karimi, The Iranian height datum offset from the GBVP solution and spirit-levelling/gravimetry data, J. Geod. 93 (2019) 1207–1225.
[24] P.L. Woodworth, C.W. Hughes, R.J. Bingham, T. Gruber, Towards worldwide height system unification using ocean information, J. Geod. Sci. 2 (4) (2012) 302–318.
[25] A. Safari, M. Sharifi, I. Foroughi, H. Amin, An approach to height datum unification based on local gravity field modeling using radial base function case study: Height datum unification of levelling network of class 1 in Iran, J. Earth Space Phys. 40 (2014) 69–81.
[26] I. Foroughi, A. Safari, P. Novák, M.C. Santos, Application of radial basis functions for height datum unification, Geosciences 8 (2018) 369.
[27] M. Mosayebzadeh, A.A. Ardalan, R. Karimi, Regional improvement of global geopotential models using GPS/Levelling data, Stud. Geophys. Geod. 63 (2019) 169–190.
[28] A. Saadat, A. Safari, D. Needell, IRG2016: RBF-based regional geoid model of Iran, Studia Geophys. et Geod. 62 (3) (2018) 380–407.
[29] J. Bouman, M.J. Fuchs, GOCE gravity gradients versus global gravity field models, Geophys. J. Int. 189 (2) (2012) 846–850.