THE BAYESIAN INFERENCE AND OUTLIERS DETECTION
IN PHOTOTRIANGULATION

J. F. Custódio da Silva
Department of Cartography
IPEA/UNESP - C.P. 957
19100 - Pres. Prudente - SP.
Brasil

J. B. Lugnani
Department of Geosciences
CPGCG/UFPr - C.P. 19010
81540 - Curitiba - PR
Brasil

ISPRS Commission III

ABSTRACT

The Bayesian inference along with the step-by-step technique allows foresee a possibility to increase the efficiency and reliability of phototriangulation. It also appears to improve the strength of the block, in order to increase the economy with respect to control points.

The used methodology consists on simulation of a 25 points mesh, from which a block of aerial photos is generated. Random and gross errors are added to the ideal photocoordinates. The LS adjustment, based on concept of Bayesian inference, is made. The Pope's method is used to detect and eliminate gross errors.

The results show that for few outliers the approach works well; when the quantity increases the, efficiency decreases. By other hand, the better the quality of a priori parameters, the better the Z-coordinates of the object points.

Content

1. INTRODUCTION

The bundle block adjustment mixes photogrammetric, statistical and data processing concepts. Usually, the data are split into observables and parameters (unknowns). The photocoordinates and control points are observations and exterior orientation of photos and pass point coordinates are unknowns. The functional model (colinearity equations) links the observations and parameters. Because of the superabundance, the adjustment of observations is based on method of least squares, adopting the parametric technique with weight constraints to control points.

In phototriangulation, one worries about reducing and preparing the data because of its considerable mass. It is almost impossible to think of phototriangulation without gross errors (outliers) because of its own process (digitation, recognition of points etc). The quality control is essential and it must be as rigorous as possible. The bundle block adjustment is sensibly damaged in presence of gross errors; the system reliability decreases and the end product becomes doubtful.

By other hand, it is fair to admit a previous knowledgement about the unknowns in despite of their approximation. In fact, in practice, one always knows something about them. In principle, the question is the nature and extent of its knowledgement. If the nature of the "a priori" knowledgement about parameters is subjective the question rents to Bayesian inference. In this case, the previous information is "got" by researcher based on his own experience (BECKMAN & COSTA NETO |2|, BOSSLER |3|). By other hand, the previous knowledgement from objective nature, given by experimental evidence (direct observation of parameter), is treated under generalized adjustment (SCHMID & SCHMID |9|. The extent of the previous knowledgement of unknowns, subjective or objective, is incorporated to the phototriangulation system as a vector of "a priori" parameters and the corresponding dispersion matrix (variance-covariance matrix) and then weight matrix.

It is important to note the realization of reference system for phototriangulation can be done by the parameter vector and its corresponding dispersion matrix. This means that the control points are avoided to inform the reference system at least temporarily. The great advantage links to removal the conjecture that gross errors, possibly existing in control points, influence the phototriangulation results while in depurating the observations. This process assures greater reliability to solution.
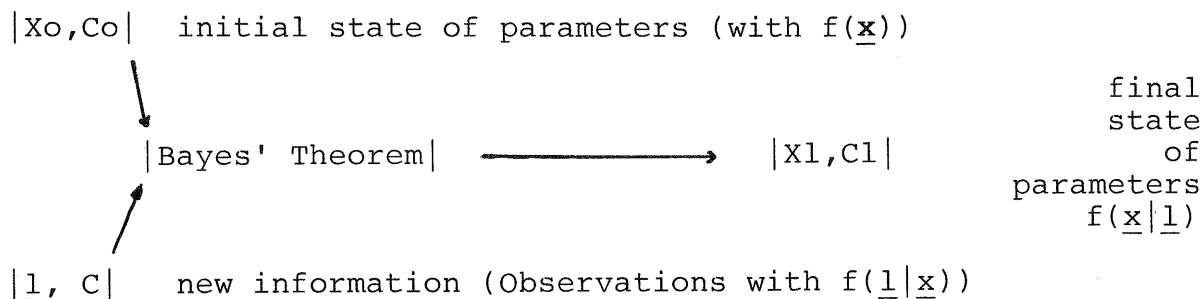
2. BAYESIAN INFERENCE

The process of statistical inference seeks to estimate parameters (particularly means and variances) from experimental data. In case of Bayesian inference, it occurs the estimate to the new state of parameters based on acquaired knowledgement (experiment realization) and on initial state of parameters. Therefore, it is an estimate revision deriving from a new state of information.

