Méthode d'Ajustement Robuste "M-Estimateurs" : Application au Réseau Cadastral de la ville d'Oran

B. GOURINE, B. GHEZALI, R. NAFFI, H. TAJBI Centre des Techniques Spatiales – Division de Géodésie CTS, BP 13, 31200 Arzew – ALGERIE, Fax: (213) 041 47 36 65 E-mail: bachirgourine@yahoo.com

Résumé: La méthode des moindres carrés demeure un estimateur peu robuste, bien qu'elle soit un procédé optimal sous le concept de la normalité. En effet, la présence de grosses erreurs ou d'erreurs systématiques dans les observations peut perturber d'une façon significative l'estimation des paramètres inconnus. Afin d'affaiblir les effets de ces erreurs et de garantir une solution fiable, nous proposons à travers cet article une méthode d'estimation robuste appelée "M-estimateurs" qui repose sur le principe des poids équivalents.

Les tests effectués ont porté sur l'ajustement d'un réseau cadastral de la ville d'Oran composé de 45 points géodésiques. Les résultats obtenus ont montré que la qualité du réseau ajusté par la méthode des M-estimateurs est nettement meilleure que celle fournie par la méthode des Moindres Carrés.

Mots clés : M-estimateurs, Robustesse. Poids équivalents, Ajustement, Réseaux géodésiques.

Abstract: The least squares method remains a non very robust estimator, although it is an optimal process under the concept of normality. Indeed, the presence of gross errors or systematic errors in the observations can disturb significantly the estimate of the unknown parameters. In order to decrease the effects of these errors and to guarantee a reliable solution, we propose through this paper a robust estimation method called M-estimators which based on the equivalent weights principle.

The tests carried out related to the adjustment of a cadastral network of Oran city composed of 45 geodetic points. The results obtained showed that the quality of the network adjusted by Mestimators is definitely better than that provided by Least squares method.

Key words: M-estimators, Robustness, Weights equivalent, Adjustment, Geodetic networks,

Introduction

Généralement, le principe des méthodes de compensation ou d'ajustement repose sur la conception d'un modèle théorique d'un vecteur de mesure noté y par une fonction f telle que : $y \approx f(x)$ où x est le vecteur des paramètres (inconnues). Le problème consiste à déterminer le vecteur x des paramètres vérifiant "au mieux" la relation $y \approx f(x)$ suivant un critère donné, Le modèle n'est en général qu'approximatif et les mesures sont très souvent entachées d'erreurs aléatoires.

La méthode des moindre, scarrés qui est basée uniquement sur l'estimation statistique, s'avère peu appropriée car comme toute autre méthode d'estimation, elle n'est pas exempte de défauts, L'inconvénient majeur de cette méthode, e'est son manque de robustesse [Coulot, 2002].

Pour pallier à ce problème, plusieurs techniques ont été développées en vue de réduire la sensibilité de l'estimateur des moindres carrés aux changements mineurs dans la population des observations traitées. Parmi ces techniques, la méthode des M-estimateurs s'est révélée un outil adéquat pour minimiser les effets des erreurs de mesures et garantir une solution de meilleure qualité. C'est une méthode généralisée du maximum de vraisemblance qui repose sur la définition des poids équivalents (poids de Andrew, de Huber, de Tukey et de l'IGG (Institute of Geodesy and Geophysics - Autriche) [Youcai et al., 1995].

Les tests effectués ont porté sur l'ajustement d'un réseau cadastral de la ville d'Oran constitué de 37 points nouveaux et de 8 points d'appui. Les résultats obtenus ont révélé que la méthode d'ajustement des M-estimateurs offre des performances beaucoup plus intéressantes en terme de robustesse que celle des moindres carrés.

