**Int. J. Nonlinear Anal. Appl. 13 (2022) 2, 2565–2572**
**ISSN: 2008-6822 (electronic)**
http://dx.doi.org/10.22075/ijnaa.2022.27462.3609

# Reliability analysis of Helmert model for Robust M-estimator

Hala Kadhum Obead[a], Aseel Mahmood Shakir[b], Saifaldin Hashim Kamar[a,*]

[a] *Department of Finance and Banking, College of Economic and Administration, Al-Iraqia University, Baghdad, Iraq*

[b] *Department of Business Administration, College of Economic and Administration, Al-Iraqia University, Baghdad, Iraq*

*(Communicated by Javad Vahidi)*

**Abstract**

Traditional techniques of the Helmert model are known to be used for the purpose of smoothing the data. Therefore, the special Helmert robust technique is adopted in this research to reach the internal and external reliability and suitability for the data using observed points and corresponding weights, which form the objective of this study. The least-squares method poorly performed in the presence of outliers, in comparison with the use of the robust traditional techniques, to reduce the outliers' impact. Geodetic reliability consists of two basic components: internal reliability and external reliability, which are critical measures of validating the model. Through the analyses, the internal reliability values represented by the minimum detectable bias become larger, increasing the reliability value and decreasing the probability of error. After comparing the two methods, the best method was chosen based on the final value of the probability of error. The values of the measuring parameters were modified accordingly. This method is essential to determine the reliability of the model and is bound by the effect of observations on the estimated variance of the parameters. Depending on the using various methods to estimate the reliability of the Helmert model, the results indicated that the suggested robust M method is the most accurate. And in this contribution, the minimum detectable bias was used to determine the outliers in the Gauss-Helmert model. The results showed that the robust M method has the ability to detect outliers and reduce their impact by calculating the value of the Mean Absolute Percentage Error.

Keywords: Helmert Model, Robust M, Minimum detectable bias, Bias ratio noise, Mean Absolute Percentage Error
2020 MSC: 62J07

## 1 Introduction

The German geodesist F. Helmert made significant contributions to the least-squares theory [1]. The precision of the Gauss-Helmert model (GHM) least-squares method has proven to be one of the most versatile techniques available for estimating unknown (constant) parameters in a nonlinear functional model. The model uses simple elevation grids and a GPS simulation study, and adjusts the reliability numbers for correlated observations. To discover outliers between unconnected observations, the relative size between the catalyst for the residuals and the corresponding observation must be verified [10]. Two common problems in weighted iterative least-squares were solved in geodesy, namely linear regression and affine transformation, using real and simulated data. The results were identical to

*Corresponding Author
*Email addresses:* hala.obead@aliraqia.edu.iq (Hala Kadhum Obead), aseel.shakir@aliraqia.edu.iq (Aseel Mahmood Shakir), dr.saifhkamar@gmail.com (Saifaldin Hashim Kamar)

*Received:* February 2022 *Accepted:* May 2022

the nonlinear Gauss-Helmert model, and the difference in unit weight was estimated based on the difference in the covariance matrix in the estimate.

This formula allows the researcher to obtain internal and external reliability and to apply a data snooping procedure that identifies remote measurements [2]. Helmert's model was analysed using a single scattering matrix that focuses on the necessary conditions in the presence of a unique solution for the remainder vector and the parameter vector as this solution holds specific statistical characteristics [8]. In order to easily obtain the total least-squares estimates (TLS) within the EIV model, as the conversion example shows the similarity in planar coordinates through accurate evaluation of Helmert's model, the weights are presented without additional restrictions Neitzel [7] and Koch [5] used the EM algorithm to detect outliers in the Gauss-Helmert model based on the variance amplification and the shift model mean that slightly changes the weights of observations from one repetition to another. The previous sequence results in a slight change in the estimated parameters as estimates can be presented in approximation of the linear iterations of the GH model.

Furthermore, one the most robust estimation techniques that was used to control the effect of outliers, especially after the least-squares estimates, have shown poor performance in the presence of major errors. Encompassing two basic components, external and internal, geodetic reliability is considered an essential tool of measurements for the validation of any model. Both components are important diagnostic tools for inferring the validity of a model. The contrast characteristics of the internal and external reliability measures of a given robust estimator are addressed. During the iterative re-weighing procedure of disjointed observations, the internal reliability measures as represented by minimum detectable bias become increasingly larger [4].

