

Synthetic Aperture Radar Interferometry: Separation of atmospheric artifacts from effects due to the topography and the terrain displacements

Paola Ballatore

Mediterranean Agency for Remote Sensing and Environmental Control, Villa dei Papi, Benevento, Italy

(Received March 8, 2006; Revised May 10, 2006; Accepted May 24, 2006; Online published September 16, 2006)

This paper reports the new concept of possibly applying independent component analysis (ICA) on Synthetic Aperture Radar (SAR) satellite images for separating atmospheric artifacts from effects due to topography and terrain displacements. Specifically, the FastICA algorithm is applied on simulations of SAR interferograms with the purpose of extracting the different independent sources. Results show the existence of significant correlations between estimated and original components, with correlation coefficients above 0.9 and statistical confidence level above 99.9%. These findings suggest that ICA might provide a useful tool in SAR data processing, with a specific crucial usefulness in cases of an absence of ground truth knowledge, as in the cases of insufficient meteorological information at specific observational times or in satellite monitoring of remote lands. Applications on real data show that the topographical component is automatically derived by the FastICA algorithm for whatever real data set. What is different is that the extraction of terrain displacements may require some a priori information for separating different kinds of landslides and that the use of possible semi-blind ICA/FastICA approach might be considered, dependent on the specific data set.

Key words: Earth observation by satellite, synthetic aperture radar interferometry, remote sensing, ICA, blind source separation, artifacts in images.

1. Introduction

Synthetic Aperture Radar (SAR) interferograms are obtained by analyzing two sets of radar images of the same areas with two slightly different viewing angles. The data for the coupled images can be obtained by two antennas located on the same platform (single-pass interferometry) or via repeated passes of the same sensor on two nearly parallel trajectories (repeated-pass interferometry). The latter is the most common situation in SAR processing, and it is always the case in any multi-satellite SAR interferometry (ERS-1/2, ERS-ENVISAT, RADARSAT or other).

Because data acquisitions are not simultaneous in repeated-pass interferometry, there can be several sources of phase shifts due to slight differences in the viewing angle or satellite location and to the temporal fluctuations in the atmosphere and ground scatters. A realistic scenario of the interferometric phase differences between the corresponding pixels of two images taken at times t_A and t_B can be expressed as the sum of components due to: (1) the topography, (2) the line-of-sight (LOS) cumulative deformations between the two times of the interferometric coupled images, (3) the atmospheric fluctuations, and (4) a random noise term which takes into account the temporal decorrelation due to random changes of the ground scatters for spatial scales of the order of signal wavelength or other random

instrumental noises:

$$\phi(t_A) - \phi(t_B) = \delta\phi = \delta\phi_{\text{topogr}} + \delta\phi_{\text{displ}} + \delta\phi_{\text{atm}} + \delta\phi_{\text{noise}} \quad (1)$$

It is not the complete phase shift $\delta\phi$ that is measurable, but only its component in the interval $(-\pi, \pi)$, which is known as the ‘wrapped phase’. The generic unwrapping procedures only involve pixels with an estimated coherence value larger than a fixed threshold, therefore regions of the images that are significantly affected by the $\delta\phi_{\text{noise}}$ are not taken into account and are estimated by interpolation of the resulting unwrapped interferograms. Therefore, if we compute an unwrapped SAR interferogram with commonly available software, we obtain a mixture of the first three phase components related to the topography, to the terrain displacement in the LOS, and to the atmosphere.

A comparison of several interferograms indicates the presence of large differences that remain even if the most evident outliers are excluded. These can possibly be attributed to the influence of strong atmospheric perturbation as the atmosphere is the most dynamic source of temporal variation of results (see Ferretti *et al.*, 1996).

The independent component analysis technique, ICA, is specifically devised for separating mixed signals into their sources, which are unknown and represent the information searched (see Comon, 1994; Cardoso, 1998). In particular, the FastICA algorithm (Hyvärinen and Oja, 2000) has previously been successfully employed in cases of separating astrophysical data acquired at different times from time changing artifacts and, in particular, atmospheric effects (see Funaro *et al.*, 2001).

Although FastICA has never been applied to SAR data,

the reasonability of its potential application on arrays of SAR interferograms is based on evidence of the independence of the three components of phase shifts involved in Eq. (1). In fact, in principle, relationships among the ground altitude, the terrain movements, and the atmospheric fluctuations are not expected to exist. Further details on the applicability of ICA are presented in detail in the following sections.

2. ICA and its Application on SAR Interferometric Data

We take into account a multiple number, m , of interferograms that are similar to that given in Eq. (1), but which are calculated by different couples of SAR images. In this case, we have an array of m observations $(\delta\phi_1, \delta\phi_2, \delta\phi_3, \dots, \delta\phi_m)$ that can be considered to be the elements of a vector X , defined:

$$X = (\delta\phi_1, \delta\phi_2, \delta\phi_3, \dots, \delta\phi_m) \quad (2)$$

Each element, $\delta\phi_i$, of this vector corresponds to an interferogram—i.e., an image of n pixels. Therefore, each $\delta\phi_i$ itself is a vector of n elements that correspond to all of the pixels ordered in a sequence of rows or columns for computational reasons.

$$\delta\phi_i = (\delta\phi_i^1, \delta\phi_i^2, \dots, \delta\phi_i^n) \quad (3)$$

This means that X defined in Eq. (2) is a matrix of dimensions $m \times n$, although it is generally named ‘the vector of the observations’ in ICA language.