In figure 1, |Xo, Co| and |Xl, Cl| represent the same parameter vector and respective covariance matrices, although in distinct states, that is, with previous and posterior estimates to the experiment realization ($\underline{1}$). Taking the posterior normal probability density function, its Bayesian estimates follow three distinct situations:
(a) unknown mean and known variance
(b) known mean and unknown variance
(c) both mean and variance unknown.

Figure 1: Ilustration of Bayesian inference concept

|Xo,Co| initial state of parameters (with $f(\underline{x})$)

|Bayes' Theorem| $\longrightarrow$ |Xl,Cl|

final
state
of
parameters
$f(\underline{x}|\underline{1})$

|$\underline{1}$, C| new information (Observations with $f(\underline{1}|\underline{x})$)

For the first case (a), the expressions for estimators are presented in abundance at specialized literature (eg: |2|, |3|, |4| and |13|). BECKMAN & COSTA NETO |2| present the solution for the (b) situation. And for the last case, that one is the nearest to photogrammetric data reality, it is cited the references |2| and |3|.

BOSSLER |3| gives the following three Bayesian estimates:

1) For the classic linear(ized) model:

$$\underline{x} = - (A'PA^{-1}).A'P\underline{1}$$
$$\hat{C}_x = \hat{\sigma}_o^2.(A'PA)^{-1} \tag{1}$$

$$\hat{\sigma}_o^2 = \underline{v}'P\underline{v}/(n-u-2)$$

2) For the linear(ized) model with relevant previous information:

$$\underline{x} = - (A'PA + P_x)^{-1}.(A'P\underline{1} + P_x.\underline{1}_x)$$
$$\hat{C}_x = \hat{\sigma}_o^2 (A'PA + P_x)^{-1} \tag{2}$$

$$\hat{\sigma}_o^2 = (\underline{v}'P\underline{v} + \underline{v}_x'P_x \underline{v}_x)/(n-2)$$

$$\hat{\sigma}_o^2 = \text{common to both set of data.}$$

The lost of two degrees of freedom to estimate the posterior "variance of unit weight" (VUP) is oughted to the two unknown parameters of the postulated function (the mean vector and variance-covariance matrix).

3) Yet for the linear(ized) model with relevant previous information but with distinct and unknown VUPs to each set of data:

$$\underline{x} = - (A'PA/\hat{\sigma}_o^2 + P_x/\hat{\sigma}_x^2)^{-1}.(A'P\underline{l}/\hat{\sigma}_o^2 + P_x\underline{l}_x/\hat{\sigma}_x^2)$$

$$\hat{C}_x = (A'PA/\hat{\sigma}_x^2 + P_x/\hat{\sigma}_o^2)^{-1} \tag{3}$$

$$\hat{\sigma}_o^2 = \underline{v}'P\underline{v} / (n-u) \quad ; \quad \hat{\sigma}_o^2 = \underline{v}_x'P_x\underline{v}_x / (n-u)$$

The above expressions for $\hat{C}_x$, $\hat{\sigma}_o^2$, and $\hat{\sigma}_x^2$ are only approximated. The rigorous solutions are not possible (VANICEK & KRAKIWSKY |12|).

3) RELIABILITY

Reliability relates to the adjustment quality with respect to detecting "errors" on mathematical model. Such errors can be gross or systematic. Systematic errors are directly related to functional model adequacy, that means the "ideal" functional model eliminates systematic errors from the system. Then the permanence of such errors means that the functional model is not qualified to represent the reality.

