FUSION OF GIS AND SAR STATISTICAL FEATURES
FOR EARTHQUAKE DAMAGE MAPPING AT THE BLOCK SCALE

Giovanna Trianni, Fabio Dell'Acqua, Paolo Gamba, and Gianni Lisini

Department of Electronics, University of Pavia
via Ferrata, 1 – I-27100, Pavia, ITALY – e-mail: -name.surname@unipv.it

SS–6: Geo-Information for Disaster Mapping

KEY WORDS: Earthquake, Damage Mapping, SAR, Urban Remote Sensing

ABSTRACT:

This work shows that multitemporal SAR data allows mapping earthquake damage in urban areas with an acceptable accuracy, once some ancillary information defining urban blocks is available. A statistical analysis of the parameters of the models representing backscatterer intensity or coherence values for each block may be used to discriminate between damaged and undamaged areas and, to some extent, to evaluate the damage. A comparison with a recently proposed supervised segmentation approach shows that the simpler, unsupervised methodology presented here can achieve comparable results.

1. INTRODUCTION

At some stages in the disaster management cycle the sheer availability of information, any kind of it, is more important than its absolute accuracy. This is the case for image interpretation tools meant to support immediate manual characterization of a scene via aerial or satellite imagery. For this reason, simple algorithms for fast data mining and analysis are welcome in applications related to disaster and, generally speaking, to civil protection management.

The use of remotely sensed imagery in such situations has been steadily growing in the past years, but these efficient and fast algorithms, although not highly accurate, are still lacking in a number of applications.

Thus, on the one hand several projects and web sites make data and maps publicly available, like UNOSAT (http://www.unosat.org) and RESPOND (http://www.respond-int.org). On the other hand, these maps are usually nothing more than geometrically and radiometrically corrected data, with geographical information overlaid. Interpretation is still left to the human expert, a visual approach is assumed and thus the maps are intended solely to help the experts understand -on their own- what has happened in the observed area; legends are usually provided which guide the interpretation of the maps. However, the training of the interpreter is not easy, and the problem is worsened by the fact that organizations such as the "International Charter on Space and Major Disasters" (Allenbach *et al.*, 2005) are compelled to consider a wide range of sources, so that visual interpretation of diverse data may require multiple skills. This is the reason why the use of SAR data is unlikely in a first phase, although these may happen to be the only available data in case of poor weather conditions.

In recent literature, however, some works have already appeared suggesting that multi-temporal SAR data may provide, at a proper temporal and spatial scale, interesting information about disasters, particularly earthquakes and floods. As for the earthquakes are concerned, most of these works employ -to a large extent- data coming from ground surveys, not only to validate but often also to initiate the process of information extraction. Subsequently, these approaches are really valuable for correlating damage patterns with ground displacements and soil properties (Yonezawa and Takeuchi, 2001, and Matsuoka and Yamazaki, 2005), or to provide very precise 3D changes of the earth crusts (Stramondo *et al.*, 2005), but offer limited validity for damage assessment, and especially for rapid mapping of areas affected by disasters. Other approaches combine SAR data with multi-spectral images (Stramondo *et al.*,2006). This is surely interesting, but the scope of this work is limited to radar data. Statistical analyses of SAR data for damage analysis have been done (Mansouri *et al.*, 2005), but they suffer from problems in interpretation, which needs to be guided in order to provide useful information to the viewer. The aim of this work is to make the interpretation process as automatic as possible, and to reduce the time elapsed from data acquisition to information delivery.

Therefore, we cannot rely on the sole use of classification and change detection methodologies that, even for the best ones available so far in technical literature, are still far from providing satisfying and immediately useful results for the final user. Instead, the integrated use of these methods and ancillary data, either already available, or easily extracted by manual interpretation of maps and/or optical images, may focus the imprecise results of multitemporal SAR analysis towards more detailed results, producing maps that are accurate enough for the proposed applications.

