Tropospheric corrections to SAR interferometry from GPS observations

Volker Janssen · Linlin Ge · Chris Rizos

Abstract Interferometric synthetic aperture radar (InSAR) techniques have been recognized as an ideal tool for many ground deformation monitoring applications. However, the spatially and temporally variable delay of the radar signal propagating through the atmosphere is a major limitation to accuracy. The dominant factor to be considered is the tropospheric heterogeneity, which can lead to misinterpretation of InSAR results. In this paper, a between-site (BS) and between-epoch (BE) double-differencing algorithm for the generation of tropospheric corrections to InSAR results based on GPS observations is tested. In order to correct the radar results on a pixel-by-pixel basis, the GPS-derived corrections have to be interpolated. Using experimental data it has been found that the inverse distance weighted and kriging interpolation methods are more suitable than the spline interpolation method. Differential corrections as large as several centimeters may have to be applied in order to ensure sub-centimeter accuracy for the InSAR result. The algorithm and procedures described in this paper could easily be implemented in a continuous GPS network data center. The interpolated image of BS, single-differenced tropospheric delays can be derived as a routine product to assist radar interferometry.

Received: 13 February 2004 / Accepted: 6 April 2004
Published online: 4 August 2004

V. Janssen · L. Ge (✉) · C. Rizos
School of Surveying and Spatial Information Systems,
The University of New South Wales, Sydney,
NSW, 2052, Australia
E-mail: l.ge@unsw.edu.au
Tel.: +61-2-93854177
Fax: +61-2-93137493

Introduction

Interferometric synthetic aperture radar (InSAR) is a technique first suggested in the early 1970 s (Graham 1974). The technique produces an "interferogram" from the phase difference between two SAR images acquired over the same region. The satellites which currently acquire SAR images are ERS-2 and Radarsat-1. Envisat was launched in March 2002, and new radar satellites are planned for launch over the next few years, e.g. the Japanese advanced land observation satellite (ALOS). The InSAR interferogram contains several types of information: (1) *topographic pattern*, a contour-like pattern representing the topography of the area, (2) *geometric pattern*, a systematic striped pattern caused by differences between the two SAR sensor trajectories, and (3) *differential pattern*, fringes associated with any change of the range between the two SAR images, the sources of which include ground displacement, change of atmospheric refraction, and phase change by reflection due to, for example, growth of vegetation. The geometric pattern can be removed by modeling the geometry of the satellite orbits and ground targets. Unless a high-resolution digital elevation model (DEM) is available, an additional radar image is necessary to remove the topographic pattern from the interferogram. In a typical ERS three-pass InSAR procedure, two repeat-pass ERS-2 images will be processed to generate the interferogram (InSAR result 1), containing all the information mentioned earlier. A third ERS-1 image which forms a tandem pair with one of the two ERS-2 images is also introduced. The tandem pair with the ERS-2 satellite following the ERS-1 satellite 1 day later can be processed to generate the topographic pattern (InSAR result 2), because the deformation and growth of vegetation, etc., within 1 day can be neglected. By differencing the two InSAR results, the *residual* interferogram will contain only the differential pattern.

Due to its high spatial resolution, ability of SAR to penetrate clouds, and cost effectiveness, InSAR has definite advantages over many conventional deformation monitoring techniques. Many earthquake rupture zones and volcanoes have been studied using InSAR (Massonnet et al. 1993; Lu et al. 1997). Studies, however, have shown that a change of atmospheric refraction (e.g., caused by a cold front moving across the region being imaged) can result in biases, which can lead to a misinterpretation of the InSAR results (Zebker et al. 1997; Hanssen et al. 1999).

 DOI 10.1007/s10291-004-0099-1

Therefore, in order to reliably derive ground displacement from InSAR results, it is crucial to correct for the atmospheric heterogeneity.

The atmospheric heterogeneity can be partitioned into tropospheric and ionospheric portions. In general, the troposphere can be divided into a wet component and a dry component. The ionosphere extends, in a number of distinct layers, from about 50–1,000 km above the Earth's surface. The SAR satellite orbit altitudes are typically in the range of 600–800 km. The effect of the variations caused by the ionospheric layers lower than the SAR satellite altitude will be much smaller than that from the troposphere because the area penetrated by the radar is much smaller. For example, for a SAR satellite at 800 km altitude the extend of the image in the ionosphere at the height of 400 km is 40 km, while it will be about 80 km within the troposphere. Therefore, the ionospheric delay on the radar signal is usually considered to be uniform within one SAR image and can mostly cancel because the SAR images are acquired at the same time of the day, and hence the residual effect can be neglected. It is the tropospheric variations that can lead to misinterpretation of InSAR results. While the dry component of the tropospheric delay is well modeled, the wet component is much more difficult to model because of the large variations of water vapor content with respect to time and space (Spilker 1996).

Since 1997, researchers have been developing methodologies to correct InSAR results for these biases using measurements from other techniques, such as GPS (e.g., Bock and Williams 1997; Ge et al. 1997; Ge 2000). However, progress has been slow because in order to integrate InSAR with GPS both datasets have to be available for the same region, at the same time, and the region under study has to be experiencing ground displacement. The establishment of continuous GPS (CGPS) arrays in many parts of the world has eased such difficulties significantly [Southern California Integrated GPS Network (SCIGN 2003); Geographical Survey Institute (GSI 2003)].

In this paper, a between-site (BS) and between-epoch (BE) double-differencing algorithm for the generation of tropospheric corrections to InSAR results based on GPS observations, as proposed by Hanssen (2001), is tested. The tropospheric parameters are interpolated in order to enable the radar results to be corrected on a pixel-by-pixel basis. Experimental results generated from data collected in two CGPS networks are presented.

GPS-derived tropospheric delay

The troposphere can be defined as the neutral (i.e. non-ionized) part of the atmosphere that stretches from the Earth's surface to a height of approximately 50 km. The dominant impact of tropospheric path delay on radio signals occurs in the lower part, typically below 10 km (Spilker 1996). The tropospheric delay is dependent on temperature, atmospheric pressure and water vapor content. The type of terrain below the signal path can also have an effect. The tropospheric effect can be divided into two components, the dry and the wet component. The dry component accounts for about 90% of the effect and can be accurately modeled using surface measurements of temperature and pressure. However, due to the high variation in the water vapor content, it is very difficult to model the remaining wet component.

Several models based on a "standard atmosphere" have been developed to account for the tropospheric delay in the absence of accurate ground meteorological data, e.g., the Hopfield (1969) model, Saastamoinen (1973) model and Black (1978) model. As recommended by Mendes (1999), the Saastamoinen model has been used in this study. This model utilizes the gas laws to deduce refractivity, and the tropospheric delay is therefore a function of zenith angle, pressure, temperature and the partial pressure of water vapor. Saastamoinen (1973) used the refractivity constant given by Essen and Froome (1951) for mid-latitudes and average conditions. The original model has subsequently been refined to include two correction terms: one being dependent on the station height (B) and the other on the height and the zenith angle (δR). Both terms can be obtained from tables.

The tropospheric delay, expressed in meters, is then given by Bauersima (1983):

$$\mathrm{dTrop}_{\mathrm{apr}} = \frac{0.002277}{\cos z}\left[p + \left(\frac{1255}{T} + 0.05\right)e - B\tan^2 z\right] + \delta R \tag{1}$$

where z denotes the zenith angle of the satellite, p the atmospheric pressure in millibars, T the temperature in Kelvin, and e the partial pressure of water vapor in millibars.

