We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,100
Open access books available

149,000
International authors and editors

185M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

Sensitivity Analysis and Modeling for DEM Errors

Mohammed El-Diasty and Rifaat Abdalla

Abstract

The Digital Elevation Model (DEM) can be created using airborne Light Detection And Ranging (LIDAR), Image or Synthetic-Aperture Radar (SAR) mapping techniques. The direct georeferencing of the DEM model is conducted using a GPS/inertial navigation system. The airborne mapping system datasets are processed to create a DEM model. To develop an accurate DEM model, all errors should be considered in the processing step. In this research, the errors associated with DEM models are investigated and modeled using Principal Component Analysis (PCA) and the least squares method. The sensitivity analysis of the DEM errors is investigated using PCA to define the significant GPS/inertial navigation data components that are strongly correlated with DEM errors. Then, the least squares method is employed to create a functional relationship between the DEM errors and the significant GPS/inertial navigation data components. The DEM model errors associated with airborne mapping system datasets are investigated in this research. The results show that the combined PCA analysis and least squares method can be used as a powerful tool to compensate the DEM error due to the GPS/inertial navigation data with about 27% in average for DEM errors produced by the direct georeferenced airborne mapping system.

Keywords: sensitivity, PCA, least squares, DTM errors, navigation

1. Introduction

The Digital Elevation Model (DEM) can be created using airborne Light Detection And Ranging (LIDAR), Image or Synthetic-Aperture Radar (SAR) mapping techniques. The direct georeferencing of DEM model is conducted using GPS/inertial navigation system. The accuracy of the developed DEM model is strongly dependent on the mapping system and the georeferencing system grades. The selection of the mapping system and the georeferencing system grades is carried out in the planning stage based on the accuracy requirements of the required DEM model.

The previous literatures focused and heavily investigated the DEM model generation based on LIDAR, Images, and SAR data cleaning and filtering techniques such as Triangulated Irregular Network (TIN)-based filtering, slope-based filtering, mathematical morphological filtering, interpolation-based

filtering, and Machine-learning-based filtering. The original TIN-based filtering was developed based on the classical progressive TIN densification (PTD) and was implemented effectively in the commercial software TerraScan [1, 2]. Then, the revised PTD was investigated and reduced the total errors by about 8% when compared with classical PTD method [3]. Afterwards, a Parameter-Free PTD (PFPTD) algorithm was developed and outperforms the classical and revised PTD methods [4]. The original slope-based filtering was derived based on height differences in the training dataset [5, 6]. Then, adaptive slope-based filtering algorithm was developed to improve the accuracy in urban applications when compared with the original slope-based filtering algorithm [7]. The original mathematical morphological filtering was proposed to filter LIDAR data [8]. Then, the progressive morphological filtering algorithm was developed to improve the original method by applying threshold condition based on the elevation differences and proposed increasing gradually the filtering window size [9]. Afterward, the spline iteration method was introduced to improve the morphological filtering algorithm [10]. All mathematical morphological filtering methods outputs are strongly dependent on adapting the filtering window size. The original interpolation-based filtering method was proposed to deal with the steep areas [11]. Then, the interpolation-based DEM generation method was developed where one of the Inverse Distance Weighted, Kriging, and Natural Neighbor (NN) can be employed for DEM generation [12]. The Natural Neighbor (NN) method was proven to provide most efficient results. Finally, Machine-learning-based filtering was investigated where this method depends on topographic characteristics of the areas under investigation [13]. The deep convolutional neural network (CNN) was proposed for develop accurate DEM model [14]. The Machine learning method was optimized using windowing method to improve the DEM model generation [15].

However, the above-mentioned methods are mainly dependent on the data cleaning and filtering techniques of the heights to develop the DEM model. These methods are associated with errors that are not considered in the DEM modeling and could be corelated with system navigation data. Therefore, this research investigated the sensitivity analysis modeling of DEM errors that are potentially correlated with navigation data to improve the overall accuracy of the DEM model.

2. Combined PCA and least squares method

The sensitivity and modeling of DEM errors are investigated in this research using PCA and least squares function modeling. The system navigation data (position, velocities, attitudes, accelerations and dopplers) are considered the inputs to the model and the DEM height error is considered the desired output of the model.

PCA is a numerical technique used to study multidimensional processes that can be used to (1) reduce the dimensionality of a dataset and (2) identify relationships between the underlying variables of the process. PCA is based on eigen or singular value analysis of the process correlation or covariance matrix. The goal of PCA is to determine the minimum number of eigenvectors that best describe the key features of the process correlation matrix. This results in a reduced-dimensionality model for the matrix which can be used for data analysis, reduction, and model synthesis. Singular value decomposition (SVD) is fundamental to PCA. More details on PCA and SVD can also be found in Jolliffe [16].