This paper first calculated the value of the least squares method and the robust estimator for the uncorrelated observations and the internal and external reliability measures for the robust estimator. Afterwards, the paper concluded with an example of a number linked to the simulation and the internal and external reliability measure that showed the smallest size of the total errors.

## 2 Model Description

### 2.1 Helmert Model

Helmert's model was used in geodesy as used in the indoor imaging camera and photographic sciences. The model can be written as follows [3], [6]:

$$w = By = A\xi + Be,\ e \sim (0, {\sigma_0}^2 P^{-1}) \tag{2.1}$$

where:

$w$ is a vector of the linear combination of a random vector y of rank $(2n \times 1)$.

$B$ is non-random matrix of rank $(2n \times 2n)$.

$y$ is a dependent variable of rank $(2n \times 1)$.

$A$ is non-random matrix of rank $(2n \times 6)$.

By using the weighted least squares method to model 2.1, the result is:

$$\xi = (A'(BP^{-1}B')^{-1}A)^{-1}A'(BP^{-1}B')^{-1}w. \tag{2.2}$$

However, the natural equations in 2.1, that leads to equation 2.2, provide a solution to the Lag-range vector $\hat{\lambda} = -(BP^{-1}B')^{-1}(w - A\hat{\xi})A$. A necessary condition of Lag-range Euler is that the observed residual vector can be directly calculated:

$$\widetilde{e} = P^{-1}B'(BP^{-1}B')^{-1}(w - A\hat{\xi}) \tag{2.3}$$

substituting $\hat{\xi}$ into 2.3 the result is:

$$\widetilde{e} = P^{-1}B'(BP^{-1}B')^{-1}(I - A(A'(BP^{-1}B')^{-1}A)^{-1}A'(BP^{-1}B')^{-1})w \tag{2.4}$$

with

$$N = A'(BP^{-1}B')^{-1}A \text{ and } P = (BP^{-1}B')^{-1}.$$

Let $I$ be the identity matrix. Then

$$\tilde{e} = P^{-1}B'\bar{P}\left(I - A\bar{N}^{-1}A'\bar{P}\right)w.$$

The residual cofactor matrix may be computed from 2.4

$$Q_{\tilde{e}} = P^{-1}B'\bar{P}\left(I - A\bar{N}^{-1}A'\right)BP^{-1} \tag{2.5}$$

whereas

$$\widetilde{e} = PQ_{\tilde{e}}y \tag{2.6}$$

where $PQ_{\tilde{e}}$ represent the Reliability matrix.

### 2.2 Robust M-estimator

In the case of extreme values, the study resorted to the robust methods based on the M-estimation to reduce the error to a minimum compared to the traditional methods.

$$\min = \tilde{e}'\tilde{e} \tag{2.7}$$

$$\tilde{e} = y - B^{-1}A\xi \tag{2.8}$$

$$\sum\sum\rho\left(y - B^{-1}A\xi\right) \tag{2.9}$$

$$\sum\sum A\psi\left(y - B^{-1}A\xi\right) = 0 \tag{2.10}$$

$$\sum\sum\rho\left(\frac{y - B^{-1}A\xi}{\sigma_{\tilde{e}}^2}\right) \tag{2.11}$$

$$W^* = \frac{\psi\left(\frac{y-B^{-1}A\xi}{\sigma_{\tilde{e}}^2}\right)}{\frac{y-B^{-1}A\xi}{\sigma_{\tilde{e}}^2}} \tag{2.12}$$

$$\sigma_{\tilde{e}}^2 = 1.483\left(Median\left|\tilde{e} - Median\left(\tilde{e}\right)\right|\right) \tag{2.13}$$

$$\rho(\tilde{e}) = \begin{bmatrix} G^2(1-\cos(\tilde{e}/G)) \quad |\tilde{e}| \le G\pi \\ 2G^2|\tilde{e}| > G\pi \end{bmatrix}$$

$$\psi(\tilde{e}) = \begin{bmatrix} G\sin(\tilde{e}/G)|\tilde{e}| \le G\pi \\ 0|\tilde{e}| > G\pi \end{bmatrix}$$

$$G = 1.339, \pi = 4.13.$$

The robust estimators are obtained when estimating the value of $\psi$ in 2.2:

$$\xi_{Rob} = \left(A'\left(BW*^{-1}B'\right)^{-1}A\right)^{-1}A'\left(BW*^{-1}B'\right)^{-1}w. \tag{2.14}$$