2. PROCESSING CHAIN FOR DAMAGE MAPPING

The goal of this work is indeed to show that a first damage mapping of limited accuracy in earthquake-stricken areas can be obtained using multitemporal SAR data via a block-wise statistical characterization of the backscattered field. To this aim, simple neighborhood of each tested pixel may be considered. The definition of this neighborhood is questionable, and its width/scale must be adaptively changed. This is

particularly true in urban environments, usually neither homogeneous nor continuously changing. In this situation one may clearly see the advantage of integrating ancillary data, when available. In fact, they might provide clues or directly indicate homogenous land use areas. By integrating the corresponding information in the extraction flow, it will become more focused and precise, as we will show in the result section.

The basic steps of the proposed algorithm are:
1. evaluation of the best approximating statistical function for the SAR feature under test;
2. extraction of the statistically significant parameters for the area of interest, possibly using ancillary data to improve the effectiveness of the analysis;
3. comparison of the pre- and post-event parameters for efficient characterization of the changes and quick and semi-automated extraction of the damage assessment.

While all these steps have a general validity for any change detection problem involving radar data, we stress here that urban areas are considered, and earthquakes are the focus of this research. This is essential to understand some of the choices detailed in the following paragraphs.

According to the first step described above, we need first to evaluate the best statistical approach useful for characterizing urban areas at the block level. Indeed, the most commonly available spaceborne SAR technology achieves a geometric ground resolution achievable between 10 and 15 m so the finest statistical analysis achievable is at urban block level. The technical literature lists, among the proposed distributions, the lognormal and Weibull ones (see for instance Oliver and Quegan, 1998). They have been compared for urban areas and polarimetric SAR in Costamagna *et al*, 2000 and found as reliable and effective. The lognormal distribution is computer according to the formula:

$$P(x) = \frac{1}{x\sqrt{2\pi c}} \, exp\left[-\frac{(\ln x - b)^2}{2c}\right] \qquad (1)$$

The ***b*** parameter is connected to the scale of the distribution, while ***c*** rules the shape. Weibull distributions are instead characterized by the following formula:

$$P(x) = \frac{c\, x^{c-1}}{b^c} \, exp\left[-\left(\frac{x}{b}\right)^c\right] \qquad (2)$$

where *b* e *c* have identical meaning as for the lognormal distribution.

Weibull distributions are commonly used for the analysis of radar images, when the intensity is selected as the analysis feature. Lognormal distributions have been used in urban areas, since they better adapt to abrupt changes in intensity due to concentrations of strong backscatterers, which is commonplace for densely built-up areas.

The steps in the procedure for the statistical evaluation of a given portion of a given image are therefore:
1. the computation of parameters *b* and *c*, initially set to the mean β and variance *V* of the logarithm of the amplitude SAR data;
2. the minimization of the Root Mean Square Error (RMSE) between the original normalized histogram and a lognormal distribution, starting from the above mentioned initial guess;
3. the computation of the *b* and *c* factor, again minimizing the RMSE, for the best fitting Weibull distribution.

An example of the results of this procedure applied to a homogeneous land use block in an urban area is shown in Figure 1, where the original histogram and the three approximating distributions are shown and may be compared. Apparently, there is no clear preference for a given distribution model against the other one; 20 areas in 3 different images have given similar results, thus confirming this consideration.
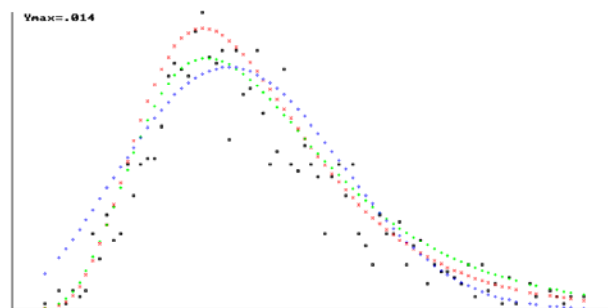

Figure 1. Statistical analysis of a test area: points define the actual amplitude value histogram, while the red curve is the lognormal distribution base on mean and variance values computed from it, the green curve the best fitting (in the mean square error sense) lognormal distribution, and the blue curve is the bets fitting Weibull distribution.

