GeoInformation in Europe, M.A. Gomarsca (ed.)
 ISBN 9789059660618

High resolution satellite imagery orientation accuracy assessment by leave-one-out method: accuracy index selection and accuracy uncertainty

M.A. Brovelli, E. Realini
DIIAR - Politecnico di Milano, Polo Regionale di Como – v. Valleggio, 11 – 22100 Como, Italy. Tel ++39-0313327517, Fax ++39-0313327519, e-mail <maria.brovelli,eugenio.realini>@polimi.it

M. Crespi, F. Fratarcangeli & F. Giannone
DITS - Area di Geodesia e Geomatica, Sapienza Università di Roma – v. Eudossiana, 18 – 00184 Roma, Italy. Tel ++39-0644585068, Fax ++39-0644585515, e-mail <mattia.crespi, francesca.fratarcangeli, francesca.giannone>@uniroma1.it

Keywords: orientation accuracy assessment, leave-one-out method, accuracy index selection

ABSTRACT: The Leave-one-out cross-validation (LOOCV) was recently applied to the evaluation of High Resolution Satellite Imagery orientation accuracy and it has proven to be an effective method alternative with respect to the most common Hold-out-validation (HOV), in which ground points are split into two sets, Ground Control Points used for the orientation model estimation and Check Points used for the model accuracy assessment.

On the contrary, the LOOCV applied to HRSI implies the iterative application of the orientation model using all the known ground points as GCPs except one, different in each iteration, used as a CP. In every iteration the residual between imagery derived coordinates with respect to CP coordinates (prediction error of the model on CP coordinates) is calculated; the overall spatial accuracy achievable from the oriented image may be estimated by computing the usual RMSE or, better, a robust accuracy index like the mAD (median Absolute Deviation) of prediction errors on all the iterations.

In this way it is possible to overcome some drawbacks of the HOV: LOOCV is a reliable and robust method, not dependent on a particular set of CPs and on possible outliers, and it allows us to use each known ground point both as a GCP and as a CP, capitalising all the available ground information. This is a crucial problem in current situations, when the number of GCPs to be collected must be reduced as much as possible for obvious budget problems.

The fundamental matter to deal with was to assess how well LOOCV indexes (mAD and RMSE) are able to represent the overall accuracy, that is how much they are stable and close to the corresponding HOV RMSE assumed as reference.

Anyway, in the first tests the indexes comparison was performed in a qualitative way, neglecting their uncertainty. In this work the analysis has been refined on the basis of Monte Carlo simulations, starting from the actual accuracy of ground points and images coordinates, estimating the desired accuracy indexes (e.g. mAD and RMSE) in several trials, computing their uncertainty (standard deviation) and accounting for them in the comparison.

Tests were performed on a QuickBird Basic image implementing an *ad hoc* procedure within the SISAR software developed by the Geodesy and Geomatics Team at the Sapienza University of Rome.

The LOOCV method with accuracy evaluated by mAD seemed promising and useful for practical cases.

1 INTRODUCTION

Interest in high-resolution satellite imagery (HRSI) is spreading in several application fields, at both scientific and commercial levels. A fundamental and critical goal for the geometric use of this kind of imagery is their orientation and orthorectification, the process able to correct the geometric deformations they undergo during acquisition.

One of the main objectives of the studies about orthorectification is the definition of an effective methodology to assess the spatial accuracy achievable from oriented imagery. Currently, the most used method (Hold-out-validation – HOV) to compute this accuracy just consists in partitioning the known ground points in two sets, the first used into the orientation-orthorectification model (GCPs – Ground Control Points) and the second to validate the model itself (CPs – Check Points); in this respect, the accuracy is just the RMSE of residuals between imagery derived coordinates with respect to CPs coordinates.

However this method has some drawbacks: it is generally not reliable and it is not applicable when a low number of ground points is available. First of all, once the two sets are selected, accuracy estimate is not reliable since it is strictly dependent on the points used as CPs; if outliers or poor quality points are included in the CPs set, accuracy estimate is biased. In addition, when a low number of ground points is available, almost all of them are used as GCPs and very few CPs remain, so that RMSE may be computed on a poor (not significant) sample. In these cases accuracy assessment with the usual procedure is essentially lost.