2.1 PCA analysis

Let X denote an m x n matrix. For convenience, we assume $m \geq n$. The elements of the ith row of X form the n-dimensional vector $\mathbf{g}_i$. The elements of the jth column of X form the m-dimensional vector $\mathbf{a}_j$. The general singular value decomposition (gSVD) of X can be written as:

$$X = U \Sigma V^T \tag{1}$$

where U is an m x m matrix, Σ is an m x n matrix containing the singular values, and V^T is an n x n matrix. The columns of U are called the *left singular vectors*, $\{\mathbf{u}_k\}$, and form an orthonormal basis for the range space, so that $\mathbf{u}_i \cdot \mathbf{u}_j = 1$ for $i = j$, and $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ otherwise. The rows of V^T contain the elements of the *right singular vectors*, $\{\mathbf{v}_k\}$, and form an orthonormal. U and V are orthonormal so that their inverses exist and are their transposes. The matrix Σ can be decomposed as:

$$\Sigma = \begin{bmatrix} S \\ 0 \end{bmatrix} \tag{2}$$

where S is an m x n diagonal matrix in which only the diagonal elements are non-zero, $S = \text{diag}(s_1, ..., s_n)$ where the diagonal elements are zero. If the rank of X is r, $s_k > 0$ for $1 \leq k \leq r$, and $s_i = 0$ for $(r + 1) \leq k \leq n$. [For problems like the one we are interested in, noise generally ensures that $r = n$.] By convention, the ordering of the singular vectors is determined by high-to-low sorting of singular values, with the highest singular value in the upper left index of the S matrix. Note that for a square, symmetric matrix X, SVD is equivalent to eigenvalue decomposition. In PCA, the right singular vectors are frequently called the *components*. While the scaled left singular vectors $\{s_i \mathbf{u}_k\}$ are called the *scores*.

Note that U can be decomposed into two submatrices, an m x n matrix U_R and an m x m-n matrix U_N where $U = [U_R \ U_N]$. U_R defines the range space of U, while U_N defines the null space. Note that $X = U_R S V^T$ so that $XV = U_R S$. This provides the reduced form of SVD often used in PCA. In practice this is the form generally used; hence, we often drop the R subscript on U. **Figure 1** illustrates the various reduced-form matrices. Note that the right singular vectors span the space of the row vectors $\{\mathbf{g}_i\}$ and the left singular vectors span the space of the column vectors $\{\mathbf{a}_j\}$.

Several relationships can be derived. The SVD equation for $\mathbf{g}_i$ is:

$$\mathbf{g}_i = \sum_{k=1}^{r} u_{ik} s_k v_k \tag{3}$$

which is a linear combination of the right singular values $\{\mathbf{v}_k\}$. The ith row of U, $\mathbf{g'}_i$, contains the coordinates of the ith entry in the coordinate system (basis) of the scaled right singular values, $s_k \mathbf{v}_k$. If $r < n$ (or if we truncate the singular values to $r = l$), this computation requires fewer variables using $\mathbf{g'}_i$ rather than $\mathbf{g}_i$, thus reducing the dimension of the problem. Similarly, the SVD equation for $\mathbf{a}_j$ (the jth column of X) is:

$$\mathbf{a}_j = \sum_{k=1}^{r} v_{jk} s_k u_k \tag{4}$$

which is a linear combination of the left singular values $\{\mathbf{u}_k\}$. The jth column of V^T, $\mathbf{a'}_j$ (see **Figure 1**), contains the coordinates of the jth column of X in the coordinate
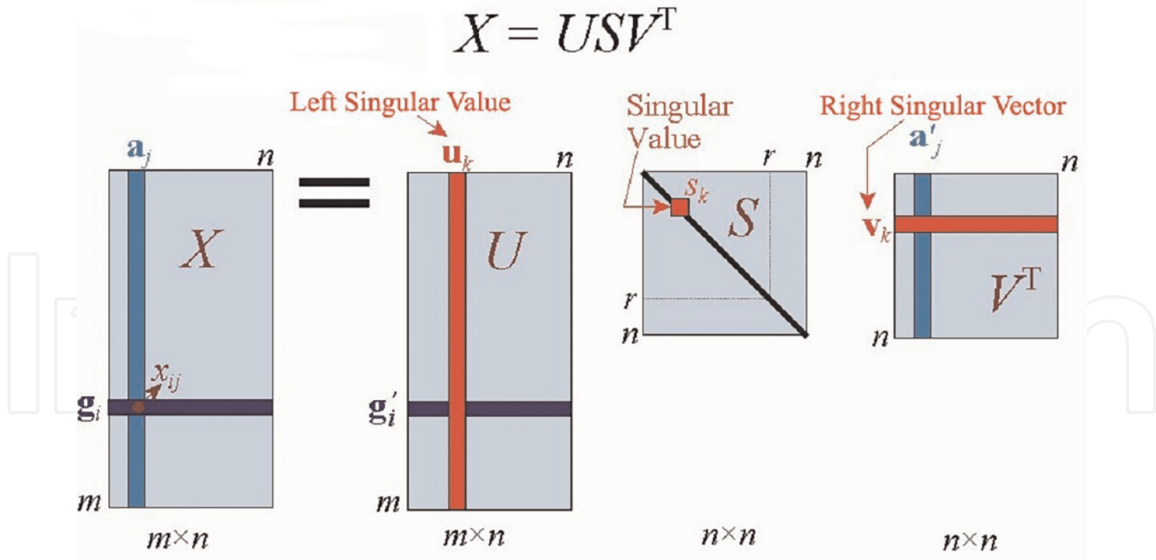
Figure 1.
Illustration of reduced SVD matrices. The right singular vectors have often termed the components while the left singular values are the scaled scores.