A closer look to RMSE values made us prefer the lognormal distribution. Even the initial guess for the distribution parameters reduces the RMSE value below the limit reachable after optimization with a Weibull distribution. On the other hand, the latter shows a more stable set of parameters for images referring to the same situation. This is also true for the initial guess of the lognormal distribution. Since these values are very easy to compute, they are the best candidates for a quick and effective model-based change analysis.

To this aim, it is mandatory to understand if and how the damage produced by the earthquake changes these parameters. Visually, the effects of the earthquake on amplitude SAR images result in a reduction of the contrast between urbanized and rural areas accompanied by a sudden change in the pixel variance. In accordance with our expectations, this change is greater in densely built-up areas, characterised by a dense concentration of strong backscatterers against a generally very dark background represented by road pixels.

The statistical analysis of block-based histogram shows –in hard-hit areas- a remarkable reduction of the mean amplitude value and, as a consequence of the increased variance, a reduction in the peak of the distributions. Therefore,

multitemporal analysis confirms the single-date analysis, in that the proposed distributions confirm themselves as suitable representations; moreover, it again points to the variance *V* as the most important parameter for urban change detection at a block level using SAR images.

The final step of the processing chain is therefore the definition of a suitable threshold for change detection. We may expect this threshold to be dependent from the particular test site at hand, and this will be indeed discussed in the results section; though, the capability of reducing the characterization to a single parameter is an interesting result of this research. In any case this threshold should be carefully chosen, to avoid too many false positives to pop up.

SAR amplitude values are however only a part of the available information, and the complex nature of the data may be also taken into account by considering the coherence between two SAR images. A procedure very similar to the one described in the preceding paragraphs shows that a normal distribution is capable of modeling both pre-pre and pre-post coherence data at a block level in urban areas. Moreover, variance values and their changes are an interesting way to characterize damage patterns in these same areas.

3. DAMAGE MAPPING RESULTS

The test sites for the proposed procedure are the towns of Bam, Iran, affected by a disastrous earthquake in December 2003, and the town of Gölcük, Turkey. To analyze the Bam area, a sequence of three acquisitions by the ASAR sensor is considered. In particular, data were recorded on 11th June, 2003, 3rd December, 2003 and 7th January, 2004. Due to the public availability of many results of interpretation work on aerial or satellite images of the area by the ZKI (ZKI, 2006), UNOSAT and RESPOND, ancillary data were reconstructed and are thus considered as available. Similarly, for the Gölcük area, where the earthquake stroke on 19th August 1999, ERS-1 and ERS-2 images are available, recorded on 12th and 13th August before the event, and on 16th September after the event. Ancillary data were collected using the information extracted from the web sites of various emergency relief agencies.

Following the procedure outlined in the previous section, a complete analysis of the Bam area was first obtained. The threshold for detecting damaged blocks was fixed during a preliminary analysis. Figure 2(a) shows the histograms of percent change in the amplitude variance for undamaged (green curve) versus damaged (blue curve) blocks in some sample test areas. The vertical axis represents the number of blocks which show a percentage change in variance between the adjacent numbers reported in the horizontal axis. There is some overlap, but 10% is in fact the threshold between the two subsets: blocks with a variance change greater than 10% are assumed to be damaged areas. This criterion will be used in the following evaluation of the overall Bam area. Similarly, Figure 2(b) shows the percent change for coherence variance of the same test blocks, with the further discrimination between slightly damaged (blue) and heavily damaged (red curve) blocks. This figure shows that a 15% threshold is a good guess for coherence-based statistical discrimination of heavily damaged versus slightly damaged or non-damaged blocks.
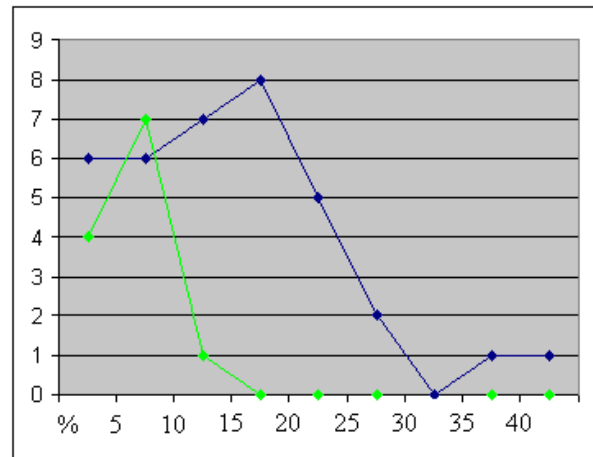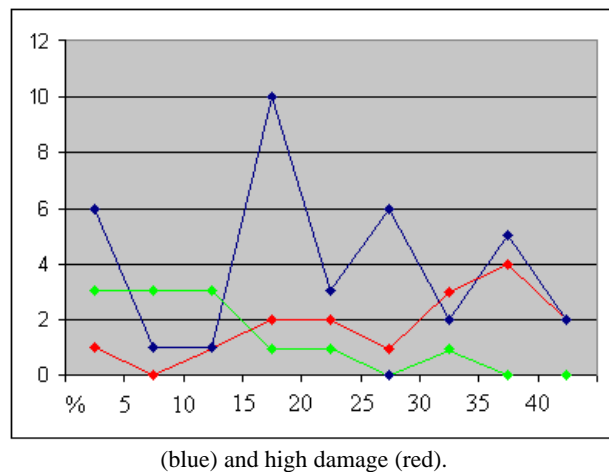