According to Eq. (1), after the unwrapping procedure, each $\delta\phi_i$ can be modeled as the sum of three contributions: $\delta\phi_{i,\text{topogr}}$, $\delta\phi_{i,\text{displ}}$, and $\delta\phi_{i,\text{atm}}$. The index i represents the interferogram number, $i = 1, 2, \dots, m$, while $j = 1, 2, \dots, n$.

$$\delta\phi_i^j = \delta\phi_{i,\text{topogr}}^j + \delta\phi_{i,\text{displ}}^j + \delta\phi_{i,\text{atm}}^j \quad (4)$$

In this situation, ICA algorithms are devised to extract sources, ‘ s_i ’, which correspond to the single contributions $\delta\phi_{i,\text{topogr}}$, $\delta\phi_{i,\text{displ}}$, and $\delta\phi_{i,\text{atm}}$, a minus of one multiplicative and one additional constant. The following equations are valid for each i :

$$s_1 = c_{i,1} \cdot \delta\phi_{i,\text{topogr}} + c_{i,2} \quad (5)$$

$$s_2 = c_{i,3} \cdot \delta\phi_{i,\text{displ}} + c_{i,4} \quad (6)$$

$$s_3 = c_{i,5} \cdot \delta\phi_{i,\text{atm}} + c_{i,6} \quad (7)$$

where s_1 , s_2 , and s_3 are n dimensional arrays (such as $\delta\phi_{i,\text{topogr}}$, $\delta\phi_{i,\text{displ}}$, and $\delta\phi_{i,\text{atm}}$); $c_{i,1}$, $c_{i,2}$, $c_{i,3}$, $c_{i,4}$, $c_{i,5}$, and $c_{i,6}$ are single value constants, whose value depends on the interferogram number i (i.e., on the observational parameters). This means that the source s_1 is a component common to all of the measurements of the topography, s_2 to all the measurements of displacements, and s_3 to all the measurements of atmosphere. In addition, this means that there are linear transformations among $\delta\phi_{1,\text{topogr}}$, $\delta\phi_{2,\text{topogr}}$, and $\delta\phi_{3,\text{topogr}}$. Similarly, this is also true for the terrain displacements and the atmosphere. As explained in the Section 2.1, all of the topographical and displacements components can

be related one to each other, while the atmosphere is treated as an image artifact (Hyvärinen *et al.*, 2001): only its component common to all observations can be extracted. In any case, the purpose of the present elaboration is the separation of the topographical and terrain displacement components and the elimination of the atmospheric spurious effects. Thus, ICA results can be of use for our processing. For example, if we consider only three interferograms, because of previous equations, the following system is valid:

$$\begin{cases} \delta\phi_1 = a_{1,1} \cdot s_1 + a_{1,2} \cdot s_2 + a_{1,3} \cdot s_3 \\ \delta\phi_2 = a_{2,1} \cdot s_1 + a_{2,2} \cdot s_2 + a_{2,3} \cdot s_3 \\ \delta\phi_3 = a_{3,1} \cdot s_1 + a_{3,2} \cdot s_2 + a_{3,3} \cdot s_3 \end{cases} \quad (8)$$

The nine elements $a_{i,j}$ and the $3 \times n$ elements s_j (recall that each one of the s_i elements is an image of n pixels) are estimated through ICA. Specifically, according to Eq. (2) for m equal to 3, we put $X = (\delta\phi_1, \delta\phi_2, \delta\phi_3)$. Similarly to X , it is possible to define the vector S as:

$$S = (s_1, s_2, s_3) \quad (9)$$

S is named the ‘vector of the sources’ in ICA language, also if it is an array of $3 \times n$ elements, or, in general, $p \times n$ elements. In ICA algorithms, the dimension of the vector of the sources is always taken to be smaller than, or equal to the dimension of the vector of the observations ($p \leq m$). Therefore, in our application for separating three sources, the minimum possible number of observations is three. The matrix A formed by the elements $a_{i,j}$ is named the ‘mixing matrix’ in ICA language, so that system 8 is equivalent to

$$X = A \cdot S \quad (10)$$

Equation 10 is the typical formulation of the problem case resolved by ICA. More precisely, given the vector X of multiple observations, ICA consists in finding a linear transformation

$$S = W \cdot X \quad (11)$$

so that the components s_i of S are as independent as possible. This means a maximizing of some function $F(s_1, s_2, \dots, s_m)$ (recall that in our case m is taken equal to 3) that can be considered to be a measure of ‘independence’. In practice, it is not possible to solve Eq. (10) exactly; in fact, the unknown variables are too frequent, being all of the elements of W and the $n \times m$ elements of S . Therefore, ICA has a statistical valence, and the maximization of $F(s_1, s_2, \dots, s_m)$ is used to estimate the best approximation of S and W . The basic principle of ICA (Hyvärinen *et al.*, 2001 and references therein) is based on the fact that for whatever the elements of W are, it is $W \cdot X = W \cdot A \cdot S$, after X given in Eq. (10). Assuming that the elements of S are independent, each element of the vector $W \cdot A \cdot S$ is more Gaussian than any of the s_i because it represents the sum of at least two independent variables that are known to be more Gaussian than the original variables. Therefore, each element of $W \cdot X$ becomes the least Gaussian when it equals one of the s_i . For example, in the case of three observations, x_1 , x_2 , and x_3 , and three sources, s_1 , s_2 , and s_3 , each source is estimated by calculating three constant values w_1 , w_2 , w_3 which maximize the non-Gaussianity of

$(w_1 \cdot x_1 + w_2 \cdot x_2 + w_3 \cdot x_3)$. Once these three w_i are estimated, then

$$s_i = w_1 \cdot x_1 + w_2 \cdot x_2 + w_3 \cdot x_3 \quad (12)$$

The classical measure of non-Gaussianity is the Kurtosis or the fourth order cumulant; otherwise, the measure of non-Gaussianity can be taken to be equal to the neg-entropy, which is always non-negative, and is zero if the variable has a Gaussian distribution. Neg-entropy has an additional interesting property: it is invariant for the invertible linear transformation (Comon, 1994). In particular, several approximations can be taken into account for the neg-entropy computation (Hyvärinen, 1998), and different algorithms are developed in the context of ICA processing.