For high precision surveys, an additional parameter can be introduced into the least squares reduction of the observations to estimate the *residual* tropospheric delay (after modeling). The total tropospheric delay correction dTrop_i^k can be expressed as (Rothacher and Mervart 1996):

$$\mathrm{dTrop}_k^i = \mathrm{dTrop}_{\mathrm{apr},k}\, f_{\mathrm{apr}}(z_k^i) + \mathrm{dTrop}_k(t) f(z_k^i) \tag{2}$$

where $\mathrm{dTrop}_{\mathrm{apr},k}$ denotes the tropospheric delay according to the a priori model, which is time-invariant (i.e., dependent on the station height only) if a standard atmosphere is used. $z_k{}^i$ denotes the zenith angle (for satellite i and station k), f_{apr} the mapping function (different for each a priori model), $\mathrm{dTrop}_k(t)$ the time-dependent troposphere parameter for station k, and is $f(z_k{}^i)$ is the mapping function used for the parameter estimation, which may be different from f_{apr} and is usually $1/\cos z$.

In this study, the Bernese-GPS processing software was used to derive tropospheric delay parameters for the individual stations of the network during parameter estimation. The user can specify the number of correction parameters to be estimated within the observation period.

Double-differencing algorithm for tropospheric delay corrections

Only the relative tropospheric delay (the tropospheric heterogeneity) between two SAR imaging points and between the two SAR image acquisitions will distort the deformation information derived by InSAR, because it is the phase difference that is used and deformation is always referenced to a stable point (site) in the image. Therefore, a BS and BE double-differencing algorithm can be used to derive the corrections to the InSAR result from GPS observations (Hanssen 2001).

Single-differences

Assume that A is a stable site in the SAR image to be used as a reference point. B is another site in the same SAR image. If the tropospheric delay estimated from GPS for A and B at SAR imaging epoch j is denoted as ${D_A}^j$ and ${D_B}^j$ respectively, the BS difference of the delays is:

$$D^j_{AB} = D^j_B - D^j_A \tag{3}$$

Using site A as the reference, single BS difference delays at other GPS sites can also be calculated using Eq. 3, which are then interpolated (see next section) to generate a tropospheric delay image product similar to the radar single-look-complex (SLC) data.

Double-differences

Assuming two sites A and B, and two epochs j (master SLC image) and k (slave image), two single-differences may be formed according to Eq. 3:

$$\begin{aligned} D^j_{AB} &= D^j_B - D^j_A \\ D^k_{AB} &= D^k_B - D^k_A \end{aligned} \tag{4}$$

A double-difference is obtained by differencing these single-differences:

$$\begin{aligned} D^{jk}_{AB} &= D^k_B - D^j_A \\ &= (D^k_B - D^k_A - D^j_B - D^j_A) \\ &= (D^k_B - D^j_B) - (D^k_A - D^j_A) \end{aligned} \tag{5}$$

Equation 5 illustrates two possible approaches to double-differencing, either BS differencing first and then BE differencing (BSBE approach), or BE differencing first and then BS differencing (BEBS approach). The BSBE approach is preferred because the BS difference can be interpolated to generate a single-difference correction product. This product will be associated with only the SLC image and hence can be used freely to form combinations for further BE differences as soon as InSAR pairs have been formed from SLC images.

Interpolating tropospheric delay corrections

Continuous GPS networks may be as dense as one station every 25 km at the national level, as is the case for the GPS Earth observation network (GEONET) in Japan (GSI 2003), or as dense as one station every few kilometers at the regional level, as is the case for the SCIGN in the USA (SCIGN 2003). However, in order to correct the InSAR result on a pixel-by-pixel basis (ERS SAR resolution ~25 m), the GPS-derived tropospheric corrections have to be interpolated.

In this section, the utility of three interpolating methods will be discussed. Each interpolation technique makes assumptions about how to determine the estimated (interpolated) values. Depending on the phenomenon being modeled (i.e. differential tropospheric delay) and the distribution of sample points (in this case, GPS stations), one interpolator may produce better models of the actual surface (the tropospheric delay correction model) than others. Regardless of the interpolator, as a rule-of-thumb, the more input points and the more even their distribution, the more reliable the results.

Inverse distance weighted (IDW) interpolation

Inverse distance weighted interpolation (Lancaster and Salkauskas 1986) explicitly assumes that things that are close to one another are more alike than those that are farther apart. To predict a value for any unmeasured location, IDW will use the measured values surrounding the prediction location. Those measured values closest to the prediction location will have more influence on the predicted value than those farther away. Thus, IDW assumes that each measured point has a local influence that diminishes with distance, hence the name "inverse distance weighted".

The general formula of IDW is:

$$\widehat{D} = (\lambda_0,\ \varphi_0) = \sum_{i=N}^{N} w_i D(\lambda_i,\ \varphi_i) \tag{6}$$

where $\widehat{D}(\lambda_0,\ \varphi_0)$ is the interpolated tropospheric delay for a location point with easting λ_0 and northing ϕ_0, N is the number of GPS stations surrounding the prediction location that will be used in the interpolation, and w_i (i=1,2,...,N) are the weights assigned to each GPS-derived delay value that will be used. For IDW these weights will decrease with distance to the interpolated location. $D(\lambda_i, \phi_i)$ is the GPS-derived delay (either single-differenced or double-differenced) at location easting λ_i and northing ϕ_i. The weights are determined as follows:

$$w_i = \frac{d_{i0}^{-p}}{\sum_{i=1}^{N} d_{i0}^{-p}} \quad \text{and} \quad \sum_{i=1}^{N} w_i = 1 \tag{7}$$

From Eq. 7 it can be seen that as the distance becomes larger, the weight is reduced by a factor of p. The quantity d_{i0} is the distance between the prediction location (λ_0, ϕ_0) and each of the GPS stations (λ_i, ϕ_i). The power parameter p influences the weighting of the GPS-derived delay on the interpolated value: as the distance increases between the GPS stations and the prediction location, the weight (or influence) that the measured point will have on the prediction will decrease exponentially. By defining a high power, more emphasis is

placed on the nearest points, and the resulting surface will have more detail (be less smooth). Specifying a lower power will give more influence to the points that are further away, resulting in a smoother surface. A power of two is most commonly used. The weights for the GPS-measured locations that will be used in the prediction are scaled so that their sum is equal to one.

Spline interpolation

This general-purpose interpolation method fits a minimum-curvature surface through the input points (Schultz 1973). Conceptually, this is like bending a sheet of rubber to pass through the points while minimizing the total curvature of the surface. It fits a mathematical function (a minimum-curvature, two-dimensional, thin-plate spline) to a specified number of the nearest input points while passing through all input points. Therefore, the idea behind a spline fit is to approximate a function by a polynomial which is defined piecewise. This method is best for gradually varying surfaces. It is not appropriate when there are large changes within a short horizontal distance because it can overshoot estimated values. Hence, it would not be applicable to correct atmospheric interference induced by extreme weather conditions that may be caused by a cold front moving across the area.

For simplicity, the 1D "basis" splines (B-splines) are described here, which became popular when De Boor (1978) developed a package of FORTRAN routines for their numerical application. For example, a cubic spline fit uses cubic polynomials which are defined over distinct, non-overlapping regions. The term *spline* means that the coefficients of the polynomial are chosen so that the following conditions are satisfied at the borders when two regions abut: (a) the values of the fit polynomials are the same, and (b) one or more of the derivatives match as well so that the slope (first derivative), etc., are continuous. For cubic splines, it is possible to match the function values and first derivatives (slopes) at both ends of the interval, resulting in a sufficiently smooth join for most purposes.