Modèles de contamination des observations

Le modèle linéaire des observations de Gauss-Markov est exprimé par :

$$Y = A.X + V \tag{1}$$

Avec :

Y: vecteur nxI des observations $(y_1, y_2, ..., y_n)$

A: matrice nxp de coefficients (a_{ij})

X : vecteur pxI des paramètres $(x_1, x_2,...,x_p)$

V: vecteur nxI des résidus $(v_1, v_2,...,v_n)$

Si nous supposons que l'ensemble des résidus suit une distribution normale f d'une moyenne E(V)=0 et d'une variance $E(V^TV)=\sigma^2$, alors l'estimateur des moindres carrés X_{MC} a la variance minimale parmi tous les estimateurs linéaires non biaisés [Youcai et al., 1995].

Dans ce cas, l'estimation des moindres carrés est l'estimateur non biaisé le plus efficace. Cependant, si la véritable distribution des résidus dévie de la distribution normale f, alors l'estimation des moindres carrés perd son efficacité et sa variance ne sera pas minimale [Youcai et al., 1995]. Cette déviation est due à la contamination des observations par des grosses erreurs et à l'existence des biais de mesures.

Pour ce qui suit, nous considérons deux modèles d'observations contaminées (relations 2 et 3). Le premier modèle (2) est une combinaison de deux distributions normales avec des variances différentes, i.e., le modèle considéré est normal $[N(0,\sigma^2)]$ au milieu de la distribution, mais il a une grande variance $[N(0,k^2\sigma^2), k>1]$ dans les queues (extrémités de la distribution). Le deuxième modèle (3) est normal au milieu et exponentiel (distribution de Laplace) dans les queues. Ce modèle, originairement proposé par Huber, est basé sur sa recherche théorique qui postule que le comportement des grosses erreurs dans les queues ressemble à une distribution de Laplace. Pour formuler les deux modèles de distribution (2) et (3), trois densités de probabilité sont considérées :

- Si y suit $N(0, \sigma^2)$:

$$f_{\sigma} = \frac{1}{\sigma \sqrt{2\pi}} \left[\exp{-\frac{(y - a^T x)^2}{2\sigma^2}} \right]$$

- Si y suit $N(0, k^2 \sigma^2)$:

$$f_{\kappa\sigma} = \frac{1}{\kappa\sigma\sqrt{2\pi}} \left[\exp{-\frac{(y - a^T x)^2}{k^2 2\sigma^2}} \right]$$

 Si y suit la distribution de Laplace avec le paramètre d'échelle σ :

$$f_{la} = \frac{1}{2\sigma} \exp \left[-\frac{\left| y - a^T x \right|}{\sigma} \right]$$

Suivant la proposition de Tukey, les deux modèles (2) et (3), sont présentés sous la forme suivante:

$$f_{(I)} = (1 - \varepsilon) f_{\sigma} + \varepsilon f_{k\sigma} \tag{2}$$

$$f_{(H)} = (1 - \varepsilon) f_{\sigma} + \varepsilon f_{l\sigma}$$
(3)

Où $\mathcal{E} \approx b/n$ représente la partie des observations contaminées $(0 \le \mathcal{E} \le 1)$ et b est le nombre des observations contaminées dans l'échantillon de taille n.

2. Estimation robuste par la méthode des Mestimateurs

Les M-Estimateurs, proposés par Huber, sont généralisés à partir du maximum de vraisemblance mais ils dérivent d'un modèle de distribution légèrement dévié du modèle gaussien. La méthode du maximum de vraisemblance (MLE: Maximum Likelihood Estimator) est à la fois l'une des plus utilisées et des plus controversées en statistique. Elle a en effet un attrait à la fois intuitif, parce que la vraisemblance semble bien contenir toute l'information fournie par les observations, et théorique, à cause des bonnes propriétés asymptotiques des estimateurs correspondants sous certaines conditions de régularité. Cependant, cette méthode a plusieurs inconvénients dont deux défauts majeurs qui sont [Huber et al., 1997]:

- L'absence de définition correcte d'un estimateur du maximum de vraisemblance lorsqu'il peut y avoir une ambiguïté sur la version de la densité à employer, ou bien lorsque le modèle n'est pas dominé.
- Le manque de robustesse des estimateurs du maximum de vraisemblance qui sont relatifs au modèle gaussien. Le prolongement des estimateurs MLE, dans un ensemble plus général, est celui des M-estimateurs qui traite ce problème. Il permet de corriger l'absence de robustesse de certains estimateurs MLE.