Figure 2. Histograms of the percent change in amplitude (a-top) or coherence (b-bottom) for no damage (green), low damage

(blue) and high damage (red).

Figure 3 shows results elaborating amplitude (left) and coherence (right) information, compared with *in situ* analysis carried out by the National Cartographic service of Iran (http://www.ngdir.ir/, section "Earthquake database") for the same blocks. Due to the very simple target of this research, the figure shows just two categories: orange (damaged) versus green (non-damaged) blocks, while the available ground truth further discriminates between slightly (yellow) and heavily damaged (red) blocks.

As a matter of fact, there is a very good similarity between the damage patterns obtained by means of amplitude-based (a) and coherence-based (b) statistical analysis at the block level. Moreover, both patterns are remarkably similar to the damage patterns in Figure 3(c). Coherence-based results are slightly more accurate (28 out of 36 lightly damaged areas instead of 24).
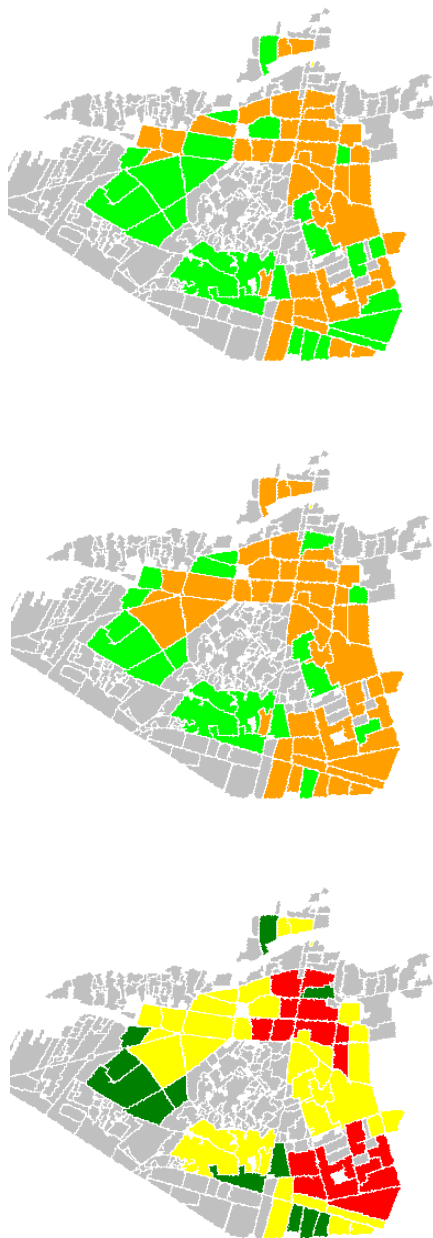
Figure 3. Damage mapping results for the Bam test site, based on amplitude (a) or coherence (b) percent change between pre- and post-event data, to be compared with (c), damaged areas as measured by in situ analysis by the National Cartographic Service of Iran [9]. Color legends are explained in the text.