## 3 Variation of reliability measures in the Robust M- estimator

Reliability refers to the range the measurements of which are free of errors. Such range helps establishing consistency between the measurements of a variable. The robust M-estimator is treated when outliers are present and are resistant to influence, where reliability measures are used to remove the effect of outliers on the estimating parameters.

When estimating the equation 2.6 in robust, the result is:

$$\tilde{e} = \bar{p}\bar{Q}_{\tilde{e}}y \tag{3.1}$$

### 3.1 Internal Reliability Measures

In reliability theory, Minimum Detected Bias (MDB) describes the size of the model's errors that can be calculated by a standardized test [4]. The Internal Reliability Measures can be formulated as:

$$MDB_i = \sigma_0 \sqrt{\frac{\lambda_0}{(PQ_{\ddot{e}})_{ii}}} \tag{3.2}$$

where $\lambda_0$ represents the non-centrality parameter and $\sqrt{\lambda_0}$ =4.13 [10].

$$\sigma_0 \sqrt{\frac{\lambda_0}{(P^{k-1}Q^{k-1}{}_{\tilde{e}})_{ii}}} \leq \sigma_0 \sqrt{\frac{\lambda_0}{(p^k Q^k{}_{\tilde{e}})_{ii}}}$$

Note that MDB values are larger after the iterative reweighting procedure. It is considered an important diagnostic tool in monitoring data quality [11].

### 3.2 External Reliability Measures

External reliability (Bias Ratio Noise BRN) has more practical value than internal reliability. If external factors that affect the estimation of the parameter are not disclosed, they are rendered to be of little importance. Then, the External Reliability Measures can be formulated as:

$$BNR_i = \sqrt{\lambda_o(1/r_{ii} - 1)} \tag{3.3}$$

where $r_{ii}$ is the diagonal element of matrix $(I - Q_e)$.

## 4 Result and discussion

### 4.1 Simulation

The data has been generated based on the Helmert model 2.1, and three samples were generated to simulate this model (5, 10, and 15 observations). Helmert's method was compared with the robust Helmert method depending on the polluting data (10%, 20%, 30%) by calculating the absolute mean percentage error (MAPE) for two methods [**?** **?** ].

Table 1: Estimating parameters using the Helmert method and the robust Helmert's method for all sample sizes with contamination ratios (10%, 20%, 30%)

| n | Parameter | Contamination ratios 10% | | 20% | | 30% | |
|---|---|---|---|---|---|---|---|
| | | Helmret | Robust Helmret | Helmret | Robust Helmret | Helmret | Robust Helmret |
| 5 | a1 | 1.5395 | 2.9075 | 10.3162 | 11.3808 | 7.4969 | 1.5017 |
| | b1 | -2.0494 | -2.6733 | -6.2168 | -6.6844 | -3.8728 | -0.9718 |
| | c1 | 53.6241 | 46.3854 | 44.8353 | 37.819 | 25.8699 | 33.7373 |
| | a2 | -0.0886 | 0.0371 | -8.6907 | -10.202 | -4.8684 | -2.349 |
| | b2 | 1.2222 | 1.215 | 5.384 | 6.1905 | 2.1319 | 1.1225 |
| | c2 | 16.9134 | 16.4198 | -9.9297 | -8.9948 | 42.0441 | 26.3691 |
| 10 | a1 | 1.7013 | -0.8727 | -2.1181 | -2.2856 | 3.7496 | 1.9046 |
| | b1 | -1.0563 | 0.1037 | 1.096 | 1.2613 | -1.9004 | -0.9788 |
| | c1 | 6.5254 | 9.3969 | 21.9121 | 14.2662 | 8.323 | 6.2181 |
| | a2 | -2.625 | 2.2091 | -3.7273 | -1.3414 | -5.5345 | -4.323 |
| | b2 | 1.7507 | -0.4659 | 1.8363 | 0.7653 | 2.6074 | 2.2798 |
| | c2 | 16.0367 | 6.5629 | 23.4835 | 13.4696 | 46.4319 | 29.3334 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 15 | a1 | -44.0441 | -12.8265 | -1.3378 | 5.9715 | 2.3563 | 2.1149 |
| | b1 | 22.0128 | 6.4862 | 0.29 | -3.1636 | -1.5974 | -1.3656 |
| | c1 | 38.6382 | 16.8788 | 34.1869 | 15.1965 | 40.1551 | 31.0723 |
| | a2 | 28.0617 | 13.1099 | -1.2527 | 0.7881 | 1.3605 | -1.686 |
| | b2 | -14.0331 | -6.3388 | 0.9699 | 0.0733 | -0.5454 | 1.1618 |
| | c2 | 2.2733 | -0.6078 | 9.0134 | 2.2448 | 17.5678 | 7.8646 |