The idea behind B-splines is to expand the function in "basis" splines $B(x)$, which are zero over most of the domain to be fitted. The $B(x)$ are splines, not simple polynomials—i.e., they are different polynomials in different regions. Consider the simplest useful B-splines, the cubic splines. The $B(x)$ will be non-zero in the region between $x[i]$ and $x[i+3]$, whereas $B(x)=0$ for $x<x[i]$ or $x>x[i+3]$. In order to be continuous, $B(x[i])=B(x[i+3])=0$. The function $B(x)$ can be written for a cubic B-spline as:

$$B(x) = \mathrm{A}(x - x[i])_+^3 + \mathrm{B}(x - x[i+1])_+^3 + \mathrm{C}(x - x[i+2])_+^3 + \mathrm{D}(x - x[i+3]_+^3 \quad (8)$$

where $(x-y)_+$ is $(x-y)$ if $(x-y)>0$ and zero otherwise. Thus, in the left-most interval, only the term proportional to A contributes, while in the right-most interval all of the terms contribute. The additional conditions on the derivatives of $B(x)$ at the end of the right-most interval, namely $B'(x[i+3])=B''(x[i+3])=0$, result in the B-spline being unique up to a normalizing constant which multiplies B. The resultant B-splines are bell-shaped functions which are non-negative. B-splines can be defined for higher degrees but cubic B-splines are generally used in practice. Beyond the endpoints of the domain there are points which are needed to define the B-spline at the edge of the domain. De Boor (1978) typically chooses $x[-3]=x[-2]=x[-1]=x[0]$ at the left-hand end, $x[0]$ being the endpoint of the domain, and similarly at the right-hand side. These points $x[i]$, at which the fit is defined, are typically called the *knots* or *breakpoints* of the splines. Often, the knot spacing is uniform within the domain, i.e., $x[i+1]-x[i]=\mathrm{d}x$ is a constant, although this is not necessary.

The process of fitting a function by splines involves determining the coefficients of the splines which satisfy the user-imposed conditions, which are typically to match the specified function and its derivatives on a set of points. This is done by building a spline approximation to the function from overlapping B-splines. The first cubic B-spline might cover the region from $x[0]$ to $x[3]$; the next would be defined over $x[1]$ to $x[4]$, and so on. Recall that each cubic B-spline, as defined above, has one free parameter, its scale factor. Performing the fit requires determining this scale factor. Because at the ends of its range the B-spline takes on the value 0, there will be three non-zero B-splines contributing to the value of the sum at each interior point. The linear system which must then be solved to fit the B-spline approximation to a set of function values is, then, a tri-diagonal system within the interior of the domain. Because of the multiple knots at the edges it is somewhat more complicated at either end. Such a system is still banded, and so generally can be solved without the complexity of a full system solver.

One interesting feature of B-splines is the locality of influence. The value of a function to be fitted influences only the coefficients of the B-splines which are non-zero over that interval. Thus, for cubic splines, only four coefficients are affected.

Kriging interpolation

This interpolation method assumes that the distance or direction between sample points reflects a spatial correlation that can be used to explain variations in the surface. Kriging fits a mathematical function to a specified number of points, or all points within a specified radius, to determine the output value for each location. Kriging is a multistep process including exploratory statistical analysis of the data, variogram modeling, creating the surface, and (optionally) exploring a variance surface (Stein 1999). This function is most appropriate when there is a spatially correlated distance or directional bias in the data.

Like IDW, kriging weights the surrounding GPS-measured values to derive a prediction for a non-measured location. The general formula for the kriging interpolator is the same as IDW, i.e., Eq. 6. However, in IDW the weight w_i depends solely on the distance to the prediction location. Kriging also takes into account the overall spatial arrangement among the measured points by quantifying the spatial autocorrelation. Thus, in ordinary kriging, the weight w_i depends on a fitted model to the measured

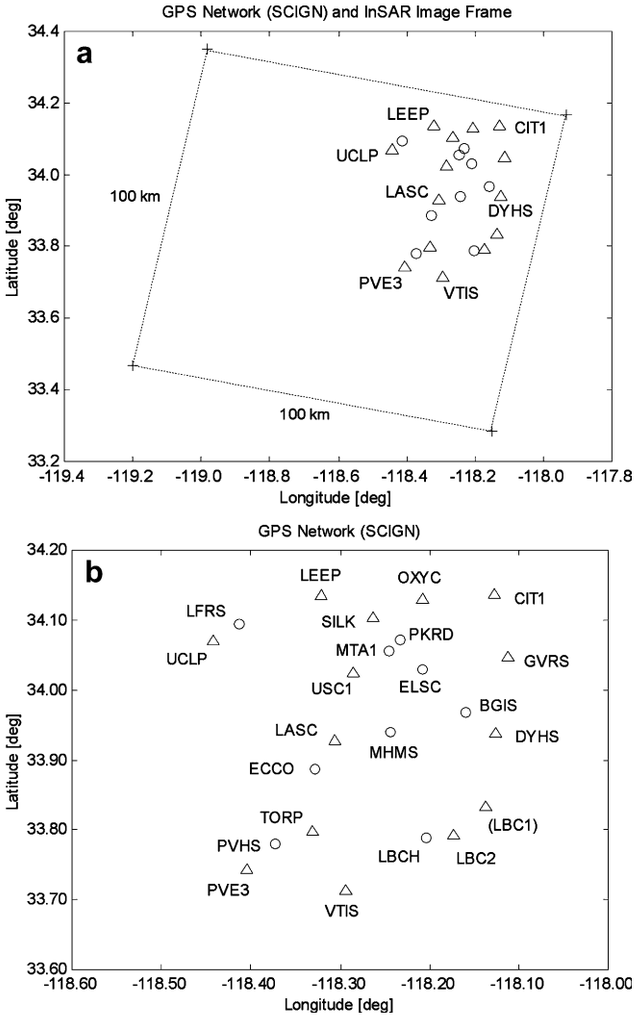

Fig. 1
SCIGN stations within the ERS SAR image frame (*top*), and a close-up showing reference stations (*triangles*) and prediction stations (*circles*) (*bottom*)

points, the distance to the prediction location, and the spatial relationships among the measured values around the prediction location.

In order to create the empirical semivariogram the distance and squared difference between each pair of locations has to be calculated. The distance d_{ij} between two locations (λ_i, ϕ_i) and (λ_j, ϕ_j) is determined by the Euclidean distance:

$$d_{ij} = \sqrt{(\lambda_i - \lambda_j)^2 + (\varphi_i - \varphi_j)^2} \tag{9}$$

The empirical semivariance s_{ij} is half the square of difference between the GPS-derived tropospheric delay for the two locations:

$$s_{ij} = \frac{1}{2}[D(\lambda_i,\ \varphi_i) - D(\lambda_j,\ \varphi_j)]^2 \tag{10}$$

With larger datasets (more GPS stations) the number of pairs of locations will increase rapidly and will quickly become unmanageable. Therefore, it is necessary to group the pairs of locations, in a process referred to as "binning". In this case, a bin is a specified range of distances. That is, all points that are $0<d_{ij} \leq 1$ kilometer apart are grouped into the first bin, those that are $0<d_{ij} \leq 2$ kilometers apart are grouped into the second bin, and so forth. The average empirical semivariance of all pairs of points in a bin is taken as the semivariance of the bin.