A first comment to the figure is obtained when the original land use legend of the blocks is used to identify difference between the satellite-derived maps and the ground survey. Worst misclassification appears in lightly built up areas, with a great extent of vegetation. Here the threshold chosen above fails, because the backscattering mechanism is dominated by the vegetation, and the change in building patterns is somehow masked out. However, it is interesting to note that this is not a problem of SAR data alone. The same discrepancy between satellite-based damage maps and in situ measurements is observable in the map provided by UNOSAT using SPOT optical data a few days after the earthquake (SERTIT 2004).
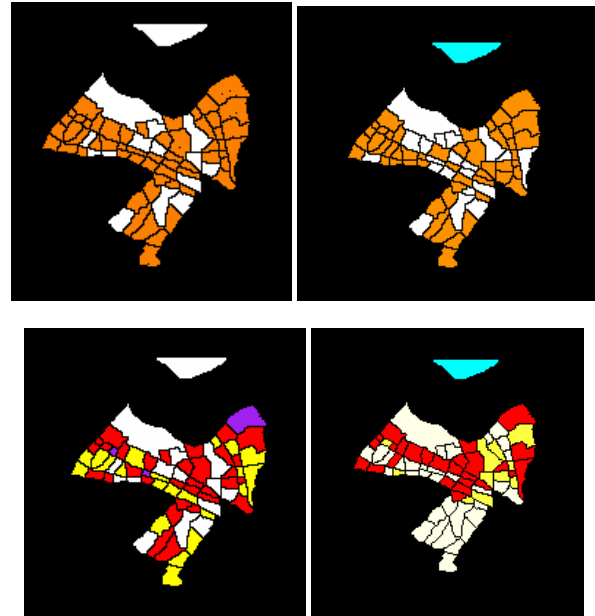

Figure 4: Damage mapping results for the Gölcük test site, based on amplitude (a-top left) or coherence (b-top right) percent change between pre- and post-event data, to be compared with (d-bottom right), ground truth for damaged areas. The map in (c-bottom left) shows that a finer subdivision of amplitude variance changes may provide an even better damage pattern recognition, with more classes of damages

The results for the second test site, the town of Gölcük, show a very similar behavior. This is peculiar because the change detection thresholds were maintained to the same levels, even if the land use in Gölcük, as well as the damage mechanisms, were very different. In fact, the damage maps are shown in Figure 4 (color legends are identical to those for Figure 3) and report a total of 30 detected blocks out of 35 for amplitude and 25 for coherence-based analysis. Please note that in this situation an even more precise characterization of the building damage is possible, by further subdividing the ranges of percentage change of the statistical parameters of interest, as shown in Figure 4(c), where two additional thresholds, 50% and 80%, were considered.

Moreover, for the Gölcük area a comparison has been carried out, between these results and those obtained from the recently proposed supervised procedure by the same authors of this work (Gamba *et al.*, 2006). The advantage of the algorithm in the present work is its unsupervised nature, while the advantage of the supervised procedure, based on a non-parametric classifier, is its ability to take into account both amplitude and coherence information as a whole. In Gamba *et al.* (2006) the technique was applied to the Bam test area, but it proved to be useful for the second test site of this work as well. The damage maps using three different combination of the original multitemporal SAR data and other information layers which may be computed from them for the Gölcük area are shown in Figure 5, were the same block boundaries used in this work have been considered to produce a per-parcel damage analysis. The features used for this classification are the pre- and post-event intensity, the coherence information and finally the "backscattering coefficient" as defined in Matsuoka (2005), basically the difference between the decimal logarithms of the average pixel