So, reliability is understood as the ability in detecting and eliminating gross errors apart from observations, with help of a statistical test, when estimating parameters. In this context, it is internal reliability, which can be measured by the outlier to be detected by the statistical test in the observation. The external reliability is described by the influence of undetectable errors on object coordinates. It is assumed that a statistical test was aplicated and all standardized residuals are under a critical value, which depends on the significance level of the test (KOCH |5|).

The utility of both (internal and external) reliability concept is perceptible when there is more than one outlier in the set of observations. If the statistical test fails, that is it occurs type II error, the meaning is that the adjustment has a certain internal reliability degree (the critic value), although at least an undetected outlier is present among observations. In this case, the investigation on external reliability indicates the effect of this outlier on object point coordinates. In this present work, the used statistical test is the Pope's method (POPE |8|), which rejects the corresponding observation to approximated standardized residual $\tilde{w}_i > c$. The critical value "c" is obtained from tau-distribution. The approximated standardized residual $\tilde{w}_i$ is computed by.

$$\tilde{w}_i = |v_i| / \tilde{s}_{v_i} \quad ; \quad \tilde{s}_{v_i} = \hat{\sigma}_o^2.(\overline{m}/p) \quad ; \quad \overline{m} = (n-u)/n \tag{4}$$

$\hat{\sigma}_o^2$: the same as before, n: number of observations, u: number of unknowns and p: weight for all photocoordinates.

4. METHODOLOGY

Firstly, it was defined a 25 idealized object points grid on a few uneven terrain and regularly distributed. Ten aerial photos were idealized at 1:10000 approximated scale, in order to guarantee circa of 60% and 30% end and side lap, respectively. In some photos angular movements lesser than $2°$ sexagesimal were permitted. There are four photos with six image points and six photos with nine image points. Therefore, there are 156 observation equations and 135 unknowns (parameters) completing 21 superabundant equations. If all parameters are weighted justified by a previous knowledgement, it arrives to an equation system with 156 degrees of freedom, exactly the number of observation equations.

The computed photocoordinates under the rigorous colinearity assumption (without systematic errors) represent ideal values to which were added random errors not greater than 3 standard deviation. One standard deviation (σ) was made equal to 5 μm.

The implemented solution was the set of equations (3). The previous variances for parameters were $10 \ m^2$ and $100 \ m^2$. Lesser variances grant to the block geometry greater strengthening and conversely.

5. RESULTS AND DISCUSSION

The table 1 synthetizes the results referent to the fourteen realized experiments, which were split in five groups as showed by column 1. The columns 2, 3 and 4 mean "value of gross error" (VEG), that is the outlier size, "previous variance of parameter" (VPP) and the "quantity of gross errors" (QEG), respectively. The headed columns by DLAEG (automatic detection and localization of gross errors) represents C corret detections, E wrong detections and F faults, that is nor correct nor wrong detection. The occurrence of E can be compared to type I error" and F to the "type II error", The columns 8 and 9 represent degrees of freedom and the "a posteriori" VUP referent to photocoordinates (as in (3)). The 10th column shows the absolute value of the greatest remainder residual, that is all residuals lesses than refer to the non rejected observations by the statistical test, what is effectively a quality control indicator. The last column shows the iteration number until to finishing the outliers detection plus (+) the iteration number after introduction the control points.

The G1 and G2 groups present the occurrence of 6σ outliers (30 μm) and G3 and G4 groups 18σ outliers (90 μm). The later represents approximately the inferior limit of detection capacity of step-by-step technique. The G1 and G3 groups and the G0A test represent the greatest geometric strength situation for the block when comparing to G2 and G4 groups and to G0B test, because of VPP. In all groups the outliers are in small, middle and great quantities.