In a recent paper (Brovelli *et al.* 2006) we proposed an alternative to the previously described method to perform a spatial accuracy assessment, that is the use of the Leave-one-out cross-validation (LOOCV) method for the accuracy assessment of the HRSI orientation. Here we discuss how well LOOCV indexes (mAD and RMSE) are able to represent the overall accuracy, that is how much they are stable and close to the corresponding HOV RMSE, assumed as a reference. In fact, in the first tests the indexes comparison was performed in a qualitative way, neglecting their uncertainty. Now the analysis has been refined on the basis of Monte Carlo simulations, starting from the actual accuracy of ground points and images coordinates, estimating the desired accuracy indexes (e.g. mAD and RMSE) in several trials, computing their uncertainty (standard deviation) and accounting for them in the comparison.

To test the proposed method, we modified the SISAR software, developed by the Geodesy and Geomatics Team at the Sapienza University of Rome to perform rigorous orientation of HRSI, integrating it with modules suited to carry out iteratively the core orientation algorithm with different points configurations and performing Monte Carlo simulations to assess the uncertainty of the accuracy indexes.

In Section 2 LOOCV and Monte Carlo simulation principles are briefly recalled; in Section 3 the SISAR orientation model is shortly outlined; in Section 4 the results of a test considering a QuickBird Basic image are presented and discussed.

2 LOOCV AND MONTE CARLO SIMULATIONS

The Leave-one-out cross-validation (LOOCV) method (Stone 1974) is a special case of the general k-fold cross-validation method, which involves the partitioning of the original data set in k subsets of equal size (approximately). The model is trained k times, using each subset in turn as the test set, with the remaining subsets being the training set. The overall accuracy can be obtained averaging the accuracy values computed on each subset.

In particular, LOOCV is k-fold cross-validation computed with $k = n$, where n is the size of the original data set. Each test set is therefore of size 1, which implies that the model is trained n times.

This method applied to HRSI involves the iterative application of the orientation model, using all the known ground points as GCPs except one, different in each iteration, used as a CP. In every iteration the residual between imagery derived coordinates with respect to CP coordinates (prediction error of the model on CP coordinates) is calculated; the overall spatial accuracy achievable from the oriented image may be estimated by calculating the usual RMSE or, better, a robust accuracy index like the mAD (median Absolute Deviation) of the prediction errors on all the iterations.

In this way we solve some of the drawbacks of the classical HOV procedure: it is a reliable and robust method, not dependent on a particular set of CPs and on outliers, and it allows us to use each known ground point both as a GCP and as a CP, capitalising all the available ground information.

Monte Carlo simulations (Millard 2001) is a well known method for investigating the distribution of a random variable by simulating random numbers. Usually the random variable of interest, say Y, is function of one or more other variables (X): $Y = h(X) = h(X_1, X_2, \ldots, X_n)$. Monte Carlo simulation involves creating a large number of realizations of the random vector X, say n, and computing Y for each of the n realizations of X. The resulting distribution of Y, or same characteristic of this distribution, is then assumed to be "close" to the true distribution or distribution characteristic of Y. Usually Monte Carlo simulation involves generating random numbers from some specified theoretical probability distribution, such as a normal. The process of describing the distribution of the output variable is called uncertainty analysis. This analysis involves describing the variability or distribution of values of the output variable Y that is due to the collective variation in the input variable X.

3 SISAR MODEL

Since 2003, the research group of the Geodesy and Geomatics Team at the Sapienza University of Rome has developed a specific and rigorous model designed for the orientation of imagery acquired by pushbroom sensors carried on satellite platforms with asynchronous acquisition mode, like EROS-A and QuickBird.

The first version of the model (Crespi *et al*. 2003) was uniquely focused on EROS-A imagery, since no commercial software including a rigorous model for this platform were available at that time. Later, the model was refined (Baiocchi *et al*. 2004) and extended to process QuickBird imagery too.