system (basis) of the scaled left singular vectors (the scores), $s_k\mathbf{u}_k$. By using the vector $\mathbf{a}'_j$, the analysis may be captured by $r \leq n$ variables, which is always fewer than the m elements in the vector $\mathbf{a}_j$, thus SVD reduces the number of variables required. Essentially, there are only r (which we can truncate to eliminate small singular values and further reduce the dimensionality) component vectors (the corresponding right singular vectors) which explain the behavior of X. The application of PCA often use the SVD property:

$$X^{(l)} = \sum_{k=1}^{l} \mathbf{u}_k s_k \mathbf{v}_k^T \tag{5}$$

where $X^{(l)}$ is the closest rank-l matrix to X, i.e., $X^{(l)}$ minimizes the sum of the squares of the difference between the elements of X and $X^{(l)}$, diff $= \sum_{ij}|x_{ij} - x^{(l)}_{ij}|^2$.

We can define the *covariance matrix* as $X^TX = \Sigma_i\mathbf{g}_i\mathbf{g}_i^T$. SVD analysis of X^TX yields V^T, which contains the principal components of $\{\mathbf{g}_i\}$, i.e. the right singular vectors $\{\mathbf{v}_k\}$ are the same as the principal components of $\{\mathbf{g}_i\}$. The eigenvalues of X^TX are equivalent to s_k^2, which are proportional to the variances of the principal components. The matrix $XX^T = \Sigma_j\mathbf{a}_j\mathbf{a}_j^T$ is proportional to the covariance matrix of the variables of $\mathbf{a}_j$. The left singular vectors $\{\mathbf{u}_k\}$ are the same as the principal components of $\{\mathbf{a}_j\}$. The s_k^2 are proportional to the variances of the principal components. The diagonal values of S (*i.e.*, s_k) are the "singular value spectrum". The value of a singular value is indicative of its importance in explaining the data. More specifically, the square of each singular value is proportional to the variance explained by each singular vector.

In PCA, X is defined and the SVD is computed. The singular values s_k are then plotted versus k. A reduced dimensionality approximation to X is computed by truncating the singular value series, i.e. by setting $s_k = 0$ for $k > K$ where K is the chosen threshold value, and using Eq. (3). Note that the squared singular values s^2_k are a

measure of the variability or "power" in the corresponding signal component specified by the corresponding to singular vector (like the frequency component in Fourier analysis).

2.2 Least squares functional modeling with PCA optimization

We ultimately seek a model that relates the inputs (navigation parameters) to the height error. The least squares method is used to model DEM errors. More details on the least squares method can also be found in Ghilani [17]. Assuming a linear forward model:

$$Xh = e \qquad (6)$$

The set of navigation data parameters is used to form the h vector. We will assume that the corresponding DEM height errors form the e vector. For each epoch, we form a single h and e vector. Since the model parameters vary with time, vectors are created by stacking the values as a function of time.

Using the matrix form, the X matrix in Eq. (6) represents the mapping between the model parameters and the height. This mapping is what we want to estimate from the DEM error training set. While there are several approaches, the following approach is attractive. Since Eq. (6) must apply for all realizations we create matrices H and E by combination of all the vectors of h and e, i.e., $H = [h_1|h_2| \ ... \ h_n]$ and similarly for E. We can then write:

$$XH = E \qquad (7)$$

Shifting the mean and scaling the vector elements affects the performance of the estimate. Assuming we have enough realizations, we can write a least-squares empirical estimate of X, X_e, as:

$$X_e = EH^T(HH^T)^{-1} \qquad (8)$$

While we can use X_e directly, SVD analysis of X_e provides us with a powerful tool to understand the relationship between H and E. The SVD analysis of X is:

$$X_e = U\Sigma V^T \qquad (9)$$

The singular value spectrum of Σ tells us the dimensionality (say l) of the model parameter vector (h) are "useful" while the first l column vectors of V provide a basis for the space of "useful" subspace of model parameters. (In the ideal case, a retained column vector of V that contained a single 1 in a particular location w and zero elsewhere would suggest that only the wth component of model vector needs to be used. In general, the l useful columns of V provide a mapping (rotation and scaling) from the original model parameter space to a restricted model parameter space. As noted previously, the value of l is either the rank or is selected by truncating the singular value spectrum. The first l columns of U form a basis of the output or the range space of X_e, that is, of the height error function. The columns of U and V greater than l can be discarded, which simplifies usage of the truncated model.

Note that while the order of the model parameters or error values within the vectors is not important, the range of cases included does matter. The model may be unable to adequately represent a case not included in the training set. Also, the model can only represent linear functions of the input model parameters. (Recall, that we can produce non-linear responses by including non-linear transforms of model parameters in the model parameter vector.)

Let X_r be the singular value reduced estimate of X_e, i.e. Eq. (9) where the diagonal elements of Σ are set to zero beyond the lth element. To apply a height correction, we form an h vector of the model parameters from a new data take, and compute an estimate of the height error vector **e**:

$$X_f h = e \qquad (10)$$

The error e is subtracted from the height map to remove the height error. Note that due to our formulation of the original model and height vectors that this is done "block-wise".