Now the average semivariance can be plotted against the average distance of the bins, to produce the empirical semivariogram. However, the empirical semivariogram values cannot be used directly because standard errors for the predictions might be negative; instead, a model must be fitted to the empirical semivariogram. Once this is done, the fitted model can be used to determine semivariogram values for various distances. For simplicity, the model to be fitted is a least squares regression line, which has been forced to have a positive slope and pass through zero. Many other models can also be used. The slope k of the regression line is then used to determine the semivariance γ_{ij} at any given distance:

$$\gamma_{ij} = kd_{ij} \tag{11}$$

where d_{ij} is the distance between two GPS stations at (λ_i, ϕ_i) and (λ_j, ϕ_j) calculated using Eq. 9. In order to interpolate the tropospheric delay at location (λ_0, ϕ_0), a matrix Γ and a vector $\mathbf{g}$ can be defined using the semivariance of Eq. 11:

$$\Gamma = \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{1N} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma_{N1} & \cdots & \gamma_{NN} & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} \gamma_{01} \\ \vdots \\ \gamma_{0N} \\ 1 \end{bmatrix} \tag{12}$$

The 1s and 0s in the bottom row and the right-most column of Γ, as well as the last element in $\mathbf{g}$, arise due to unbiasedness constraints. Now that the matrix Γ and the vector $\mathbf{g}$ have been defined, the kriging weights vector $\mathbf{w}$ can be solved for:

$$w = \begin{bmatrix} w_1 \\ \vdots \\ w_N \\ m \end{bmatrix} = \Gamma^{-1}\mathbf{g} \tag{13}$$

where Γ^{-1} is the inverse matrix of Γ . The m is an unknown to be estimated, arising from unbiasedness constraints. Therefore, the interpolation can now be carried out using Eq. 6. It should be noted that kriging uses the GPS-derived delay data twice: the first time to estimate the spatial autocorrelation of the data, and the second to make the predictions.

Experimental data analysis: SCIGN

Data from the SCIGN (2003) were used to investigate the feasibility of the above methods to derive tropospheric delay corrections from GPS observations. Of the 23 sta-

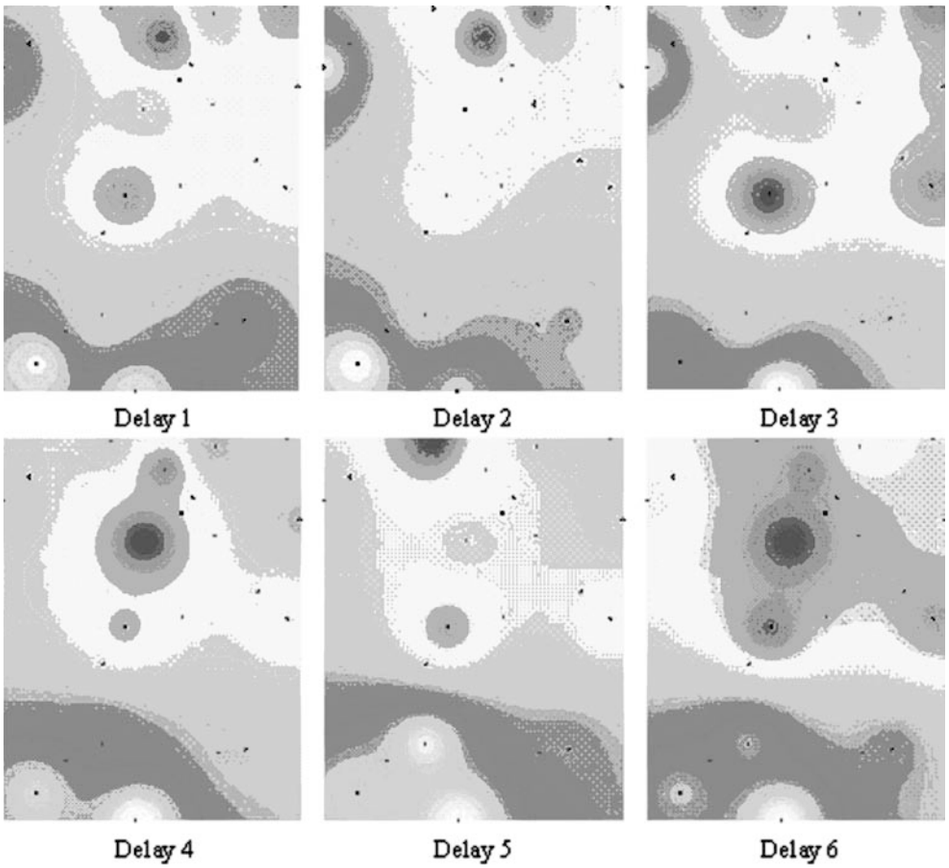

Fig. 2
Interpolation images for double-differenced tropospheric corrections (IDW)

tions considered, 14 were treated as measured locations (reference stations) and nine were used as prediction locations for which tropospheric delay corrections had to be determined and compared with their GPS-derived delays. A 2-h session was observed on August 2, 2001 (DOY 214) and again on September 6, 2001 (DOY 249), simulating a typical ERS SAR satellite single repeat cycle of 35 days. Data were collected at a 30-s sampling rate for a period of 1 h before and after the flyover of the radar satellite. Figure 1 shows the location of the GPS sites within a typical ERS SAR image frame (the dashed lines) for this area. A close-up of the GPS sites is also shown, where the reference stations are denoted by triangles, while the sites to be interpolated are indicated by circles. For all sites precise coordinates were obtained using the scripps coordinate update tool (SCOUT) provided by the scripps orbit and permanent array center (SOPAC 2003). This service computes the coordinates of a GPS receiver (whose data are submitted to the website) by using the three closest SCIGN reference sites and precise GPS ephemerides. In this case the coordinates were determined by taking the mean of six 24-h solutions obtained in two blocks of three successive days (DOY 213–215 and 248–250). The average baseline lengths ranged from 2 to 7 km. The repeatability of these six coordinate solutions was at the sub-centimeter level for all but one GPS site, indicating a solid, stable network. Site LBC1 showed relatively large coordinate variations indicating lower quality data or a possible displacement of 3.5 cm and has therefore been left out of the subsequent interpolation.

GPS-derived tropospheric delay corrections

The Bernese-GPS processing software (Rothacher and Mervart 1996) was used to process the network on both days, the coordinates of CIT1 being held fixed as the primary reference station. Baseline lengths vary from 7 to 49 km, and the largest height difference is 270 m. For each site tropospheric delay corrections were determined every 20 min, resulting in six parameters per site throughout the 2-h observation span. Single-differenced tropospheric corrections (Eq. 3) were then obtained by forming the differences relative to CIT1. These corrections range from −6.1 to +2.2 cm, and in some cases show variations of a few centimeters within the 2-h observation span.
Radar interferometry applications use two images of the same area in order to detect any ground deformation that might have occurred between the two satellite flyovers. To correct such an InSAR image for the effect of the tropospheric delay, the relative *change* in the tropospheric conditions is of great importance. Hence double-differenced tropospheric corrections are obtained by forming the BE difference of the single-differenced values derived in the previous step (Eq. 5). A comparison of the single- and double-differenced corrections revealed that almost all the double-differenced delay is smaller than the single-differenced delay (except for stations OXYC, MTA1 and PKRD). The double-differenced corrections range from