Les M-estimateurs sont donc une méthode généralisée de l'estimateur du maximum de vraisemblance. Ils ont l'avantage d'hériter toutes les propriétés des bons estimateurs MLE, c'est-à-dire, l'efficacité et la normalité asymptotique, plus celle d'être peu sensible aux écarts des observations par rapport au modèle.

Ils sont basés sur le principe de la minimisation de la fonction des résidus, $\rho(v)$, qui contrôle l'influence des observations aberrantes sur la procédure de l'estimation. Ce qui n'est pas le cas pour la méthode des moindres carrés qui est basée sur la minimisation des carrés des résidus.

D'après [Youcai et al., 1995], la méthode des moindres carrés re-pondérés est une technique de calcul des M-estimateurs. Le modèle des moindres carrés re-pondérés est formé par introduction des poids équivalents dans la formule des moindres carrés classiques [Vanicek et al., 1995] tel que :

$$\hat{x} = (A^T W A)^{-1} A^T W y \tag{4}$$

Où W représente la matrice poids dans laquelle l'élément diagonal est $w(v_i)$ et ses éléments non diagonaux sont nuls.

Le poids équivalent $w(u_i)$ est exprimé par :

$$w(u_i) = \frac{\psi(u_i)}{u_i}, i = 1, 2, \dots, n$$
 (5)

Avec:

 $u_i = \frac{v_i}{\sigma_{vi}}$: Résidu normalisé.

σνi: Écart type du résidu.

 $\Psi(u_i) = \frac{\partial \rho(u_i)}{\partial X}$: Fonction dérivée de la fonction de

vraisemblance des résidus normalisés.

 $\rho(v_i) = \ln f(y_i - a_i^T x)$: Fonction de vraisemblance des résidus [Rehailia, 1995].

Si (n-b) observations suivent la distribution normale $N(0, \sigma^2)$ et les b observations suivent la distribution normale $N(0, k^2\sigma^2)$ ou la distribution de Laplace, par conséquent le modèle linéaire (1) devient [Youcai et al., 1995]:

$$\begin{bmatrix} y_{(n-b)} \\ \dots \\ y_{(b)} \end{bmatrix} = \begin{bmatrix} A_{(n-b)} \\ \dots \\ A_{(b)} \end{bmatrix} x + \begin{bmatrix} v_{(n-b)} \\ \dots \\ v_{(b)} \end{bmatrix}$$

La fonction ψ pour le MLE peut être écrite sous la forme suivante :

$$\psi(u) = -\frac{g'(u)}{g(u)} = (-\ln g)'$$
 (6)

Où g'(u) représente la première dérivée de la fonction de densité de probabilité g des résidus normalisés u.

Les poids équivalents pour le MLE deviennent donc:

$$w(u) = -\frac{g'(u)}{g(u)u} \tag{7}$$

- Pour le modèle de distribution (2) :

$$w(u) = \frac{\psi(u)}{u} = \begin{cases} 1, & |\mathbf{u}| < k \\ \frac{1}{k^2}, & |\mathbf{u}| \ge k \end{cases}$$

L'estimation des moindres carrés re-pondérés des paramètres X et σ^2 est donnée par :

$$\hat{x} = (A^{T}_{(n-b)}A_{(n-b)} + k^{-2}A^{T}_{(b)}A_{(b)})^{-1}.$$

$$(A^{T}_{(n-b)}y_{(n-b)} + k^{-2}A^{T}_{(b)}y_{(b)})$$
(8)

$$\hat{\sigma}^{2} = \frac{1}{n-p} \left\{ y_{(n-b)} - A_{(n-b)} \hat{x}^{T} (y_{(n-b)} - A_{(n-b)}^{T} \hat{x}) \right.$$

$$\left. k^{-2} \cdot (y_{(b)} - A_{(b)} \hat{x}^{T} (y_{(b)} - A_{(b)}^{T} \hat{x}) \right\}$$
(9)