In this research, the sensitivity analysis is carried out using PCA and the modeling is investigated the least squares function modeling where the system navigation data are considered the inputs to the model and the DEM height error is considered the desired output of the model.
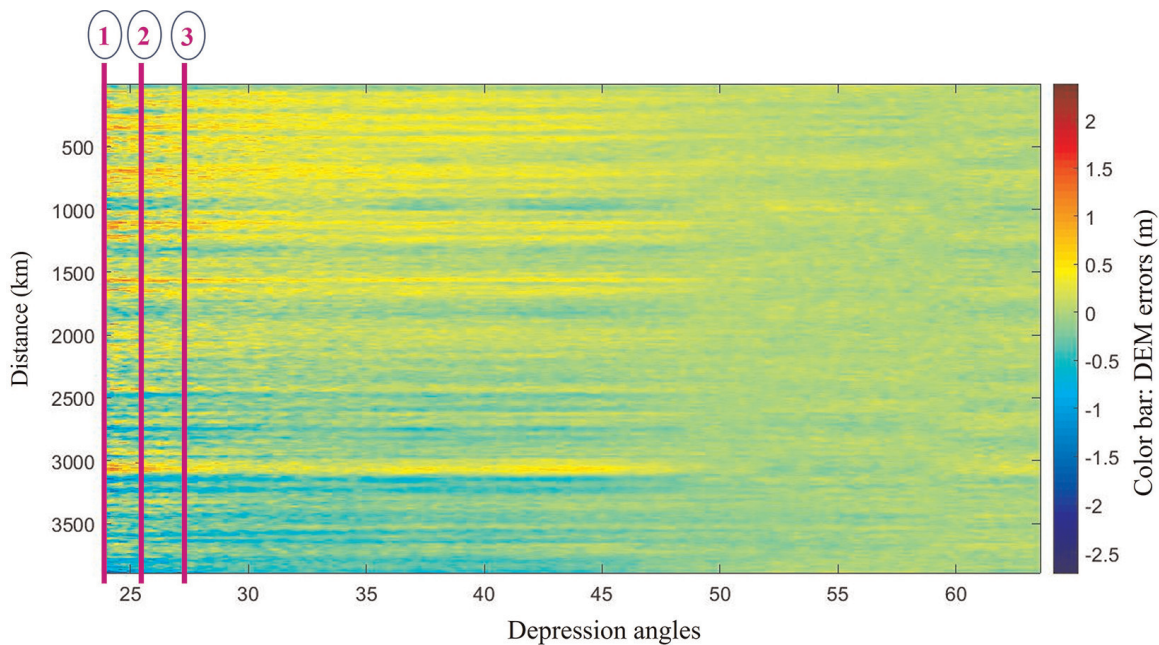

Figure 2.
DEM error map along with the locations of the across-track sections in purple solid lines (1, 2, and 3).
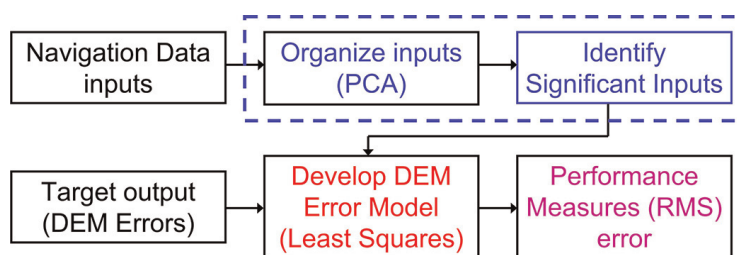

Figure 3.
Combined PCA analysis and least squares methodology.

3. Data and methodology

The DEM model errors associated with airborne mapping system datasets are investigated in this research. **Figure 2** shows an example of DEM errors case study. It can be seen that the DEM errors map shows large errors in the left side associated with low depression angles and small errors in the right side associated with high depression angles. The across-track DEM errors sections at three different depression angles are investigated (sections 1 to 3) as shown in **Figure 2** where the DEM errors are the highest in these locations.

The methodology for the proposed combined PCA analysis and least squares method is shown in **Figure 3**. The PCA analysis is utilized to identify the significant inputs from multiple navigation data and the least squares method is implemented to estimate the DEM errors models. The root mean squares (RMS) errors are used to quantify the accuracy of the developed DEM errors models.

4. Results and discussion

Three across-track sections (1 to 3) as shown in **Figure 2** have been investigated to test the performance of combined PCA analysis and least squares method. The navigation data (3 positions, 3 velocities, 3 attitudes, 3 accelerations, 3 attitude rates, 3 attitudes accelerations, 1 doppler, and 1 doppler rate) represent the input parameters for DEM errors modeling were investigated using PCA analysis method. Out of 20 inputs, 10 inputs were found significant (2 positions, 2 accelerations, 2 attitudes, 2 attitude rates, 1 doppler and 1 doppler rate) in the three across-track sections.

