

Méthode d'Ajustement Robuste "Rewlse" des Données Gravimétriques

Fateh TOUATI

Centre des Techniques Spatiales – Division de géodésie Spatiale
BP 13, 31200, Arzew, Algérie
fateh73@gmail.com

ملخص : السهوب الجزائرية تمثل مجالاً متميزاً مقارنة مع الأقاليم الطبيعية الأخرى للبلاد. ي موقعها بين الأطلسيين النابي و الصحراوي، تمت هذه السهوب على مساحة تقدر بـ 20 مليون هكتار. خصائصها البيوجغرافية (طبوغرا菲 ، مناخ ، نبات طبيعى ، ...) ليست ملائمة لتعزيز سكاني كثيف ، مع ذلك ، هناك عدة ثبايات يمكن استخلاصها مابين مختلف أجزاء هذا الإقليم الواسع. المطلوب هنا هو توضيح هذه الاختلافات مع التركيز على المكانة التي يشغلها إقليم السهوب الغربية مقارنة مع باقي إقليم السهوب. طريقة العمل المرتكزة على الكارتوغرافيا كانت أساس منهاجيتنا بربط عدد من المعابر الطبيعية (طبوغرا菲 ، مناخ ، نبات طبيعى) ، والديمغرافية (عدد السكان ، معدل النمو السكاني) ، والاقتصادية (بطالة ، تجهيزات ، هيكل).

Résumé : L'ajustement gravimétrique par la méthode des moindres carrés standard ne prend pas en compte d'éventuelles mesures aberrantes. Les techniques classiques de détection d'erreurs qui sont basées sur des tests statistiques sont souvent inefficaces en présence de plusieurs aberrations. Nous présentons dans cet article, une nouvelle technique statistique d'estimation à la fois robuste et efficace (au sens statistique) appelée "REWLSE" (Robust and Efficient Weighted Least Squares Estimator). C'est une méthode qui permet d'une part, de réduire l'influence des aberrations, et d'une autre part, de garder la totalité de l'information initiale. L'étude est basée sur la robustesse et l'efficacité en utilisant une valeur de coupure telle que développée en statistiques et en proposant une nouvelle fonction du poids en vue de l'adopter aux observations gravimétriques. Les tests effectués ont porté sur l'ajustement d'un réseau gravimétrique Taïwanais composé de 16 stations relatives et de 6 points absolus. Les résultats obtenus ont montré l'efficacité et la pertinence de la technique REWLSE.

Mots-clés : Pesanteur, Moindres carrés pondérés, Aberration, Estimation robuste, Point de rupture, Efficacité.

Abstract : The gravimetric data adjustment by least squares doesn't take into account the outliers in observations. One of typical approaches to handling this problem is to use the robust estimation techniques. In this paper, we introduce a new procedure of estimation which is in same time efficient (in statistical sense) and robust called REWLSE (Robust and Efficient Weighted Least Squares Estimator).

This method allows the reducing of outliers' impact and makes more use of the information provided by the data. In order to adapt this technique to the relative gravity data, a new robust weight function is developed by keeping the same adapted cut-off value as suggested by Gervini and Yohai. A sample of real and semi-simulated gravity data is tested.

The obtained results showed the pertinence of REWLSE techniques.

Key words : Relative gravity, Weighted least square, Outlier, Robust estimation, Efficiency, Breakdown point.

1. Introduction

Le problème majeur d'ajustement géodésique ou gravimétrique par la méthode des moindres carrés (LS) réside dans la présence d'aberrations dans les observations. En gravimétrie, ceci est dû, d'une part, au fait que les gravimètres sont des instruments de haute sensibilité induisant diverses erreurs systématiques, et d'autre part, aux modèles gravimétriques développés qui ne répondent pas toujours aux conditions réelles du terrain [Becker, 1989].

En géodésie comme en gravimétrie, deux tests statistiques sont généralement utilisés pour la détection d'aberrations : "Data Snooping Technique" de Baarda (1968) et le " τ -test" de Pope (1975). Cependant, dans la plupart des cas, la procédure basée sur de tels tests statistiques ne peut pas détecter et localiser des aberrations de façon correcte. En effet, les résidus obtenus par la méthode des moindres carrés ne correspondent pas souvent aux observations aberrantes existantes [Junhuan, 2005].

