

ANALYSE DE LA QUALITE DES RESEAUX GEOMETRIQUES APPLICATIONS AUX RESEAUX DU CERN

C. Lasseur
CERN - Genève

1. INTRODUCTION

Les déterminations de formes et de dimensions d'un objet soit directement soit à partir de points auxiliaires disposés suivant un réseau, doivent être liées à des critères caractérisant la qualité du modèle mesuré : les résultats attendus sont les coordonnées des points de l'objet et les paramètres fixant les points "fixes" (le datum) mais aussi les précisions des coordonnées qui sont corrélées avec les précisions des observations : ces paramètres doivent être également traités en tenant compte de leur fiabilité c'est à dire des contrôles mutuels des observations.

Les ordinateurs actuels permettent de développer des modèles mathématiques et statistiques qui sont des compléments obligatoires à tout calcul normal d'ajustement.

Les travaux géométriques nécessitant, en général, de telles analyses se ramènent à 2 types principaux :

- ceux se référant à la qualité de l'objet : ce dernier est connu géométriquement et on cherche à déterminer le "tel que construit" ("as built") par rapport au "tel que conçu" ("as designed").

- ceux se référant aux caractéristiques géométriques de l'objet : ce dernier n'est pas connu géométriquement et les mesures sont à la base de calculs de droites, de plans, de surfaces, de volumes ("best - fit") afin d'en restituer une image fidèle ("reverse engineering").

Les concepts et les définitions d'indicateurs de fiabilité ont été développés dès 1967 par Baarda : un système sera jugé fiable s'il peut détecter les erreurs grossières d'un ordre de grandeur donné avec un coefficient statistique de sécurité choisi à l'avance ... un système dit précis pouvant, en fait, ne pas être fiable du tout. Le programme L.G.C (Logiciel Général de Compensation) utilisé au CERN présente certaines facilités permettant l'analyse de la précision et de la fiabilité des résultats.

2. ANALYSE DES RESULTATS ET DES MESURES

2.1 analyse de la précision et intégration dans L.G.C

Les éléments de la matrice des covariances C_x , produit de la matrice des cofacteurs Q_x par le facteur de variance a posteriori S_0^2 issu des résidus après compen-



sation, sont à la base du calcul des erreurs standard des coordonnées, des ellipses et ellipsoïdes d'erreur pour chaque point calculé ; la précision des résultats dépend donc de la configuration des mesures, de la position des points par rapport au datum, et des observations.

La direction des demi-axes des ellipses absolues peut montrer certaines faiblesses globales du réseau : une direction du demi grand axe pointant vers les points fixes indique une possible faute de l'échelle alors qu'une direction du demi petit axe pointant vers les points fixes indique

une possible faute en azimut. Les grandeurs des demi-axes des ellipses sont les erreurs standard s maximum et s minimum, la situation idéale (cercle) étant obtenue pour des valeurs équivalentes et petites : l'ellipse d'erreur décrit en fait une région de confiance dans laquelle les coordonnées estimées du point sont contenues suivant une probabilité connue de 39.4 % pour une erreur standard de $1 S_0$. Pour l'ellipsoïde, la situation idéale est la sphère et la probabilité est de 19.9% pour une erreur standard de $1S_0$.

Les valeurs propres standard maximum et minimum sont les carrés des erreurs standard maximum et minimum pour une ellipse, pour un ellipsoïde ce sont les carrés des axes ; la moyenne géométrique des valeurs propres pour un point indique une valeur relative de la précision de ce dernier, permettant ainsi des comparaisons entre tous les points calculés d'un même réseau et, par exemple, de dresser des courbes d'égale précision. Pour l'ellipsoïde, les orientations spatiales de chaque axe - en fait son orientation dans chaque plan - sont représentées par les vecteurs propres.

La variance S_0^2 est un multiplicateur et sa valeur doit être testée afin d'estimer si elle est significativement différente de l'unité. La statistique F de $K = S_0^2 / s_0^2$ (s_0^2 variance unité) suit une loi de Fisher réduite à une distribution en CHI^2 ; pour un seuil de signification α , on calculera $F(m) = S_0^2$ si $S_0 > 1$ ou $F(m) = 1 / S_0^2$ si $S_0 < 1$ (m est le degré de liberté ou redondance).