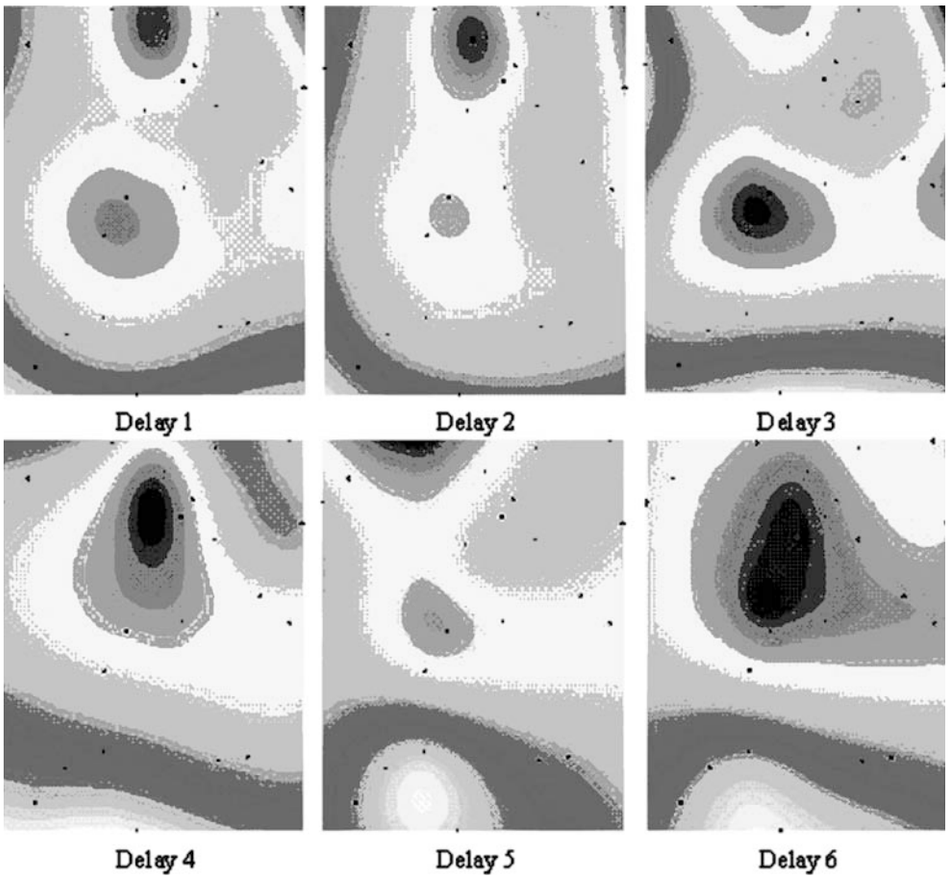

Fig. 3
Interpolation images for double-differenced tropospheric corrections (spline)

−5.0 to +3.3 cm although the 23 stations spread over only a quarter of the SAR image frame (Fig. 1). Therefore, it is crucial to apply such corrections in order for InSAR to achieve sub-centimeter accuracy.

Interpolation of tropospheric delay corrections

For each of the nine prediction sites shown in Fig. 1, the tropospheric delay corrections were interpolated using the three methods described earlier: IDW interpolation, spline interpolation and kriging interpolation. Both the single-differenced tropospheric corrections relative to CIT1 for days 214 and 249, and the double-differenced tropospheric corrections between these two epochs were investigated by comparing the interpolated values to the "true" values obtained directly using the Bernese software. This was done for each of the six 20-minute time intervals (Delay 1 through to Delay 6) within the 2-h observation span. Figures 2, 3 and 4 show the interpolation images obtained for the different interpolation methods in the double-differenced case, which is most important and can be directly used for the correction of InSAR results. The dots indicate the locations of the 22 GPS stations used in the analysis (refer to Fig. 1 for their codes) and the color/grey step interval is 1 mm. The main areas of tropospheric activity can be recognized in all three figures, and the temporal and spatial variability of the tropospheric delay is obvious. The double-differenced interpolation values obtained with the different interpolation methods only differ by small amounts and are generally below or just above the centimeter-level. However, they *do* reach values of up to 3 cm in some cases.

Which interpolation method is the most suitable?

In order to determine which interpolation method gives the best results, the standard deviations of the results compared to the "true" values obtained using the Bernese software were computed. The left graph of Fig. 5 shows the standard deviations for the single-differenced case on days 214 (top plot) and 249 (middle plot), as well as for the double-differenced case (bottom plot). It is obvious that all three interpolation techniques deliver results with the same accuracy in this particular case, which is mostly at the sub-centimeter level. For the fourth time interval the accuracy is considerably lower compared to the rest of the observation span, almost reaching the 2-cm level. This may have been caused by a short-term tropospheric event on day 249, which again highlights the importance of applying the differential tropospheric delay corrections to InSAR results. The tropospheric delay corrections are to be used to correct a set of InSAR images obtained from two SAR satellite flyovers. Hence, it is important that the reference stations (GPS-measured locations) do not undergo any deformation between these two epochs. In practice, however, small movements may still occur. These can be due to minor tectonic events, nearby construction work, or if the GPS site is in fact just inside the

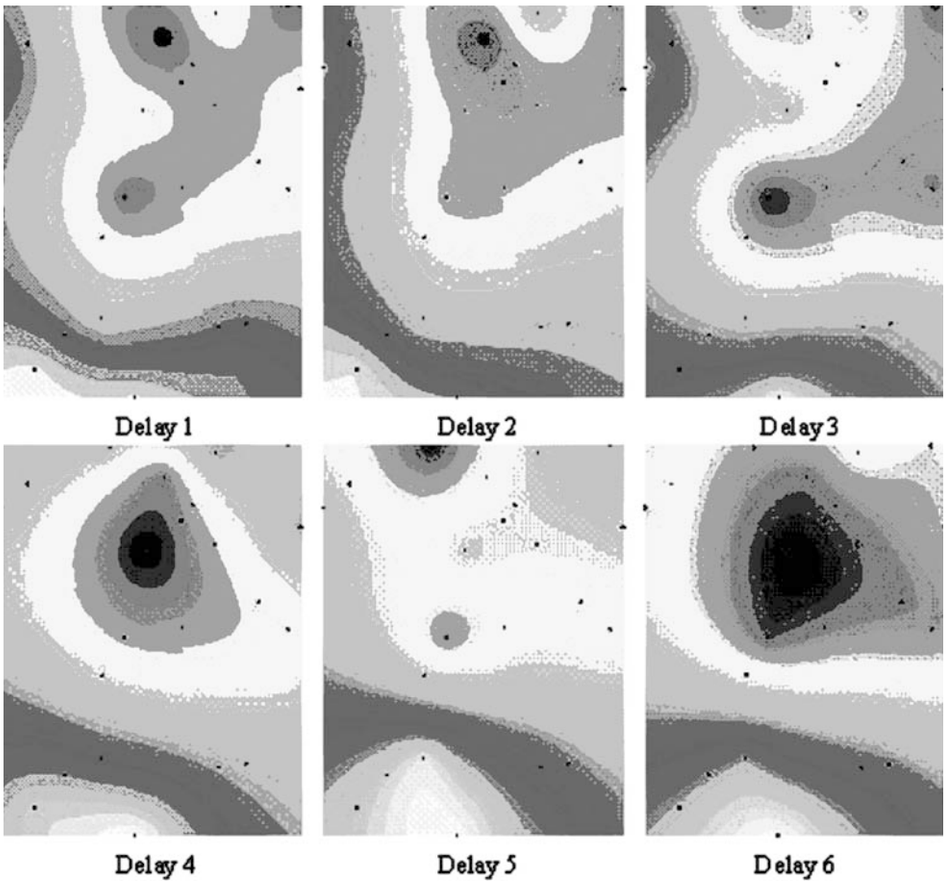

Fig. 4
Interpolation images for double-differenced tropospheric corrections (kriging)

deformation zone to be monitored. It is therefore useful to test the susceptibility of the interpolation techniques to outliers caused by small displacements in the reference stations or by reduced data quality. LBC1, a site that had earlier been identified as having a problem, was now included as a reference station in the interpolation process. The data were then processed again. The standard deviations of the resulting tropospheric corrections for the single-differenced case on days 214 (top plot) and 249 (middle plot), as well as for the double-differenced case (bottom plot), are shown in the right graph of Fig. 5.
It is obvious that the spline interpolation method has difficulties coping with such an "outlier" in the reference station network. Standard deviations reach values of up to 4 cm in the double-differenced case. The values for the IDW and kriging interpolation techniques remain unchanged compared to the "clean" reference network used in the previous case. Only the sixth time interval of the IDW interpolation on day 249 shows a change for the worse. However, this does not influence the double-differenced result (bottom right graph of Fig. 5), which indicates the robustness of the method. It is therefore suggested that either the IDW or the kriging interpolation method be used to determine tropospheric delay parameters from GPS observations. On the other hand, the two techniques can be used as a mutual check.