In this paper, we apply the FastICA to SAR interferograms, which maximizes the non-Gaussianity of Eq. (12) by an approximative Newton iteration with appealing convergence properties (Hyvärinen and Oja, 1997; Hyvärinen *et al.*, 2001). In particular, FastICA allows a preliminary whitening (through principal component analysis (PCA)) of the data, transforming the vectors x_i into $z_i = V \cdot x_i$ so that the new vectors x_i have uncorrelated and unit variance elements. The mixing model then becomes

$$z_i = V \cdot x_i = V \cdot A \cdot s_i = W \cdot s_i \quad (13)$$

where the matrix W is orthogonal. To compute W , FastICA calculates its individual columns w_n through updating by iteration

$$w_n := E \{ z \cdot g(w_n^T \cdot z) \} - E \{ g'(w_n^T \cdot z) \} \cdot w_n \quad (14)$$

followed by orthonormalization of W after each updating step. The function g determines the updating rule and g' is its derivative. This function can be chosen according to criteria discussed in (Hyvärinen and Oja, 1997) or (Hyvärinen *et al.*, 2001). In our case, we used $g(u) = u^3$, which is the function that FastICA uses by default if no different specification is given.

2.1 Topographical, terrain displacement and atmospheric components

We note that the validity of Eq. (5) requires that the topographical component contributing to one specific interferogram can be linearly related to the topographical component measured in any other interferogram. In fact, this is true and, according to repeated-pass SAR interferometry, the phase shift in each interferogram can be calculated as (see Zebker and Goldstein, 1986; Rocca *et al.*, 1997):

$$\delta\phi_{\text{topogr.}} = (4 \cdot \pi \cdot B_n \cdot q) / (\lambda \cdot R_0 \cdot \sin \theta) \quad (15)$$

'Baseline' is denoted as the distance between the satellite orbit at the two times of acquisition of the interferogram couples; in Eq. (15), for example, B_n is the baseline component that is normal to the slant direction (the slant direction is perpendicular to the direction of the satellite speed). Moreover, in Eq. (15), q is the relative elevation of the resolution cell, R_0 is the sensor-target distance of reference, λ is the wavelength, and θ is the off-nadir angle. All of these previous parameters can be considered to be constant in all the elements of the vector f the observations X , except for the baseline B_n , which depends on the specific two orbits of

the two passes selected for the acquisition of the two images of the interferometric couples. This baseline B_n determines the proportionality among the topographical contributions (Eq. (15)) in different interferograms.

Similarly, Eq. (6) is also verified for the phase shift component due to possible terrain displacement. In fact, this phase shift is modeled as (Massonnet and Feigl, 1998)

$$\delta\phi_{\text{displ}} = 4 \cdot \pi \cdot z / \lambda \quad (16)$$

where z is the displacement's length in the LOS direction. Assuming a constant speed for the landslide observed by SAR interferograms, the displacement z is larger or smaller for larger or smaller time lags, respectively, between the two acquisitions of the two images of the interferometric couple. Apart from this proportionality between z and the acquisition time lag, a common source can be identified for $\delta\phi_{\text{displ}}$. In fact, in a specific region, the zones of subsidence or accumulation are fixed, unless 'catastrophic' events occur that are not intended to be measured as interferometric $\delta\phi_{\text{displ}}$. These 'catastrophic' events are considered as changes in the topography that can be detected in a following sequence of interferograms. If these events happen within the period of acquisitions of the data for the set of interferograms that are considered in one single series, these are omitted, such as spurious signal.

It is important to highlight that terrain displacements of crustal deformation (earthquake or volcanic activity) have a close relation to the topography; in fact, historical repetition of many similar earthquakes has likely made the topography (see Fujiwara *et al.*, 2000). This has to be taken into account in ICA processing. In fact, if the observed terrain displacements are of this kind and a significant correlation occurs between topography and terrain displacement, a blind ICA application might be not appropriate and a semi-blind approach could be necessary. In fact, sources containing correlations among sources can be processed by ICA methods, but only until certain levels of correlation coefficients (Hyvärinen *et al.*, 2001): for highly significant correlations, it is no longer possible to process the sources as independent components.

The component $\delta\phi_{\text{atm}}$ takes into account the different propagation conditions of the radar sounding signal through the atmosphere at the two specific times of observations of the interferometric couple. For example, a more humid atmosphere can be associated to a phase shift in the phase signal that can be misinterpreted as a larger distance of the target from the radar. It is not the atmospheric effect itself that influences the interferograms, but its variation between the times t_A and t_B of each observation $\delta\phi_i$. The atmospheric effect can somehow be deterministic in terms of its relationship with the topography. For example, the atmospheric total pressure is known to decrease with increasing elevation, and the opposite is true for the temperature at certain levels of altitudes. However, the meteorological differences between relatively distant times can be considered to be casual, although close regions have to have similar atmospheric parameters. Therefore, Eq. (7) is not valid for the total atmospheric phase shift component in SAR interferograms. However, a source, s_3 , can be calculated as a component common to all atmospheric effects. This source can

represent a constant ratio between atmospheric variations recorded over specific different locations. In these terms of the horizontal atmospheric variability, the interest in this component within the context of geophysics is related to the study of the existence of areas of micro-meteorology/micro-climates.

With respect to the vertical atmospheric stratification, which can be quite similar to the topography (Hanssen, 2001), this perturbation is included into the casual atmospheric perturbation discussed above, which is a component independent of the topography unless a part of it is correlated with the topography and can be expressed as a function of the topography. This part of the atmospheric perturbation is not separately detected and it is not seen; in fact, it is not an independent component as with re-scaling of the topographical source extracted by FastICA, the effect of any source that is correlated with the topography is eliminated.

In any case, a pure mathematical significance of this third source does not represent an obstacle to the application of FastICA on SAR interferograms. In fact, $\delta\phi_{\text{atm}}$ can be treated as an artifact that affects our capability of observing the other two components of interest, as in previous ICA applications (Funaro *et al.*, 2001).