Dans LGC (voir tableau 'Analyse des Précisions') les erreurs standard SX, SY, SZ suivant les axes de coordonnées sont systématiquement données ce qui permet une estimation plus directe de la précision des points. Les éléments de l'ellipse et de l'ellipsoïde (axes et vecteurs propres), les valeurs propres et la moyenne géométrique associée, sont fournis à la demande, certains paramètres tels que les orientations des axes de l'ellipsoïde n'étant pas interprétables facilement (ex : les

valeurs -0.510, -0.151 et 0.847 correspondant respectivement au plan XY, YZ et XZ montrent que l'axe est orienté plutôt vers les Z+, incliné vers les Y+ et plus légèrement vers les X- ...).

Le point de pourcentage F de la répartition en χ^2 est calculé pour le niveau de confiance de 95% : si $S_0 < \sqrt{F}$ (valeur critique), S_0 n'est pas significativement différent de 1. Dans le cas contraire, la matrice des cofacteurs Q_x sera multipliée par la variance calculée et, dans l'hypothèse d'absence de fautes importantes, on pourrait conclure que le modèle calculé n'est pas correct ou incomplet (erreur d'échelle non corrigée) ou que les mesures n'ont pas la précision attendue.

Pour chaque type d'observation et en fonction du degré de liberté et du niveau de certitude $1-\alpha$, une estimation de l'intervalle de certitude de la moyenne des résidus après compensation (valeur théorique nulle) permet de qualifier un systématisme éventuel ; de même une estimation de l'intervalle de certitude de la variance propre au type d'observation considéré permet de qualifier la présence de résidus importants. La largeur de ces intervalles peut fluctuer assez largement suivant la taille de l'échantillon et le résultat du test peut ne pas être assez significatif pour accepter ou rejeter l'hypothèse initiale (présence de fautes).

Les intervalles de certitude permettent de situer le type d'observations comportant des écarts importants : un examen plus fin des résidus (grandeur et répartition) et du quotient par l'erreur standard a priori (RES/SIG) pour chaque observation peut aider à la recherche de fautes.

L'examen, à la demande, des ellipses (et ellipsoïdes) d'erreurs relatives permet d'estimer l'homogénéité de la configuration : des valeurs importantes peuvent traduire soit des précisions moindres sur les points considérés soit des faiblesses locales.

Ces critères, basés sur la redondance générale, constituent des facteurs d'évaluation de la précision pour chaque point mesuré et pour chaque type d'observation sans pour cela caractériser la précision du modèle proposé, réseau ou objet. Trop globaux, ils ne permettent pas d'estimer l'influence de chaque observation sur la précision des points et doivent être complétés, sinon remplacés, par d'autres caractérisant chaque observation.

2.2 définition et analyse de la fiabilité

La fiabilité permet de tester, à partir de paramètres statistiques, la signification de l'évaluation de la précision et décrit la sensibilité, sa résistance, d'un réseau face à sa propre structure et face aux grosses fautes de mesures.

Baarda (1968) identifie 2 sortes de fiabilité :

- fiabilité interne : elle caractérise la possibilité de détection d'une faute pour chaque observation d'erreur standard a priori s_i ; elle s'exprime par le facteur $T_i = s_i / s_{vi}$ (s_{vi} erreur standard du résidu estimée à partir de la matrice des covariances des résidus C_v) et par la probabilité que l'observation correspondante ne contienne pas à un niveau de confiance donné une erreur supé-

rieure à une limite fixée à l'avance, par exemple 4 fois S_i .

- fiabilité externe : elle caractérise pour chaque observation l'influence d'une faute non détectée sur la position relative de 2 points ; elle s'exprime par le facteur $Y_i = s_{li} / s_{vi}$ (s_{li} erreur standard de l'observation estimée à partir de la matrice des covariances des observations C_l). La relation $Y_i^2 = T_i^2 - 1$ (Cross 1983) permet de calculer facilement Y_i et montre bien qu'une bonne fiabilité interne reflète une bonne fiabilité externe ; la relation établie par Pelzer (1979) $T^2_{\min} = 1 / n \sum (i, n) Y_i^2$ (n nombre d'observations) caractérise la fiabilité globale d'un réseau par un terme unique et permet de comparer la fiabilité de plusieurs réseaux entre eux.