L'utilisation de la technique d'estimation robuste devient une alternative pour remédier au problème d'observations aberrantes. Le but de cette nouvelle technique est de réduire l'influence des aberrations sur les paramètres estimés. Les méthodes robustes sont fondamentalement développées par les travaux de Huber (1964 et 1981) et Hampel (1973). Leur utilisation en géodésie a été initiée par Carosio (1979) et appliquée, pour la première fois, aux observations gravimétriques par Mäkinen (1981).

Nous présentons dans cet article une nouvelle procédure d'estimation robuste et efficace baptisée REWLSE. La méthode dans son origine est proposée par Gervini et Yohai (2002). Elle est basée sur le calcul du seuil de coupure adapté. Afin qu'elle soit applicable aux mesures relatives du champ de pesanteur, nous suggérons une nouvelle fonction robuste, permettant ainsi d'atteindre simultanément une valeur maximale de point de rupture et une efficacité statistique élevée sous le modèle Gaussien des erreurs. Des tests ont été effectués à travers un jeu de données gravimétriques et les résultats obtenus ont montré les performances et la puissance de cette technique.

2. Ajustement par les moindres carrés pondérés

L'estimation des valeurs du champ de pesanteur et des paramètres inconnus du gravimètre par la méthode des moindres carrés est basée sur le modèle de Gauss Markov.

L'équation d'observation s'écrit sous la forme [Touati, 2004] :

$$\Delta L_{i,j} + V_{i,j} = (\Delta F(z_j) - \Delta F(z_i)) + (D(t_j) - D(t_i)) + \Delta G \quad (1)$$

$V_{i,j}$: Résidu de $\Delta L_{i,j}$

$\Delta L_{i,j} = L_j - L_i$: Observation relative entre i^{th} et j^{th} observations

t_i, t_j : Temps de mesure

$\Delta F(z)$: Fonction de calibration inconnue

z : Lecture gravimétrique en unité de comptage (CU)

$D(t)$: Dérive inconnue du gravimètre

ΔG : Différence de gravité entre la j^{th} et la i^{th} station.

Supposant m observations, n stations et u paramètres gravimétriques. Soit W la matrice du poids de ΔL , la formulation matricielle des équations d'observation est donnée par :

$$\Delta L + V = AX \quad (2)$$

Avec, A la matrice de configuration et X vecteur de p inconnus ($p = n + u$).

La solution par la méthode des moindres carrés est impossible sans l'introduction de contraintes supplémentaires. Ceci est dû au problème de déficit du rang [Hwang et al., 2002]. Dans le but d'obtenir une solution unique au sens des moindres carrés, deux méthodes peuvent être utilisées : la complémentation gravimétrique et les contraintes pondérées. Cette dernière est donnée par :

$$\hat{X} = (A^T W A + A_g^T W_g A_g)^{-1} (A^T W \Delta L + A_g^T W_g L_g) \quad (3)$$

Où A_g, L_g, W_g , sont respectivement la matrice de configuration des contraintes, le vecteur des valeurs de la gravité absolue des stations fixes et la matrice diagonale des contraintes.

3. Méthodes d'ajustement robuste

Les méthodes robustes procurent une procédure alternative, sans avoir besoin à identifier les aberrations ou de les exclure.

Elles utilisent toute l'information fournie par les données. De plus, en cas d'absence d'aberrations, elles donnent des résultats similaires à ceux obtenus par les moindres carrés. Nous citons essentiellement :

- Le M-Estimateur dont le principe consiste à minimiser la fonction score $\rho(v)$ des résidus telle que [Huber, 1981] :

$$\hat{X}_M = \arg \min_X \sum_{i=1}^m \rho(v_i) \quad (4)$$

- L'estimation par la méthode LTS (Least Trimmed Squares) qui est basée sur la sommation des carrés des h plus petits résidus [Rousseeuw et all., 1987] :

$$\hat{X}_{LTS} = \arg \min_X \left(\sum_{i=1}^h v_{(i)}^2 \right) \quad (5)$$

- Les MM-Estimateurs qui sont basés sur la combinaison de deux techniques différentes : le M-Estimateur avec un autre estimateur de point de rupture élevé (ex. LTS).

Estimation du facteur d'échelle

L'estimation robuste du facteur d'échelle est obtenue par la méthode de la médiane donnée par [Yang et al., 1999] :

$$\hat{\sigma} = \frac{1}{\Phi^{-1}(0.75)} \text{med} \{ |v_i| \} \quad (6)$$

Où $\Phi^{-1}(0.75) = 4.18$ permet d'avoir un biais nul sur l'estimation de σ lorsque la distribution des résidus est Gaussienne.