Then the least squares were employed to model the functional relationship between the 10 significant navigation inputs and the targeted DEM error output. The RMS errors were estimated before and after modeling to test the performance of the developed model to compensate the DEM errors. **Figure 4** shows the targeted DEM errors and the modeled DEM errors where **Figure 5** shows the differences between the
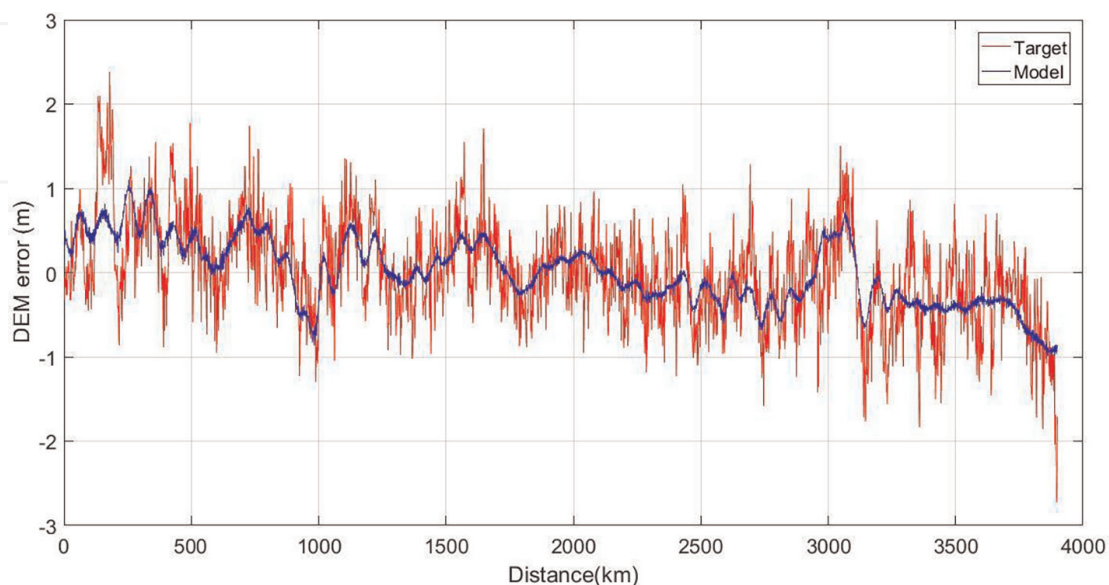

Figure 4.
Section (1) DEM errors where blue line represents the targeted DEM error and red lines represents the modeled DEM error.
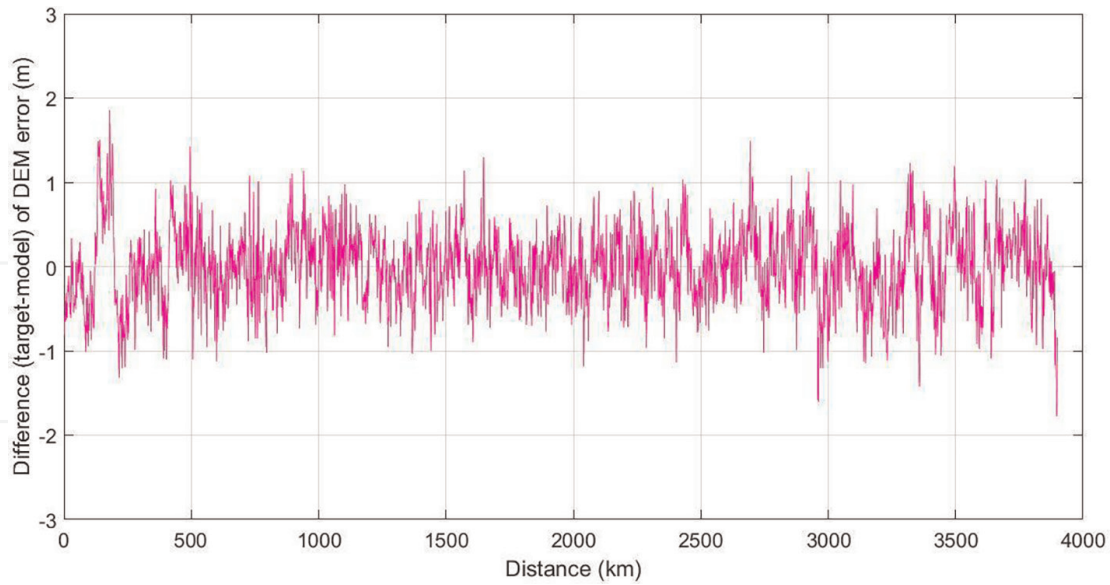
Figure 5.
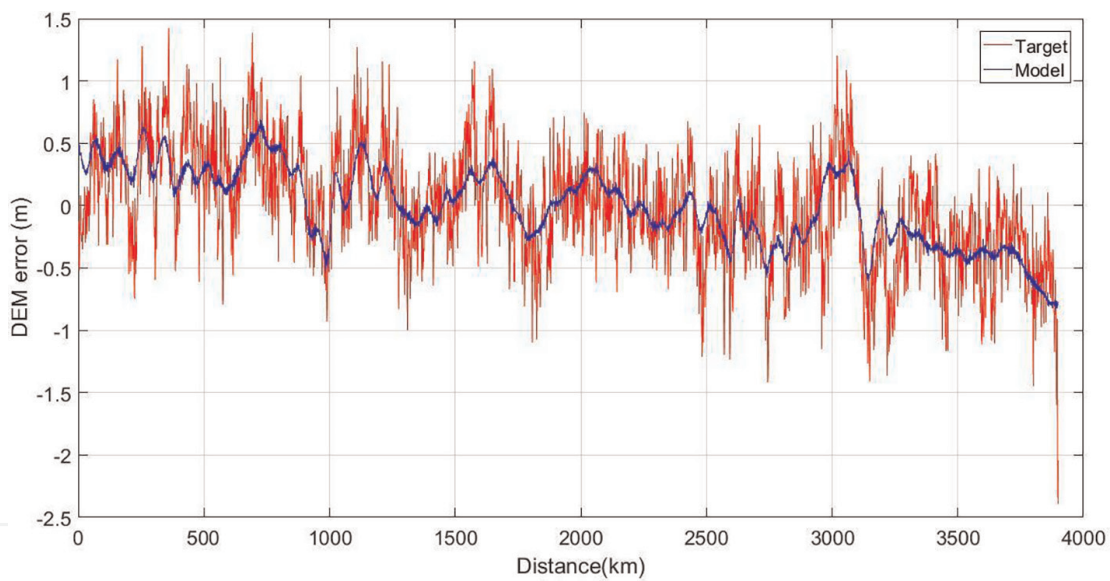*Section (1) difference between the targeted DEM error and modeled DEM errors.*