2.3 principe de la détection d'erreurs et intégration dans L.G.C

Les tests d'analyse de fiabilité s'accompagnent de procédures de détection des grosses erreurs ("data snooping") ; elles reposent sur les méthodes statistiques permettant de tester si l'hypothèse H_0 suivant laquelle l'observation l_i ne contient pas de grosse erreur G_i est vraie avec une certaine probabilité ; cette dernière est composée en fait de 2 types de probabilités correspondant à 2 types de "fautes" du test :

- α est la probabilité de rejeter H_0 quand elle est juste ; $\alpha = 5\%$ signifie que l'on peut accepter H_0 à un niveau de confiance de 95%.

- β est la probabilité d'accepter H_0 quand elle est fautive ; $\beta = 10\%$ signifie que la puissance du test de détecter une grosse erreur est de 90%.

Appliquer les tests de détection d'erreurs en utilisant ces valeurs de α et β signifie que, à une probabilité de 95%, 90% des observations contenant une grosse erreur seront détectées.

Baarda, Pope et Cross ont démontré que la statistique $w_i = v_i / s_{vi}$ suivait une distribution en tau (pour une variance unité) définie à partir d'une distribution de Student ; w_i est comparé à la valeur critique de tau calculée pour un taux de rejet de 99% : quand $w_i >$ valeur critique, l'observation correspondante est susceptible d'une faute et la probabilité qu'elle n'en contienne en fait pas est 1%.

Dans LGC, les tests de fiabilité sont introduits sous forme de plusieurs indicateurs :

- ZI : indicateur de la fiabilité locale de l'observation ou redondance partielle, on a $ZI = (s_{vi} / s_i)^2 * 100 \% = 100\% / T_i^2$ et $\sum ZI =$ redondance totale. Ainsi $ZI = 0\%$ indique que l'observation n'est pas du tout contrôlée et une faute ne pourra pas être détectée quelque soit sa grandeur, $ZI = 100\%$ correspond à une mesure totalement contrôlée et superflue (cas d'une mesure entre 2 points fixes), $ZI = 50\%$ correspond à une mesure double. La valeur de ZI dépend de la géométrie du réseau (ex : chaque direction d'un triangle mesurée avec une égale précision a un ZI de 1/6 soit 16.6%) et de la précision des mesures (ex : une mesure très précise sera plus difficilement contrôlable et aura un ZI petit). Dans un bon réseau, les ZI doivent se situer entre 25% et 60% ; un grand nombre de ZI plus grands indique que les mesu-

res surabondantes sont trop nombreuses et que le réseau peut être optimisé et coûter moins cher! Dans le cas contraire, il faut compléter par des mesures complémentaires de même nature ou autre (distances, écarts).

- WI : erreur résiduelle normée calculée pour chaque observation, la valeur critique W_{lmax} est calculée suivant α (5% ou 1%).

- GI : grandeur probable d'une erreur grossière calculée ($GI = -vi / ZI$) quand $WI > W_{lmax}$; GI est d'autant plus "précis" que l'observation est mieux contrôlée (ZI grand).

- NABLA : plus petite faute que le test WI peut détecter à $(1-\beta)$ % avec un seuil de signification de α %, $NABLA = (svi / ZI) * \delta$ avec $\delta = \text{centile de } \alpha + \text{centile de } \beta$. Cette valeur sert de base dans le calcul des indicateurs de fiabilité externe.

- T : indicateur de fiabilité interne, 1 est la valeur maximum (voir plus haut).

- DELTY : utilisé dans le calcul de l'influence d'une faute ; $DELTY = \delta * Yi$ (Yi facteur de fiabilité externe) et $\Delta i = DELTY * S\Delta i$ est l'influence maximale d'une faute égale à NABLA sur la quantité dérivée ($S\Delta i$ erreur standard de cette quantité dérivée).