How many troposphere parameters should be determined?

The Bernese-GPS processing software allows the user to specify the number of tropospheric delay parameters to be determined. The estimation of about six to 12 parameters for a 24-h observation session is recommended by Rothacher and Mervart (1996). Estimating one parameter for every 2–4 h may be sufficient for geodetic control surveys where a set of coordinates is derived from a long observation session, taking into account all possible atmospheric effects. However, a special situation arises when one is dealing with GPS-derived tropospheric corrections for InSAR. The SAR satellite will pass over the area of interest at a certain epoch and we are specifically interested in estimating the tropospheric delay as accurately as possible at this epoch within the observation span. It is therefore necessary to determine how many parameters should be estimated in order to obtain an accurate representation of the tropospheric conditions at any point in time.
A sub-network involving three GPS sites from the original network (Fig. 1) was used. The baselines CIT1-UCLP and CIT1-VTIS are 30 and 49 km in length with height differences of 104 and 156 m, respectively. The 2-h session observed on September 6, 2001 (DOY 249) was processed several times incorporating a different number of

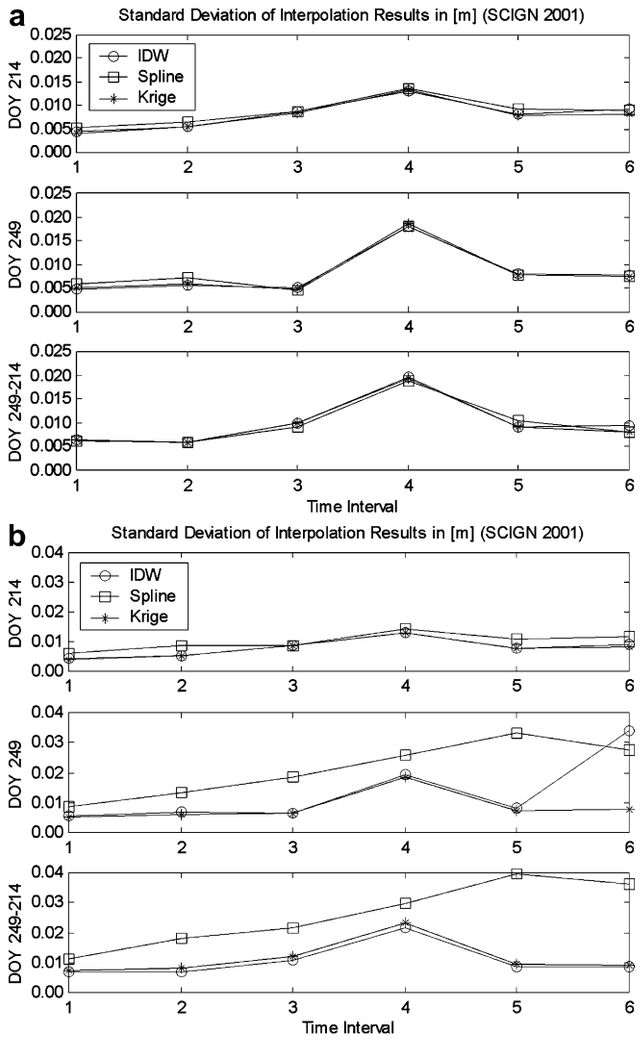

Fig. 5
Standard deviation of the interpolation results obtained by different methods for a "clean" reference network (*top*) and including an "outlier" (*bottom*)

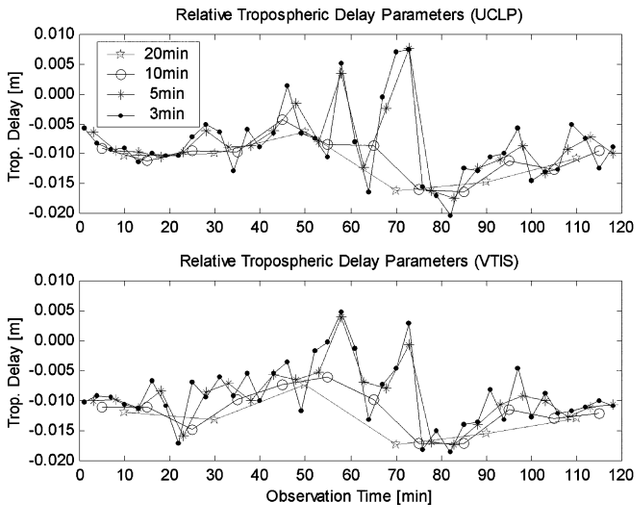

Fig. 6
Relative tropospheric delay parameters over 2 h

estimable troposphere parameters. Tropospheric delay corrections were estimated for time intervals of 20, 10, 5 and 3 min in length, corresponding to 6, 12, 24 and 40 parameters per site, respectively. Figure 6 shows the (single-differenced) tropospheric delay parameters for the sites UCLP (top) and VTIS (bottom), both relative to CIT1. The figure indicates that both the 3- and 5-min cases generate a rather detailed record of the variations in the troposphere. Short-term fluctuations are visible and values range from about +1 to −2 cm, even for the relatively small height differences of 100–150 m between the stations. The 10- and 20-min cases produce a smoothed record of the tropospheric delay, which is obviously less likely to represent the correct conditions present at a specific SAR time epoch. The resulting coordinates are practically the same for both the 3- and 5-min tropospheric parameter estimation, with variations at the sub-millimeter level. If compared to the results obtained using 10- and 20-min intervals, the coordinate differences are at the few-millimeter level. This corresponds to a difference of a few millimeters in the troposphere parameters between the 3- and 5-min cases on the one hand and the 10- and 20-min cases on the other (Fig. 6). It should be noted, however, that the short-term fluctuations could also in part represent noise. Independent data are needed in order to distinguish noise from the signal. This is currently under investigation. These results suggest that by estimating tropospheric delay parameters for 5-min time intervals during a 2-h observation session, the short-term variations of the troposphere can be reliably modelled. At the same time, the number of additional parameters to be estimated is still kept at a reasonable level.

Experimental data analysis: GEONET

Based on the above findings a second dataset from Japan's GEONET (GSI 2003) was analyzed. Of the 37 stations considered, 29 were treated as measured locations (reference stations) and eight were used as prediction locations for which tropospheric delay corrections had to be determined and compared with their GPS-derived delays. A 2-hour session was observed on June 17, 2002 (DOY 168) and on July 22, 2002 (DOY 203), again simulating a typical ERS SAR satellite single repeat cycle of 35 days, and covering the satellite flyover epoch. Figure 7 shows the location of the GPS sites, evenly distributed across a typical ERS SAR image frame (the dashed lines) for this area. The reference stations are denoted by triangles, while the sites to be interpolated are indicated by circles. Precise coordinates for all sites were provided by the GSI of Japan.