Figure 6.
Section (2) DEM errors where blue line represents the targeted DEM error and red lines represents the modeled DEM error.

targeted and modeled DEM errors for section (1) case. In section (1) case, the estimated RMS error before modeling is 0.59 m and after modeling is 0.44 m which means that about 25% of DEM errors can be compensated in section (1) case. **Figure 6** shows the targeted DEM errors and the modeled DEM errors where **Figure 7** shows the differences between the targeted and modeled DEM errors for section (2) case. In section (2) case, the estimated RMS error before modeling is 0.45 m and after modeling is 0.33 m which means that about 27% of DEM errors can be compensated in section (2) case. **Figure 8** shows the targeted DEM errors and the modeled DEM errors where **Figure 9** shows the differences between the targeted and modeled DEM errors for section (3) case. In section (3) case, the estimated RMS error before modeling is 0.35 m and after modeling is 0.25 m which means that about 28% of DEM errors can be compensated in section (3) case.
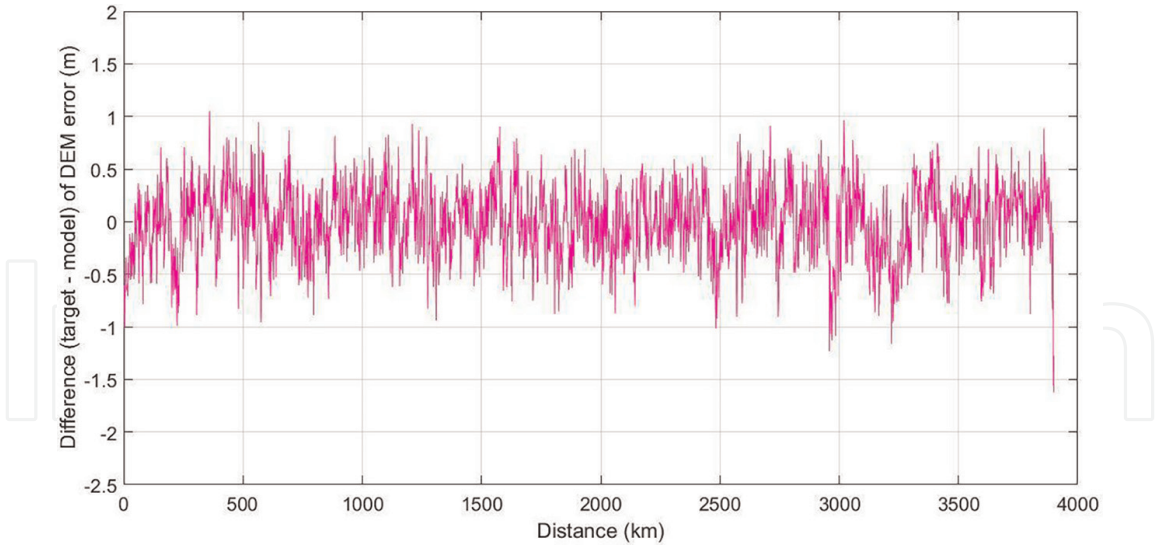
Figure 7.
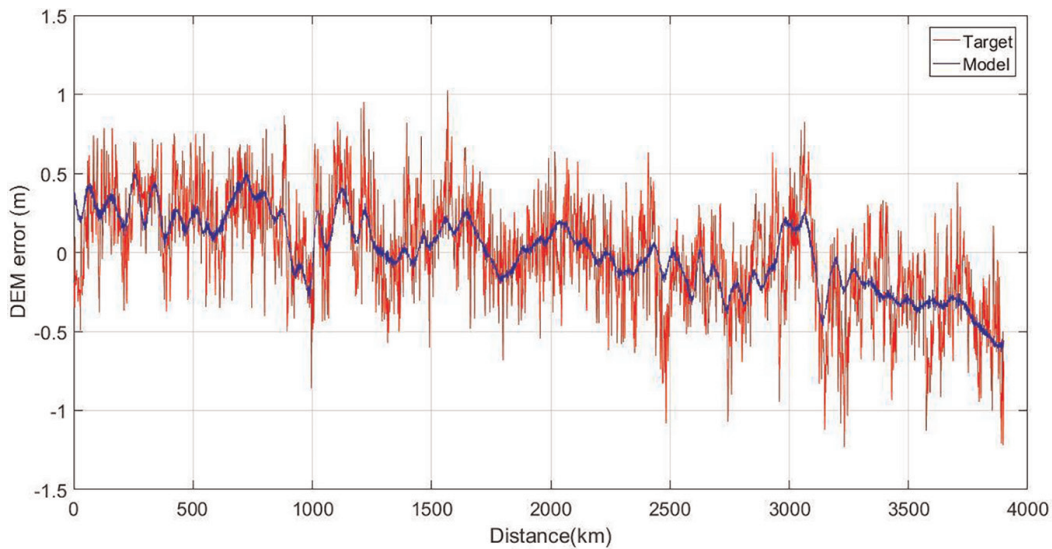*Section (2) difference between the targeted DEM error and modeled DEM errors.*