The model, implemented in the SISAR software, bases the indirect orientation of the imagery on the well known collinearity equations, including different subsets of parameters for the satellite position, the sensor attitude and the viewing geometry (internal orientation and self-calibration). In particular, the satellite position is described through the Keplerian orbital parameters corresponding to the orbital segment during the image acquisition; the sensor attitude is supposed to be represented by a known time-dependent term plus a 2nd order time-dependent polynomial, one for each attitude angle; moreover, atmospheric refraction is accounted for by a general model for remote sensing applications (Noerdlinger 1999). For further details about the implementation of the rigorous orientation model in SISAR, see Crespi *et al*. (2006).

In 2006 the SISAR package was enriched by the group of Laboratorio di Geomatica of the Politecnico of Milan with the introduction of the LOOCV. The latest improvement, here presented, consists in the implementation of an automatic procedure to generate Monte Carlo simulations and to process in turn all the possible configuration of available GCPs.

4 TESTS AND RESULTS

The experimental tests were carried out on one QuickBird Basic image; its technical details, extracted from the attached metadata-file, are (GSD – Ground Sample Distance, γ – off-nadir angle):

- 04JAN06093307-P1BS-000000130187_01_P002 [Augusta (Sicily) – Italy, min(lat,lon): 37.0813, 15.0636; max(lat, lon): 37.2702, 15.2812]: QuickBird, 2004-06-06; |mean $\gamma = 28.2°$; area 20×20 km^2, mean GSD 0.75 m; 39 available ground points

All the ground points were carefully GNSS (GPS and Glonass) surveyed with horizontal and vertical accuracy ranging from 10 to 20 cm; in this respect, no outliers have to be suspected in ground points coordinates.

The modules that were added to SISAR to perform the tests are based on the following procedure: the image is tiled and the number of GCPs that fall in each tile is computed; in order to select the GCPs that will be used for the uncertainty analysis, all the possible combinations of GCPs sets are computed, provided that only one GCP for each not empty tile is chosen, so that the GCPs are well distributed on whole image (Figure 1).

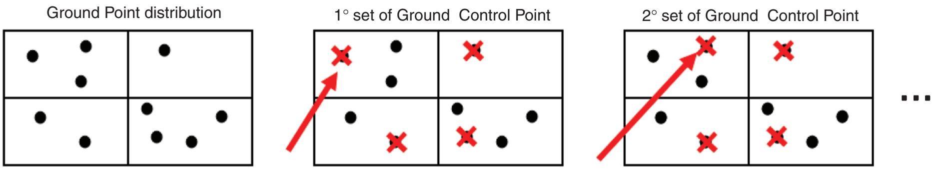

Figure 1. Example of GCPs combinations in different iterations.

It is necessary to underline that the number of selected GCPs has to be equal or greater than the number of GCPs needed to obtain a good accuracy assessment (Crespi *et al.* 2006). Then the LOOCV is computed iteratively for all the combinations of GCPs, in order to calculate the RMSE and mAD of CP residuals and their standard deviation over all the iterations.

Even if 39 ground points were available for the Augusta image, only 25 were actually used because of the high computational time needed by the software to calculate all the possible GCPs combinations; the image was split in 5 × 5 tiles (Figure 2).

The results show that the RMSE for each set of GCPs remains nearly constant and mAD trend remains constant too, even if lower (Figure 3). This consideration is confirmed by the low RMSE and mAD standard deviations (Table 1).

The Monte Carlo simulations were then performed by adding new points, generated by "dirtying" the original data with random values, to each original GCPs combination. In order to obtain proper random values, an algorithm (based on the ziggurat method by Marsaglia & Tsang (2000)) was implemented that generates random variables extracted from a normal distribution with null mean and a standard deviation equal to that of the original GNSS measured coordinates.