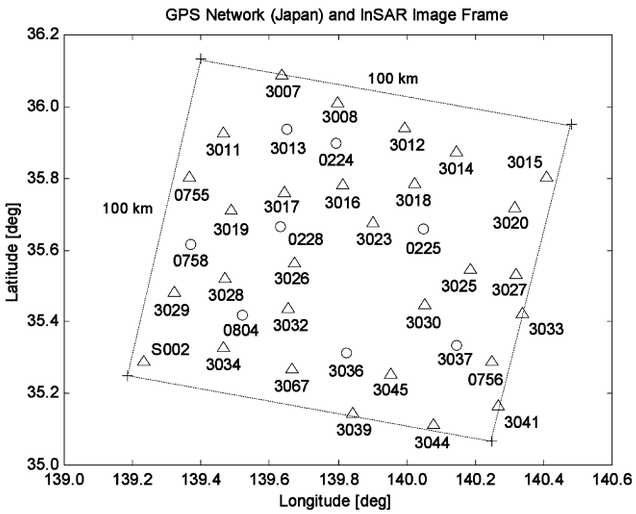

Fig. 7
GEONET stations within the ERS SAR image frame

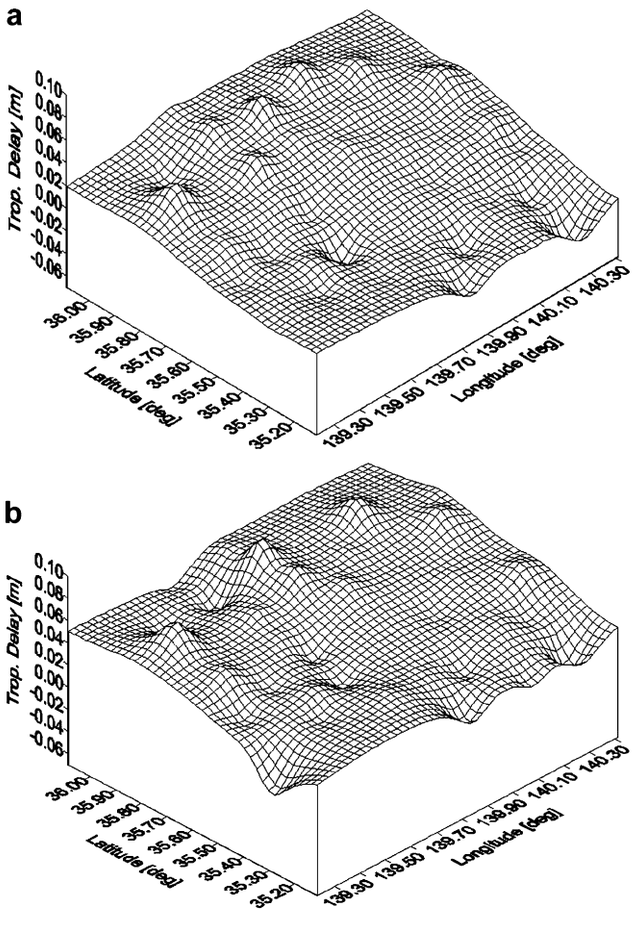

Fig. 8
Double-differenced interpolation maps for the 11th (*top*) and the 24th (*bottom*) time interval (IDW interpolation)

GPS-derived tropospheric delay corrections

Again, the Bernese-GPS processing software was used to process the network on both days, the coordinates of S002 being held fixed as the primary reference station. Baseline lengths vary from 22 to 121 km, and the largest height difference is 321 m. For each site tropospheric delay corrections were determined every 5 min, resulting in 24 parameters per site throughout the 2-h observation span. It should be noted that in practice the primary reference station should be situated in, or close to, the center of the SAR image frame in order to keep the baseline lengths to a minimum. In this analysis, however, the results obtained over longer baselines are also of interest.

Single-differenced tropospheric corrections (Eq. 3) were determined by forming the differences relative to S002. These corrections range from −9.5 to +4.2 cm, showing variations of up to a few centimeters within the 2-h observation span. Double-differenced tropospheric delay corrections were then obtained by forming the BE difference of the single-differenced values derived in the previous step (Eq. 5). The double-differenced corrections range from −6.7 to +10.9 cm, indicating significant changes in the tropospheric conditions (see also Fig. 8).

Interpolation of tropospheric delay corrections

For each of the eight prediction sites shown in Fig. 7, the tropospheric delay corrections were interpolated using the IDW interpolation method. Both the single-differenced tropospheric corrections relative to S002 for days 168 and 203, and the double-differenced tropospheric corrections between these two epochs were investigated by comparing the interpolated values to the "true" values obtained directly using the Bernese software. This was done for each of the 24 5-min time intervals within the 2-h observation span.

As an example, Fig. 8 shows two 3D interpolation "maps" obtained in the double-differenced case, for the 11th and 24th time interval (65 min apart). The spatial and temporal variability of the troposphere can easily be recognized. Similar "maps" can be generated for the single-differenced case and distributed as a routine CGPS product, with minimum effort, to aid radar interferometry.

Figure 9 shows the double-differenced corrections for the eight prediction sites, obtained for each of the 24 time intervals. The graphs show the parameters determined by the Bernese software, the interpolated values using the IDW method, and the differences between the two. It can be seen that the interpolation results agree very well with the "true" values. The standard deviations of the differences are all (with one exception) at the sub-centimeter level, even for baselines of 85 km in length (Table 1).

Conclusions

Tropospheric heterogeneity (differential tropospheric delay) can lead to misinterpretation of InSAR results. A BS and BE double-differencing algorithm to derive tropo-