Table 1: Synthesis of the results of realized experiments

| | | | | | RESULTS | | | | | |
| CHARACTERISTICS | | | | DLAEG | | | | STATISTICS | | |
Group (1)	VEG (2)	VPP (3)	QEG (4)	C (5)	E (6)	F (7)	GL (8)	VUP (9)	max V (10)	NI (11)
G0A	0	10 m²	0	–	–	–	21	1.89	9.4	3+2
G0B	0	100 m²	0	–	–	–	21	3.68	9.8	3+2
G1 A			2	2	0	0	19	2.51	7.4	5+2
G1 B	6σ	10 m²	6	3	1	3	17	3.31	7.3	11+2
G1 C			10	1	0	9	20	11.22	16.8	5+2
G2 A			2	2	0	0	19	1.27	7.3	7+2
G2 B	6σ	100 m²	6	3	0	3	18	4.20	9.0	9+2
G2 C			10	1	0	9	20	10.98	16.4	5+2
G3 A			2	2	0	0	19	1.32	7.4	5+2
G3 B	18σ	10 m²	6	5	5	1	11	1.24	4.8	17+2
G3 C			10	5	0	5	16	12.01	20.1	11+2
G4 A			2	2	0	0	19	1.27	7.4	5+2
G4 B	18σ	100 m²	6	4	3	2	14	2.11	7.5	13+2
G4 C			10	5	0	5	16	11.87	20.1	11+3

The two first lines of the table present the characteristics and results of experiments without outliers in distinct situations of previous knowledgement of parameters. The remarkable difference is the VUP result, which doubles when VPP increases ten times. In other words, the VUP varies inversely to geometric strength of the block.

The present method of removal of probable inconsistent observations may be interpreted as an automatization of the ideia of rejecting the associated observation to the largest residual in a repetitive process until the solution appears satisfactory to the researcher. However, because of automatization it is possible that more than one suspicious observation can be eliminated. More important is the fact that the rejection rule substitutes the researcher to decide on satisfactory values to finish searching possible outliers.

When the number of outliers increases, DLAEG becomes worse. It is possible to conclude that Pope's method, used after Bayesian solution, functions when there are few outliers, it functions partially when the number is mean and it does not function when the quantity is great.

When the VPP decreases, it corresponds to the increasing the previous information about the parameters and the consequence is that DLAEG improves. However, it is necessary to pay attention to B and C situations of table 2. The B situation is undesirable because it represents a fault of an observation in a superabundant set and the C situation is undesirable and intolerable because it represents the inclusion of an outlier as a "quasi-observation".

Table 2: "Quasi-observation" à weight versus quality

"quasi-observation	quality	
	good	bad
weight great	A	C
weight small	B	D

The increasing of the outlier value in the sense of leaving the system resolution limit of DLAEG to sensibility regions of the statistical teste is benefic, because the detection is improved. By otler hand, the occurrence of very great outlier value is prejudicial to the statistical tests. Therefore, the step-by-step technique is necessary.

The following tables present the experiments referring to object points coordinates. The table 3 synthetizes the results bsed on concepts of "mean square error" (EMQ) and "estimated standard deviation" (DPE).

The first column of both tables 3 and 4 identifies the experiments, which characteristics are the same as table 1. The columns 8, 9 and 10 of table 3 expresses the "EMQ/DPE" ratio, that means the deterioration, or conservation, of external precision (accuracy) to internal precision. By conservation it is understood the ratio value lesser than 1. This could perhaps be consequence of the approximated estimate of C. Other aditional reason could be the neglecting of statistical correlations of the parameters.

The table 4 presents the root square in X- and Y-axes (EMQ-X and EMQ-Y) in standard deviation (σ = 5 μm) units of photocoordinates. The EMQ-Z, in this table, is presented in units to a thousand meters of height flight (Hv).