3. Simulations and Results

3.1 Simulations

It is important to emphasize that one simulation cannot cover all possible observational configurations, and that each simulation requires approximations that are not always verified in the experimental situation. For example, the basic approximation that is generally assumed in SAR processing theory is the monochromatic approximation for the sounding signal, which is known to be not monochromatic (it is a frequency band). In the following, other approximations are applied, but these do not affect the significance of the simulated interferograms in terms of their separability under ICA treatment and they are clearly stated and discussed. In particular, this specific simulation is intended to present a possible experimental case that could occur in an experimental context and would be not processed by: (1) the PS method, because only two data points are supposed to be known on the ground, or (2) by the SBAS, because the baselines cannot be approximated to be zero (for PS and SBAS meaning see Section 3.2).

The minimum number of interferograms necessary for separating the three contributions from topography, terrain displacements, and atmospheric effects is three. Therefore, we simulate three different interferograms (taken as elements of X). We then apply the FastICA algorithm to them to derive the mixing matrix A and the vector of the sources S . This is done for purposes of verifying if the experimental values of SAR observational and instrumental parameters are suitable for obtaining a sufficient precision in the estimation of topographical and displacement components by FastICA.

For the topographical component, we calculate the phase shift according to Eq. (15), in which the pixel values of q correspond to a synthetic digital elevation model (DEM) of 256×256 pixels, representing an area of 5.12×5.12 km

with a spatial resolution (pixel dimension on the ground) equal to 20×20 m and an elevation ranging from 10 m to 550 m.

The other specifications in Eq. (15) are taken from satellite and sensor parameters of the ERS-SAR mission: i.e.: $\lambda = 5.66$ cm, the satellite altitude is 780 km, and the incidence angle θ at the center of the field of view is equal to 23° . In each one of the three interferograms, the baseline component, B_n , is different. In particular, for the example reported here, it is 50 m, 100 m and 70 m, respectively. Of course, results shown in the following are independent of these specific values.

More specifically, in the real case, the perpendicular baselines are actually varying along the track; however, their variability can be neglected, although it might be taken into account as, for example, a condition for restricting the sub-image considered for each single FastICA input. Specifically, the perpendicular baseline variation is of the order of 10^{-1} m over 256×256 pixels in the case of ERS-SAR data. Therefore, it might be neglected or not, according to the desired observational precision. In any case, to give a quantitative indication of this approximation, we recall that the perpendicular baseline is approximated to be zero in the SBAS techniques, even in cases in which it reaches length of the order of few tens of meters.

The terrain displacement component is modeled according to Eq. (16), where λ is the same as given above and z is an image of 256×256 pixels calculated as a random auto-correlated two-dimensional array. This auto-correlation is necessary to ensure that close pixels have similar values. In fact, it is experimentally true that areas closer to zones in movements are more probably within movements, while areas closer to terrains which are very stable are more probably stable. In particular, a randomly normally distributed random array of 256×256 numbers is sorted. The auto-correlation is then included by filtering this signal with a filter designed using a two-dimensional Gaussian window, with a pass-band smaller than 0.8 of the normalized frequency (1.0 corresponds to half of the sampling spatial frequency). We can simply refer to the Matlab for help about the function 'wind2' for more details about the generation of this filter. A realistic absolute value of terrain displacements can be of the order of few tens of millimeters per year. This value has to be compared with the radar sounding signal wavelength of 5.6 cm, and it indicates that the maximum value of the simulated phase shift is much smaller than the one representing the topography (hundreds of meters), which is of the order of many 2π cycles. We highlight that the array of terrain displacements is equal in all three interferograms, except for a multiplicative constant which indicates the proportionality between the time lag between the acquisitions of the two interferometric images and the amplitude of linear displacement (z in Eq. (16)). In fact, in general, the speed of terrain displacements can be known only as an approximated constant speed in any SAR processing technique. The specific coefficients of proportionality for the specific three interferograms considered here are 1, 1.5 and 3.1, respectively. Again, results are independent of these specific coefficients. Moreover, we recall that the possible ground movements that do not belong to a regu-

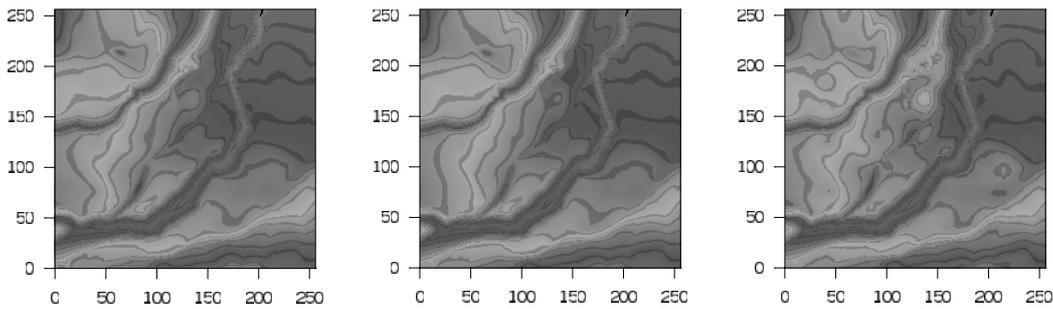


Fig. 1. Maps of the phase shifts (i.e., interferograms) given by the sum of topographical, terrain displacement, and atmospheric/spurious components, as described in Section 3.1. The same scale of grays is applied in all the three panels. The axis label indicates the pixel number.

lar common field present in all three interferograms and are totally casual are treated as spurious signals and included in the third artifact component.

More specifically, there can be cases when landslides move according to different time-functions (the specific function is not important here: linear, exponential, logarithmic, or other can be taken into account) and appear in one single image on which ICA is applied. One solution is the restriction of the sub-image to include only one kind of landslide. Otherwise, there is the possibility of their extraction as separate ICA components (see Section 4), but more than three interferograms are then necessary. In the simulation presented here, only one kind of terrain displacement is taken into account, together with other possible casual movements that are not extracted and are treated as image artifacts within the effects of the third component.

This third component represents the atmospheric artifact and can be represented by an auto-correlated two-dimensional random array similar to that described for the terrain displacements. However, in this case the normally distributed random array of 256×256 is re-sorted for each interferogram. In fact, in principle there is no relationship or proportionality between the atmospheric fluctuation that occurred between two times t_1 and t_2 and other different times t_3 and t_4 , unless all these points are very close in time; however, this is not the case of SAR repeated-pass interferometry. In this sense, this third component of phase shift represents a noise, and it can also take into account any other similar source of unpredictable fluctuations in SAR observations.