- OVERALL NETWORK RELIABILITY FACTOR : valeur numérique de la formule de Pelzer (voir plus haut) ; si un seul des indicateurs est indéterminé, cette valeur n'est pas donnée.

Voir les tableaux "Contrôle fiabilité réseau et fiabilité externe" et "Détection de fautes dans un réseau bien contrôlé".

3. UTILISATION DES INDICATEURS DE FIABILITE

Les méthodes statistiques de détection se basent sur l'hypothèse qu'il n'existe qu'une seule grosse faute mais elles sont indispensables lorsque, en particulier, les techniques automatiques de prise de données et de calcul en ligne rendent difficiles le contrôle et la recherche de fautes grossières sur le terrain.

Les indicateurs ZI, NABLA, T, DELTY et le facteur global ne se rapportent pas directement à de vraies observations et peuvent être utilisés comme facteurs de pré-analyse et d'optimisation d'un réseau. WI et GI se rapportent uniquement aux mesures, souvent déformées par les contraintes entre les points de rattachement : l'analyse

de ces indicateurs, en particulier la recherche de fautes, doit être effectuée sur un réseau libre ou à contrainte minimale ou encore dans un réseau où les coordonnées de rattachement sont introduites comme observations.

Ils ne donnent des valeurs correctes que si le choix des erreurs moyennes a priori est réaliste et que la redondance est suffisante : les tests statistiques ne peuvent qu'identifier les fautes provenant soit d'une erreur importante dans les mesures, soit d'un mauvais modèle mathématique (datum incorrect) sans pouvoir séparer l'une ou l'autre raison. Il est prudent de rechercher la cause des erreurs avant d'envisager le rejet total des observations fausses, le risque, dans ce cas, étant de manquer d'autres sources éventuelles de fautes : déformation du datum, facteur d'échelle incorrect etc ... Neptune a été découverte grâce à des observations a priori fausses de la trajectoire d'Uranus! La principale difficulté est lorsque plusieurs fautes s'annulent ou se compensent mutuellement : des procédures difficiles à mettre en oeuvre ont été proposées mais pas introduites jusque là (Stefanovic et Kok).

4. CONCLUSION

L'étude de la fiabilité d'un réseau s'établit donc à partir de ses éléments constitutifs, du modèle stochastique retenu (définition du datum et des poids moyens), de l'hypothèse alternative (fautes éventuelles du modèle) et du test statistique, complété par le niveau de confiance choisi, avec lequel le modèle sera testé après les mesures et les calculs.

BINGO (Kruck), CAP (Kotowsky), LOKAL (Gruendig), PANDA (Niemeier) sont d'autres programmes d'ajustement tridimensionnel permettant le calcul des indicateurs de fiabilité et la recherche d'erreurs.

L'enseignement et la vulgarisation de ces notions permettraient de développer l'utilisation systématique et raisonnée de ces indicateurs sans qu'ils aient à faire l'objet d'une normalisation.

BIBLIOGRAPHIE

P.A Cross Advanced least squares applied to position-fixing Polytechnic of East London

E.T.H Zurich Fiabilité dans la mensuration Journée d'étude du 16 mars 1990 rapport 169

C.Lasseur / C.MacCain Reliability testing and LGC note interne CERN/SU/EXP Mai 1991

ANALYSE DES PRECISIONS (mm et cc)

S0 A POSTERIORI = 0.71359 (VAL. CRIT = 0.8168) ECARTS-TYPES CALCULES / S0 A POSTERIORI

ELLIPSOIDES ABSOLUS (XY XZ YZ)

	<u>SX</u>	<u>SY</u>	<u>SZ</u>	<u>AXES</u>	<u>VAL.PROPRES</u>	<u>VECTEURS PROPRES</u>		
				0.16	0.0257	0.229	-0.973	-0.035
R1733	0.08	0.16	0.08	0.08	0.0057	-0.510	-0.151	0.847
MOY. GEO. = 0.010				0.08	0.0058	-0.829	-0.176	-0.531

ANGLES HORIZONTALS

		<u>OBS</u>	<u>SIGMA</u>	<u>DIR.CALC</u>	<u>RESIDU</u>	<u>RES/SIG</u>
L12	L1733	238.8000	10.00	199.367	0.4	0.0
L120	R173	187.0897	10.00	147.6566	0.8	0.1