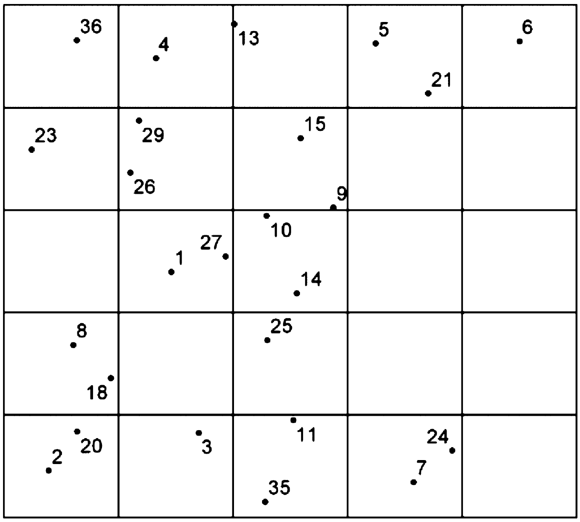

Figure 2. Example of GCP distribution for the Augusta image in 5 × 5 tiles.

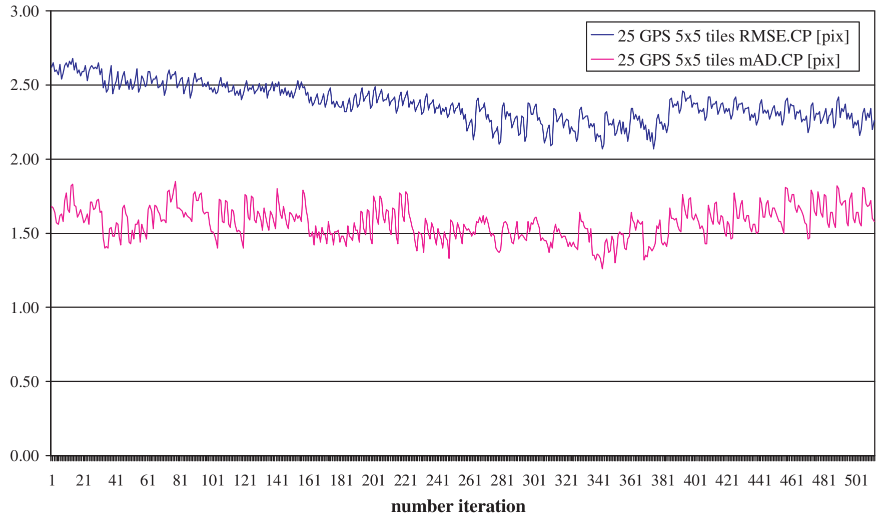

Figure 3. RMSE and mAD trends without Monte Carlo simulation.

Table 1. Mean and standard deviation without Monte Carlo simulations.

	RMSE [pix]	mAD [pix]
mean	2.373	1.570
standard deviation	0.126	0.111

Table 2. RMSE and mAD standard deviation with varying spawn rate.

Spawn rate	0	1	2	3
RMSE st. dev. [pix]	0.126	0.098	0.093	0.088
mAD st. dev. [pix]	0.111	0.115	0.128	0.099

In particular, for each original GCP, n "spawned" GCPs were generated, with the spawn rate n ranging between 0 and 3. The whole iterative LOOCV procedure was then applied again using all the available data (both the original and the newly generated ones). The RMSE, mAD and their standard deviations were calculated (Table 2).

In this case the RMSE mean and the mAD mean can be considered approximately equal to the results obtained when only the original GCP coordinates were used (Table 3) (Figure 4).

In order to underline the robustness of an accuracy index like mAD in respect to RMSE, outliers in ground coordinates or in image coordinates were introduced in two distinct tests, in which the ground and the image coordinates of point 1 were increased of 5 m and 5 pix respectively. These tests confirm the mAD as a more robust index than the RMSE for all the iterations (in Table 4 the values obtained in the first iteration as an example).

Table 3. RMSE and mAD mean with spawn rate 3.