Pour le modèle de distribution (3) :

$$w(u) = \begin{cases} 1, & |\mathbf{u}| < c \\ \frac{csign(u)}{u} = \frac{c}{|u|}, & |u| \ge c \end{cases}$$

Dans ce cas, l'estimation des moindres carrés re-pondérés de X et σ^2 est exprimée par :

$$\hat{x} = (A^{T}_{(n-b)}A_{(n-b)} + A^{T}_{(b)}A^{T}_{(b)}W_{(b)}A_{(b)})^{-1}.$$

$$(A^{T}_{(n-b)}Y_{(n-b)} + A^{T}_{(b)}W_{(b)}Y_{(b)})$$
(10)

$$\hat{\sigma}^{2} = \frac{1}{n-p} \left\{ y_{(n-b)}^{-} - A_{(n-b)}^{-} \hat{x}^{T} (y_{(n-b)}^{-} - A_{(n-b)}^{T} \hat{x}) \right. \\ \left. \cdot (y_{(b)}^{-} - A_{(b)}^{-} \hat{x}^{T} \cdot W_{(b)}^{-} (y_{(b)}^{-} - A_{(b)}^{T} \hat{x}) \right\}$$

$$(11)$$

Avec:
$$w_{(n-b)} = I_{(n-b)}$$

$$w_{(b)} = (c/|u_j|)I_{(b)}, j = n-b+1,...,n$$

3. Poids équivalents

En pratique, la diversité des M-estimateurs réside dans la définition des poids équivalents et leur application est basée sur la méthode des moindres carrés re-pondérés. En effet, la formulation générale des M-estimateurs est donnée par l'équation suivante [Yang Y., 1999]:

$$\hat{X}_{p,1} = \left(A_{p,n}^T \cdot \overline{P} \cdot A_{n,p}\right)^{-1} \cdot A_{p,n}^T \cdot \overline{P} \cdot y_{n,1} \quad (12)$$

Où P représente la matrice des poids équivalents. La construction de cette matrice est basée sur la définition des poids équivalents qui peuvent être formulés comme suit [Gui et al, 1998] :

- Poids de Andrew :

$$\overline{P}_i = \begin{cases} P_i \frac{\sin\left(\frac{u_i}{c}\right)}{\frac{u_i}{c}} & |u_i| \le c\pi \\ 0 & |u_i| > c\pi \end{cases}, \text{ avec } c = 0.339$$

- Poids de Huber :

$$\overline{P}_i = \begin{cases} P_i & |u_i| \le c \\ P_i \frac{c}{u_i} & |u_i| > c \end{cases}, \text{ avec } c = 2.0$$

- Poids de Tukey:

$$\overline{P}_i = \begin{cases} P_i [1 - (\frac{u_i}{c})^2]^2 & |u_i| \le c \\ 0 & |u_i| > c \end{cases}$$

Avec : c = 4.685 si l'écart type à priori de l'unité du poids σ_0 est connu, ou c = 6 si σ_0 est inconnu.

- Poids de IGG (Institute of Geodesy and Geophysics, Autriche) [Yang Y., 1999]:

$$\overline{P}_{i} = \begin{cases} P_{i} & |u_{i}| \leq c_{0} \\ P_{i} \frac{c_{0}}{|u_{i}|} (\frac{c_{1} - |u_{i}|}{c_{1} - c_{0}})^{2} & c_{0} \prec |u_{i}| \leq c_{1} \\ 0 & |u_{i}| \geq c_{1} \end{cases}$$

Avec: $1.0 \le c_0 \le 1.5$ et $2.5 \le c_1 \le 3.0$

4. Application : Ajustement du réseau cadastral de la ville d'Oran

L'application a porté sur l'ajustement du réseau cadastral de la ville d'Oran composé de 37 points nouveaux et de 8 points d'appui, Figure (1).