It is worth mentioning that the absolute value of the atmospheric phase shift component is of the order of few cycles; therefore, it is much smaller than the topographical contribution. Consequently, if we look at the three resulting interferograms, they are very similar if plotted as images. This is shown in Fig. 1, which illustrates the three mixtures representing the three simulated interferograms: although the images look very similar (similar to the dominant topographical component), significant differences exist among the numerical values of corresponding pixels.

3.2 Motivation of ICA application

We suppose that experimental observations give our three interferograms, simulated as described in the previous section and shown in Fig. 1. Then, to derive the topography, it is supposed that the terrain displacements can be approximated to zero (otherwise we do not know how to

separate them, unless further specific data are available). Therefore, it is possible to apply Eq. (15) on each one of the three interferograms to derive the relative elevation q . It is then assumed that the atmospheric effect is uniform around some pixels whose altitude is known *a priori*. These ground points of known information are chosen so that they can be identified in the corresponding image pixels for their particularly coherent scatter. For this reason, they are defined as 'permanent scatter' (Ferretti *et al.*, 2001). For example, these points can be parts of buildings in cities or rocks on mountains. The distribution of these scatters on the observational area has to be quite dense for a precision generally considered to be sufficient, with a minimum of five to ten permanent scatters per square kilometer or even much more. In fact, especially on mountainous areas, the ground parameters can vary within a few tens of meters, while the correction of elevation (generally attributed to the atmosphere) for each permanent scatter is applied all around until next permanent scatter zone. In our case, the three q values of corresponding pixels can be re-scaled by taking into account one point with an elevation equal in all three interferograms, similar to a permanent scatter. Figure 2 shows that the three sequences of altitudes can differ just in the neighbor pixel of the permanent scatter. This is what can generally happen in experimental SAR observations as well. Having three interferograms, the final elevation corresponds to the average of the three values plotted, and the error is the difference between the maximum minus the minimum of these three values. Figure 2 shows relatively large errors associated with the determination of the topography (i.e., the relative altitude of pixels). The correction of these errors requires further data or a knowledge of many more real ground values to re-scale the atmospheric effect differently in each sub-region. In the case where we had many interferograms and only few of these were different from the others, these could be excluded as bad measurements, and the statistics would improve together with the potential precision. In the case of only three interferograms the statistics is quite small, except in cases of flat terrain and uniform atmospheric parameters over the observational area. In Figure 2, the error ranges from zero to more than 5 m, and the relative error reaches about 25% of the altitude (taken as the average among the three altitudes). However, these pixels represent only a fraction of $1/256$ of the whole image, which is 256×256 pixels, and the maximum relative error all over the image become much larger unless many addi-

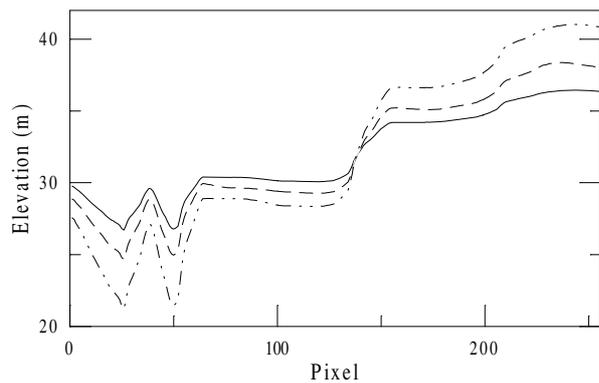


Fig. 2. Altitudes calculated for a number of corresponding pixels of the three simulated interferograms (each line style refers to one interferogram) as described in Section 3.2.

tional permanent scatter points are considered. In general, our case does not represent the worst precision that is observable. In fact, in experimental data, the typical error can be of the order of tens of meters, even reaching hundred of meters at specific times. An example of these errors (typically attributed to atmospheric artifacts) is shown in figure 9 of the paper by Rocca *et al.* (1997).

A detailed study of the meteorological parameters at the specific times of image acquisitions can be used in several models for representing the atmospheric delay on the HF sounding signal propagation (e.g., Saastamoinen, 1972). This can allow the subtraction of the atmospheric fluctuations and an increase in topographical precision. However, the availability of meteorological information over the entire observed area and at the specific times of the observations is not always realistic. In view of the significant dynamics of atmospheric perturbation, this can represent a limitation in the possible precision of the removal of atmospheric effect from SAR interferograms.

In addition to the difficulty of estimating and subtracting the atmospheric effects, it is worth emphasizing that the useful information about terrain displacements is impossible to be extracted from observations of any random three interferograms. For example, one successful algorithm for identifying this displacement component consists of an elaborated statistics applied to many interferometric couples that have been sorted to correspond to the basic requirement that their interferometric phase shift due to the topography can be approximated to zero for exceptionally small baselines (Berardino *et al.*, 2002). In fact, the topographical component given in Eq. (15) can be approximated to be zero if B_n is very small compared with the distance sensor-target (R_0). In this case, the only role of the 'small baseline algorithm' is to remove the spurious atmospheric artifacts. However, this algorithm can not be applied to just any random set of interferograms (requiring the small baseline condition) and, in any case, the specific case of only three available interferograms (as in the example given here) would represent a too small statistical data set.