4. Estimation par REWLSE

Considérant un couple de solutions initiales robustes pour les paramètres et le facteur d'échelle, \hat{X}_0 et $\hat{\sigma}_0$ respectivement. Les résidus normalisés sont définis comme suit :

$$\bar{v}_i = \frac{l_i - A_i \hat{X}_0^T}{\hat{\sigma}_0} \quad (7)$$

Une large valeur de \bar{v}_i suggère que la i^{th} observation soit aberrante.

4.1 Seuil de coupure adapté

Afin de maintenir le point de rupture de l'estimateur initial, Gervini et Yohai (2002), ont proposé une stratégie de calcul du seuil de coupure adapté :

$$t_n = \min \{t : F_n^+(t) \geq 1 - d_n\}^+ \quad (8)$$

Où F_n^+ est la distribution empirique des résidus absolu et normalisé et d_n représente une mesure de la fraction des aberrations des observations donnée par :

$$d_n = \sup_{t \geq \mu} \{\phi^+(t) - F_n^+(t)\}^+ \quad (9)$$

ϕ^+ est la distribution hypothétique (loi normale dans notre cas) et μ est le seuil de coupure initial, souvent pris entre 2.5 et 3.

4.2 Détermination du poids

Avec le seuil de coupure adapté, la forme du poids W et l'estimateur REWLSE sont définis comme suit [Gervini et Yohai, 2002] :

$$T_{ln} = \begin{cases} (A^T WA)^{-1} A^T WL & \text{si } S_n > 0 \\ T_0 & \text{si } S_n = 0 \end{cases} \quad (10)$$

Le poids W est choisi de façon à avoir un rejet sévère sur les données aberrantes :

$$w_i = \begin{cases} 1 & \text{si } |\bar{v}_i| < t_n \\ 0 & \text{si } |\bar{v}_i| \geq t_n \end{cases} \quad (11)$$

$$W = \text{diag}(w_1, w_2, \dots, w_n)$$

Notons qu'avec ce poids, l'estimateur T_{ln} garde la même valeur du point de rupture de l'estimateur initial T_0 .

4.3 Adaptation de la technique REWLSE

aux données gravimétriques Afin d'obtenir une seule matrice de poids et de configuration, le système

décrivant le modèle d'ajustement par la méthode des moindres carrés pondérés (équation 3) a été modifié de la manière suivante :

$$A^T WA + A_g^T W_g A_g = M^T PM \quad (12)$$

M et P sont respectivement les matrices de configuration et de poids modifiées.

L'estimateur des moindres carrés s'écrit alors :

$$T = (M^T PM)^{-1} M^T PL = qPL \quad (13)$$

L'estimation des données relatives du champ de pesanteur par la technique REWLSE en appliquant la fonction de poids précédente mène à un système singulier. L'objectif est alors de développer des fonctions de poids plus adaptées aux données gravimétriques afin d'avoir non seulement un système soluble, mais aussi d'obtenir une estimation robuste et efficace des paramètres inconnus. Dans le cas d'observations entachées d'erreurs ($L_\delta = L + \delta$), toutes les estimations robustes utilisées dans l'ajustement des réseaux géodésiques ou gravimétriques peuvent s'écrire comme suit (Jianjun, 1996) :

$$T_\delta = (M^T P_\delta M)^{-1} M^T P_\delta L \quad (14)$$

Où P_δ représente une fonction des résidus v .

En posons $q = (M^T P_\delta M)^{-1} M^T$ et en supposant que q est approximativement invariable, nous obtenons :

$$T_\delta = q P_\delta L_\delta \quad (15)$$

La différence entre T_δ et T s'écrit :

$$\Delta = T - T_\delta = qPL - qP_\delta L = q(PL - P_\delta L_\delta) \quad (16)$$

La distance moyenne entre les deux estimateurs :

$$Dm = E(\Delta^T \Delta) \quad (17)$$

En minimisant la distance Dm , on obtient la fonction de poids suivante :

$$P_\delta(i) = \begin{cases} P_i & \text{si } |\bar{v}_i| \leq t_n \\ P_i \cdot \frac{r_u}{S_n^2 \bar{v}_i^2} & \text{si } |\bar{v}_i| > t_n \end{cases} \quad (18)$$

r_u est la redondance locale de la i^{th} mesure.

5. Application

L'application a porté sur l'ajustement d'un réseau gravimétrique Taïwanais composé de 16 stations observées avec le gravimètre Lacoste & Romberg modèle G.

Les données gravimétriques utilisées sont constituées de 56 observations relatives d'écart-type s'étalant entre 30 et 50 µgal. Le canevas comporte aussi 06 points absous obtenus avec une précision de 2-3 µgal (FG5).