RMSE mean [pix]	2.24
mAD mean [pix]	1.34

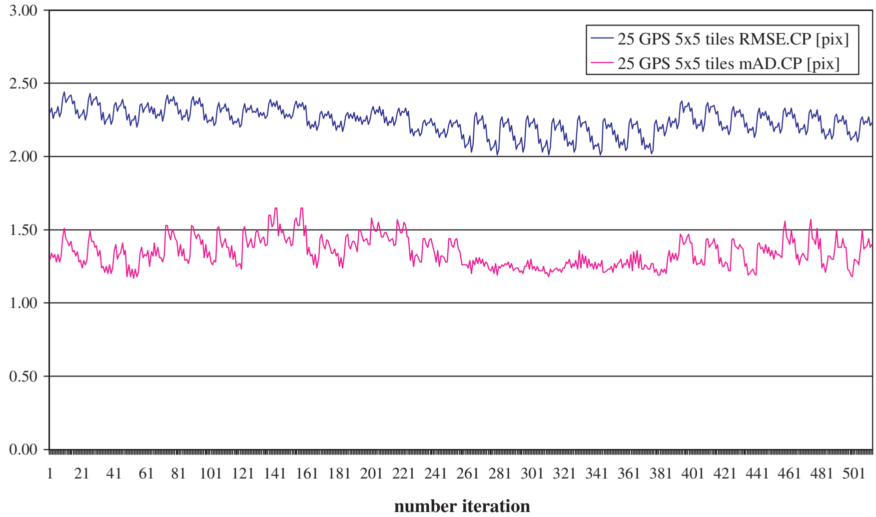

Figure 4. RMSE and mAD trends with spawn rate 3.

Table 4. RMSE and mAD without and with outliers.

No outlier		**5 pix outlier**		**5 m outlier**	
RMSE [pix]	mAD [pix]	RMSE [pix]	mAD [pix]	RMSE [pix]	mAD [pix]
2.62	1.68	3.75	1.98	4.73	2.10

5 CONCLUSIONS

The proposed method, implemented in the SISAR software, allows now to calculate the accuracy indexes by iterating all the available ground data. With the LOOCV all points can be used both as GCPs and CPs, capitalising all the available ground information. The implemented cyclic and iterative method allows to test all the possible GCPs combinations and, in absence of gross errors, all the iterations generate sufficiently stable indexes (mAD and RMSE). It was also verified that if we introduce a single outlier the mAD slightly increases, as we expected, but it stays close to the value obtained without gross errors in the input data. Vice versa the RMSE increases more significantly, pointing out the presence of the outlier value. With only one outlier and using the LOOCV the error can be identified and removed. In future works tests with more than one outlier will be performed.

ACKNOWLEDGEMENTS

This research was partially supported by grants of the Italian Ministry for School, University and Scientific Research (MIUR) in the frame of the project MIUR-COFIN 2005 – "Analisi, comparazione e integrazione di immagini digitali acquisite da piattaforma aerea e satellitare " – National Principal Investigator: S. Dequal.

The Authors thank very much: Valerio Caroselli (Informatica per il Territorio S.r.l., Rome, Italy) and Fabio Volpe (Eurimage S.p.A., Rome, Italy), who kindly supplied the EROS-A1 and Quickbird imagery respectively.

REFERENCES

Baiocchi, V., Crespi, M., De Vendictis, L. & Giannone, F. 2004. A new rigorous model for the orthorectification of synchronous and asynchronous high resolution imagery. *Proceedings of the 24th EARSeL Symposium, 5–27 May*, Dubrovnik: 461–468.

Brovelli, M.A., Crespi, M., Fratarcangeli, F., Giannone, F. & Realini, E. 2006. Accuracy assessment of High Resolution Satellite Imagery by Leave-one-out method. *Proceedings of the 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, 5–7 July*, Lisboa: 533–542.

Crespi, M., Giannone, F. & Poli, D. 2006. Analysis of rigorous orientation models for pushbroom sensors. Applications with Quickbird. *Accepted for the Reviewed part of the Proceedings of the ISPRS Commission I, WG V Meeting*, *4–7 July*, Paris.

Marsaglia, G., Tsang, W.W. 2000. The ziggurat method for generating random variables, *Journal of Statistical Software 5*.

Millard, S., Neerchal, N. 2001. *Environmental Statistic.* CRC Press.

Noerdlinger, P.D. 1999. Atmospheric refraction effects in Earth remote sensing, *ISPRS Journal of Photogrammetry & Remote Sensing* 54, 360–373.

Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions (with discussion), *Journal of the Royal Statistical Society B* 36: 111–147.