In conclusion, up to the present time the availability of precise and extensive ground-based information is generally required for the good quality of topographical and ter-

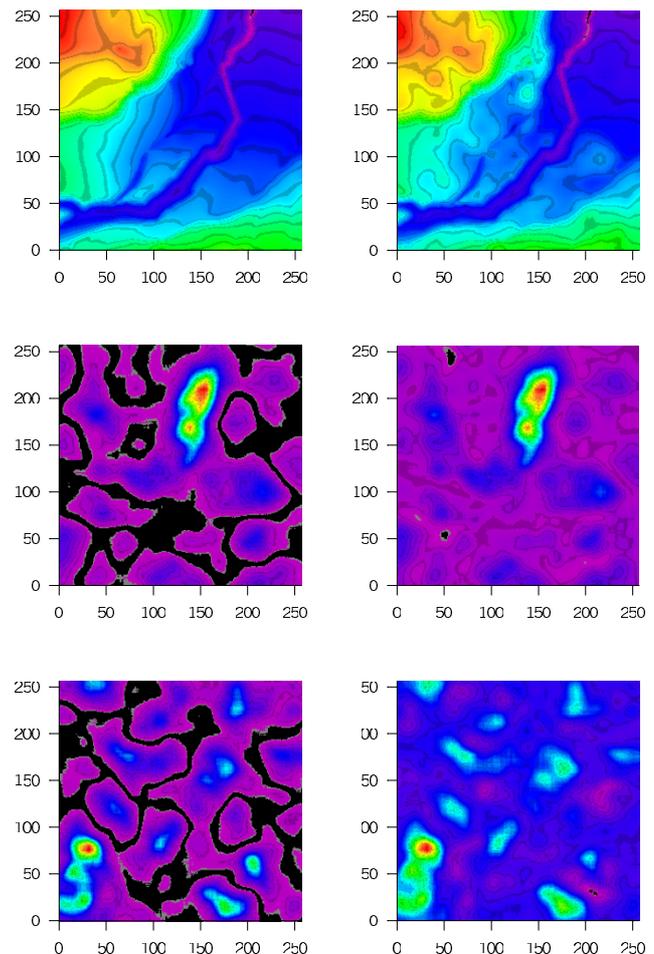


Fig. 3. Maps of the interferometric components used in the mixtures of Fig. 1 (left columns) and of the independent components derived by FastICA (right column): topographical components are shown in the top two panels, terrain displacement ones are shown in the middle panels, and spurious/atmospheric ones are shown in the bottom panel.

rain displacement observations using SAR interferograms (see Ferretti *et al.*, 2001). However this ground information is not always available, as in the case of the most remote area (for example, large forests, remote islands or deserts) that are not accessible via ground transportation. For these kinds of SAR observations, the blindness of the ICA algorithms represents a very useful processing tool.

3.3 Processing and results

The three images corresponding to the three phase shifts due to the topographical, terrain displacements, and atmospheric components are reported in the left column of Fig. 3, respectively, from the top to the bottom. The sum of these three components is represented in the first of the images on the left of Fig. 1. The other two panels of Fig. 1 correspond to the two sums of components that are proportional to the topographical and terrain displacement components shown in the left column of Fig. 3, respectively, plus different atmospheric components, as explained in Section 3.1.

After input of the three mixtures shown in Fig. 1 into FastICA, the output gives the three elements of the vector S of the sources. The corresponding three images (s_1 , s_2 , and s_3) are shown on the right column of Fig. 3. The values reported in the left and right columns of Fig. 3 are dif-

ferent because the estimated sources are extracted as linear functions of the original sources; therefore, the colors are differently scaled in each image. However, the parameter that is significant for the precision of the estimation is the linear correlation of s_1 , s_2 , and s_3 with the original $\delta\phi_{i,\text{topogr}}$, $\delta\phi_{i,\text{displ}}$, and $\delta\phi_{i,\text{atm}}$. These correlation coefficients are 0.987 for the topography, 0.989 for the terrain displacements, and about 0.897 for the atmosphere, with the latter correlation coefficient being slightly dependent on the specific interferogram number i . It is worth mentioning that all of the correlation coefficients correspond to a statistical confidence level above 99.9%. For the reason of completeness, we recall that the confidence level or statistical significance for a given correlation coefficient ρ is given by the difference $(1 - \beta)$, where β is the probability that correlation coefficients higher or equal to ρ can be obtained for uncorrelated variables. To be expressed as a percentage, it has to be multiplied by 100. An example of a computation with statistical significance is give by the sub-routine 'Pearsn', given in Press *et al.* (1990).

These good correlations mean that Eqs. (5), (6), and (7) are valid for the vector of the sources S given by FastICA. Therefore, we can compute the absolute values of topography and terrain displacements by re-writing Eq. (5) together with Eq. (15), and Eq. (6) together with Eq. (16)

$$s_1 = k_1 \cdot q + k_2 \quad (17)$$

$$s_2 = k_3 \cdot z + k_4 \quad (18)$$

The unknown parameters k_1 , k_2 , k_3 , k_4 can be assumed for all the corresponding pixels of s_1^j and q^j or of s_2^j and z^j , once these have been calculated for only two pixels. We made this calculation, and we can check its precision by using the real q and z used for the simulations. This precision can be expressed as relative errors $|(q - \underline{q})/q|$ and $|(z - \underline{z})/z|$, where the underline indicates the values estimated by Eqs. (17) and (18). For all the 256×256 pixels of our simulations, the maximum relative error is 12% for the topography and 15% for the terrain displacements. In particular, for the topography we obtain a relative error smaller than, or equal to 3% for 79% of the pixels. These values can be compared with those reported in Section 3.2 and indicate the usefulness of applying a FastICA processing to SAR interferograms.

Apart from the vector of the sources S , the FastICA also estimates the mixing matrix A . An interesting observation can be derived by considering the elements of this matrix, whose elements are reported in Table 1, together with the baselines B_n and the coefficient of proportionality among the terrain displacements components in our simulations. The sequence of the elements $a_{i,j}$ associated with the source s_1 are larger or smaller similarly to the sequence of the baselines taken into account for the simulation and, in fact, s_1 corresponds to the topographical component. This characteristic of the elements of the mixing matrix A can be useful for recognizing which one is the topographical component among the three sources of the vector S .

Funaro *et al.* (2001) carried out a similar use of the mixing matrix and associated the elements of the elements of A with the temporal evolution of the corresponding sources at the different times of the observations.

Table 1. Table of the coefficients of the mixing matrix A ($a_{i,j}$), of the satellite normal baselines (B_n) and of the coefficient of proportionality among terrain displacement components in each one of the three interferograms considered (C_{displ}).