Table 2 shows the parameters estimated according to the two methods used. In general, it is noted that the values of the estimates are close to each other according to the pollution rates and the sample size adopted in generating the data. It can also be inferred that the robust Helmert method performed better, according to Table 3 based on MAPE. Hence, it is noted that MAPE increases. In most cases, it is possible to see an increase in MAPE and an increase in rates of pollution based on the sample size.

Table 2: MAPE using the Helmert method and the robust Helmert method for all sample sizes with contamination ratios (10%, 20%, 30%)

| **n** | **Contamination ratios** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **10%** | | | **20%** | | | **30%** | | |
| | **Helmret** | **Robust Helmret** | **Best** | **Helmret** | **Robust Helmret** | **Best** | **Helmret** | **Robust Helmret** | **Best** |
| 5 | 5.8463 | 5.0766 | Robust Helmret | 5.5161 | 5.1547 | Robust Helmret | 13.9561 | 13.9965 | Helmert |
| 10 | 4.2001 | 3.5087 | Robust Helmret | 11.8601 | 9.5952 | Robust Helmret | 16.3532 | 13.7911 | Robust Helmret |
| 15 | 8.9792 | 7.0931 | Robust Helmret | 12.6245 | 9.4607 | Robust Helmret | 18.4402 | 16.095 | Robust Helmret |

### 4.2 Real Data

Depending on observed points and corresponding weights according to [9], the parameters were estimated using the Helmert method and the robust Helmert method, the data of which is represented in Table 4, and the differences in the estimated values can be seen.

Table 3: Neri's Data for 2-D line-fitting: coordinate pairs $(x_i, y_i)$

| **Point No.** | **Observation** | |
|---|---|---|
| | **x** | **y** |
| 1 | 5.9 | 0.0 |
| 2 | 5.4 | 0.9 |
| 3 | 4.4 | 1.8 |
| 4 | 4.6 | 2.6 |
| 5 | 3.5 | 3.3 |
| 6 | 3.7 | 4.4 |
| 7 | 2.8 | 5.2 |
| 8 | 2.8 | 6.1 |
| 9 | 2.4 | 6.5 |
| 10 | 1.5 | 7.4 |

Table 4: Parameter estimators using the Helmert method and the robust Helmert method

| **Parameter** | **Helmret method** | **Robust Helmret method** |
|---|---|---|
| a1 | -2.0301 | 0.3440 |
| b1 | -5.8994 | -0.0730 |
| c1 | 40.2287 | 10.4663 |
| a2 | 0.1698 | 0.9626 |
| b2 | -0.8842 | 1.0768 |

| | | |
|---|---|---|
| c2 | 6.7133 | -3.3392 |

It can be seen that the estimated values using the robust Helmert method approach the original values of (y) better than the Helmert method through Table 5. This is confirmed by the spread of the estimated (y) values shown in Figure. 1.