Les données géodésiques utilisées sont constituées de 204 observations de géodésie classique (194 directions et 10 distances).

Les coordonnées des points sont exprimées dans le système géodésique Algérien (Nord Sahara 1959), en projection UTM (fuseau 30).

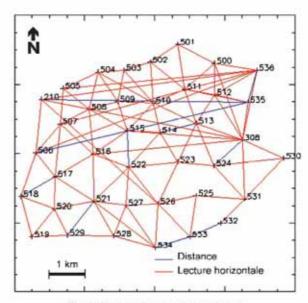


Fig. 1 Configuration du réseau cadastral de la ville d'Oran.

La méthodologie adoptée à travers cette application porte sur une étude comparative des résultats obtenus par les deux méthodes d'ajustement : Mestimateurs et Moindres carrés.

Le choix du modèle des poids équivalents appliqué dans la technique d'estimation M-estimateurs est celui de l'IGG (Institute of Geodesy and Geophysics).

Pour illustrer la puissance (robustesse) des Mestimateurs par rapport aux Moindres Carrés, des tests ont été effectués en bruitant volontairement quelques mesures.

Le tableau (1) représente les six observations qui ont été altérées (contaminées) arbitrairement par des erreurs :

Tableau 1. Valeurs simulées des grosses erreurs introduites dans le vecteur des observations.

Station	Pt Visé	Nº obs	Angle(Grades)	Erreur(dmgr)
502	501	9	180.3239	+200
510	513	50	104.8518	-400
308	513	96	9.6268	+100
519	520	131	38.1261	-200
Station	Pt Visé	Nº obs	Distance (m)	Erreur(m)
518	517	187	1132.765	-0.7
535	536	201	962.410	+0.5

5. Résultats et analyse

Les résultats obtenus sont analysés suivant la précision des coordonnées des points du réseau, les paramètres ajustées, les résidus et les ellipses d'erreurs.

5.1 Influence des mesures erronées sur la précision du réseau obtenue par les deux méthodes

Les écarts-types sur les paramètres ajustés (position et orientation) par les deux méthodes, Moindres carrés et M-estimateurs, sont fournis par les tableaux suivants :

Tableau 2. Ecarts-types sur les paramètres ajustés. Méthode des Moindres carrés.

	Cas d'observations non contaminées			Cas d'observations contaminées		
	σ _{v0} (dmgr)	σ _x (m)	σ _y (m)	σ _{v0} (dmgr)	σ _x (m)	σ _y (m)
Min.	5.4	0.017	0.015	16	0.049	0.042
Max.	22.4	0.117	0.101	64	0.333	0.288
Moy.	11.1	0.039	0.037	32	0.112	0.105

Tableau 3. Ecarts-types sur les paramètres ajustés. Méthode des M-estimateurs.

	Cas d'observations non contaminées			Cas d'observations contaminées		
	σ _{v0} (dmgr)	σ _x (m)	σ _y (m)	σ _{v0} (dmgr)	σ _x (m)	σ _y (m)
Min.	2.8	0.009	0.009	3	0.009	0.011
Max.	10.2	0.050	0.044	11	0.051	0.045
Moy.	5.6	0.020	0.019	6	0.021	0.021

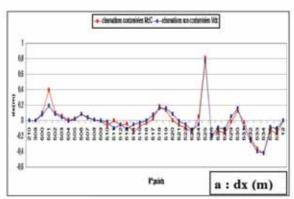
Les tableaux (2) et (3) montrent que dans le cas des traitements d'observations non contaminées, le réseau ajusté par les Moindres carrés ou Mestimateurs est de qualité suffisante pour les besoins du cadastre.