	$i = 1$	$i = 2$	$i = 3$
$a_{i,1}$	3.104313	6.225242	4.34766
$a_{i,2}$	-0.335854	-0.7481157	-0.172066
$a_{i,3}$	-0.159158	-0.216163	-0.173795
B_n	50 m	100 m	70 m
C_{displ}	1.0	1.5	3.1

Regarding the elements of A corresponding to the terrain displacements, we note that the third coefficient is higher than the other two both in the FastICA estimation and in the simulation. The second coefficient is slightly smaller than the first one in the FastICA estimation, but not in parameters used for the simulation, which are, respectively, 1 and 1.5. This could be interpreted in terms of the small difference between the coefficients 1 and 1.5 or to the small absolute value of the phase shift associated to the terrain displacements compared with the phase shift associated to the topography. Further studies are ongoing (also on real data) to establish specific criteria for an automatic recognition of the nature (topographical, terrain displacements, artifacts) of the independent components estimated by FastICA. We anticipate that, in addition to the elements of the matrix A , the distribution of pixel values can also be of use for this purpose.

In any case, for the artifact or atmospheric component, we note that the coefficient of the matrix A does not exhibit any specific modulation in the function of the interferogram number i and is always about equal to -0.2 .

Modifications of the mixtures used for the simulations indicate that all of the results described above are not significantly dependent on the specific example presented here.

4. Applications on Real Data: Present Limitations and Further Studies

The aim of this section is to demonstrate some of the practical difficulties arising in real data applications and the motivation for further studies on the use of ICA techniques on SAR interferograms. An additional aim is to highlight the fact that many different kinds of FastICA applications on real data are possible.

The data considered here are ERS-SAR data (provided by the European Space Agency) for the geographical area around the city of Benevento (geographic coordinates $41^\circ 08'N$ and $14^\circ 46'E$), in southern Italy during the year 1995. In particular, the interferograms are computed on ERS-Tandem couples of images, which correspond to the same ground area and taken at about a 1-day delay by ERS-1 and ERS-2. More specifically, the acquisition dates considered here are: (1) July 13 and 14, 1995; (2) October 26 and 27, 1995; (3) November 30 and December 1, 1995. The input data into FastICA are the three unwrapped interferograms reported in Fig. 4, which are extracted from frame no. 819 of the ERS acquisitions specified above. Figure 4 is derived by applying GAMMA R.S. software on raw ERS SAR data.

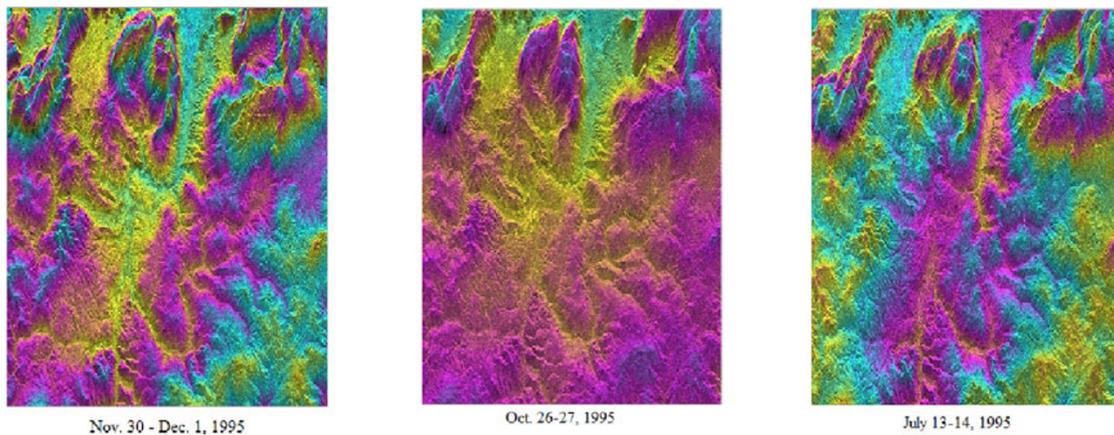


Fig. 4. Images extracted from three experimental interferograms of the ERS SAR data, representing the mixture signals given into FastICA input.

One of the three independent components automatically extracted by FastICA looks very similar to each one of the three images shown in Fig. 4. Therefore, this is selected to be the topographical component. This component has been transformed into altitudes (in meter units) using values of two ground points and compared (through interpolations in coordinates) with a standard DEM available for the area under study. This DEM is provided by the United States Geographical Survey SRTM DTED Level 1 (3 arc second), and it has a nominal vertical accuracy ≤ 18 m. The comparison between extracted and available DEM indicates that 100% of the data coincide within the accuracy error of 18 m, although the interferograms reported in Fig. 4 are significantly different one from the other, even by visual inspection.

The other two independent components extracted by FastICA might eventually be associated with atmospheric or terrain displacement effects. However, their significance is still being studied, and the criteria for possible automatic differentiation between atmospheric and landslide effects are going to be defined in a following paper. The probability-distribution-functions of the atmospheric variations and of the landslides are certainly different. However, other sources different from the landslides may appear as dominant independent components, and more than three interferograms may be necessary for a good precision of the blind separation FastICA technique. In any case, to verify the accuracy of absolute values of terrain movements, it would be necessary to have certain specific information about terrain movements that occur at least at two ground point locations. If the ongoing study indicates that this further information is crucial for the identification of the terrain displacement component, the FastICA technique would be modified into a semi-blind application (see Valpola and Särelä, 2002) through post-processing of extracted sources.

The specific interferogram computation considered in this section (using couples of images at 1-day delays) is slightly different from the simulation presented earlier. In fact, in each interferogram, the effects due to 1-day movements are about null ($\delta\phi_{i,\text{displ}} \sim 0$, for $i = 1, 2, 3$), and the differences induced by landslides that occurred between July 13 and December 1 affect the interferograms as an ad-

ditional artifact, similar to those of the atmospheric contamination. In particular, the landslides in the specific area under study are of the order of few millimeters per year. Therefore, between July 13 and December 1 of the year 1995, their contribution to topographical changes is relatively small.