Table 5: Estimators (y) using the Helmert method and the robust Helmert method

| **Observations** | **y** | **Helmret method** | **Robust Helmret method** |
|---|---|---|---|
| 1 | 5.9 | 9.8373 | 8.2891 |
| 2 | 5.4 | 9.0703 | 7.5377 |
| 3 | 4.4 | 6.5734 | 5.2800 |
| 4 | 4.6 | 7.0409 | 5.7692 |
| 5 | 3.5 | 3.2558 | 3.1835 |
| 6 | 3.7 | 3.9636 | 3.9261 |
| 7 | 2.8 | 3.2373 | 2.1459 |
| 8 | 2.8 | 2.8749 | 1.8146 |
| 9 | 2.4 | 0.4489 | 0.9301 |
| 10 | 1.5 | -1.2651 | -0.7380 |
| 11 | 0.0 | -2.6397 | -3.4147 |
| 12 | 0.9 | -1.6143 | -2.3476 |
| 13 | 1.8 | -0.4166 | -0.0847 |
| 14 | 2.6 | 0.6315 | 1.0170 |
| 15 | 3.3 | 3.5734 | 3.1937 |
| 16 | 4.4 | 4.6500 | 4.3230 |
| 17 | 5.2 | 6.8103 | 6.8875 |
| 18 | 6.1 | 7.5577 | 7.6927 |
| 19 | 6.5 | 8.1434 | 9.2485 |
| 20 | 7.4 | 9.0500 | 10.2245 |

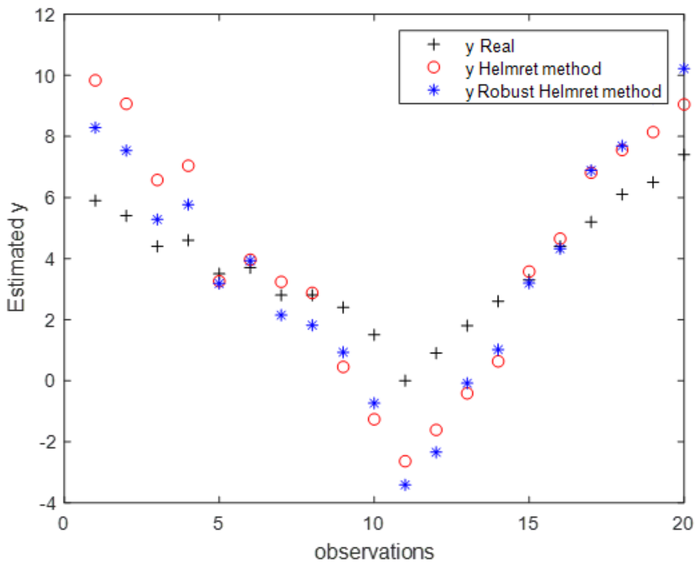


Figure 1: Estimated (y) values

Relying on calculating MAPE can be ascertained that the robust Helmert method is better compared to the Helmert method where MAPE is 1.7091 for Helmret method and 1.5816 for Robust Helmret method. In addition, it can be observed in Figure. 2 that the high internal reliability values are dependent on the robust Helmert method while the external reliability is low, as shown in Figure. 3.