Table 3: Synthesis of EMQ and DPE of object point coordinates

	EMQ (m)			DPE (m)			EMQ/DPE		
G	X	Y	Z	X	Y	Z	X	Y	Z
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
G0A	0.11	0.10	0.70	0.10	0.11	0.55	1.1	0.9	1.3
G0B	0.10	0.10	2.16	0.10	0.11	0.56	1.1	0.9	1.4
G1A	0.11	0.12	0.60	0.10	0.11	0.55	1.1	1.1	1.1
G1B	0.07	0.11	0.60	0.10	0.12	0.55	0.7	0.9	1.1
G1C	0.14	0.10	0.82	0.10	0.11	0.55	1.4	0.9	1.5
G2A	0.11	0.12	1.38	0.10	0.11	1.59	1.1	1.0	0.9
G2B	0.10	0.12	1.39	0.10	0.12	1.59	1.0	1.1	0.9
G2C	0.14	0.11	2.88	0.10	0.11	1.56	1.4	1.0	1.9
G3A	0.11	0.12	0.60	0.10	0.11	0.55	1.1	1.1	1.1
G3B	0.21	0.69	0.51	0.24	0.72	0.60	0.9	1.0	0.9
G3C	0.16	0.21	0.68	0.10	0.11	0.55	1.6	1.9	1.2
G4A	0.11	0.12	1.38	0.10	0.11	1.59	1.1	1.0	0.9
G4B	0.14	2.26	1.32	0.11	2.24	1.60	1.2	1.0	0.8
G4C	0.16	0.23	1.41	0.10	0.12	1.59	1.6	2.0	0.9

Again, the two first lines of tables present the results of experiments without outliers, modifying the previous knowledgement of parameters. The EMQ and DPE values referent to X- and Y-coords (table 3) are practically equivalent. Meantime, younder differing between themselves, the EMQ-Z and DPE-Z are significantly greater than those for X and Y, as expected by the way. The mentionned difference is greater when the comparison is done between EMQ and DPE of the experiments. In this case, the deterioration ratio of external precision is approximately 3.1 and 2.9 for internal one, when the VPP increases ten times. The table 4 confirms this verification.

Table 4: Synthesis of EMQ-X e Y in units of σ and EMQ-Z in ‰ Hv.

Group	EMQ in units of σ		EMQ in ‰ Hv
	x	y	
(1)	(2)	(3)	(4)
G0A	2.3	1.9	0.7
G0B	2.2	1.9	2.2
G1A	2.1	2.5	0.6
G1B	1.5	2.1	0.6
G1C	2.8	1.9	0.8
G2A	2.1	2.4	1.4
G2B	2.0	2.4	1.4
G2C	2.7	2.2	2.9
G3A	2.1	2.5	0.6
G3B	4.3	13.8	0.5
G3C	3.2	4.2	0.7
G4A	2.1	2.4	1.4
G4B	2.8	45.3	1.3
G4C	3.2	4.5	1.4

The figure 2 shows the Z-discrepancies at each object point because of VPP change. It is seen that under greater geometric strength of the block the discrepancies are lesser. It is also seen that error propagation (discrepancies) is cross to the strips, as expected. This occurs because there are only 4 Z-control points ate corners. Meantime, the stablished geometric strength by VPP can control partially the error propagation.

Figure 2: Z-discrepancies at 25 object points
     a: ... when VPP = 10 m^2.
     b: ... when VPP = 100 m^2.

a	0.0	0.1	0.0	0.2	0.0 m
	0.6	0.5	0.5	0.7	0.6
	1.2	1.1	1.1	1.1	1.1
	0.7	0.6	0.6	0.5	0.5
	0.0	0.1	0.2	0.0	0.0
b	0.0	0.1	0.1	0.2	0.0 m
	1.8	1.7	1.7	1.9	1.8
	3.6	3.5	3.5	3.4	3.5
	1.9	1.8	1.8	1.7	1.7
	0.0	0.1	0.2	0.0	0.0

In general, when VPP is lesser, the trend to conserve the external and internal precision is perceptive in table 3, considering only Z-coords. When X-and Y-coords, the results show that the change of parameter state is indifferent for external or internal precision, although there are exceptions (G3B, G3C, G4B and G4C, tables 3 and 4).