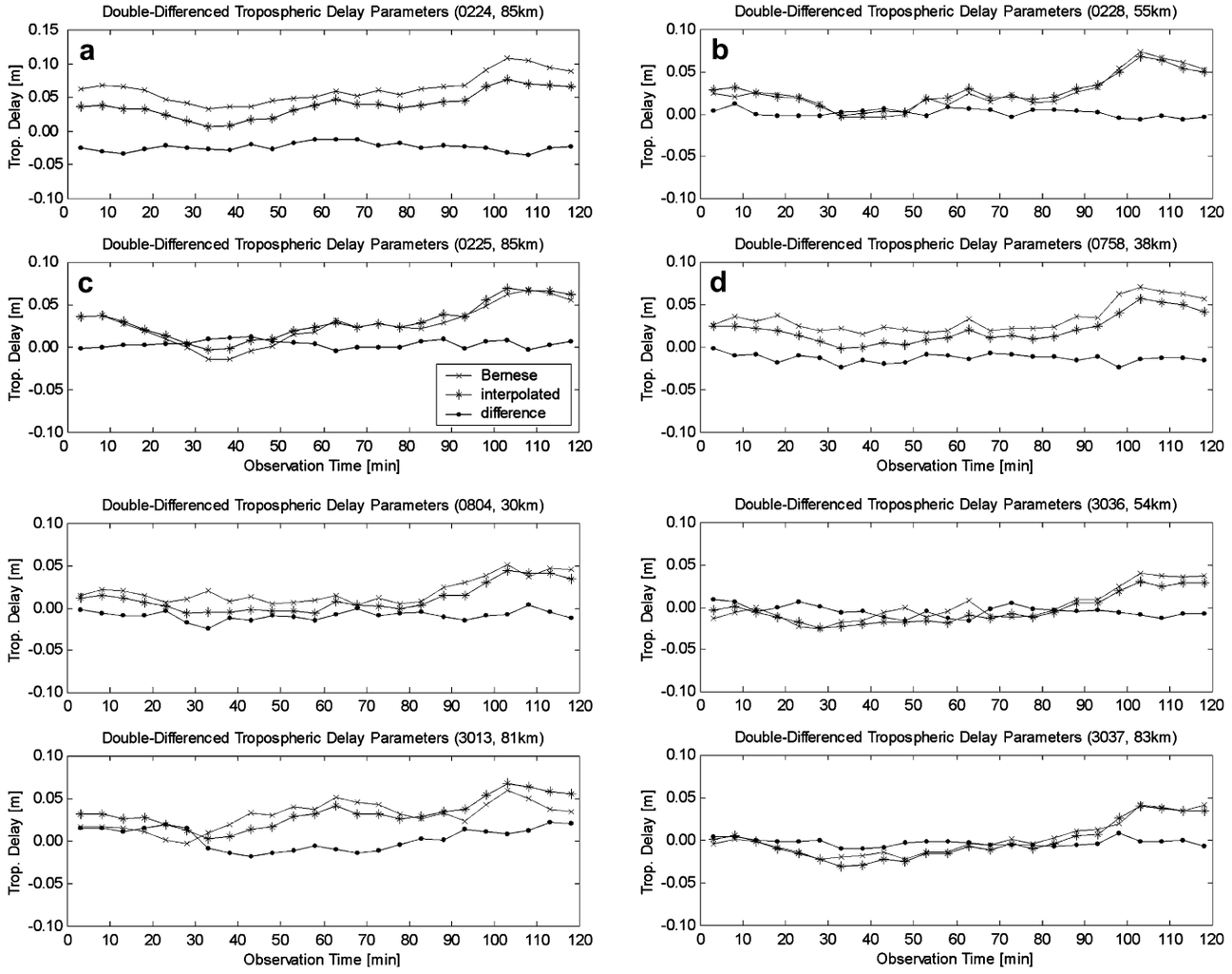

Fig. 9
Comparison of Benese-derived and interpolated double-differenced tropospheric corrections

Table 1
Standard deviations of the differences between Bernese-derived and interpolated troposphere corrections

Site	STD (m)	Baseline length (km)
0224	0.00625	85
0225	0.00445	85
0228	0.00472	55
0758	0.00510	38
0804	0.00597	30
3013	0.01323	81
3036	0.00697	54
3037	0.00450	83

spheric corrections to radar results from GPS observations has been tested. These GPS measurements can be collected by either a network of CGPS stations or GPS campaigns synchronized to the radar satellite flyover. In order to correct the radar result on a pixel-by-pixel basis, the GPS-derived corrections have to be interpolated. Three interpolation methods, namely the IDW, spline, and kriging techniques, have been investigated. Using GPS data from two test networks, it has been found that the IDW and kriging interpolation methods are more suitable. Differential corrections as much as several centimeters may have to be applied in order to ensure sub-centimeter accuracy for the radar result. It seems optimal to estimate the tropospheric delay from GPS data at 5-min intervals. How many GPS sites are actually required to achieve a given level of accuracy for the corrections depends on the tropospheric conditions present in the area. This includes the geographic location, the extend of the area under investigation, and the height differences between the GPS network sites. These factors determine the maximum distance between the GPS stations that would still allow the troposphere to be adequately modeled. Obviously, the reference sites should be evenly distributed across the area to maximize the quality of the interpolation results. The algorithm and procedures described in this paper could easily be implemented in a CGPS network data center. The interpolated grid of BS, single-differenced tropospheric delays can be generated as a routine product to assist radar interferometry, in a manner similar to the SLC radar images.

Acknowledgements SCIGN and its sponsors, the W.M. Keck Foundation, NASA, NSF, USGS and SCEC, as well as Japan's Geographical Survey Institute are acknowledged for providing the GPS data used in this study. Mr. Yufei Wang is thanked for his assistance with the interpolation algorithms.

References

Bauersima I (1983) NAVSTAR/global positioning system (GPS) II: radiointerferometrische satellitenbeobachtungen, vol 10. Mitteilungen der Satellitenbeobachtungsstation Zimmerwald, Berne

Black HD (1978) An easily implemented algorithm for the tropospheric range correction. J Geophys Res 83(B4):1,825–1,828

Bock Y, Williams S (1997) Integrated satellite interferometry in Southern California. Eos 78(29):293

De Boor C (1978) A practical guide to splines. Springer, Berlin Heidelberg New York

Essen L, Froome KD (1951) The refractive indices and dielectric constants of air and its principal constituents at 24,000 mc/s. Proc Phys Soc 64(B):862–875

Ge L (2000) Development and testing of augmentations of continuously-operating GPS networks to improve their spatial and temporal resolution. UNISURV S-63, The University of New South Wales, Australia

Ge L, Ishikawa Y, Fujiwara S, Miyazaki S, Qiao X, Li X, Yuan X, Chen W, Wang J (1997) The integration of InSAR and CGPS: a solution to efficient deformation monitoring. In: Wuhan PR (ed) Proceedings of the international symposium on current crustal movement and hazard reduction in East Asia and South-East Asia, China, 4–7 November 1997, pp 145–155

Graham LC (1974) Synthetic interferometer radar for topographic mapping. Proc IEEE 62(6):763–768

GSI (2003) http://mekira.gsi.go.jp/ENGLISH/index.html

Hanssen RF (2001) Radar interferometry: data interpretation and error analysis. Kluwer, Dordrecht

Hanssen RF, Weckwerth TM, Zebker HA, Klees R (1999) High-resolution water vapor mapping from interferometric radar measurements. Science 283:1,295–1,297

Hopfield HS (1969) Two-quartic tropospheric refractivity profile for correcting satellite data. J Geophys Res 74(18):4,487–4,499

Lancaster P, Salkauskas K (1986) Curve and surface fitting: an introduction. Academic, London

Lu Z, Fatland R, Wyss M, Li S, Eichelberger J, Dean K, Freymueller J (1997) Deformation of new trident volcano measured by ERS-1 SAR interferometry, Katmai National Park, Alaska. Geophys Res Lett 24(6):695–698

Massonnet D, Rossi M, Carmona C, Adragna F, Peltzer G, Feigl K, Rabaute T (1993) The displacement field of the Landers earthquake mapped by radar interferometry. Nature 364:138–142

Mendes VB (1999) Modeling the neutral-atmosphere propagation delay in radiometric space techniques. Department of Geodesy and Geomatics Engineering, Technical Report No. 199, University of New Brunswick, Fredericton

Rothacher M, Mervart L (eds) (1996) Bernese GPS Software Version 4.0. Astronomical Institute, University of Berne, Switzerland

Saastamoinen J (1973) Contributions to the theory of atmospheric refraction. Bull Géodésique 107:13–34

Schultz MH (1973) Spline analysis. Prentice-Hall, Englewood Cliffs

SCIGN (2003) http://www.scign.org/

SOPAC (2003) http://sopac.ucsd.edu/cgi-bin/SCOUT.cgi

Spilker JJ (1996) Tropospheric effects on GPS. In: Parkinson BW, Spilker JJ (eds) Global positioning system: theory and applications I, vol 163. American Institute of Aeronautics and Astronautics, Washington, pp 517–546

Stein ML (1999) Interpolation of spatial data: some theory for kriging. Springer, Berlin Heidelberg New York

Zebker HA, Rosen PA, Hensley S (1997) Atmospheric effects in interferometric synthetic aperture radar surface deformation and topographic maps. J Geophys Res 102(B4):7,547–7,563