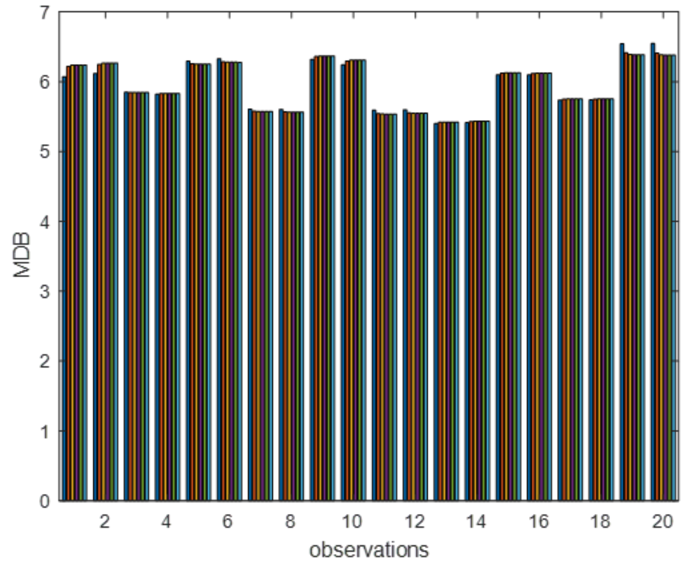


Figure 2: Internal Relability using Robust Helmert mothed

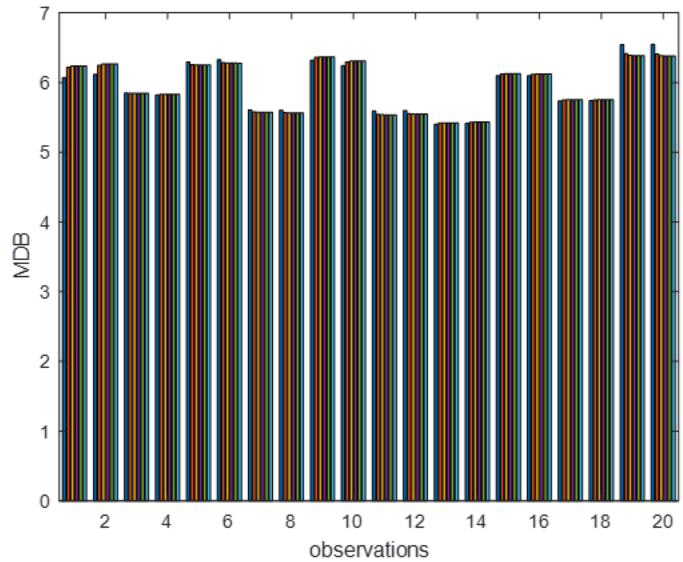


Figure 3: External Relability using Robust Helmert mothed

## 5 Conclusion

This study used robust evaluation techniques on a large scale to deal with the effects of outliers. To achieve that, reliability scales were used to determine the likelihood of outliers being detected so that their influence can be suppressed. The results showed that the internal reliability values increased dramatically during the iteration process against a decrease in the external reliability. The robust M method contributed to achieving good results in terms of adjusting the data and calculating the internal and external reliability by comparing mean values of absolute percentage error for both methods, The Minimum Detected Bias and Bias ratio noise showed the smallest size of the total errors when using the robust M method.

## Acknowledgement

This article is original and has neither been published before nor it is submitted for publication elsewhere. This is acknowledgement that all authors have contributed significantly and are in agreement with the contents of the manuscript.

## References

[1] *Adjustment computations based on the least- squares principle (in German.*

[2] A. Amiri-Simkooei and S. Jazaeri, *Weighted total least squares formulated by standard least squares theory*, J. Geodetic Sci. **2** (2012), no. 2, 113–124.

[3] J.D. Cothren, *Reliability in constrained Gauss-Markov models: an analytical and differential approach with applications in photogrammetry*, The Ohio State University, 2004.

[4] J.-F. Guo, J.-K. Ou, and Y.-B. Yuan, *Reliability analysis for a Robust m-estimator*, J. Survey. Engin. **137** (2011), no. 1, 9–13.

[5] K.-R. Koch, *Robust estimations for the nonlinear gauss helmert model by the expectation maximization algorithm*, J. Geodesy **88** (2014), no. 3, 263–271.

[6] D. Mihajlović and Ž. Cvijetinović, *Weighted coordinate transformation formulated by standard least-squares theory*, Survey Rev. **49** (2017), no. 356, 328–345.

[7] F. Neitzel, *Generalization of total least-squares on example of unweighted and weighted 2d similarity transformation*, J. Geodesy **84** (2010), no. 12, 751–762.

[8] F. Neitzel and B. Schaffrin, *On the Gauss–Helmert model with a singular dispersion matrix where bq is of smaller rank than b*, J. Comput. Appl. Math. **291** (2016), 458–467.

[9] F. Neri, G. Saitta, and S. Chiofalo, *An accurate and straightforward approach to line regression analysis of error-affected experimental data*, J. Phys. E: Sci. Instrum. **22** (1989), no. 4, 215.

[10] B. Schaffrin, *Reliability measures for correlated observations*, J. Survey. Engin. **123** (1997), no. 3, 126–137.

[11] Z. Zhou and Y. Wu, *System model bias processing approach for regional coordinated states information involved filtering*, Math. Prob. Engin. **2016** (2016).