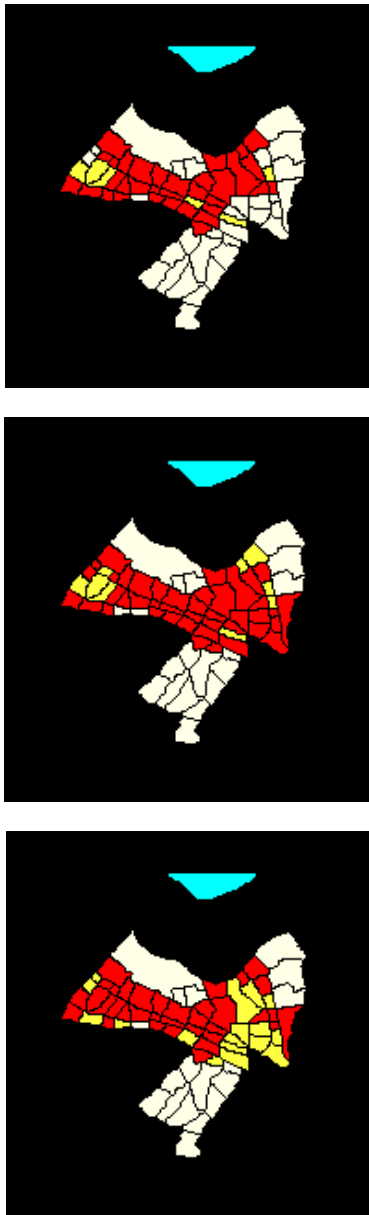values computed in two corresponding, sliding windows over the two figures.

Figure 5. Damage mapping results for the Gölcük test site, based on pre- and post-event intensity (a-top), the same as in (a) plus the Matsuoka backscattering coefficient (b-mid) or the same as in (b) plus coherence information (c-bottom).

The qualitative comparison of Figure 4 and Figure 5 shows that the damage patterns revealed by the two techniques are similar, even if in some parts of the map it is clear that they suffer from different misclassification errors. In particular, the supervised procedure is more capable of detecting the damage, but is it somehow overestimates the damage levels. On the contrary, the statistical analysis shows a more imprecise characterization of the damages, especially for the low damage level. This suggests that a further step in this research may be the definition of a methodology to fuse at the information level the maps available by the two approaches.

A quantitative comparison of the results of the unsupervised and supervised procedure for the Gölcük area is provided in Table I, where the confusion matrices at the block level for some of the maps proposed in the two figures are compared. Note that the four rows correspond from top to bottom to no damage (white blocks), low damage (yellow blocks), high damage (red blocks), and the sea (light blue area).

Fig. 4(a)

Damaged	81.08%	30	7	0
Undamaged	21.21%	26	7	0
Water	100.00%	0	0	1
		53.57%	50.00%	100.00%

Overall accuracy: **53.52%**

Fig. 4(b)

Damaged	67.57%	25	12	0
Undamaged	24.24%	25	8	0
Water	100.00%	0	0	1
		50.00%	40.00%	100.00%

Overall accuracy: **47.89%**

Fig. 5(b)

Lightly Damag.	45.45%	15	7	11	0
Heavily Damag.	18.18%	2	2	7	0
Undamaged	92.31%	2	0	24	0
Water	100.00%	0	0	0	1
		78.95%	22.22%	57.14%	100.00%

Overall accuracy: **59.15%**

Fig. 5(c)

Lightly Damag.	45.45%	15	4	14	0
Heavily Damag.	72.73%	1	8	2	0
Undamaged	73.08%	2	5	19	0
Water	100.00%	0	0	0	1
		83.33%	47.06%	54.29%	100.00%

Overall accuracy: **60.56%**

Table 1. Confusion matrices for some of the damage maps in Figure 4 and 5.