RESIDU MOYEN = 0.0 CC : LIMITES DE CONFIANCE A 95.0 = (-1.3, 1.3)
 ECART-TYPE = 3.2 CC : LIMITES DE CONFIANCE A 95.0 = (2.5, 4.4)

DISTANCES ZENITHALES

		<u>OBS</u>	<u>SIGMA</u>	<u>CALCULE</u>	<u>RESIDU</u>	<u>RES/SIG</u>
L120	60Z151	100.9541	10.00	100.9526	14.5	1.4
L120	L173	102.6921	10.00	102.6925	-4.7	-0.5

RESIDU MOYEN = -1.3 CC : LIMITES DE CONFIANCE A 95.0 = (-2.5, 2.5)
 ECART-TYPE = 6.7 CC : LIMITES DE CONFIANCE A 95.0 = (5.3, 9.0)

DISTANCES THEODOLITE

		<u>OBS</u>	<u>SIGMA</u>	<u>CALCULE</u>	<u>RESIDU</u>	<u>RES/SIG</u>
L120	40Z151	11.3208	0.26	11.3211	-0.34	-1.33
L120	L1733	4.9998	0.22	4.9998	0.00	0.00

RESIDU MOYEN = 0.00 MM : LIMITES DE CONFIANCE A 95.0 = (-0.08, 0.08)
 ECART-TYPE = 0.21 MM : LIMITES DE CONFIANCE A 95.0 = (0.16, 0.28)

ERREURS RELATIVES

		<u>SIGMA L</u>	<u>SIGMA G</u>	<u>SIGMA R</u>	<u>SIGMA Z</u>	<u>SIGMA V</u>
R173	R120	0.06	7.8	0.06	0.08	10.0
R173	L173	0.0	6.5	0.05	0.08	9.4

ELLIPSOIDES RELATIFS

		<u>AXIS</u>	<u>VECTEURS PROPRES (XY XZ YZ)</u>		
R173	R120	0.0633	-0.025	-0.001	1.000
		0.0491	0.994	0.108	0.025
		0.0503	0.10	-0.994	0.002

CONTROLE FIABILITE RESEAU ET FIABILITE EXTERNE

SEUIL DE SIGNIFICATION TEST WI, ALPHA = 1.0 %
 PUISSANCE DU TEST POUR NABLA ET DELTY, (1-BETA) = 90.0 %
ZI PARTOUT BIEN CONTROLES ... S0 A POSTERIORI = 0.7136 (VAL.CRIT. = 0.8168)
 *** OVERALL NETWORK RELIABILITY FACTOR = 0.5137 ***

R1733 et L1733 'SIMPLEMENT' DETERMINES RISQUES DE FAUTES

OBSERVATIONS ANGL

		<u>OBS</u>	<u>SIGMA</u>	<u>RESIDU</u>	<u>ZI</u>	<u>WI</u>	<u>GI</u>	<u>NABLA</u>	<u>I</u>	<u>DELTY</u>
R120	R1733	130.1000	10.00	-0.43	6.4 **	-0.24		215.5	3.95	14.75
L120	L1733	238.8000	10.00	0.40	5.6 *	0.24		230.5	4.23	15.85

OBSERVATIONS ZENI

		<u>OBS</u>	<u>SIGMA</u>	<u>RESIDU</u>	<u>ZI</u>	<u>WI</u>	<u>GI</u>	<u>NABLA</u>	<u>I</u>	<u>DELTY</u>
R120	R1733	101.8700	10.00	0.00	0.0 **			INDETERMINEE	I N F I N I	
L120	L1733	102.6500	10.00	0.00	0.0 **			INDETERMINEE	I N F I N I	

OBSERVATIONS DTHE

		<u>OBS</u>	<u>SIGMA</u>	<u>RESIDU</u>	<u>ZI</u>	<u>WI</u>	<u>GI</u>	<u>NABLA</u>	<u>I</u>	<u>DELTY</u>
L120	L1733	4.99980	0.22	0.00	0.2 **	-0.24		41.9	25.61	98.71