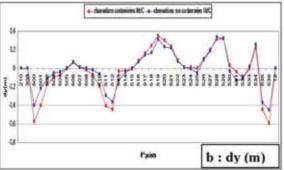
D'après le tableau (3), nous remarquons que la précision des paramètres ajustés à partir des observations non contaminées est presque identique à celle obtenue à partir des observations contaminées. Ceci est dû à l'effet des poids équivalents qui ont atténué l'influence des grosses erreurs.

Par contre, dans le cas de l'estimation par la méthode des Moindres carrés, les valeurs des écarts-types sur les paramètres ajustés (tableau 2) à partir des observations non contaminées sont différentes à celles fournies à partir des observations contaminées. La valeur moyenne des écarts-types sur la position et sur l'orientation est respectivement de \approx 11 cm et \approx 32 dmgr (cas des observations contaminées), et de \approx 4 cm et \approx 11 dmgr (cas des mesures non erronées).

5.2 Influence des mesures erronées sur les paramètres ajustés par les deux méthodes

Les figures suivantes illustrent les paramètres ajustés (position x, y et orientation V_0) par les deux méthodes, Moindres carrés (figures 2a, 2b et 2c) et M-estimateurs (figures 3a, 3b et 3c) :





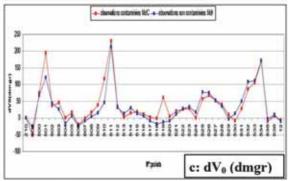
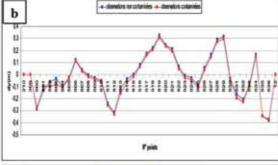


Fig. 2 Ajustement des paramètres (dx, dy, dV₀) par les moindres carrés. Cas des observations contaminées (rouge) et non contaminées (bleu).

La figure (2) montre que la méthode des moindres carrés est très sensible aux observations aberrantes. La différence entre les valeurs des paramètres ajustés dans les deux cas (observations contaminées et non contaminées) est de ≈ 20 cm en position et de ≈ 73 dmgr en orientation.

En revanche, les paramètres ajustés par la méthode des M-estimateurs (figure 3) n'ont subi aucune importante altération en présence d'observations aberrantes (contaminées).





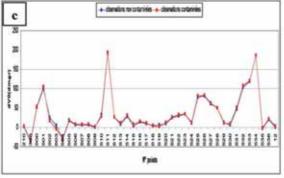


Fig. 3 Ajustement des paramètres (dx, dy, dV₀) par les M-estimateurs. Cas des observations contaminées (rouge) et non contaminées (bleu).

Les figures (3a, 3b, 3c) montrent bien qu'il n'y a pas de différences significatives entre les paramètres estimés par la méthode des Mestimateurs, dans les cas des observations non contaminées et contaminées. Ces différences sont au maximum de l'ordre de 4cm en coordonnées et de 8 dmgr en orientation.

5.3 Influence des mesures erronées sur les résidus obtenus par les deux méthodes

Les figures suivantes illustrent la répartition des résidus des observations non contaminées et contaminées obtenus par les moindres carrés (figure 4a) et les M-estimateurs (figure 4b).

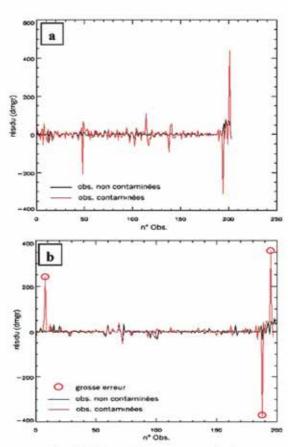


Fig. 4 Résidus des observations contaminées (rouge) et non contaminées (noir): (a) Moindres carrés (b) M-estimateurs.

La figure (4.a) montre clairement que les résidus obtenus dans le cas des observations contaminées sont plus importants que dans le cas d'observations non contaminées.

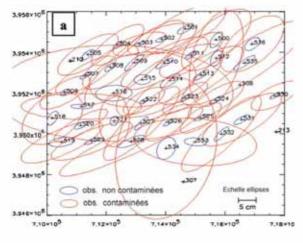
Ceci est dû au fait que la méthode des moindres carrés est très sensible aux obser-vations aberrantes.