Figure 8.
Section (3) DEM errors where blue line represents the targeted DEM error and red lines represents the modeled DEM error.

Section	RMS error before modeling (m)	RMS error after modeling	Percentage of DEM error compensation (%)
Section (1)	0.59	0.44	25
Section (2)	0.45	0.33	27
Section (3)	0.35	0.25	28
Overall percentage of DEM error compensation			27

Table 1.
RMS errors for all sections before and after DEM error modeling.

Table 1 and **Figure 10** summarize the RMS errors and the percentage of DEM error compensation from the three sections where the overall percentage of DEM error compensation is about 27% on average. The results show that the combined PCA
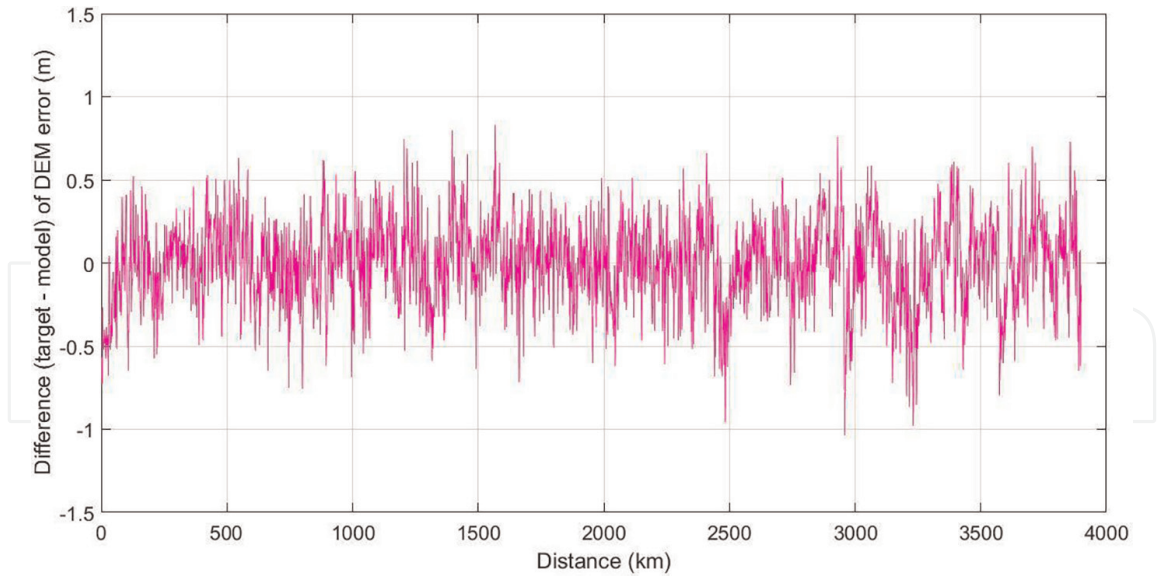
Figure 9.
Section (3) difference between the targeted DEM error and modeled DEM errors.
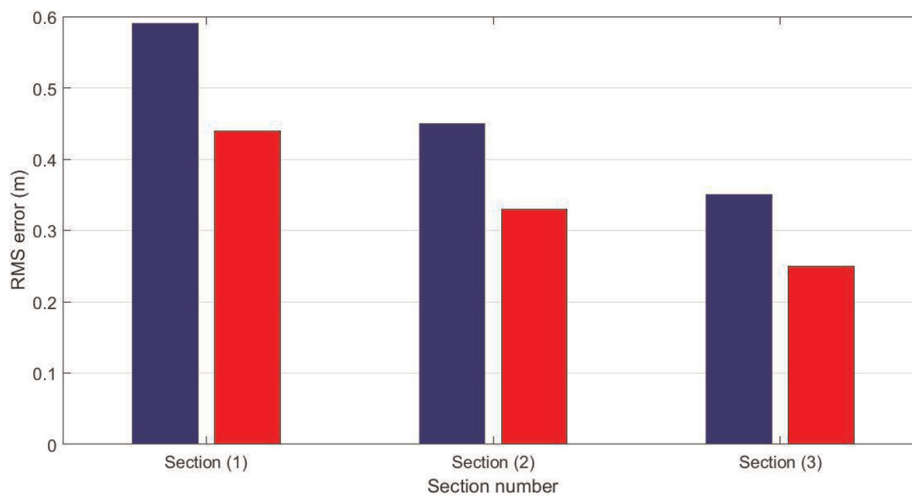

Figure 10.
RMS errors for all sections before and after DEM error modeling.

analysis and least squares method can be used as a powerful tool to compensate the DEM error produced by the navigation data with about 27% in average for the case study investigated in this research.