**** INDETERMINATE OVERALL NETWORK RELIABILITY ****

INFLUENCE SUR LES QUANTITES DERIVEES

DIRECTIONS CALCULEES APRES COMPENSATION

			<u>SIGMA L</u>	<u>SIGMA G</u>	<u>SIGMA R</u>	<u>SIGMA Z</u>	<u>SIGMA V</u>
		L173		R120	51.2632		
		R173		R120	399.9014		
<u>POINT 1</u>	<u>POINT 2</u>						
L173	R120		0.07	6.19	0.07	0.07	6.40
R173	R120		0.06	7.84	0.06	0.08	10.10

OBSERVATIONS ANGL

POINT 1	POINT 2	RES	ZI	WI	NABLA	I	DELTY	INFLUENCE DE NABLA
L173	R120	0.75	61.9 !	0.12	63.7	1.27	3.03	Δ NABLA = 6.19 * 3.03 = 19CC
R173	R120	5.61	47.6	1.06	72.6	1.45	4.05	Δ NABLA = 7.84 * 4.05 = 32CC

CALCUL APRES COMPENSATION AVEC FAUTE = NABLA SUR OBSERVATIONS

L173	R120	51.2657	"EFFET"	25 C	19 CC CALCULE
R173	R120	399.9053	"EFFET"	39 CC	32 CC CALCULE

EFFETS SUR LES COORDONNEES

	COORD.COMPENS		COORD.COMPENS.AVEC FAUTE			
R120	2.82071	-12.37513	2.82106	-12.37523	Dx = 0.25 mm	Dy = 0.10 mm
L173	-2.71111	-17.69166	-2.71105	-17.69164	Dx = 0.05 mm	Dy = 0.02mm
R173	2.82877	-17.55613	2.82877	-17.55609	Dx = 0.00 mm	Dy = 0.04mm

R120 MOINS RESISTANT ...

DETECTION DE FAUTES DANS UN RESEAU BIEN CONTROLE

SEUIL DE SIGNIFICATION TEST WI, ALPHA = 1.0 %
 PUISSANCE DU TEST POUR NABLA ET DELTY, (1-BETA) = 90.0 %

MESURES VRAIES			MESURES FAUSSES			"FAUTES"
*ANGL	10.		*ANGL	10.		
R120	40Z151	59.9438	R120	40Z151	59.9408	- 30 cc
R173	40Z151	82.4375	R173	40Z151	82.4305	- 70 cc
*ZENI	10		*ZENI	10.		
R173	R120	102.7441	R173	R120	102.7501	+ 60 cc
*DTHEO	0.2	5.	*DTHEO	0.2	5.	
R173	R120	5.1858	R173	R120	5.1868	+ 1.0 mm

SO A POSTERIORI = 1.23661 > VALEUR CRITIQUE FAUTES ?
GI ... CORRECTIONS A APPORTER

	OBS	SIGMA	RESIDU	ZI	WI	GI	NABLA	I	DELTY
R120	40Z151	59.9408	0.00	-12.28	44.5	-1.46	45.9	1.50	4.31

FAUTE < NABLA ... DIFFICILEMENT DETECTABLE

	OB	SIGMA	RESIDU	ZI	WI	GI	NABLA	I	DELTY
R173	40Z151	82.43050	10.00	-22.00	33.8	-3.06 **	65.1	53.7	1.72 5.40

FAUTE DETECTEE - 65 CC POUR - 70 CC (RESIDU -22.00 FAUTE?)

OBSERVATIONS ZENI

	OBS	SIGMA	RESIDU	ZI	W	G	NABLA	I	DELTY
R173	R120	102.75010	10.00	28.30	42.0	3.53 **	-67.3	48.1	1.54 4.53

FAUTE DETECTEE + 67 CC POUR + 60 CC (RESIDU 28.3 FAUTE?)

OBSERVATIONS DISTTHE

	OBS	SIGMA	RESIDU	ZI	WI	GI	NABLA	I	DELTY
R173	R120	5.18688	0.23	0.80	91.6 !	2.98 **	-0.9	0.7	1.04 1.17

FAUTE DETECTEE + 0.9 MM POUR + 1.0 MM (RESIDU 0.80 FAUTE...)