La figure 1 illustre le diagramme de QQ-Normal des résidus robustes (à droite) et des résidus standard (à gauche).

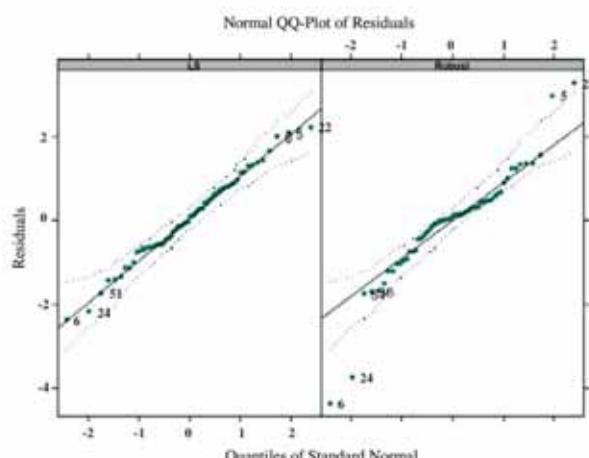


Fig. 1 Diagramme quantile-normale.
Données réelles.

La figure (1) montre que la méthode robuste est très sensible aux observations aberrantes (5, 6, 22 et 24) et qu'elles peuvent être facilement détectées. En revanche, avec la méthode LS aucune observation aberrante n'a pu être décelée.

Pour illustrer les performances de la technique REWLSE, des tests ont été effectués en bruitant quelques observations telles que :

$$L_\delta = L + K\sigma, \text{ avec } K=1, 2, 3 \text{ et } \sigma = 0.03 \text{ mgal}$$

Pour chaque valeur de K, on calcule les 3 estimateurs suivants :

T1 : LS pondérée sans aberrations

T2 : LS pondérée avec aberrations

T3 : REWLSE avec aberrations

Ainsi que les deux distances :

Dm1 : entre T1 et T2

Dm2 : entre T1 et T3

Pour se prononcer sur l'efficacité statistique et sur la robustesse de REWLSE et du facteur d'échelle proposé, on calcule pour chaque valeur de K :

- l'erreur moyenne quadratique pondérée EMQ qui est une bonne indicatrice de l'efficacité statistique,
- le facteur d'échelle rob $\hat{\sigma}_{rob}$
- les distances Dm1 et Dm2 qui sont des indicatrices de robustesse de l'estimateur, alors que la première,

Les tables 3, 4 et 5 fournissent les statistiques des résultats obtenus par les 3 estimateurs (T1, T2 et T3) pour chaque valeur de K.

Les distances Dm1 et Dm2 sont calculées pour 2 cas : le 1^{er} cas (*valeur en gras*) a été obtenu en intégrant tous les paramètres inconnus (champ de pesanteur et les paramètres du gravimètre) et le 2nd cas a été déterminé uniquement en utilisant les valeurs du champ de pesanteur estimées.

Table (3). Statistiques des résultats obtenus par les 3 estimateurs (K=1).

EMQ	$\hat{\sigma}_{rob}$	Dm1	Dm2
T2	1.88		
T1	0.99	0.028 0.008/0.007	0.008/0.003
T3	1.02		
$t_0 = 2.5, t_n = 2.64, \% \text{ d'erreurs détectées} = 8/10$			

Table (4). Statistiques des résultats obtenus par les 3 estimateurs (K=2).

EMQ	$\hat{\sigma}_{rob}$	Dm1	Dm2
T2	3.08		
T1	0.99	0.030 0.031/0.023	0.015/0.014
T3	1.00		
$t_0 = 2.5, t_n = 5.48, \% \text{ d'erreurs détectées} = 10/10$			

Table (5). Statistiques des résultats obtenus par les 3 estimateurs (K=3).

EMQ	$\hat{\sigma}_{rob}$	Dm1	Dm2
T2	4.52		
T1	0.99	0.026 0.0728/0.573	0.003/0.001
T3	0.97		
$t_0 = 2.5, t_n = 8.76, \% \text{ d'erreurs détectées} = 10/10$			

Les tables (3, 4 et 5) montrent clairement que l'EMQ obtenu par les estimateurs T1 et T3 est proche de l'unité, alors que celui déterminé par l'estimateur T2 est affecté par les aberrations. Le facteur d'échelle robuste rob $\hat{\sigma}_{rob}$ est presque le même pour tous les cas (de 0.026 à 0.030 mgal) et il reflète parfaitement la précision du gravimètre utilisé (LCR G838).