A first comment is that, as the visual analysis suggested, the overall accuracy values are similar to the Bam case. However, there is a larger misclassification in the Southern part of the urban area. An inspection of Gölcük Municipality Website (http://www.golcuk.bel.tr/en/, section "city map", category "new housing parcels", accessed 21st November 2006 3:10 PM CET) has revealed significant development plans for the southern part of the town, which may have already been in progress between the two pre- and post-event images, resulting in significant backscattering changes. This may have "forged" a modification in the features similar to those arising from a collapse of the buildings, in an area where no significant, actual damage was reported (see ground truth in figure 4d) This in turn may have been the cause of the confusion with damage classes observed in the southern part of the town. Another interesting point is that coherence and amplitude-based analysis provide similar overall accuracy, but some distinctions are to be made. The inclusion of coherence produces a limited improvement in the overall accuracy, which is a result of a substantially better classification of the slightly damaged areas contrasted by worse classification of the two other categories. In other words the coherence appears to be the best contributing feature to distinguish slightly damaged areas Finally, the supervised approach is somehow superior, having the inherent possibility to exploit the whole complex radar measurements. The introduction of a decision fusion step even for the unsupervised procedure proposed in this paper would be, as noted above, an interesting research field, while not equally straightforward.

4. CONCLUSIONS

This work is focused on using multitemporal SAR data for urban damage assessment after a disaster, and shows that a fast statistical analysis, performed at block level using ancillary data, may provide enough information to detect damage patterns. Additional research is required to characterize the damages in a more precise way, even if the methodology looks promising so far.

Future working lines, besides the fusion of amplitude and coherence-based damage evaluation, are aimed at validating the

procedure using different data sets and to investigate the possibility to introduce different views in this scheme to reduce the time required for acquiring useful data.

REFERENCES

Allenbach, B., R. Andreoli, S. Battiston, C. Bestault, S. Clandillon, K. Fellah, J.-B. Henry, C. Meyer, H. Scius, N. Tholey, H. Yesou, and P. de Fraiport, 2005, Rapid EO disaster mapping service: added value, feedback and perspectives after 4 years of Charter actions, Proc. of *IGARSS'05*, Seoul, Korea.

Costamagna, E., P. Gamba, P. Lombardo, G. Chinino, 2000. Statistical analysis and neuro-fuzzy classification of polarimetric SAR images of urban areas, Proc. of the *ERS/ENVISAT Symposium*, Gotheborg, Sweden, unformatted CD-ROM.

Gamba, P., F. Dell'Acqua, G. Trianni, 2006. Rapid damage detection in Bam area using multi-temporal SAR and exploiting ancillary data, *IEEE Trans. Geoscience and Remote Sensing*, *in press*.

B. Mansouri, M. Shinozuka, C. Huyck, and B. Houshmand, 2005. "Earthquake-induced change detection in the 2003 Bam, Iran, earthquake by complex analysis using Envisat ASAR data", Earthquake Spectra, vol. 21, n. 81, pp. 8275-8284, 2005.

Matsuoka, M., and F. Yamazaki, 2005. Building damage mapping of the 2003 Bam, Iran, Earthquake using Envisat/ASAR intensity imagery, *Earthquake Spectra*, **21**(81), 8285-8294.

Oliver, C., and S. Quegan, S. 1998. *Understanding Synthetic Aperture Radar Images*, Artech House, Boston.

SERTIT 2004. Available on-line at: http://unosat.web.cern.ch/unosat/freeproducts/iran/BAM/doc07_bam_damages_and_land_use.jpg). Last accessed on 28th April 2008, 14:45 PM CET.

Stramondo, S., M. Moro, C. Tolomei, F.R. Cinti and F. Doumaz, 2005. InSAR surface displacement field and fault modelling for the 2003 Bam earthquake (southeastern Iran)', *Journal of Geodynamics*, **40**(2-3), 347-353.

S. Stramondo A1, C. Bignami A2, M. Chini A3, N. Pierdicca A2, A. Tertulliani, 2006. Satellite radar and optical remote sensing for earthquake damage detection: results from different case studies. *International Journal of Remote Sensing*, Volume 27, Number 20 / 20 October 2006, pp. 4433 – 4447.

Yonezawa, C., and. S. Takeuchi, 2001. Decorrelation of SAR data by urban damages caused by the 1995 Hyogoken-nambu earthquake, *Int. J. Remote Sens.*, **22**(8), 1585-1600.

ZKI, Center for Satellite Based Crisis Information. Available on line at http://www.zki.caf.dlr.de/intro_en.html Last accessed on 28th April 2008, 14:44 PM CET.