En revanche, nous constatons que dans le cas des Mestimateurs (figure 4.a), l'effet de la présence de mesures aberrantes est affaibli par l'intégration des poids équivalents.

Nous remarquons aussi, que les observations contaminées ne sont pas noyées dans les résidus comme dans le cas des moindres carrés et qu'elles peuvent être facilement détectées.

5.4 Influence des mesures erronées sur les ellipses d'erreurs obtenues par les deux méthodes

Les ellipses d'erreurs des points du réseau fournies par les deux méthodes d'ajustement, Moindres carrés et M-estimateurs, sont représentées par les figures suivantes:



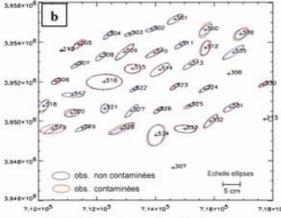


Fig. 5 Ellipses d'erreurs des points du réseau obtenues par : (a) Moindres carrés (b) M-Estimateurs. Cas des observations contaminées (rouge) et non contaminées (bleu).

La figure (5.b) qui illustre les ellipses d'erreurs obtenues lors de l'ajustement des paramètres par la méthode des *M-estimateurs* montre bien que la qualité des points du réseau, en terme d'ellipses d'erreurs, n'a pas été altérée par la présence de grosses erreurs observations contaminées.

Par contre, les ellipses d'erreurs obtenues lors du traitement des observations *contaminées* par la méthode des Moindres carrés (figure 5.a) ont une amplitude beaucoup plus importante à celle des ellipses d'erreurs fournies à partir des observations *non contaminées*.

6. Conclusion

Dans ce présent article, nous avons présenté la méthode des M-estimateurs et ses avantages par rapport à l'estimateur des moindres carrés.

Les résultats obtenus ont montré que dans le cas d'observations aberrantes, la qualité du réseau obtenue par la méthode des M-estimateurs est nettement meilleure que celle fournie par les moindres carrés. Cette dernière est très sensible aux grosses erreurs qui sont noyées dans les résidus. C'est une méthode peu stable en présence d'observations aberrantes et la qualité des paramètres estimés se détériore.

Par contre, dans le cas de la méthode des Mestimateurs, l'effet de la présence de mesures erronées est atténué par l'introduction des poids équivalents et la précision des paramètres ajustés n'est pas dégradée. En outre, cette technique d'estimation permet de détecter les observations aberrantes. Néanmoins, cette méthode qui est basé sur un processus itératif peut présenter des inconvénients en terme de délais d'exécution, dans le cas d'ajustement de gros réseaux,

Références bibliographiques

Coulot D., 2002: Initiations théorique et pratique à la méthode d'estimation par moindres carrés. Notes de cours, école d'été du GRGS, France, Septembre 2002.

Gui Q., Zhang J., 1998: Robust biased estimation and its applications in geodetic adjustments, Journal of Geodesy (1998) vol.72, pp. 430-435.

Huber C., Nikulin M., 1997: Remarques sur le maximum de vraisemblance, Questiio, vol. 21, 1 i 2, pp: 37-58, 1997.

Rehailia M.H., 1995: Modèles linéaires statistiques: modèles à effets fixes, Office des publications universitaires, Alger.

Vanicek P., Krakiwsky E.J., 1995: Geodesy: The concepts. Edition Elsevier Science B.V., second edition, 697p,1995. ISBN 0-444-87777-0.

Yang Y., Cheng MK., Shum CK., Tapley BD., 1999: Robust estimation of systematic errors of satellite laser range. Journal of Geodesy, Vol.73 n°7, pp: 345-349, 1999.

Youcai H., Metrikas S. P., 1995: On the design of robust regression estimators, Manuscripta geodetica n°3, vol. 20, pp:145-160, 1995.