In order to better measure the movement of regular landslides in the area under study, the interferograms would have to have been produced by using couples of images acquired at more distant months—possibly years—so that the $\delta\phi_{i,\text{displ}}$ components would have been larger and better differentiable from the atmospheric or other spurious artifacts. If we had taken the same six images, specified these but sorted them in different couples with monthly delays (July-October, July-November, October-July), the images of Fig. 4 would have shown a much smaller coherence. Even in this case, the topographical component is identifiable and in good agreement with the available DEM. However, the images of the other two independent components extracted by FastICA are different from the ones extracted earlier. In this case, the landslide components $\delta\phi_{i,\text{displ}}$ (for $i = 1, 2, 3$) are different from zero and their extraction is expected to have a different statistical meaning. However, we do not have accurate data on terrain movements (with a specification of possible errors) in the region under study during the specific considered dates. Some estimates are carried out by the Service for Landslide Monitoring (whose www site is presently reachable at <http://www.slamservice.info>), but these are average speeds of maximally 29 mm/year that were computed considering data spaced over period of years, so that there is no possibility of any quantitative comparison with the independent components extracted here by FastICA. Future comparisons are planned to be quantitatively more precise because in 2005 several GPS were installed on the ground in critical areas around the city of Benevento.

In particular, in past years this area has been affected by terrain movement of multiple origins: abundance of rain associated with critical slopes, earthquakes, and others. Therefore, if the landslides have to be also differentiated one from the other according to their different nature, more than three independent components (therefore more

than three interferograms) would be necessary.

In conclusion, the calibration and the automatization of the FastICA technique for real data application is still being completed for what concerns the identification and differentiation of different kinds of landslides in automatic mode. Once these studies are completed, it may also be possible to extract landslides, similarly to topography, as blind sources or, eventually, in semi-blind automatic modes.

5. Conclusion

The application of the FastICA algorithm on an array of three simulated SAR interferograms indicates its capability in deriving estimates of terrain altitude and displacements that are linearly correlated with the original data. The corresponding correlation coefficients are above 0.9, with a statistical confidence level above 99.9%.

This finding suggests the possibility of estimating the relative ground altitude and displacements in satellite SAR images of remote lands, with whatever value of the baselines and in the total absence of ground truths. Specifically, the subtraction of image artifacts due to the atmospheric fluctuations is automatically performed, without any knowledge of the local meteorology.

The advantage of this ICA technique compared with the PS technique is in the significant smaller amount of ground-based information necessary. Moreover, compared with the SBAS technique, the present ICA approach can be useful to study time intervals for which the perpendicular baselines in the available data sets cannot be approximated to be zero.

Preliminary results on real data show that the topographical component is of immediate and automatic extraction by FastICA application. Differently, the extraction and differentiation of different kinds of landslide is under study with respect to the automatization and generalization of procedures.

Acknowledgments. The author thanks the European Space Agency for providing ERS-1 and ERS-2 SAR data within the context of the Category-1 Scientific Project no. 3193.

References

- Berardino, P., G. Fornaro, R. Lanari, and E. Sansosti, A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms, *IEEE T. Geosc. Remote*, **40**, 2375, 2002.
- Cardoso, J. F., Blind signal separation; statistical principles, *Proc. IEEE*, **86**, 2009, 1998.
- Comon, P., Independent component analysis—a new concept?, *Signal Process.*, **36**, 287, 1994.
- Ferretti, A., A. Monti Guarnieri, C. Prati, and F. Rocca, Multibaseline interferometric techniques and applications, Proc. Fringe '96, ES-RIN/ESA, Frascati (Italy), 1996.
- Ferretti, A., C. Prati, and F. Rocca, Permanent scatters in SAR interferometry, *IEEE T. Geosc. Remote*, **39**, 8, 2001.
- Fujiwara, S., T. Nishimura, M. Murakami, H. Nakagawa, and M. Tobita, 2.5-D surface deformation of M6.1 earthquake near Mt. Iwate detected by SAR interferometry, *Geophys. Res. Lett.*, **27**, 2049, 2000.
- Funaro, M., E. Oja, and H. Valpola, Artefact detection in astrophysical image data using independent component analysis, Proc. 3rd Int. Conf. ICA Blind Source Separation, edited by T. W. Lee *et al.*, San Diego, Calif., Dec. 9–12, 2001.
- Hanssen, R. F., *Radar Interferometry*, Kluwer Academic Publishers, Dordrecht, 2001.
- Hyvärinen, A., New approximations of differential entropy for independent component analysis and projection pursuit, in *Advances in Neural Information processing Systems*, vol. 10, pp. 273–279, 1998.
- Hyvärinen, A. and E. Oja, fast fixed-point algorithm for independent component analysis, *Neural Comp.*, **9**, 1483, 1997.
- Hyvärinen, A. and E. Oja, Independent component analysis: algorithms and applications, *Neural Networks*, **13**, 411, 2000.
- Hyvärinen, A., J. Karhunen, and E. Oja, *Independent Component Analysis*, Wiley Interscience, New York, 2001.
- Massonnet, D. and K. L. Feigl, Radar interferometry and its application to change in the Earth's surface, *Rev. Geophys.*, **36**, 441, 1998.
- Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, Cambridge Univ. Press, New York, 1990.
- Rocca, F., C. Prati, and A. Ferretti, An overview of SAR interferometry, Proc. 3re ERS Symp, Florence, <http://florence97.ers-symposium.org>, 1997.
- Saastamoinen, J., Contribution to the theory of atmospheric refraction, *Bull. Geodesique*, **105**, 106, 1972.
- Valpola and Särelä, A fast semi-blind source separation algorithm, Tech. Rep. A66, Lab of Computer and Information Science, Helsinki University of Technology (Finland), 2002.
- Zebker, H. and R. Goldstein, Topographic mapping from SAR observation, *J. Geophys. Res.*, **911**, 4993, 1986.