6. CONCLUSION

The bayesian in inference is presented as a tool to solve the bundle block adjustment, based on a certain previous information degree about parameters, in general, available after step-by-step technique for automatic detection and localization of outliers (DLAEG). It allows to stop using control points at least temporarily. The great advantage is the residual analysis of photocoordinates without the influence of control points and its probable outliers. Under the minimum quantity of control points, the Bayesian inference contributes to establish the geometric strength of the block and to control partially the error propagation, specially Z-coordinates of object points. The Pope's method is efficient for DLAEG if the quantity of outliers is small (circa of 1.3%). The form that Pope's method was implemented represents the automatization of the idea of rejecting the associated observation to the largest residual in a repetitive process.

The external (EMQ) and internal (DPE) precision are worse as the number of outliers and VPP increase. These increments can influence the EMQ/DPE ratio when the values vary differently. Therefore, the reliability of phototriangulation under Bayesian inference is mostly influenced by the number of outliers and VPP.

Although the results can show high reliability (the greatest remainder residual from 1 to 2σ) if the algorithm fails that will be a false information.

The external and internal precision referent to Z-coords trend to be conservatives, if the geometric strength of the block is greater, that is, VPP lesser, independent from the number and the size of the outliers. The same is worth when there are not outliers.

Some questions remaind to be investigated: Bayesian inference along with Pope's method and robust estimation, in order to improve DLAEG, using for example $\sigma_{v_i} = f(p_i)$ and $p_i = g(v_i)$; variation of "a priori" VUP of the parameters, σ_x^2 to turn compatible the relantionship between the weights of parameters and photocoordinates. Idem for σ^2; to improve the approximated solution to the estimates of C_x, $^\circ\sigma_x^2$ and $\sigma_\circ^2$, in order to reach better estimates for external precision.

7. BIBLIOGRAPHIC REFERENCES

|1| BAYES, T. A essay towards solving a problem in the doctrine of chances. **Biometrika**, 45:293à315. 1958.

|2| BEKMAN, O.R. & COSTA NETO, P.L.O. **Análise estatística da decisão.** São Paulo, Edgard Blucher, 1980. 124 p.

|3| BOSSLER, J.D. **Bayesian inference in geodesy.** s.l. 1972. 79p. Tese, Doutorado. The Ohio State University.

|4| HOGG, R.V. & CRAIG, A.T. Estimation. In: ___ . **Introduction to the mathematical statistics.** 4 ed. New York, Macmillan, 1978. p.200-34.

|5| KOCH, K.R. **Statistical analysis of geodetic data with applications for the deformation analysis.** Curitiba, Universidade Federal do Paraná, Curso de Pós-Graduação em Ciências Geodésicas, 1985. 125p.

|6| LUGNANI, J.B. **Introdução à fototriangulação.** Curitiba, UFPr, 1987. 134p.

|7| MIKHAIL, E.M. **Observations and least squares.** New York, DunDonnelley, 1976. 497p.

|8| POPE, A.J. The statistics of residuals and the detection of outliers. **NOAA Technical Report NOS 65 NGS** 1:1-133, 1976.

|9| SCHMID, H.H. & SCHMID, E. A generalized least squares solution for hybried measuring systems. **The Canadian Surveyor,** 19(1):27-41, 1965.

|10| SILVA, J.F.C. **A inferência bayesiana e a detecção e localização automática de erros grosseiros em fototriangulação por feixes de raios;** Curitiba, 1987. 101 p. Tese, doutorado. Universidade Federal do Paraná. Curso de Pós-Graduação em Ciências Geodésicas.

|11| STEFANOVIC, P. Pitfalls in blunder detection techniques. In: INTERNATIONAL CONGRESS OF PHOTOGRAMMETRY AND REMOTE SENSING, 14., Hamburg, 1980. **International Archieves...** Hamburg, International Society of Photogrammetry and Remote Sensing, 1984. v.25, t.A3a, p.10-5.

|12| VANICEK, P. & KRAKIWSKY, E.J. Formulation and solving of problems. In: ___ . **Geodesy:** the concepts. Amsterdam, North-Holland, 1982. p. 242-83.

|13| WONNACOTT, T. & WONNACOTT, R. **Introductory statistics.** 3 ed. New York, John Willey, 1977. 650p.