Nous constatons d'après les tables 3, 4 et 5 que les valeurs de Dm1 s'amplifient au fur à mesure que le nombre et la magnitude des aberrations augmentent. Par contre, les valeurs de Dm2 restent toujours petites et insignifiantes. Cette distance dépend principalement du nombre d'aberrations détectées et de leurs positions dans les observations.

Ceci est bien sûr lié directement à l'efficacité de l'estimateur initial (LTS dans notre cas). Nous remarquons aussi, que la valeur de Dm^2 est légèrement inférieure à celle obtenue en incluant tous les paramètres estimés ; Ce qui explique l'influence du modèle d'ajustement sur la robustesse de l'estimateur.

Notons que les valeurs du seuil de coupure adapté t_n augmentent avec la magnitude des aberrations.

6. Conclusion

Sur la base des statistiques calculées, l'étude menée dans ce présent travail nous a permis de conclure sur l'efficacité des méthodes robustes dans la réduction de l'influence des aberrations et les limites de la méthode des moindres carrés en présence d'observations erronées.

Avec l'estimateur REWLSE, nous avons pu estimer efficacement avec moins de biais les paramètres inconnus. En utilisant une valeur de coupure bien adaptée, l'estimateur REWLSE a bien gardé la puissance de l'estimateur initial (LTS) dans la détection des observations aberrantes. D'autre part, la fonction du poids que nous avons proposé a permis non seulement de surmonter le problème de déficit du rang, mais aussi d'estimer, avec fiabilité et moins de biais, les valeurs du champ de pesanteur et les paramètres du gravimètre. La difficulté de cette méthode réside dans sa solution initiale qui est basée sur des méthodes robustes avec un point de rupture élevé que quelques fois n'assurent pas toujours la détection efficace et l'identification correcte des aberrations lorsque la magnitude de celles-ci est relativement petite. La structure des données relatives de gravité complique aussi la tâche puisque une mesure aberrante sur le terrain conduit à un doublet d'observations aberrantes.

Remerciements

Mes remerciements au Pr. Cheinway Hwang de National Chiao Tung University de Taiwan, pour avoir mis à ma disposition les données gravimétriques de Taiwan. Je tiens aussi à remercier Monsieur Daniel Gervini de l'université de Zurich pour son aide.

Références bibliographiques

- M. Becker, 1989 : *Adjustment of microgravimetric measurements for detecting local and regional vertical displacements*. Gravity, Gradiometry, and Gravimetry, symposium No.103 : 149-161 Springer-Verlag.
- G. Csapo, M. Kis, L. Völgyesi, 2003: *Different adjustment methods for the Hungarian part of the unified Gravity Network*. XXIII General Assembly of the International Union of Geodesy and Geophysics, Sapporo, Japan.
- D. Gervini, V.J. Yohai, 2002: *A class of robust and fully efficient regression estimators*. The annals of Statistics, Vol. 30, No.2, pp.583-616.
- G. Harnisch, 1993: *Systematic errors affecting the accuracy of high precision gravity measurements*. Geodesy and physics of the earth, International Association of Geodesy Symposia. Symposium N°112. Springer-Verlag.
- P.J. Huber, 1981: *Robust statistics*. Wiley, New York.
- C. Hwang, C. Wang, L. Lee, 2002: *Adjustment of gravity measurements using weighted and datum-free constraints*. Computers & Geosciences, Vol.28,pp.1005-1015.
- P. Junhuan, 2005: *The asymptotic variance covariance matrix, Baarda test and the reliability of L1-norm estimates*. J Geod DOI: 10.1007/s00190-004-0433-y.
- P.J. Rousseeuw, A.M. Leroy, 1987. *Robust regression and outlier detection*. Wiley, New York.
- F. Touati (2004) : *Méthodologie d'ajustement des mesures relatives de gravité*. Mémoire de magister en techniques spatiales et applications, CNTS.
- F. Wicki, 2001: *Robust estimator for the adjustment of geodetic networks*. Proceeding of the first international symposium on robust statistics and fuzzy techniques in geodesy and GIS. Report N°. 295, Institute of Geodesy and photogrammetry, ETH Zürich, pp 53-60.
- Y. Yang, MK. Cheng, CK. Shum, BD. Tapeley, 1999: *Robust estimation of systematic errors of satellite laser range*. J Geod 73:345-349.
- J. Zhu, 1996: *Robustness and the robust estimate*. J Geod 70: 586-590.