5. Conclusions and recommendation

The sensitivity analysis of DEM errors to the GPS/inertial navigation data was investigated in this research. It was concluded that the sensitivity analysis of the DEM errors can be performed using the PCA to identify the significant GPS/inertial navigation data components that are strongly correlated with DEM errors. Also, it is concluded that the least squares method can be rigorously utilized to establish the functional relationship between the DEM errors and the significant GPS/inertial navigation data components. The combined PCA and least squares method were validated using the DEM model errors associated with airborne mapping system datasets using

three across-track sections of the DEM errors datasets. The results show that the combined PCA analysis and least squares method can be used as a powerful tool to compensate for the DEM error produced by the GPS/inertial navigation data with about 27% on average. Therefore, it is recommended to use the combined PCA analysis and least squares method to reduce the DEM errors associated with the DEM model produced by the direct georeferenced airborne mapping system.

Conflict of interest

The author declares no conflict of interest.

Author details

Mohammed El-Diasty[1]* and Rifaat Abdalla[2]

1 Civil and Architectural Engineering Department, College of Engineering, Sultan Qaboos University, Muscat, Oman

2 Department of Earth Sciences, College of Science, Sultan Qaboos University, Muscat, Oman

*Address all correspondence to: m.eldiasty@squ.edu.om

IntechOpen

References

[1] Axelsson P. DEM generation from laser scanner data using adaptive TIN models. International Archives of Photogrammetry and Remote Sensing. 2000;**33**(B4/1; PART 4):111-118

[2] Vincent G, Sabatier D, Blanc L, Chave J, Weissenbacher E, Pélissier R, et al. Accuracy of small footprint airborne LiDAR in its predictions of tropical moist Forest stand structure. Remote Sensing of Environment. 2012;**125**:23-33. DOI: 10.1016/j.rse.2012.06.019

[3] Nie S, Wang C, Dong P, Xiaohuan X, Luo S, Qin H. A revised progressive TIN densification for filtering airborne LiDAR data. Measurement. 2017;**104**:70-77. DOI: 10.1016/j.measurement.2017.03.007

[4] Shi X, Hongchao M, Chen Y, Zhang L, Zhou W. A parameter-free progressive TIN densification filtering algorithm for Lidar point clouds. International Journal of Remote Sensing. 2018;**4**:1-14. DOI: 10.1080/01431161.2018.1468109

[5] Sithole G. Filtering of laser altimetry data using slope adaptive filter. International Archives of Photogrammetry & Remote Sensing. 2001;**34**(3/W4):203-210

[6] Vosselman G. Slope based filtering of laser altimetry data. International Archives of Photogrammetry and Remote Sensing. 2000;**33**(B3/2; PART 3):935-942

[7] Susaki J. Adaptive slope filtering of airborne LiDAR data in urban areas for digital terrain model (DTM) generation. Remote Sensing. 2012;**4**(6):1804. DOI: 10.3390/rs4061804

[8] Kilian J, Haala N, Englich M. Capture Andevaluation of Airborne Laser Scanner Data. International Archives of Photogrammetry & Remote Sensing; 1996

[9] Zhang K, Chen SC, Whitman D, Shyu ML. A progressive morphological filter for removing nonground measurements from airborne LIDAR data. IEEE Transactions on Geoscience & Remote Sensing. 2003;**41**(4):872-882. DOI: 10.1109/TGRS.2003.810682

[10] Mongus D, Borut Ž. Parameter-free ground filtering of LiDAR data for automatic DTM generation. ISPRS Journal of Photogrammetry and Remote Sensing. 2012;**67**:1-12. DOI: 10.1016/j.isprsjprs.2011.10.002

[11] Kobler A, Pfeifer N, Ogrinc P, Todorovski L, Oštir K, Sašo D. Repetitive interpolation: A robust algorithm for DTM generation from aerial laser scanner data in forested terrain. Remote Sensing of Environment. 2007;**108**(1):9-23. DOI: 10.1016/j.rse.2006.10.013

[12] Polat N, Uysal M, Toprak AS. An investigation of DEM generation process based on LiDAR data filtering, decimation, and interpolation methods for an urban area. Measurement. 2015;**75**:50-56. DOI: 10.1016/j.measurement.2015.08.008

[13] Shi W, Zheng S, Tian Y. Adaptive mapped least squares SVM-based smooth fitting method for DSM generation of LIDAR data. International Journal of Remote Sensing. 2009;**30**(21):5669-5683. DOI: 10.1080/01431160802709237

[14] Hu X, Yuan Y. Deep-learning-based classification for DTM extraction from

ALS point cloud. Remote Sensing. 2016;**8**
(9):730. DOI: 10.3390/rs8090730

[15] Cai Z, Ma H, Zhang L. Feature
selection for airborne LiDAR data
filtering: A mutual information method
with Parzon window optimization.
GIScience & Remote Sensing. 2020;**57**
(3):323-337. DOI: 10.1080/
15481603.2019.1695406

[16] Jolliffe IT. Principal Component
Analysis. 2nd ed. New York, NY:
Springer; 2002

[17] Ghilani C. Adjustment
Computations: Spatial Data Analysis.
Sixth ed. John Wiley & Sons, Inc; 2017